PREFERENCES, WELL-BEING, and VALUE: 
A Critique of the Normative Foundations of the Economic Approach to 
Valuing Non-Market Goods

Ty Raterman

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill 
in partial fulfillment of the requirements for the degree of Doctor of Philosophy 
in the Department of Philosophy.

Chapel Hill
2006

Approved by:

Advisor: Douglas MacLean
Reader: Richard N. L. Andrews
Reader: Thomas E. Hill, Jr.
Reader: C. D. C. Reeve
Reader: Susan Wolf
ABSTRACT

(Under the direction of Douglas MacLean)

This dissertation assesses the theory of value underpinning cost-benefit analysis. This theory of value consists of two main claims. The first is that the satisfaction of individual preferences is what counts as a benefit and their frustration is what counts as a cost. This is standardly purported to be the case because satisfying individual preferences promotes individual well-being and, in turn, social welfare. I argue that there is a close connection between well-being and preference satisfaction: an individual’s well-being consists in the satisfaction of her fully-informed and rational preferences. However, I maintain that there may be occasions where a person’s actual preferences would, if satisfied, diminish her well-being and where the state is a better judge than the person herself of what would enhance her well-being. On such occasions, the state should strive to make people informed and rational (in the relevant respects), and then after doing so, attend to and promote the satisfaction of the preferences they then have. This has the effect enhancing individual well-being, and does so in a way that respects individual autonomy. This is fundamentally compatible with CBA.

The second claim is that all costs and benefits – including human lives lost or preserved, as well as changes affecting human health and comfort, the cleanliness of air and water, the welfare of non-human species, and ecological services rendered – can be expressed monetarily. I maintain that an individual’s willingness to pay or to accept a certain amount of money for a good does not suffice to show that this good’s value can be
monetized. With respect to certain goods, an individual can have preferences that differ from
one another along multiple irreducible dimensions. Different preferences will sometimes
incorporate fundamentally distinct attitudes such as appreciation, fascination, respect,
agitation, disgust, and shame. Thus, there will be pairs of goods of which it can be said
neither that one member is simply more valuable than the other nor that the members are
equal in value. This means that rational preferences cannot generate a single ordering of
value, monetary or otherwise.
ACKNOWLEDGMENTS

This project would never have been completed without significant help from a number of people. I am pleased to have the opportunity to acknowledge some of my debts. I only wish I could so easily repay them.

Professor Douglas MacLean, my advisor throughout the project, gave very generously of his time in meeting with me as well as reading and offering feedback on multiple drafts of all of my written work during the numerous years over which the project stretched. He had high but not impossible expectations, was encouraging and patient without being coddling or a pushover, and was firm but not dictatorial or pedantic. I mean it in the most complimentary way when I say that he always saw this dissertation as my project, not his. It is a cliché, but a fitting one in this case: where there are shortcomings in the dissertation, it is surely not because he did not reveal these to me or try to point me down the right path. Thank you very much, Doug.

I am also sincerely grateful to the other members of the dissertation committee, all of whom I know are very busy people but in spite of this spent considerable amounts of time reading drafts (of what is plainly a hefty document). More often than I am proud to admit, they were asked to do so without sufficient lead time. Professors Thomas Hill, Jr. and Susan Wolf both kindly made themselves available for numerous meetings with me where they shared truly insightful thoughts and helped massage the arguments into shape. Professors C. D. C. Reeve and Richard Andrews also offered genuinely valuable feedback that benefited
the dissertation significantly. It is a rare honor to be able to work with such distinguished scholars. That they are all such kind, thoughtful people is an added bonus!

The Department of Philosophy at the University of North Carolina at Chapel Hill was an extraordinary place to work. I benefited from its remarkable financial support. I am especially grateful to Professor Geoffrey Sayre-McCord, a tireless and thoughtful department chair who found a way to give me financial assistance in this final year that I required in order to finish the project. I would like to thank all those professors, graduate students, and even undergraduates who expressed interest in my work, made a suggestion, or fielded a question.

My parents, Meg and John Raterman, as well as my parents-in-law, Barbara and Sheldon Greenspan, have always had faith in me. Their support has been unwavering. These four people have always made it clear to me that they would help me in any way I needed, and that was a tremendous comfort. I am truly fortunate.

Finally, and (if I may say so) most importantly, thanks are due to my wife, Melanie. I have not always been a great deal of fun during this process. I do not want to paint too dreary a picture, but working on the dissertation has more than occasionally been frustrating, stressful, and tedious. I stayed in many days and nights when she would have liked to be out with me seeing the world! But Melanie always believed in my abilities and the importance of my work. Frequently, in fact, she believed more strongly than I did. She understood and supported my desire to complete this degree. She has kept my spirits up and encouraged me to keep moving forward. She truly worked a “second shift” while I sat in front of the computer typing away. I am forever grateful.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: INTRODUCTION</td>
<td></td>
</tr>
<tr>
<td>1: Overview and Thesis of This Project</td>
<td>1</td>
</tr>
<tr>
<td>2: A Basic Picture of Cost-Benefit Analysis</td>
<td>14</td>
</tr>
<tr>
<td>3: Theory of Right Action Underpinning Cost-Benefit Analysis</td>
<td>33</td>
</tr>
<tr>
<td>2: ON BENEFITS, PREFERENCES, AND WELL-BEING</td>
<td></td>
</tr>
<tr>
<td>1: Introduction</td>
<td>39</td>
</tr>
<tr>
<td>2: Why, According to Proponents of CBA, Should the State Aim to Satisfy Individual Preferences?</td>
<td>41</td>
</tr>
<tr>
<td>3: Preference Satisfaction Versus Well-Being, and Other Important Distinctions</td>
<td>56</td>
</tr>
<tr>
<td>4: Flawed Accounts of Well-Being</td>
<td>62</td>
</tr>
<tr>
<td>5: The Informed-Preference Account of Well-Being</td>
<td>71</td>
</tr>
<tr>
<td>3: INDIVIDUAL PREFERENCES AND THE ULTIMATE AIM OF PUBLIC POLICY DECISIONS</td>
<td></td>
</tr>
<tr>
<td>1: Introduction</td>
<td>94</td>
</tr>
<tr>
<td>2: Each Individual Is the Best Judge of Her Own Well-Being: Four Interpretations</td>
<td>99</td>
</tr>
<tr>
<td>3: The Role of Individual Preferences When Benefits are Ultimately Identified with Gains in Well-Being</td>
<td>113</td>
</tr>
</tbody>
</table>
4: PROVIDING INFORMATION AND ASSISTING IN ITS PROCESSING: THEORY AND PRACTICE

1: Introduction........................................................................................................ 141
2: One Basic, Concrete Example........................................................................... 142
3: Transformative Value and Existence Value..................................................... 152
4: More on Providing Information and Helping to Ensure Rationality............. 164
5: Responses to Concerns About Practicability.................................................. 178

5: WHY CITIZEN PREFERENCES AND SOCIAL VALUES DO NOT THREATEN COST-BENEFIT ANALYSIS

1: Introduction........................................................................................................ 189
2: Citizens Versus Consumers............................................................................. 192
3: A Comparison of Voting and Cost-Benefit Analysis..................................... 208
4: Individual Preferences Versus Social Values................................................ 224

6: A PLAUSIBLE META-ETHICAL UNDERPINNING FOR NON-MARKET VALUATION

1: Introduction........................................................................................................ 231
2: Unsuccessful Criticisms of the Meta-Ethics of Economic Valuation.......... 234
3: Why Objectivism and Intersubjectivism Should Be Rejected.................... 248
4: Why a Limited Subjectivism Is Not Self-Defeating.................................... 261

7: ANIMALS, PREFERENCES, AND VALUE

1: Introduction........................................................................................................ 266
2: On Animals’ Preferences and Well-Being....................................................... 269
3: Taking Animal Preferences to Count in Their Own Right............................. 279
4: The Policy Implications of Counting Animal Preferences......................... 288
8: WHY NOT ALL COSTS AND BENEFITS CAN BE MONETIZED (PART 1)

1: Introduction.................................................................305

2: Flawed Arguments Against Monetizing Costs and Benefits...........308

3: Well-Established Problems for Monetizing Costs and Benefits........317

9: WHY NOT ALL COSTS AND BENEFITS CAN BE MONETIZED (PART 2)

1: Introduction.................................................................327

2: The Substitutability of Money for Other Goods........................328

3: Substitution, Preferences, and Commensurability......................338

4: Incommensurability, Choice, and Practical Reason....................353

BIBLIOGRAPHY........................................................................364
CHAPTER 1: INTRODUCTION

Section 1: Overview and Thesis of This Project

Cost-benefit analysis is a tool employed frequently by public decision makers in order to evaluate public policy proposals, especially those concerning regulations and investment projects. Its proponents believe its range of applicability to be extremely wide. Cost-benefit analysis (henceforth, CBA) has been used to evaluate policies and projects related to:

- Roadways, including where to build them, when, and how wide, when to repair them, whether to build a bridge or a tunnel, and how big these should be;
- Waterways, including whether to dam them, whether to irrigate from them, and at what volume, whether to dredge or trench them, how extensively they can be fished, and whether receding beaches should be reclaimed;
- Utilities infrastructure, viz., water, sewage, electricity, and phone, television, and internet services, including where access to these will exist, and at what rates;
- Pollution of air, land, and water, including how much of it factories and individuals should be able to produce, and when, where, and to what extent it should be cleaned up;
- Waste disposal, including whether to build landfills or incinerators, and where to locate these, as well as where and how to store nuclear and toxic waste;
- Public lands, including whether to purchase them in the first place, and at what price, whether, how often, and how extensively they can be logged, mined, or
tapped for oil, whether and how many off-road recreational vehicles are permitted, and how animals on these lands are treated.

The list could go on and on. If CBA truly is applicable in all of these cases, and really does, when well done, provide decision makers with critical information, then it is an incredibly valuable tool.

I have my doubts, and this project is devoted to exploring some of these doubts.

In Section 2 of this chapter, I provide a description of CBA that, despite being highly abbreviated and non-technical, aims to provide a reasonably clear understanding of CBA’s basic purpose, methods, and putative justification. Prior to drawing this picture, a rough sketch of CBA can be offered. In short, the proponent of an ambitious form of CBA endorses the following three basic claims:

1. A policy is justified if it results in more benefits than costs to the affected individuals, with the gainers being able to compensate the losers such that the losers are no worse off, but the gainers still better off, than they were originally. It is the mere potential for compensation that justifies a policy; actual compensation from the gainers to the losers is not required.

2. Whether a change counts as a benefit or as a cost to an individual depends on the individual’s assessment of the change. If the change satisfies an individual’s preference, then the change counts as a benefit for that individual. If the change frustrates the individual’s preference, it counts as a cost for the individual.

3. Any benefit or cost can be accurately expressed in monetary terms. (Indeed, all costs and benefits much be expressed in a common unit in order to be weighed against one another rationally.)
I have never seen CBA distilled down in quite this way. Presenting it as consisting of three basic claims, and these three in particular, is my doing. And yet the three claims taken together do capture the core elements – do give a generally accurate, if highly simplified, picture – of CBA. As I say, however, it is a picture of an ambitious form of CBA. By that I mean that it makes claims that are bolder than at least some people who count themselves as proponents of CBA would want to endorse. For example, some proponents are quick to concede that some values cannot be represented in the cost-benefit calculus. These people say that CBA tells us something very important, but that it does not by itself – even when done perfectly – tell us everything that making the appropriate policy decision demands we know. When this kind of claim is not based merely in logistic considerations, but in some more principled consideration, this person is not a proponent of an ambitious form of CBA. Those who endorse an ambitious form of CBA do not take there to be any principled reason why a CBA cannot – at least if it gets over logistical obstacles – accurately capture and weight all pertinent values and straightforwardly generate the correct policy recommendation. Some less ambitious form of CBA may in fact be the tenable version, but typically it feels as if those who back away from the ambitious form do so too quickly. Thus, this project looks at the ambitious form in order to see if it is in fact credible, for if it is, this is surely important to know, and if it is not, we can see exactly why it is not, what the more reasonable, less ambitious form looks like, and why. Given the important place that CBA has occupied and continues to occupy in policy analysis at all levels of the government, all three claims deserve careful scrutiny. Despite many appraisals by scholars from an array of disciplines, CBA itself still stands in need of analysis. Philosophers have long held views and engaged in discussions that are relevant to CBA. It is relatively infrequently, though,
that philosophers openly and directly engage with theorists from economics and public policy in a discussion of the merits and demerits of CBA. I aim in this project to contribute to the discussion a number of criticisms, as well as some important defenses, of CBA.

Philosophers will regularly take a moral theory to be comprised by (at least) a theory of the right (i.e., a theory of “right action”) and a theory of the good (i.e., a theory of value). It seems to me that we can do precisely the same thing with CBA, and that doing so will help us to organize our inquiry and precede more clearheadedly with our analysis into CBA’s merits and demerits. Claim 1 is essentially CBA’s theory of right action. Claims 2 and 3 together roughly comprise CBA’s theory of value. There is an enormous amount to say about CBA’s theory of right action, but very little of that is included in this project. Instead, my purpose here is to explore CBA’s theory of value. My more specific interest – as evinced by the title of this project – is in appraising the normative foundations of the economic approach to valuing non-market goods. The term “normative foundations” is used in order to differentiate this project from one that considers just technical problems with carrying out the non-market valuational procedures economists recommend. The appraisal is delivered in two basic parts, the first dealing with Claim 2, of which I defend a version, and the second concerning Claim 3, which I oppose. It could thus be said that I endorse part, but not all, of CBA’s theory of value. As I will explain at length in the pages that follow, I want to defend a version of Claim 2 that is not exactly the one the typical economist seems to endorse, but that is nonetheless reasonably and importantly similar to that while also being much more plausible. I hold on to a connection between policy-relevant gains and the satisfaction of individual preferences, arguing that – subject to some very important qualifications – the economic approach is right to count the satisfaction of individual preferences as benefits.
This argument segues into a general discussion about the nature of value, and more specifically about the relationship between the value of a good or outcome and people’s preferences regarding that good or outcome. I maintain that what has value, and how much, is indeed importantly connected to individuals’ preferences; and for this reason non-market valuational techniques are fundamentally on the right track in aiming to gauge people’s preferences regarding various goods and outcomes. However, this approach to valuing non-market goods goes wrong in maintaining that all gains and losses are commensurable. They cannot all be measured on a single scale, monetary or otherwise. I argue that value is not monistic, and this does not rest on the empirical claim that there are certain goods that individuals will not voluntarily trade for any other good. If this argument is successful, it challenges the project of monetizing the value of all of the various policy-induced gains and losses; but I do not believe that commensurability among costs and benefits is required in order to make coherent, justifiable policy decisions.

This project engages with a variety of issues central to moral philosophy, political philosophy, environmental philosophy, and public policy. It explores what certain abstract philosophical notions look like when they are viewed in a very practical context. This is a difficult exercise, for philosophical discussions change when they are forced to reckon with the constraints of practical-minded public policy. My interest in pursuing this project is partly rooted in a general practical concern to see good environmental policy be implemented and in a commitment to contributing in at least a small way to such policy. But I am also extremely interested in many of the theoretical issues that connect up with environmental policy. I have probably been drawn to this particular project primarily by an interest in the question of whether the value of any good – and specifically of any non-market good – can
be accurately expressed in monetary units. Among the non-market goods I have in mind here are: human life, health, and comfort; the lives of individual animals and certain individual non-animal living things (a 3,000 year old giant sequoia, for example); the existence of an entire species; clean air, water, and soil; wilderness and other undeveloped tracts of land; and so on. I cannot help but believe that the question of whether all value can be monetized is among the most important in the field of environmental ethics. It would seem to matter not just as a theoretical issue, but as a practical one too. Which environmental policies are approved, and indeed how policy analysis fundamentally operates in general, would seem to turn on the issue of whether all of the gains and losses involved with a particular regulation or investment decision can (at least in theory) be monetized. But I will admit now that what exactly the implications are of the falsity of Claim 3 is something that will not be fully dealt with in this project. This is regrettable, but the fact is that there is only so much time one is permitted for a project of this nature.

One of the general concepts at issue here will be the commensurability of value, which I believe is an underexplored one in philosophy (if any topic can genuinely count as underexplored at this point). I take commensurability to be the concept at the heart of Claim 3. It is important to note that the three claims comprising CBA all stand or fall independent of one another. With respect to all three of the claims, showing one of them to be true does not prove anything about the truth or falsity of the other two. The same goes for showing one of the claims is false. While Claim 3 is logically independent of Claim 2, it is nonetheless the case that how one criticizes or defends Claim 3 – in fact, how one even understands Claim 3 – will turn on how one assesses and understands Claim 2. For example, how one addresses the issue of commensurability will depend on whether one is a
subjectivist or objectivist about value. No particular view on whether all value is commensurable is immediately and obviously entailed by any particular conception of value, but it will be difficult to make an argument one way or the other about commensurability without first defending a particular conception of value. To put the point another way: an answer to the question of whether the satisfaction of an individual preference counts as a benefit will leave open the question of whether each cost and benefit can be expressed monetarily. If there is a connection between satisfied preferences and policy-level benefits, then we need to go on to ask whether each preference can be monetized; but if there is no such connection, then it will not be nearly as imperative to discuss whether each preference can be monetized. I engage with the issue of the commensurability of preferences, and do so after – and precisely because – I defend a connection between satisfied preferences and policy-level benefits (though a connection that is looser than treating the two as strictly identical).

In assessing Claim 2, I frequently use Mark Sagoff’s views as a foil. I tend on the whole to oppose Sagoff, for my part arguing that Claim 2 is much more reasonable than Sagoff believes, though it can nonetheless undoubtedly be clarified and improved upon. The criticisms that Sagoff makes are generally not ones that he alone has expressed; and I do not discuss Sagoff alone. Nonetheless, I focus on Sagoff’s work to the extent that I do because of its influence in the philosophical literature on economic valuation over the past quarter of a century, and also because it is philosophically rich. While I find his work to be remarkably engaging, I also take it to be full of flaws. Focusing on Sagoff’s arguments and diagnosing their shortcomings brings me into contact with many of the important philosophical topics
and questions connected to CBA and economic valuation of environmental goods, and indeed with large and important issues central to moral and political philosophy.

In Chapter 2, I begin exploring the question of what role individual preferences ought to play in shaping public policy. The general question is: What counts as a cost, what counts as a benefit, and why? I explore the relationship between benefits, preferences, and well-being. Economists often maintain that the satisfaction of a preference counts as a benefit because an individual’s well-being, and in turn social welfare, are enhanced when the individual’s preference is satisfied. Each individual, it is claimed, is the best judge of her own well-being. In the process of assessing this view, I defend a detailed account of individual well-being – one I call an informed-preference account – according to which an individual’s well-being consists in the satisfaction of the preferences she would have if she were fully informed and rational. This kind of account is by no means brand new, but I add my own wrinkles to it in order to make it as plausible as it can be, which is, I claim, more plausible than any of its competitors.

In Chapter 3, I assess the implications of my informed-preference account for CBA. While it initially seems congenial to CBA, given that it does maintain an important connection between well-being and preferences, I argue that the consequences of accepting this account include having to acknowledge that it is not always true that each individual is the best judge of her well-being and that it is not always true that the best way to promote someone’s well-being is by satisfying her preferences. Some preferences are such that their satisfaction diminishes a person’s well-being, and the state may sometimes be in a better position than a particular individual herself to know this. Insofar as benefits (to an individual) are ultimately connected with the promotion (of that individual’s) well-being,
there may thus be reason to count the satisfaction of a preference as a cost and the frustration of a preference as a benefit. However, I argue that the ultimate identification of benefit simply with the promotion of well-being is an oversimplification. The state’s aim through its public policy decisions is surely partly to promote well-being, but also to respect individual autonomy and to promote experiences of satisfaction, neither of which is strictly entailed by promoting well-being, and both of which may provide some reason for identifying benefits with the satisfaction even of welfare-reducing preferences. We should see the state as having multiple ultimate aims in making policy decisions, and these may in fact compete with one another. I then note that the state can simultaneously achieve all of these aims by providing people with policy-relevant information and helping them to organize and process it, and then after doing so promoting people’s now informed and rational preferences. This is precisely what I recommend that it does. Doing so has the effect of helping to ensure that in promoting the satisfaction of people’s actual preferences, the state is thereby increasing people’s well-being, and this is important. But in addition, it helps to ensure that when a person’s well-being is promoted, it is promoted in a way that respects the person’s autonomy and the proper boundaries of the citizen/state relationship (which is different from at least certain other types of relationships).

In Chapter 4, I further explain and defend my view on the role that rational and informed preferences should play in public policy decisions. I begin by focusing on one fictional but realistic scenario in which a CBA might be performed and consider the related questions of whose preferences ought to be gauged and what information ought to be provided to these individuals. I discuss complications for CBA based first in the fact that certain policy decisions preclude the possibility of people having particular experiences and
thus developing preferences in the future, and also in the fact that individual preferences are wide ranging and can extend to things that the individual does not use or benefit from in any direct, physical sense. In the literature these are known as problems concerning transformative value and existence value. I then discuss what the provision of information requires in cases where, for example, there is scientific uncertainty regarding the facts; and my suggestion is that something akin to the model of journalistic neutrality be used. I explore the role the state has in helping people to understand and process the information they are receiving, and I discuss the role that public debate has in this, as well as in the very provision of information. Finally, I respond to concerns that my suggestion is such as to overwhelm analysts and citizens. I point out, among other things, that I am not committed to the view that individuals’ preferences must be sought out in each and every policy case. Sometimes the well-informed and rational preference that is sought out is the preference regarding who individuals would like to elect to make public decisions on their behalf.

Chapter 5 continues to explore the role that individual preferences ought to play in public policy decisions by explaining and then rejecting one particularly popular criticism of CBA, namely that there are different kinds of preferences, namely consumer preferences and citizen preferences, and that CBA captures only the former. It inappropriately treats people just as consumers, while voting, which is on this view preferable as a public-decision mechanism, rightly engages people in their role as citizens. Against this criticism, I argue that the examples that support the claim that there are these two kinds of preferences can be better explained by appeal to the idea of a Prisoners’ Dilemma than to the different types of preferences. I contend that CBA and voting are both, at base, ways of registering and making policy responsive to people’s preferences. I do, however, readily concede that CBA
asks that all preferences be monetized, while voting does not. This is a serious difference, and one I turn my attention back to later in the project (specifically in Chapters 8 and 9). But it is, I argue, a problem for Claim 3, not Claim 2. It does not show that there are citizen and consumer preferences. In connection with this, I also consider and reject an argument that maintains that an approach like CBA that makes policies in response to individual preferences cannot make room for social values.

In Chapter 6, I tackle some concerns about the meta-ethical commitments of CBA. I slide from working to articulate a conception of costs and benefits to trying to give an account of the value of goods. The account I offer is, I say, quasi-subjectivist and piggybacks on the account of well-being from Chapter 2 and on the arguments in Chapters 3, 4, and 5 for aiming to promote well-informed, rational preferences. The view is roughly that the value of a good (including non-market goods), relative to an individual, is identified with and fixed by the informed, rational preferences that individual has regarding the good. I defend this view against several forms of objectivism about value (arguing that objectivism regarding value has empirical and ontological problems that render it unworkable, at least at the level of public policy), as well as against naïve forms of subjectivism. Although I suspect that mine is not as purely subjectivist an account as most economists would want, I maintain it is one CBA can and should accommodate. In the process, I consider the well-known “last-man” argument, where I concede that were there no individuals left to have preferences, there would be nothing of value.

This issue segues into Chapter 7, where I discuss animals’ preferences. I argue that it is not just human preferences, but the preferences had by any kind of creature capable of having them, that fixes the value of a thing. There are good reasons for thinking that at least
some animals have preferences, and that where these can be compared to humans’ preferences, there are good reasons for thinking that even if no humans have preferences regarding them, the satisfaction of an animal’s preference should count as a benefit that gets factored into the cost-benefit calculus (and that the frustration of such a preference should get factored in as a cost). This is recommended by some of the same considerations that were at work when we were talking about humans, namely the importance of promoting well-being and promoting experiences of satisfaction. It is in this chapter that I begin to touch on Claim 3, which has it that all benefits and costs can be monetized. I suggest that at least some animal preferences and at least some human preferences can be reasonably compared (though not all of them can), though it will be extremely difficult to monetize animal preferences.

With the general preference-based account of costs and benefits in place, I then proceed in Chapter 8 more squarely to confront Claim 3, i.e., the view that all costs and benefits can be expressed monetarily. I begin by considering objections to Claim 3 that I believe are not successful, but then quickly move on to others that are surely more potent. The arguments here are generally not new, but it seems appropriate to give them airtime in this project. In short, conceptual problems plague efforts to monetize an individual’s preferences by gauging her willingness to pay or her willingness to accept. In order to measure a preference using willingness to pay, we are required to suppose, implausibly, that someone who has nothing with which to pay has no preferences. The willingness to accept measure does not require quite the same supposition, but still ties what preferences one expresses to one’s financial condition in an implausible way. Both these “measures” of a preference fail to take significantly seriously the fact that a dollar means different things to different people, and so will tend to exaggerate how beneficial or costly a change is for those
who are wealthy; and as a result, both fail to accommodate the “one person, one vote” principle and in turn the demands of democracy.

Chapter 9 continues the critique of Claim 3, and moves into somewhat more novel territory. I concede the empirical claim that there is no good that people would not rationally give up (or at least give up some of) in exchange for at least some amount of money. Still, I maintain, this does not mean that all benefits and costs can be expressed monetarily; an individual’s willingness to pay or to accept a certain amount of money for a good does not suffice to show that this good’s value can be monetized. What one prefers is certainly connected to what one will trade for what, but one’s preference does not simply consist in trading behavior, or any kind of behavior for that matter. Preferences consist in part of evaluative attitudes, and we are capable of irreducibly distinct attitudes such as admiration, fascination, respect, agitation, disgust, and shame. Because these attitudes differ from one another along multiple irreducible dimensions, so too will the preferences they comprise. This forms the basis of an argument for incommensurability that is inimical to CBA. CBA’s call for expressions of willingness to pay, I argue, inappropriately constrains our ability to express a full range of distinct kinds of evaluative attitudes that constitute our preferences. None of this amounts to a proof that policy decisions should not be made by appeal to willingness to pay and willingness to accept, but it does mean that none of the things I said in earlier chapter in defense of the connection between benefits and preferences will justify the continued use of these “measures.”
Section 2: A Basic Picture of Cost-Benefit Analysis

Proponents of CBA claim that it is a policy tool that enables us to make decisions on the model of the market on occasions where, for one reason or another, the market fails. Perfect markets, we are told, guide us toward the most highly valued outcomes. As we know, however, markets are generally not perfectly functioning. Not all information is free and perfect, not everyone is a price taker, not all decisions are private (which means that there are externalities), and not all goods are private. Market failures result in inefficiencies. Goods and services are misallocated. Among the primary aims of public policy, it is often said, is to guide individual decisions so that the overall outcomes of these decisions avoid the misallocations and inefficiencies. In other words, public policies ought to help deliver the distribution of goods that would have resulted if the market failures had not interfered, or at least a distribution that closely resembles this. Some will say, in fact, that the entire purpose of the state is to correct for market failures (which may include enforcing contracts in cases where considerations of self-interest are not sufficient to ensure that each party upholds its end).

Under conditions of a perfect market, a transaction between two people will take place, i.e., goods will change hands, services will be rendered, etc., only if both people find the transaction to be beneficial and as a result consent to it. We say that a distribution (of goods, services, etc.) is Pareto optimal if no alternative arrangement of these things could make one person better off without making another worse off. If Abe would gladly trade his necktie for Bob’s cufflinks and Bob would gladly trade his cufflinks for Abe’s necktie but each hangs on to his item, then the current distribution of goods is not Pareto optimal. Here,
swapping tie for cuff links would be preferred by both Abe and Bob, and thus would seem to make each better off. Doing so would represent a Pareto improvement. A Pareto improvement is achieved in those transactions where at least one party to the transaction is made better off, and no party to the transaction is made worse off, by the transaction. The market is supposed to work by a series of Pareto improvements. I am gladly willing to pay $3.00 for a fresh, tasty loaf of bread, and the baker is quite willing to give up his fresh, tasty loaves of bread for $1.50 each. There exists between me and the baker the opportunity for a Pareto improvement. But it is most convenient for me to get my bread, my milk, and my vegetables all in one stop, and the baker does not want to step away from his baking every time someone wants to buy a loaf of bread. The grocer recognizes this, and so serves as a middleman between me and the baker. The grocer is willing to pay the baker $1.50 for each loaf of bread, and to sell these loaves to me (and anyone else who wants one) for $3.00. Swapping the $1.50 for the bread is preferred by both the baker and the grocer, and swapping the $3.00 for the loaf of bread is preferred by both me and the grocer. The baker, the grocer, and I thus all seem to be better off after this set of transactions than any one of us was beforehand.

If a transaction would make at least one person better off and no one worse off, but the transaction does not occur, an inefficiency exists. An inefficiency, so described, is apparently in no one's interests. Economists view individuals as self-interested and with unlimited wants. More of a good is always better than less, though marginal utility diminishes, i.e., the utility of each unit acquired is less than the utility of the previously acquired unit. In a perfectly functioning market, these inefficiencies will not exist—not for long at least. Self-interested individuals will recognize the mutual advantage in carrying out
the transactions, and the swaps will take place. But this is not the only way in which a market failure generates an inefficiency. An inefficiency can also arise when transactions take place among people not all of whom are consenting. The results can be very disconcerting—for those who suffer harms without consenting or being compensated, of course, but also for economists considering such a situation.

To take a trite but informative example, imagine that a chemical company, ChemiCo, needs to dispose of its toxic chemical waste. This company could hire another company, DispoCo, whose business it is to provide environmentally safe storage or disposal of the chemicals. But this service, let us imagine, is expensive: DispoCo would charge $10,000.\(^1\) Alternatively, ChemiCo could dump the chemical waste in the river that flows nearby. There does not seem to be anyone (apart from the manager at ChemiCo responsible for such decisions) – and in the absence, of course, of public policies regulating this kind of thing – who would need to consent to the dumping: the river is not exactly owned by anyone, and so ChemiCo does not trespass or violate anyone’s property rights by dumping in it. The rationally self-interested ChemiCo prefers cheaper to more expensive disposal, and so opts to dump in the river. ChemiCo is now better off than it was with the dangerous chemicals in its hands, and better off than it would have been had it paid the hefty price to have the chemicals disposed of by DispoCo. But others are worse off than they were prior to ChemiCo’s dumping. The chemicals proceed to seriously poison many people downstream. The chemicals pollute the drinking water, and when water from the river downstream is diverted for irrigation, the chemicals pollute crops that people will soon be eating. Any uncompensated harm (in this case, the harm inflicted by ChemiCo on the people

---

\(^1\) You can imagine, if you like, that DispoCo is only one of several companies that provide this kind of safe disposal service, and that DispoCo’s charge for the service is comparable to its competition.
downstream) is called a negative externality. A benefit that someone receives but does not pay for (in this case, ChemiCo’s benefit of having unloaded a pollutant without either paying for its proper disposal or paying those downstream who are harmed by the dumping) is a positive externality. Externalities, negative or positive, are indicative of — indeed constitutive of — market failure.

ChemiCo’s decision to transfer the chemicals into the river would not have been such an easy one if everyone affected by the dumping had to consent to it before it could take place, or in other words, if all of the costs and benefits of dumping in the river were internalized. We were imagining that prior to dumping ChemiCo had the option of disposing of its chemicals by hiring DispoCo for $10,000. Given this option, ChemiCo would presumably not have been willing to offer more than $10,000 total to the people downstream who would be affected in order to get them to agree to the transfer into the river of chemicals that would come downstream to poison them. Let us just imagine that each of the 100 families downstream would have required at least $5,000 in order to consent to the dumping. (Perhaps the people downstream are all made very sick by the drinking water contaminated with toxic chemicals. It is not at all implausible to imagine medical bills of $5,000 per family.) ChemiCo, thus faced with a decision to pay $10,000 to DispoCo or $500,000 to the downstream families, would obviously have paid $10,000 to DispoCo and not dumped in the river.

Property rights, many economists will say, were not adequately assigned in this case, though there is a reason for this. The river seems by nature (as it were) to be a public good. While the banks of the river could be owned, the water in it just runs by, making it essentially impossible for private individuals to own the water. If someone owns less than all of the
water in the river, how do we keep track of which particular “bits” of water that person owns? Even if she owns all the water in the river, what happens when new water comes in from a stream, or when water leaves when the river dumps into a lake? What happens when water is added by rain? Does she own any of that added water too? What about when the water is dispersed on those occasions when the river floods its banks, or when water evaporates from the river? Moreover, it is difficult to exclude people from using the water in the river. It is difficult to prevent people’s actions upstream from affecting the water. After a hard rain, a local farmer’s pesticides run into the river, and mud from a construction site silts up the river. If I cannot prevent these people from using (and harming) the river, then I have little reason to pay to acquire a property right in the river. In a perfect market, there are no public goods. Because there are many rivers – and indeed many other goods that are like rivers in the relevant respects – the market is highly imperfect.

Again, the economists’ claim is that market failures, if not properly managed, often result in outcomes that differ from, and are inferior to, those that perfectly functioning markets would deliver. This is where government must get involved, though (for all the reasons mentioned earlier) it will likely not be able to assign or enforce property rights in the river (or other similar goods). What criterion or criteria should be employed in order to arbitrate between the various possible government interventions? Should ChemiCo be permitted to dump as it pleases? Perhaps it is just bad luck for the people downstream that ChemiCo wants to dump. Perhaps the people downstream should simply either learn to deal with the pollution or move. Or maybe ChemiCo should be utterly and absolutely forbidden from dumping in the river. Perhaps they should be required to either pay the $10,000 to DispoCo or just hold on to those toxic chemicals. Perhaps a little dumping is okay, though
there is a limit on it. Perhaps the people downstream would consent to a little dumping in exchange for a certain amount of money. And perhaps this amount is one that ChemiCo would be willing to pay. Should ChemiCo be permitted to dump only if it actually approaches every individual (or family) that would be affected by the dumping and secures an agreement with them (that, e.g., those affected by the dumping will be paid a certain amount of money for tolerating a certain amount of dumping)? And then what if one family holds out, either driving an extremely hard bargain or altogether refusing to consent to the dumping, no matter what the amount of compensatory payment? Resolving these problems, answering these questions, is among the tasks of public policy; and proponents of CBA recommend that policy makers employ CBA precisely to carry out this task.

Appealing to CBA is what we might call a quasi-market approach to making public investment decisions. For reasons that I have identified, the market has failed to deliver the efficient outcome in the case of ChemiCo’s waste disposal, and the cost-benefit analyst suggests that public decision makers try to mimic the market, i.e., to model policies after the market. If this is the aim, then one possible policy criterion is the Pareto Criterion, which says accept only those policies that benefit at least someone while harming no one. However, in virtue of the fact that many policy proposals are far reaching, there are very few policies – essentially none of those that really merits serious attention – that will satisfy the Pareto Criterion. A policy that would flatly forbid companies (in our county, say) from dumping in the river would, for example, seem to make those families living downstream better off, but would make ChemiCo worse off, than before the policy was implemented. This policy would thus not satisfy the Pareto Criterion. The general worry here is that the Pareto Criterion is simply too restrictive; it in effect locks us into the status quo. As the
economist A. Myrick Freeman says, this criterion is “a very stringent [one] in practice. There are very few policy proposals which do not impose some costs on some members of the society… The Pareto Criterion is not widely accepted by economists as a guide to policy. And it plays no role in what might be called ‘mainstream’ environmental economics.”

For this reason another criterion is appealed to. This is the so-called Kaldor-Hicks Criterion (KHC). The KHC says that a policy or project should be endorsed if and only if its benefits outweigh its costs by a greater margin than would be the case on any other policy/project, with the gainers being able to compensate the losers such that the losers are no worse off than they were originally and the gainers are still better off than they were originally. This means neither that ChemiCo should by necessity be altogether forbidden from dumping, nor that it should categorically be permitted to dump all that it wants. Rather, it means that ChemiCo should be able to dump up to, but not beyond, the point at which the benefits of dumping outweigh the costs, taking into account all people affected. If additional dumping would produce more costs (to the people downstream, apparently) than it would benefits (to ChemiCo), then to put any more chemicals into the river would be to over-pollute it. But if the costs of additional dumping would be exceeded by the benefits, then – in an odd-sounding but real sense – there is not enough pollution in the river. If a governmental regulation restricted dumping of these chemicals in such a way that a certain amount of dumping above the regulated limit would produce more benefits than costs, then the restriction is too tight. It is inefficient and wasteful. And of course if an amount of pollution right at (or anywhere under) the regulated limit produces more costs than benefits, then the restriction is inefficient and wasteful in virtue of its looseness.

For a slightly more detailed discussion of the KHC, including some common infelicities in stating the criterion, see Section 3 of this chapter. In any case, we should be clear on the fact that according to the KHC, actual compensation is not required; what is required is the mere possibility of compensation. For this reason, the KHC is sometimes called the Potential Compensation Criterion. Why supporters of the KHC do not maintain that compensation must actually be paid is something I will address later in this section. For now, note that if the compensation were paid out, then the state of affairs resulting from the policy would represent a Pareto improvement over the state of affairs prior to the implementation of the policy. In other words, if the compensation were paid out the Pareto Criterion would be met; and for this reason, the KHC is also sometimes called the Potential Pareto Criterion.

Importantly, the KHC does not entail any particular accounts of costs and benefits. It simply says (roughly) “Go ahead with a project when its benefits outweigh its costs.” So, what exactly counts as a gain, and what counts as a loss? I have already more than hinted at the standard economic answer to this question, but let me address it more directly here. Roughly put, gains are those outcomes that one prefers, and losses are those outcomes that one does not prefer. Gains and losses are relative to each individual affected; in theory, an outcome one person counts as a gain could be one that someone else counts as a loss. Among the roles of the policy analyst, then, will be discovering what outcomes are most strongly preferred, and conveying this information to public decision makers. The question of how economists define or understand the notion of a preference is a surprisingly difficult

---

3 Similarly, the Pareto Criterion on its own does not specify what counts as a benefit or as a cost. Rather, it simply says (in effect) go ahead with a project if it results in no costs but at least some benefits. To put the point in different terms, the Pareto Criterion makes reference to individuals being made better off and worse off but does not entail any particular account of “better off” or “worse off.”
one to answer. I have heard one well-known economist use the classic line: “I cannot define it, but I know one when I see it.” It is typical to hear economists described – in this context at least – as staunchly behaviorist. This is too strong a characterization. In fact, there are ways in which the economic conception of a preference is definitely not behaviorist.

On the standard economic view, to prefer something is not just to be put into any positive mental state by the thing. It is not just to have any positive attitude towards the thing, or to think consistently or exclusively positive thoughts about the thing. To this extent, the economic view is consistent with behaviorism. In fact, it is not just any kind of behavior, or even any kind of behavior that can be construed as positive, that will be identified with having a preference for something. It is not enough, in order to prefer something, that one speak positively about, or advocate for, the thing, though speaking and advocating surely count as kinds of behavior. Trading behavior is the relevant kind of behavior. But it is not the case that how much one prefers something is identified strictly with how much she actually does or does not trade for it. In fact, one’s preference for a thing is distinct from what she actually trades for it, and is instead identified with how much she is willing to trade for it. If I only have to give up one unit of good A in order to get good B, but would have been willing to trade five units of good A, my preference for good A is stronger than my actual trading behavior indicates. I resisted the characterization of the economic conception of preference as staunchly behaviorist because the economist is intent on talking not so much about actual trading behavior as about inclination or willingness to trade; but such talk could not be accommodated by a strict behaviorism.

As I have indicated, preferences are typically taken by the economic view to be comparative. The view is something like this: I prefer X to Y only when I am willing to give
up $Y$ in order to get $X$ (in a case where I start with $Y$) or when I am not willing to give up $X$ in exchange to get $Y$ (when I start with $X$). Though it would in practice require a great deal of empirical work, a map could ideally be made, for each individual and with respect to all things, detailing what he or she prefers to what. CBA works most smoothly on the assumption that preferences are not lexicographic. A person’s preferences are not lexicographic if, for any good $X$, there is some amount and/or combination of other goods $Y$ that the person would take in exchange for $X$. If there is no good or combination of goods $Y$ that an individual would accept in exchange for $X$, then this individual’s preferences are lexicographic. Whether or not preferences are lexicographic is purported by economists to be an empirical question, one answered by how people actually behave. The proponent of an ambitious form of CBA maintains – perhaps on the basis of empirical evidence, perhaps on the basis of essentially a priori arguments, or perhaps just on the basis of assumptions – that preferences are in fact not lexicographic. The move from here, then, is to say that all preferences can be expressed in monetary terms. One’s preference for a thing thus becomes identified with that amount of money for which one will exchange the thing. If one is willing to accept $Y$ for $X$, then one prefers $Y$ to $X$. If one will not give up $X$ for $Y$, then one does not prefer $Y$ to $X$. So, the claim made by proponent of an ambitious form of CBA that all costs and benefits can be expressed monetarily is effectively the claim that for any good and with respect to each person, there is some amount of money for which the person would exchange the good.

Cost-benefit analysts, we are now in a position to see, actually have two measures of, or proxies for, individual preferences (both of which I mentioned in the first section of this

---

4 Since my intention in this section is not to give the most detailed account of CBA possible, I will push past several somewhat complicated and thorny issues here, but will take these issues up in more detail throughout the remainder of this project.
chapter). One measure is willingness to pay (WTP), which is the amount of money that someone would give up in order to get or keep something good or to get rid of something bad. The other is willingness to accept (WTA), which is the amount of money that one would need to receive in order to give up something good or to accept something bad. Both are supposed to give us not just the direction of the preference, but also the strength. They purport to allow for preferences not just to be ordered, but to be quantified. In the words of James Griffin, they are meant to measure “the distance of a desire from indifference.”5 And both seek to report this distance in monetary terms. For the sake of conciseness, and in accordance with common practice, I will throughout this project tend to say just WTP, even in cases where WTA is technically the measure being employed. As I will elaborate on in Chapter 8, WTP and WTA actually fail to line up in the ways in which they are supposed to.

WTP is supposed to allow for an ordering of preferences not just intra-subjectively, but inter-subjectively as well. In a policy situation that affects both Bob and Abe, i.e., regarding which both Abe and Bob have preferences – perhaps, for example, it is a proposed policy that Abe prefers but that Bob does not – the analyst wants to compare Abe’s preferences to Bob’s. Suppose that Abe’s preference map is entirely drawn, and that Bob’s preference map is also entirely drawn. Even so, what tells us about the strength of Abe’s preference for $X$ relative to the strength of Bob’s preference for (or against) $X$? The analyst wants to put each person’s preference for $X$ on the same scale. In real policy situations, the analyst aims to put the preferences of all the people who are affected onto one scale and compare them. The proponent of CBA maintains that even this can be done via WTP. Whoever is willing to pay more has the stronger preference. If Abe is willing to pay $250

---

(for the policy to proceed, say) and Bob is willing to pay $200 (in order for it not to go ahead), then Abe has the stronger preference. If the policy were to go ahead, Abe would gain by a measure of $250 and Bob would lose by a measure of $200. If these two were the only ones affected by the policy, then the Kaldor-Hicks Criterion would endorse going ahead with the project, as Abe could compensate Bob by giving him $200 – thus leaving Bob no worse off than he was prior to the project – and Abe would still be better off (by $50) than he was before, though again, KHC does not require that this compensation actually be paid. (The reader may be worrying here about some apparent problems with using WTP to compare preferences inter-subjectively. I will discuss these in more detail later.)

This much enables us to say that, on the economic view, a good’s value to an individual is tantamount to that individual’s WTP for the good. But pinning down what exactly the relationship is, on the economic view, between preference and value is harder than expected. Among other things, it is unclear whether CBA proponents maintain that people’s preferences define what is of value, or whether people’s preferences are just incredibly adept at picking out (i.e., responding to) what is of value. Does preferring X to Y make it the case that X is more valuable than Y? Economists tend not to address the issue clearly. Perhaps they do not recognize the questions; or perhaps I am alone in thinking that the answers are not obvious. While these issues are not clearly addressed, economists are fairly clearly committed to the view that it is not possible for someone to prefer X to Y even though Y is in fact more valuable than X. In any case, a good’s economic value is not the amount that is being charged for the good, or the amount that one actually pays for the good. Rather, it is (roughly) the maximum amount that one would be willing to pay in order to obtain it or the minimum amount one would be willing to accept in order to give it up. This
is an extremely important point. The economist will point out something like water as an example of an item whose market price does not reflect the item’s true economic value. We pay very little for it, but since we cannot live without it, we would pay a great deal for it if doing so were necessary in order to acquire it. This process can be applied to goods that trade in a market, as well as to those that do not. Some say that non-market valuation is the measuring in monetary terms of the value people place on an item they may care for but that cannot be purchased through a market. To me, this definition seems too restrictive. Sometimes we want to re-value, as it were, a good that can be purchased through the market. We think that the price the good is assigned through the market is not the one it should have, and so economists attempt to ascertain the price it would have if the market were functioning properly. Thus, as I intend it, non-market valuation involves not just the attempt to price goods that cannot be purchased through a market, but more generally the pricing of a good in some way other than simply taking its market price. A non-market good, as I understand the term, is broadly construed to be a good that under typical market behavior (which is to say, imperfect market behavior) has no price or whose value we suspect is not identical with its price.

A significant part of the work involved with a typical cost-benefit analysis is trying to measure accurately people’s preferences. The analyst works carefully to figure out how much money an individual is willing to pay for something. For a variety of reasons, doing so is not as easy as it might initially seem. There are two basic kinds of methods (techniques, approaches) for valuing non-market goods, namely the revealed preference method and the stated preference method. These can be viewed as competing valuation techniques, or can be understood “as complementary, where the strengths of each approach can be used to provide
more precise and possibly more accurate benefit estimates.” Roughly put, the revealed preference approach attempts to read off an individual’s preferences from her or his real-world demand behavior. Consider radio broadcasts (and I mean broadcasts on standard AM or FM frequencies, not satellite radio broadcasts). They are public goods, which is to say that the marginal cost of production for a radio broadcast is zero (meaning that adding an additional listener does not create an additional cost for the radio station) and exclusion is costly (if not impossible). It is extremely difficult to assign property rights to such broadcasts. I pay essentially nothing to listen to my favorite radio station. I am not stealing; the music simply comes to me free of charge through the airwaves. It is essentially a positive externality. The stations advertisers pay to power the broadcast. All I have to do is turn on my radio and dial it to the station. The fact that I pay no money to listen does not, however, mean that I have no preference for listening to the radio, or that listening to the radio has no value to me. If I were made to discontinue my radio listening, I would feel a loss. If I were asked to discontinue my radio listening, I would want some compensation for doing so. Though I listen for free, I would not give my listening up for free. The market price of radio listening is zero, but if I had to pay something to listen, I would. The fact is, you will have an extremely difficult time discerning what this amount of money is – i.e., what (economists think of as being) the value of radio listening is to me – by simply observing me and my consumer behavior. For this reason, the revealed-preference method looks at demand

---

6 Joseph A. Herriges, Catherine L. Kling, Christopher Azevedo, “Linking Revealed and Stated Preferences to Test External Validity,” provided by Center for Agricultural and Rural Development at Iowa State University, Center for Agricultural and Rural Development Publications, number 99-wp222, and available online at http://www.econ.iastate.edu/faculty/herriges/Papers/LinkRPS.pdf.

7 Of course, if I did not ever want to listen to the radio and also preferred not to be bombarded with radio waves, it would count as a negative externality.

8 It should be noted that this actually makes for a certain sense in which I pay to listen to the radio, namely by listening to annoying advertisements that I certainly would prefer not to hear.
behavior regarding certain marketed goods that are related in important ways to the non-market goods which the analyst is trying to price. The analyst utilizing the revealed preference method to assess my preference for radio listening could, e.g., look at the amount of money that I spend on similar goods, including compact discs, vinyl records, and television.

Consider another case, one in which the revealed preference method is more likely to be employed, namely in establishing the economic value, i.e., the price, of clean air. There is not a market for clean air in the same sense that there is a market for vending machine potato chips. Vending machine chips are private goods; there are marginal production costs, and exclusion is relatively cheap and easy. Air, on the other hand, like radio broadcasts, is a public good. If there is, in a much looser sense, a market for clean air, the market price of clean air will be nowhere near as obvious as the market price of potato chips. Still, economists might try – have tried, in fact – to infer people’s willingness to pay for clean air by, for example, comparing housing prices in neighborhoods with clean air to housing prices in neighborhoods with dirtier air. Holding all other variables constant is meant in principle to show how much an additional unit of cleanliness in air is preferred, i.e., is worth monetarily, to various people. In practice it will of course be very difficult to hold all other variables constant, though a skilled analyst will be assisted in doing so by performing sophisticated multiple regression analyses.

There are at least two additional problems plaguing the stated preference method – and they are fairly pernicious ones at that – that should be canvassed here. The first is based in the very fact that gave rise to the usage of this method: to observe individuals in the real world is to observe individuals operating in imperfect markets, and under such conditions, a
good’s price may be different from the good’s economic value. Imperfect markets mean that externalities exist, and “in the presence of externalities, market transactions do not fully capture preferences.” So, the analyst infers a person’s preferences for some non-market good from how the person behaves with respect to some relevantly similarly good, but the analyst will need to ensure that the price of the relevantly similar good does not itself reflect externalities. This is difficult to do, however, because externalities and their effects are pervasive in the market. In Chapter 5, for example, I will discuss how collective action problems interfere with inferences that go from observations of a person driving a gas-guzzling, air-polluting sport-utility vehicle to that person’s preferences regarding clean air, taxes on gasoline, and so on. There are externalities associated with driving such vehicles – driving any vehicle, for that matter – and so the price a person pays to drive such a vehicle does not represent the true value to that person of the vehicle. It will, in turn, not tell the analyst what she wants to know about the true value – or more precisely, the disvalue – to this person of, e.g., the diminishment in air quality that this vehicle effects.

The second problem is that when the analyst looks at the individual’s behavior with respect to that relevantly similar good, the analyst can often see only that a person is, or is not, willing to pay a certain amount, but is not able to thereby infer precisely how much that person would have been willing to pay. Suppose a vending machine charges $1.00 for chips. You see me walk up to the machine, pause, scan the machine with a deliberative look on my face, and then walk away. (Perhaps I even say aloud that I was considering buying a bag of chips, but a bag costs more than I want to spend.) In doing so, I indicate that the chips are worth less than a dollar to me, but I have not given an indication as to whether they are, to

---

me, worth $0.01, $0.99, or somewhere in between. In this particular case, I presumably would not have even gone to the vending machine if I were only willing to spend $0.01, since I knew that no chips in the machine would be on sale for that little. In other cases, not even this much about my preferences will be able to be inferred. With respect to the chips, perhaps the truth is that I thought I could get them for $0.75, and would have paid that much but no more. If the difference between $0.75 and $1.00 were scaled to a big public investment project that affected many people, it becomes very significant. At other times, one is willing to pay more than one is required to pay. When this occurs, there is consumer surplus. Consumer surplus is the difference between the maximum value the consumer would be willing to pay and the market price, i.e., the amount that consumer is actually required to pay in the market; it can be defined more technically as “the area under the demand curve, between zero units and the quantity actually purchased, minus the total payments the consumer gives up to acquire the commodity.”

If you watch me at the vending machine and see me purchase a candy bar for the market price of $0.70, you may infer that the candy bar was worth at least $0.70 to me, but your knowledge of my willingness to pay is still incomplete. You do not know whether I would have refrained from buying it if it had been one cent more, or whether I would have bought it even if it had been twice as expensive. If I am willing to pay $1.00 for a Snickers but the vending machine only charges $.07, my actual preference for the Snickers is stronger than my behavior reveals. To return to the issue of clean air, the point of this discussion – and here let us entirely set aside the aforementioned problems, and even some others that I have not mentioned at all but will discuss later in the project – is to illustrate that observing me buy, or refrain from buying, a house at a certain dollar value in a neighborhood with a certain level of air

---

pollution gives you a very incomplete picture of my preferences regarding this house, and in turn a very incomplete picture of my preferences regarding the cleanliness of the air I breathe.

In cases where my preferences (and thus a good’s economic value) are not able to be safely and fully ascertained by observing my consumer behavior, the analyst’s job is more difficult. In such cases, an analyst can employ the stated-preference method of valuation. Here, the basic idea is that the analyst can devise ways – more and less tricky – of asking and soliciting answers to the question “How much is the good (service, outcome, etc.) in question worth to you?” This is also called the contingent valuation (CV) method. Proposed by S. V. Ciriacy-Wantrup, the idea is that individuals should be interviewed and “asked how much money they are willing to pay for successive additional quantities of a collective market good.”

Clearly, how a contingent valuation survey is conducted is of critical importance. In the nearly 60 years that these valuation surveys have been used, there have been marked improvements towards making CV studies more accurate and reliable; these improvements “touch on all aspects, including sampling, instrument development, formulation of the valuation scenario, questionnaire structure, and data analysis.”

Today analysts understand that the surveys should “confront subjects with a specific and realistic situation rather than with an abstraction, and… use a closed-ended question which frames the valuation as voting in a referendum.” While respondents are not, at the time of the survey, actually required to pay the amount that they specify, the surveyor should create scenarios in which what is recorded is the amount the respondent really would be willing to pay. Open-ended questions

---


12 Hanemann, ibid, p. 21.

13 Ibid, p. 22.
– e.g., “What is the most you would be willing to pay for…?” – and closed-ended questions – e.g., “If it cost $x for…, would you be willing to pay this amount?” – can elicit different responses. Respondents typically have a much tougher time answering the open-ended WTP questions. Questions should be closed ended for this reason, and because a close-ended question will more closely model choices that confront individuals in the market, where choices are typically more discrete (e.g., “here is an item, it costs $x, will you take it?”). Other things to which practitioners of CV need to be sensitive include ways in which: small changes in wording of a survey can cause dramatic changes in survey responses; the different surveying techniques can produce different results; the survey might tap into any general dislike of big business; what pieces of information are or are not given to respondents bears on responses; responses can manifest anchoring effects; and perceived pressure or expectation from the surveyor can affect responses.

One key remaining problem, however, is that for a number of reasons the economist cannot be assured of a truthful answer. In fact, because of economists’ own suppositions about human nature, namely that people are self-interested rational maximizers, the economist expects that people asked this question would misrepresent their own preferences. Go back to the ChemiCo example. Suppose that the government is considering whether to permit ChemiCo to dump in the river. They want to know how badly the dumping would hurt those living downstream, and will bar the dumping if the benefit to ChemiCo is less than the harm to those living downstream. If I live downstream, I have an incentive to overstate how badly the dumping would hurt me. And I have this incentive whether or not I expect that the amount that I specify will actually be paid to me as compensation. At least some economists miss this point. Leonard and Zeckhauser, for example, write that “it may not

14 Ibid, p. 23.
even be desirable to construct full compensation systems, since losers will generally have an
incentive under such systems to overstate their anticipated losses in order to secure greater
compensation.”15 Whether or not I actually expect – or even dream of the remote possibility
of – compensation, if I do not want the project to proceed, I will have an incentive to
overstate my losses. Doing so elevates the project’s costs and makes it less likely that these
will outweigh its benefits. In cases where I stand to be benefited by a proposed project, I will
have incentive to overstate the amount I would be willing to pay for it. None of this is to
deny, however, that there are real obstacles in the way of actually paying out compensation
to those who are harmed by a project. Nor is it to deny the claim made by many economists
(Leonard and Zeckhauser included) that if losers were to be actually compensated, “taxes and
direct expenditure packages represent a far more efficient means of effecting redistribution
than… [do] attempts to redistribute within every project.”16 (With respect to this latter issue,
however, there is obviously ample reason – particularly given the current political climate –
to doubt that such redistributive taxes would actually be levied.)

This now satisfies me as a viable summary of CBA for the purposes of this project. It
is at least sufficient to get a philosophical discussion regarding CBA off the ground. More
details can of course be provided where needed as I move forward.

Section 3: Theory of Right Action Underpinning Cost-Benefit Analysis

As indicated in Section 1 above, addressing the strengths and weaknesses of CBA’s
theory of right action, i.e., Claim 1, is not among my interests in this project. A great deal

15 Herman Leonard and Richard Zeckhauser, “Cost-Benefit Analysis Defended,” QQ: Report from the Center
for Philosophy and Public Policy, University of Maryland at College Park, Vol. 3, No. 3 (Summer 1983), pp. 6-
9.

16 Ibid, p. 7.
has been said about it by others. A great deal more could be said still. I am comfortable leaving most of this for someone else or at least for some other day, but it feels a bit inappropriate not even to mention certain obvious causes for concern with Claim 1.

For starters, KHC is not as easy to precisely state as it is commonly taken to be. I believe it was correctly stated it in the preceding section of this chapter, but it is regularly formulated in an infelicitous manner. It is worth pausing to note some of flaws in the formulations that one regularly encounters in the literature. For example, on Layard and Glaister’s construal, KHC maintains that a policy or project, A, is justifiable “if [A’s] benefits exceed those of the next best alternative course of action, and not otherwise.”\(^\text{17}\) This appears to be not quite the right formulation, because while A’s benefits could exceed the benefits of some alternative policy/project B, A’s costs could also exceed B’s costs, and in such a way that A results in a less optimal balance of benefit versus cost than B does. A. Myrick Freeman’s definition is problematic too. He says that the criterion “asks whether the aggregate of the gains to those made better off measured in money terms is greater than the money value of the losses of those made worse off. If the gains exceed the losses, the policy is accepted by this criterion.”\(^\text{18}\) The worry here is that there might be some competing policy – call this policy 2, and call the one Freeman is describing policy 1 – whose aggregate gains also outweigh its costs, and by a bigger margin than exists for policy 1. Layard and Glaister thus offer the following consideration: “It seems quite natural to refer to the ‘benefits of the next best alternative to A’ as the ‘costs of A’. For if A is done, those alternative benefits are


\(^{18}\) Freeman, ibid, p. 296.
lost. So the rule becomes: do A if its benefits exceed its costs, and not otherwise.”

This is better, though still unsatisfactory because what counts as the next best alternative (and why) is among the things the definition of KHC is intended to help us understand.

Munger’s construal is better still, but nonetheless imperfect. He says:

The Kaldor-Hicks criterion might be thought of as having two parts:
· Compare two states of the world, S1 and S2. If one is Pareto superior, no further analysis is needed: the Pareto superior one is preferred.
· If both states are Pareto optima, add up the gains and losses to all citizens from choosing each alternative. Select the policy that maximizes the difference between the gains to the gainers and the losses to the losers. Importantly, it is not required that the gainers compensate the losers, but only that the excess value created by the policy choice is maximized.

Note, however, that on Munger’s formulation, KHC seems just to be the Utilitarian theory of right action, when in fact these two things differ slightly but importantly from one another. Consider, for example, that on the Utilitarian theory of right action – and on Munger’s formulation of KHC – the losers could actually be harmed so badly as to die. As long as the gainers’ gains exceed these losses (and by a greater margin than gains exceed losses on another state of affairs), this is justifiable. On a proper understanding of KHC, such a scenario would not be justifiable. It is not possible, even in theory, to compensate someone who dies; but KHC is itself supposed to be defensible on the grounds that it recommends just those policies/projects on which the gainers could compensate the losers, at which point the losers would be no worse off than before the policy was implemented and the gainers would still be better off than they were prior to the implementation of the policy. The appropriate formulation of KHC will, in addition to requiring that benefits outweigh costs (and by a greater margin than on any other outcome), make specific mention of the requirement of a

---

19 Layard and Glaister, ibid.

20 Munger, ibid, p. 103.
possibility of compensation.\textsuperscript{21} Jenkins-Smith’s formulation seems to capture this; he says, “The Kaldor-Hicks criterion allows redistributions that increase net welfare such that those who gain from the distribution could compensate those who lose, restoring the losers to their prior level of well-being, while the winners retain enough of their gains to be better off than they would have been without the redistribution.”\textsuperscript{22} Unfortunately, he then goes back to the oversimplified version – he says, “In the comparison of policies, the relevant criterion is: which policy option serves to create the \textit{largest net gain in social well-being}?”\textsuperscript{23} – which makes one wonder if he understood what exactly his first formulation got right!

In light of all of these concerns I again offer my favored interpretation of KHC: A policy or project should be endorsed if and only if its benefits outweigh its costs by a greater margin than would be the case on any other policy/project, with the gainers being able to compensate the losers such that the losers are no worse off than they were originally and the gainers are still better off than they were originally.

This is the right wording of KHC, but simply wording it properly does not eliminate all causes for concern with the criterion. One major worry is that it is far from obviously true that what is \textit{best} is always what is \textit{right}. Whether CBA is correct in aiming to satisfy desires or preferences is a complicated question, and one that I will explore in detail later in this project. But even if CBA is correct on this count, the claim that the best (i.e., most good, most preferred) outcome is the right one will be quite contentious. Relatedly, one might ask why the mere \textit{possibility} for compensation justifies a policy. If I am one of the losers and I

\textsuperscript{21} Thanks to Doug MacLean for helping me see this point. It is an important one, and too infrequently appreciated.


\textsuperscript{23} Ibid.
am not compensated, it does not seem unreasonable to at least wonder whether I have reasonable grounds for complaint. Does the issue not turn at least in part on what kind of loss I have suffered, and how it was produced? Will there not be occasions – particularly but not exclusively when compensation is not actually paid to those who incur the costs – where a project produces more benefits than costs but nonetheless tramples individual rights? Can it be right to burden someone merely to benefit (by a greater amount) someone else?

I have explained that KHC is not exactly the utilitarian theory of right action; but it is nonetheless very close to it. Steven Kelman seems to be among those who go beyond relating KHC and the utilitarian theory of right action; he seems actually to identify them. In any case, what he says is informative:

How do we decide whether a given action is morally right or wrong and hence, assuming the desire to act morally, why it should be undertaken or refrained from? Like the Molière character who spoke prose without knowing it, economists who advocate use of cost-benefit analysis for public decisions are philosophers without knowing it: the answer given by cost-benefit analysis, that actions should be undertaken so as to maximize net benefits, represents one of the classic answers given by moral philosophers—that given by utilitarians. To determine whether an act is right or wrong, utilitarians tote up all the positive consequences of the action in terms of human satisfaction. The action that maximizes attainment of satisfaction under the circumstances is the right act. That the economists’ answer is the answer of one school of philosophers should not be surprising. Early on, economics was a branch of moral philosophy, and only later did it become an independent discipline… Utilitarianism is an important and powerful moral doctrine. But it is probably a minority position among contemporary moral philosophers. It is amazing that economists can proceed in unanimous endorsement of cost-benefit analysis as if unaware that their conceptual framework is highly controversial in the discipline from which it arose—moral philosophy.  

Given the close relationship between KHC and the utilitarian theory of right action, it would seem outrageous to endorse KHC without addressing the arguments against Utilitarianism made in what I take to be the 20th century’s two most important pieces of

political philosophy, namely Rawls’ *A Theory of Justice* and Nozick’s *Anarchy, State and Utopia*, both of which are hostile to Utilitarianism. There are worries that KHC illicitly treats people as mere means rather than as ends in themselves, that it fails to take seriously the separateness of persons, and that it is not a principle of justice – or not consistent with principles of justice – that would be chosen from a position of fairness, i.e., by people behind the veil of ignorance. There are important replies – I stay neutral on whether the replies are also *successful* – to these worries, but pursuing them will take us not just into the work of Pareto and Kaldor and Hicks, but deep into the literature on the debate between Utilitarians and deontologists, between, for example, Bentham and Mill on the one hand, and Kant on the other, as well as between the likes of Harsanyi, Rawls, and Nozick. There is much, much more to say about Claim 1, but it is not my purpose in this project to assess the claim any further. Instead, I move on to CBA’s theory of value, on which I will focus for the remainder of this project.
CHAPTER 2: ON BENEFITS, PREFERENCES, AND WELL-BEING

Section 1: Introduction

If CBA is to be tenable, its defenders will need to give and support a general characterization of costs and benefits. Suppose that factories in a particular political jurisdiction are emitting pollution into the air. A policy proposal might be to tighten regulations on pollution of some particular sort so that less of it is permitted to be released. Alternatively, the proposal might be to relax the regulation. Or suppose that the policy issue concerns not a regulation but rather a potential public investment, in floodwalls perhaps, which will make it less likely that people’s possessions, and indeed their lives, would be lost when the next severe hurricane strikes the area. Perhaps the proposed regulation is one that would reduce the amounts of lead that can be in public housing paint. It seems natural to evaluate policies such as these – policies regarding public investments and regulations – at least in part according to what their costs and benefits are expected to be. What role considerations of costs and benefits ought ultimately to play, how they are aggregated, what rights people have and how these relate to considerations of costs and benefits—these are largely issue regarding which I stay neutral in this project. But if considerations of costs and benefits are relevant at all to a policy decision, then we need to know what counts as a cost, what counts as a benefit, and why. This is the first of the two major questions I take up in this project. (The second, to be taken up later, will be whether each cost and benefit can be expressed monetarily.)
The view typically taken by economists is that whether a change counts as a benefit or as a cost is in the first place always relative to each individual affected. In particular, it depends on each individual’s assessment of the change. Roughly put, if the policy-induced change would satisfy a person’s preference, then the change counts as a benefit for that person. If the change would frustrate the individual’s preference, it counts as a cost for the individual. This is the claim that I earlier labeled Claim 2. Some proponents of CBA may disagree with it entirely, but in doing so they would be situating themselves clearly outside of the mainstream. Typically something very close to Claim 2, if not it exactly, is endorsed.

What are we to make of Claim 2? Mark Sagoff, a frequent and vocal critic of CBA, argues against the claim, conceding that “having a preference is a reason for the person who has it to try to satisfy it,” but then immediately asking, “Why should it be a goal of public policy, however, to satisfy that preference?” 25 This is a very important question. A large part of this whole project is devoted to exploring this question and possible answers to it. Sagoff’s answer, in short, is: it should not! And he is not the only one to have given this general answer. Many others have made arguments the basic upshot of which is that it is a mistake to count the satisfaction of an individual’s preferences as a policy-level benefit. Some of these arguments move along lines similar to Sagoff’s, while others move along quite different lines. As it turns out, none of them convinces me, though I believe nonetheless that the economists’ view has a number of problems.

In this chapter I begin working towards a compelling answer to the question of whether public policy should aim to satisfy individual preferences, or – what I take to be essentially the same, whether the satisfaction of individual preferences should count as a benefit and the frustration of individual preferences as a cost. I explore various possible

---

reasons why the state might aim to promote the satisfaction of individual preferences. The
suggestion that will garner the most attention is the one that prevails among economists,
namely that to promote the satisfaction of an individual’s preferences is to promote that
individual’s well-being, and that to promote individual well-being as far as possible is to
maximally promote social welfare. I consider a number of competing accounts of individual
well-being, ultimately defending an informed-preference account, which I show to preserve a
relationship between preferences and well-being without strictly identifying the satisfaction
of an individual’s preference with an increase in that individual’s level of well-being. In the
next chapter, the implications for CBA of this account of well-being are assessed.

Section 2: Why, According to Proponents of CBA, Should the State Aim to Satisfy
Individual Preferences?

The notion of a preference plays so central a role in CBA, and in turn in this project,
that it would be a mistake to move forward with a discussion of the relationship between
preferences and benefits/costs without ensuring that the notion has been sufficiently
elucidated. It is not quite a primitive, but what can be said as a characterization of it is
limited. As indicated in Chapter 1, many of the economists who employ the term preference
– in the context of CBA or elsewhere – will admit (at least in their particularly forthcoming
moments) that it is a bit fuzzy. But it is not obviously fuzzier than many other commonly
used terms – not much fuzzier than “chair,” to use that old clichéd example – and not too
fuzzy to be reasonably and coherently used. I want to work with as intuitive, non-technical
an understanding of preference as possible. Given that there are more and less technical
(“specialized,” we might say) senses of the term, it is not obvious exactly how Sagoff intends
for his question to be understood. My proposal is that we at least start by interpreting the
question such that we could restate it, while still preserving its meaning, in at least the following ways: Why think that the aim of public policy decisions should be to deliver those outcomes that people want, desire, like, or recommend?

The reformulation of the question shows basically how I understand the concept of a preference. But a bit more clarification is in order. While I want to understand it in as intuitive as way as possible, I also want to talk about preferences in essentially the way that economists do. That is, it will accomplish little to assess the role of preferences in CBA if in doing so, the concept of a preference is understood in a way much different from how the average proponent of CBA intends it. Economists long favored a behaviorist understanding of preference according to which one has a preference for X over Y if and only if one who is confronted with the choice is observed to buy X when one could have within one’s budget bought Y. But as was explained earlier, behaviorist understandings of preference have rightly lost favor, generally even among economists.  

Economists who use the stated-preference method to solicit preferences cannot possibly subscribe to a behaviorist understanding, for stating how one would be willing to behave is not the same as actually performing an action or demonstrating a behavior. I join those who think that the behaviorist understanding of the notion is flawed, and indeed I would permit preferring to be loosened from behavior more than the stated-preference method of ascertaining preferences does. This is not to say, however, that the notion should be understood to have no behavioral component. That would

---

26 It seems to me that the behaviorist conception has lost favor even among economists, though this is just a sense I have. (I have not, e.g., taken a poll.) For a seminal treatment of the concept of preferences in economics, see Paul Samuelson, “Consumption Theories in Terms of Revealed Preference,” *Economica*, Vol. 15 (1948), pp. 243-253. For a discussion of problems with Samuelson’s approach, see Amartya Sen, “Behaviour and the Concept of a Preference,” in *Choice, Welfare, and Measurement* (Oxford, UK: Basil Blackwell Publisher, 1982), pp. 54-73. I think that particularly forthcoming economists who currently work with the notion of a preference will admit that it is a somewhat fuzzy concept, but clear enough to be passable. I heard one well-known economist say of preferences – only partly in jest, I suspect – “You know them when you see them.”
be a mistake. A claim to having a preference for, say, doing philosophy will be dubious at best if the person making it never actually reads or writes any. The claim will lack all credibility if the person is never even disposed to read or write any. I will say more in later chapters about how I understand the relationship between preferences and behavior.

The above reformulation of Sagoff’s question suggested at least a close relationship between (on the one hand) preferences and (on the other hand) desires and wants, and indeed I do think they are closely related, though they are not identical. One can prefer something that one does not want or desire. It seems perfectly coherent to say that person P does not desire either X or Y, though P prefers X to Y. The claim that “P does not desire either X or Y” means something like “If P could avoid both X and Y, she would do so”; but P might still prefer X to Y in the sense that if she were, say, required to get/pursue one of these, she would get/pursue X. Moreover, consider the oddness of saying “P desires X, but does not prefer it.” In fact, it is somewhat odd to say that some X is (or is not) preferred without invoking some other thing to which X is (or is not) preferred—though the oddness is eliminated when one substitutes “desired” for “preferred.” So, part of the difference between preferences and desires is that former seem to be necessarily comparative while the latter do not. When one attests to preferring something, one is saying something not just about the thing, but about at least some other thing as well. One never prefers X simpliciter, but rather prefers X to Y. Now, I am comfortable talking about a preference for X simpliciter, i.e., without explicitly invoking some alternative or comparative thing Y, but that there is an alternative is implicit, and typically there are contextual cues that indicate what less-favored alternatives exist. Anytime I talk about a preference for X simpliciter, I could, if pressed, make the alternative to X explicit, but talking simply about a preference for X – as opposed to

---

27 I will make no distinction between wants and desires.
a preference for $X$ over $Y$ – is less cumbersome. None of this is to deny that desires can be comparative in a sense too. One can desire that $X$ and not $Y$ obtains, or that one get $X$ rather than $Y$. Still, we an frequently make sense of desires in non-comparative terms, though we cannot do so when it comes to preferences.

There is another respect in which preferences are distinct from desires and wants, and from beliefs and judgments too—at least as all these are crudely understood. Here, however, my view is perhaps somewhat more idiosyncratic. Desires are sometimes thought to be mere urges or inclinations not connected in any way to reason, which includes being immune to criticism by reason. Many others who understand desires as more substantial than this still maintain that they differ in kind from, and are at most mediated by, beliefs or judgments. Beliefs and judgments, on the other hand, are taken to subsist in one’s rational faculty, but are said to be motivationally impotent.\footnote{I claim neither that everyone subscribes to this sharp distinction between desires and beliefs/judgments, nor that it is a convincing distinction; but it is clearly influential.} I see preferences as having characteristics of both desires and beliefs. Preferences differ from desires (crudely understood) in virtue of the fact that preferences have some cognitive content. They are not mere urges, like an urge to sleep or go to the bathroom. As I will detail later in this project, it can make sense to criticize a preference, at least when the criticism is of a certain sort. At least sometimes, it will be possible to offer someone reasons for why she ought to change her preferences. Indeed I would say that a preference can consist partly in a belief or beliefs. Part of what it is to prefer, say, going for a walk over watching television right after dinner tonight is to believe certain things about the walk and the television watching (though the identity of the preference is not hinged to the truth of those beliefs).\footnote{I claim neither that everyone subscribes to this sharp distinction between desires and beliefs/judgments, nor that it is a convincing distinction; but it is clearly influential.} But the preference consists in more
than beliefs, and does have what we might call a noncognitivist element. Having a preference for a thing involves being in some sense attracted to that thing, though this does not include cases where one hates or feels hostage to this attraction. Rather, one must in some important sense endorse (embrace, assent to) the attraction. (That preferences consist partly in beliefs, and that they involve attraction that one endorses, together provide the basis for a distinction between preferring and being physically addicted.) It is intuitive to say that preferences have motivational power; we do not need to appeal to some other psychological notion in order to understand how someone could be moved to do/pursue that which she prefers. I know that not everyone will agree with me about all this. Many, for example, will find it odd to say that the preferences consist partly in beliefs. I take this to be the most plausible way to understand preferences, but let me be clear: I do not believe that anything I go on to say in this project really depends on the legitimacy of the conception of preferences that I have urged in this paragraph. By understanding preferences in the relatively cognitivist manner that I do, there is nothing to which I am trying to help myself.

What precisely a preference is, and what must be true of someone in order to say that she has a preference, will reappear as a central issue in Chapters 8 and 9 (and to some extent Chapter 7 as well). Whether or not preferences are lexicographic, and whether they all differ from indifference along some one single dimension, are also topics that will be explored later. For now, I move forward with the discussion understanding the idea of a preference in the way explained in the preceding few paragraphs. Where I come to a situation where differences between this intuitive understanding and any more technical understanding make

---

29 It should be noted that there may be an understanding of desire that is less crude – and more specifically, that is more cognitivist – than the one I have described, and that could thus be treated as synonymous with preference. Even so, I would use ‘preference’ rather than ‘desire’ in order to keep contact with the terminology that proponents of CBA favor. By understanding preferences in the substantially cognitivist manner that I do, there is nothing to which I am trying to help myself.
a difference, I will pause for careful discussion. In any case, my hope it that the lay person, the philosopher, and the economist alike will all find Sagoff’s question intelligible, and will all more or less agree with me on the basic elements of a characterization of the notion of a preference. All should see, moreover, that Sagoff’s question is relevant to an assessment of the normative foundations of CBA. Whether or not everyone takes it at this point also to be an interesting question – a question that has real bite – is something I do not know. I will try to make its importance apparent.

However it is worded, the question is considerably more difficult to answer than expected. Sagoff rightly notes that if one claims “that we should strive to satisfy preferences because that is what the people who have those preferences prefer or want,” then one merely invites an infinite regress.\(^{30}\) That people prefer for their preferences to be satisfied (as far as possible) counts as a reason to satisfy people’s preferences only if there is a reason to satisfy (as far as possible) people’s preferences; but whether there is a reason to satisfy people’s preferences is exactly what is at issue. The argument for promoting individual preferences, and more specifically for counting satisfied preferences as policy-level benefits, will have to be more sophisticated than this. But Sagoff believes that more sophisticated arguments are in trouble too. There is, he claims, no empirically or conceptually necessary connection between having a preference satisfied and receiving a benefit. The satisfaction of preferences, he contends, is neither clearly good in itself nor always good in virtue of its consequences.

Let us take these one at a time, starting with the possibility that the satisfaction of preferences is good in itself. According to Sagoff:

\[^{30}\text{Sagoff, ibid, p. 303.}\]
It cannot be argued that the satisfaction of preferences is a good thing in itself, for many preferences are sadistic, envious, racist, or unjust. Preferences may also be coerced or adapted to coercive circumstances; these do not express the free choice of the individual but a process that preempts that freedom. Many preferences, for example, the urge for a cigarette, are despised by the very people who have them. Why should the satisfaction of these preferences, which are not autonomous, be considered a good thing in itself? It may be good in itself that certain preferences be satisfied, namely preferences that are good in themselves. Ordinary consumer preferences, however, are usually not regarded as intrinsically good but as arbitrary from a moral point of view. People should be free to try to satisfy these preferences, to be sure, but why should it be an intrinsically good thing that they do satisfy them? Why is it good in itself that a person who wants a Mercedes succeed in getting one?31

Showing that something is “a good thing in itself,” or that it is “intrinsically good,” is notoriously difficult. Truth be told, however, Sagoff’s task, namely showing that something is not good in itself, is likely no less difficult. I am actually surprised by the great comfort Sagoff seems to have with using these notions. The fact of the matter is that they are among the most muddled in all of moral philosophy. One important question – and one that Sagoff does not address in any even semi-straightforward fashion – is: what needs to be true of something in order for it to properly count as being good in itself, or as intrinsically valuable? Sagoff seems to count “good in itself” and “intrinsically valuable” as synonymous. It is actually not clear that these are in fact the same. Perhaps a thing is good in itself when it possesses an objective though non-natural property, where this property is called “intrinsic value.” On this view, intrinsic value is something that could be detected by a properly attuned human being, though not exactly through any of the traditional five senses. It is real, though not detectable by any of the scales, scopes, or spectrometers of physics, chemistry, or biology. This was, in effect, the view of the famous early-20th-century philosopher G. E. Moore. More will be said in Chapter 6 about the Moorean view; but for now, suffice it to say that if the Moorean analysis is correct, then generally rejecting a non-naturalistic ontology –

as I do, though the reasons for doing so will not be articulated or defended here – permits one to reject any and all claims that there is a thing that is “intrinsically valuable,” and in turn to deny that there is a thing that is good in itself. But an additional interesting problem for the Moorean approach is that even if it could be shown that something is intrinsically valuable in Moore’s sense, nothing about what claim this value makes on us would be entailed. An additional argument would need to be given in order to show that the promotion of that which is intrinsically valuable ought to be the aim (or among the aims) of public policy.\textsuperscript{32} I have seen no such argument, do not have one myself, and cannot even imagine what one would look like.

But not all who call a thing good in itself or intrinsically valuable have this kind of Moorean notion in mind. The notions need not involve non-natural properties or anything of the sort. Mill, for example, argued that happiness was desirable as an end.\textsuperscript{33} Perhaps to say that a thing is good in itself is just to say that it is desirable as an end, not (or not merely) as a resource, not merely as a means to some other end. One can seemingly eschew all talk of non-natural properties and still affirm the existence of things that are in this sense good in themselves. We could even re-appropriate the term “intrinsic value” from Moore and give it this new naturalistic interpretation; so, a thing is intrinsically valuable if it is desirable as an end. (On this view, “good in itself” and “intrinsically valuable” are again synonymous.) But as Mill himself discovered, meaningful arguments about whether something is in this sense

\textsuperscript{32} In fact, an additional argument would need to be given even in order to show that the promotion of that which is intrinsically valuable ought to be the aim (or among the aims) of an \textit{individual}. That is to say, concluding that the satisfaction of a preference is intrinsically good in Moore’s sense does not even entail that the individual whose preference it is has reason to try to satisfy the preference.

\textsuperscript{33} See John Stuart Mill’s \textit{Utilitarianism}. 
good in itself quickly run aground.\textsuperscript{34} Is it enough, in order to show that something is desirable as an end, to show that it is capable of being desired as an end? (What if it is capable of being desired as an end but no one actually does desire it as an end? And is there anything that is actually completely \textit{incapable} of being desired as an end?) Is it enough that people do actually desire the thing as an end? (What if there is a bit of disagreement? And what if the people desiring the thing as an end are themselves morally contemptible?) Or do we have to show not just that the thing is desired as an end, but also that the thing is \textit{properly} desired – that is, that it should be desired, or that it is appropriate to desire the thing – as an end? (And how would we show that?) Moreover, it is not clear how one who takes these terms to have importantly to do with desirability avoids running smack into the charge, outlined several paragraphs back, of inviting an infinite regress.

A further complication is introduced by the distinction between intrinsic and conditional value. This is a distinction to which Kant was attuned.\textsuperscript{35} On many interpretations of “intrinsic goodness,” it seems possible to maintain, for example, that intelligence is intrinsically valuable \textit{and} that an intelligent criminal is worse than an unintelligent criminal. This means, in other words, that intelligence is intrinsically good but that the goodness of intelligence is nonetheless conditional (namely upon being possessed by someone with a good will). If the distinction between “intrinsically good” and “unconditionally good” holds, then though Sagoff is correct that “many preferences are sadistic, envious, racist, or unjust,” perhaps this does not show that the satisfaction of a


\textsuperscript{35} See Section I of Kant’s \textit{Groundwork for the Metaphysics of Morals}. 
preference is not intrinsically good, but rather means just that the satisfaction of a preference is not unconditionally good.

This preceding discussion has certainly not exhausted all the relevant (or potentially relevant) distinctions. The following is a list of predicates that all seem to get jumbled around. For any of a large number of goods $X$, philosophers might potentially say that $X$: is good in itself; is intrinsically valuable; is valuable as an end; is desirable as an end; is valuable for its own sake; is desirable for its own sake; is unconditionally valuable; is inherently valuable; is valuable beyond its resource value; is morally considerable; is morally significant; is objectively valuable; is valuable in isolation; is a member of the moral community; is an object whose worth is not conditioned by human interests; is an ultimate object of concern; is, ought to be, or could plausibly be, a direct beneficiary of a statute or policy; is a bearer of rights; is a bearer of interests; is an object of duty; is extrinsically valuable; has existence value; and/or, has contributive value. Some of these things – for example, “is a bearer of rights” – seem less applicable as predicates when the subject is something like a preference than when the subject is something like my uncle Wendell. But still, they are all predicates that creep up in the literature, and it is far from clear how (or even if) they are all distinct from one another. This list of predicates all bearing at least some resemblance to one another illustrates the difficulty in simply asking, with respect to some thing $X$, whether $X$ is intrinsically good, or good in itself. I will try to sort some of these predicates out as I go, but many of them I will simply aim to avoid altogether.

In short, I side with Sagoff in finding it to be far from obvious that the satisfaction of a preference is good in itself or intrinsically valuable. But I hope to have shown that Sagoff’s dismissal of the possibility is much too quick. I am unconvinced that the satisfaction of
preferences is good in itself or intrinsically valuable. That I am so unconvinced just reflects the fact that I have a great deal of trouble in making sense of those notion, and is not an out-and-out indictment of the promotion of preference satisfaction. Suppose, therefore, that we eschew the terms “good in itself” and “intrinsically valuable,” and indeed many of the other terms in the list in the preceding paragraph, and simply ask: Is there good reason for the state to promote the satisfaction of individual preferences even if it happens to be the case that no good consequences come from the satisfaction of individual preferences? An alternative wording is: Is there good reason for the satisfaction of people’s preferences to be the ultimate aim of public policy? These questions are intelligible. Surely something must be aimed at for its own sake and not just for the sake of something else. This is true both for individual decisions and public policy decisions—though nothing guarantees out of hand that the appropriate ultimate aim for an individual’s decisions will be the same as it is for state-level decisions. In any case, if nothing is aimed at for its own sake, then we get either a circle (e.g., one pursues A for the sake of B, and B for the sake of C, and C for the sake of A) or an infinite regress (e.g., one pursues A for the sake of B, B for the sake of C, C for D, D for E, \(ad infinitum\)). Both the picture of the circle and of the regress are dreary ones. They make life seem rather meaningless. The question I have suggested is of course not altogether free of terms that are difficult to interpret – “good reason,” for example, will potentially be messy – but it is a clearer question than the about whether the satisfaction of preferences is intrinsically good, or good in itself.

Despite its improved clarity, the question as I have worded it is still very difficult to answer. In fact, it is surprisingly difficult even to identify what answer the average proponent of CBA would want to give to even this question. Of course, at least some of
CBA’s proponents are not careful or philosophically minded enough to recognize that it is a question that needs answering. On the whole, though, it seems to me that most of the defenders of CBA who address the question seem to answer it in the negative. They maintain that preference satisfaction is *not* the appropriate ultimate aim of public policy decisions, i.e., that it should *not* be aimed at for its own sake. Rather, it seems that the economists’ “textbook” line is that an individual’s preferences should be satisfied in order to increase the individual’s well-being. That is to say, many of the most influential proponents of CBA seem actually to believe that the promotion of well-being, *not* the satisfaction of individual preferences, is the ultimate good, i.e., that at which the state ought ultimately to aim, and that this aim is achieved by promoting the satisfaction of individual preferences.

There are plenty of examples to support my contention. Consider Edith Stokey and Richard Zeckhauser, who maintain that “the purpose of public decisions is to promote the welfare of society,” and that “the welfare levels of individual members of society are the building blocks for the welfare of society.”\(^{36}\) They continue by saying that “in the United States we usually take the position that it is the individual’s own preferences that count, that he is the best judge of his own welfare.”\(^{37}\) A. Myrick Freeman, seems to concur, saying:

> We begin with the basic premises that the purpose of economic activity is to increase the well-being of the individuals who make up the society, and that each individual is the best judge of how well off he or she is in a given situation. To give this premise some operational content, we assume that each individual has preferences over alternative bundles of economic goods and services. In other words, the individual can rank all of the alternative combinations of goods and services he [or she] can consume from most preferred to least preferred. Of course there may be ties in this ranking. We assume that individuals act so as to obtain the most preferred (to them) bundles given the constraints imposed by technology and the availability of the means of production.\(^{38}\)


\(^{37}\) Ibid, p. 263.
The same tack is taken by Richard Layard and Stephen Glaister, who claim that the task of the analyst is to “find out how the [policy] decision would affect that welfare of each individual concerned.” They maintain that “to judge this effect we must ultimately rely on the individual’s own evaluation of his mental state. So, the broad principle is that ‘we measure a person’s change in welfare as he or she would value it.’”

It is hard to imagine that anyone who is engaged in a serious exploration of the merits of the normative underpinnings of CBA will be satisfied with these explanations of why CBA seeks to satisfy individuals’ preferences. That they were given by prominent economists does not in itself make them any more compelling. The economists’ arguments – if they can even be called arguments – are much too offhanded. Significantly more will need to be said in order to vindicate the claim that the satisfaction of an individual’s preference ought to count as a policy-level benefit, and that the frustration of such a preference ought to count as a cost. There are at least three significant problems that demand more careful consideration. The first is one I will not attempt to answer in this project. It is that the relationship between individual well-being and social welfare would seem to be more complicated than the economic account lets on. There seems to be a fallacy of composition lurking inside the account: just as it is not the case that an entrée tastes good if any only if each ingredient comprising it tastes good, so too is it not the case that the “welfare of society” is promoted if and only if the “welfare levels of individual members of society” are promoted. What the precise relationship is between individual and social well-being is

38 Freeman, ibid, p. 294.

39 Layard and Glaister, ibid, p. 2.

40 Ibid.
complicated. In any case, this issue has more to do with CBA’s theory of right action than with its theory of value – it is, in effect, a question about the right way to aggregate, if it is even the case that social welfare is built somehow out of individual well-beings – and so is not within my purview in this project. Thus, any time I ask about what should count as a benefit, I am talking about a benefit relative to an individual, and am simply supposing that the question of how these are aggregated is one for another occasion.

The remaining two problems are ones I will explore in detail in this project, and should be clearly identified here. First among these is that more needs to be said in support of the view that it is increasing well-being at which the state ought ultimately to be aiming. That view is actually significantly less obvious than it may initially appear to be. Second, there is need for a much more rigorous analysis of the relationship between the satisfaction of one’s preferences and an improvement in one’s well-being. I cannot address these problems at the same time. I will discuss the latter problem first, expecting that a better understanding of the very notion of well-being will cast light on the issue of whether promoting well-being is indeed the appropriate ultimate policy aim.

It may be objected that these questions should be addressed in the order opposite the one I propose. Griffin, for example, says:

> We cannot just ask, What is the best account of well-being?, as if ‘best’ could mean ‘most accurate.’ Our job is not to describe an idea already in existence prior to our search. Before we can properly explain well-being, we have to know the context in which it is to appear and the work it needs to do there. It may be that different notions of well-being are needed in different theoretical contexts.⁴¹

With all due respect to Griffin – and much respect is surely due – I think his point is confused. Griffin is right that our job is not to describe an idea already in existence prior to

⁴¹ Griffin, ibid, p. 1.
our search. However, this is right not because the idea of well-being is not already in existence prior to our search – whatever it is for an idea to be in existence, the idea of well-being surely is in existence – but rather because our job is not merely to describe the idea. I take there to be good reason for endorsing what Richard Brandt terms the “method of reforming definitions” (as opposed, say, to the “method of linguistic intuitionism”).\(^{42}\) Brandt approvingly cites R. B. Perry, who rightly “says that the question for philosophers is not how normative words are used, for they are used confusedly, but how they are best to be used.”\(^{43}\) None of this, however, means that we cannot ask what the best account of well-being is, as if best means most accurate; nor does it mean that we cannot do so until we know the context in which it is to appear. I want to know what individual well-being is – what it consists in formally, how it should be understood – and once I am sufficiently confident that I have such knowledge, I will ask whether well-being does the work in the context of CBA that economists take it to do there.

Before moving on carefully to consider those two main problems, it should be noted that among those defenders of CBA who recognize the importance of the question, “Is there good reason for the state to promote the satisfaction of individual preferences even if it happens to be the case that no good consequences come from the satisfaction of individual preferences?” at least some want to answer yes. More specifically, some attempt to justify the aim of promoting preference-satisfaction on the grounds that doing so is required in order to respect individuals as “free to choose.” It is, in other words, a part of respecting individuals as autonomous agents. This, I take it (with a dose of charity), is what is going on when defenders of CBA mention “consumer sovereignty.” This is, as I understand it, a

---


\(^{43}\) Ibid, p. 5.
(putative) kind of deontological reason for satisfying preferences (as opposed to the aforementioned textbook economic reason, which is consequentialist). Relatedly, some maintain, again seemingly on deontological grounds, that connecting up benefits with preference-satisfaction is a part of what it means to give people the input into governmental decisions that democracy demands people have. These claims will be addressed in due time. But we will start by exploring the aforementioned problems that arise when benefits (to an individual) are identified with the satisfaction of (that individual’s) preferences on the grounds that the satisfaction of an individual’s preferences promotes that individual’s well-being.

Section 3: Preference Satisfaction Versus Well-Being, and Other Important Distinctions

Is an afternoon at a philosophy lecture or one spent on a hike in the woods more conducive to a person’s well-being? Is the well-being of a 35-year-old better promoted through a life of monogamous marriage and children or sexually-unrestricted singlehood and childless independence? Would the well-being of a rural Indonesian be better enhanced by getting a water-sealed toilet in her home or by getting satellite television? Would my neighbor be better off drinking contaminated but inexpensive well water or paying for and drinking more expensive but quite clean bottled water. In order to answer these questions, as well as to assess properly the economists’ claim that the best way to increase a person’s well-being is by promoting the satisfaction of her preferences, we need a better understanding of the notion of well-being.

It does not seem terribly difficult to establish that there is a conceptual difference between the satisfaction of an individual’s preference and an increase in the individual’s
well-being. Some people consistently prefer artery-clogging pork rinds to fresh fruit. Others prefer smoking two packs of cigarettes a day, or gambling under terrible odds with what little money they have, or wrestling alligators, or even playing Russian Roulette. It is far from obvious, though, that actually doing any of these things really does increase the well-being of every person who prefers to do these things. I cannot show that it is incoherent to insist on definitions of “well-being” and “preferences” such that the two always and necessarily hang together. In other words, I do not think it is logically impossible to define these two terms in such a way that one’s well-being is increased when and only when her preferences are satisfied. While strictly speaking not logically impossible, doing so is certainly an extremely counterintuitive, implausible way of understanding the concepts. It is much more plausible to say of those people who have any of the aforementioned preferences not only that it is not a conceptual truth, but that it is not a truth of any kind, that their well-being is promoted when these preferences are satisfied. If you want to promote someone’s well-being, you will probably not assist her in playing Russian Roulette, even if you know she prefers to play it. Instead, you would reasonably say that this person has some problems, and does things that are bad for herself (that diminish her well-being, that are contrary to her own interests, etc.).

In short, promoting a person’s well-being and promoting the satisfaction of that person’s preferences are not always identical. The fact that some person $P$ prefers $X$ is neither necessary nor sufficient for establishing that $X$ promotes $P$’s well-being. One does not always, in promoting the satisfaction of $P$’s preference, promote $P$’s well-being; and one does not always, in preventing the satisfaction of $P$’s preferences, act in a way that diminishes $P$’s well-being.
As an aside, albeit an important one: while I have suggested that preference-satisfaction and well-being are conceptually distinct, and while I will go on to make a number of other distinctions, there are a handful of concepts that I do want just to identify. I have already bounced a bit between “welfare” and “well-being.” For example, in the passages quoted above, Stokey and Zeckhauser, as well as Layard and Glaister, used the term “welfare,” while Freeman used “well-being”; but these two terms, as they are used by these various authors, as well as according to the most intuitive ways of understanding the terms, seem to mean the same thing. I treat them as conceptually identical. Others, at least some of whom will be explicitly discussed at one point or another in this project, talk in terms of an individual’s “interests” – not in the sense of “taking an interest,” but rather of one’s “best interests.” Others talk of an individual’s “good” or “future good,” or of an individual’s “quality of life.” I do not insist that no one could come up with sensible distinctions between these notions. I do believe, however, that my trying to do so here would badly muddy the water rather than clear it. I will not distinguish these concepts – an individual’s “well-being,” “welfare,” “interests,” “good,” or “quality of life” – but instead treat them as synonymous. “Well-being” will be my term of choice, but no sleight-of-hand is intended by an occasional use of “interests,” for example.

---

44 Stokey/Zeckhauser and Layard/Glaister, for example, use “welfare.”

45 Freeman, for one, tends to use the term “well-being.”


47 This is used, for example, by Sidgwick and Railton in works from which I will later quote.

It is, however, important to hang on to several key distinctions. One of these, as I have explained, is between individual well-being and social welfare. I also of course distinguish the aforementioned cluster of synonymous terms from the concept of “preference satisfaction”; and I will return to that discussion soon. Before doing so, it is worth pausing briefly to note that to the list of concepts worth distinguishing from one another we could also add happiness and second-order preferences. What follows is not intended as rigorous conceptual analysis. I merely mention and comment briefly upon several examples which I believe suffice to demonstrate the distinctness of the four notions.

(i) Preference Satisfaction vs. Happiness:

\( P \) may prefer to spend all her money on philosophy books as soon as she is paid each week. Ultimately, however, satisfying this preference undermines \( P \)’s happiness. When \( P \) blows her paycheck on books, she lacks money for well-rounded meals, and is in turn unhealthy, uncomfortable, and finally unhappy. Moreover, \( P \) is driven to psychological exhaustion, depression, and eventually a nervous breakdown by all of her philosophical reading, which surely is not consistent with \( P \)’s happiness.\(^{49}\) Conversely, \( P \) might, to her own surprise, discover herself to be happy despite the fact that her individual preference for new philosophy books has not been and is not being satisfied. \( P \) orders a philosophy book, and to her chagrin a schmaltzy romance novel is mistakenly sent instead. \( P \)’s preference for a new philosophy book has been frustrated. She is upset, but begins reading the romance novel and is surprised to find that she loves it. She snuggles under a blanket and reads long into the night. In this novel, much more than in any philosophy book, she finds happiness. The empirical disconnect between getting what one wants and being happy has been well

\(^{49}\) Those familiar with John Stuart Mill’s personal history will know that this is not entirely fantastical. See Ch. 5, “A Crisis in My Mental History,” from Mill’s Autobiography.
documented by the likes of Daniel Kahneman, Tim Wilson, Daniel Gilbert, and George Loewenstein. Again, one could insist on definitions of “happiness” and “preferences” such that one is happy when and only when her preferences are satisfied. But the move here is no more compelling than was the similar move to identify “well-being” and “preference satisfaction.”

(ii) Well-Being vs. Happiness:

Here it is not easy to give uncontroversial examples, since what specifically contributes to a person’s well-being is (at this point) controversial. (My account of well-being will be offered in the next section of this chapter, but I am just running through plausible pre-theoretic intuitions here.) It seems plausible enough to say that at least a moderate amount of exercise contributes to $P$’s well-being. Exercising helps $P$ be physically healthy, and while $P$’s well-being need not be identical to her physical healthiness, $P$ is better off when she is physically healthy than when she is physically unhealthy, ceteris paribus. Still, exercising is neither necessary nor sufficient for $P$’s happiness. Not every person whose well-being is being increased by exercise will be happy to be exercising. This suggests that not every person whose well-being is promoted will thereby be happier. There is also no reason to think that one is always promoting another’s well-being by helping to make that other happy. If $P$ is a smoker and you give $P$ a cigarette, you may very well make $P$ happy (assuming that $P$ is not, e.g., trying to quit smoking), but you will not thereby be promoting $P$’s well-being (which, whatever it consists in exactly, is presumably not improved through the bad health effects concomitant with smoking).

(iii) Satisfaction of (First-Order) Preferences vs. Satisfaction of Second-Order Preferences:
On Sunday afternoons in autumn, \textit{P} always prefers sitting inside and watching football on TV rather than going for long bicycle rides with \textit{P}'s friends. However, she does not prefer to have this preference. \textit{P} wishes, as it were, that she did not care so much about football, an excessively violent sport played by overpaid athletes, and cared more for adventurous and healthy activities like bicycling. Now, one worry is that in light of my general characterization of preferences from earlier in this chapter, \textit{P} should not count as preferring to watch football. But as I imagine it, \textit{P} does not hate her “attraction” to watching football, she is not quite hostage to it, is not \textit{addicted} to it, etc. It actually seems permissible to call this a preference, though one with which she is not entirely uncomfortable. By encouraging \textit{P} to watch football, you are promoting \textit{P}'s preference – her \textit{first-order} preference – but not her \textit{second-order} preference. If my reader finds the term “deep preference” more apt than “second-order preference,” he or she should feel free to use it. I do not insist on the terminology that distinguishes preferences from second-order preferences. In fact, I do not insist that second-order preferences literally are of a different order. What I am most interested to show is simply that one can be conflicted in one’s preferences, such that she has a preference for \textit{X} and another for \textit{not-X}; and in turn that there is no guarantee – certainly no \textit{conceptual} guarantee – that by helping to satisfy one’s preference you are not thereby also frustrating another of one’s preferences, and perhaps one to which one is more deeply attached.

(iv) \textit{Well-Being vs. Satisfaction} of \textit{Second-Order Preferences}:

Suppose again that \textit{P}'s well-being is promoted through regular exercise. Now imagine that \textit{P} is lazy through and through; she is not plagued by one of those nasty consciences that makes a person feel guilty for not exercising. \textit{P} will sit on the couch and
vegetate all day, as doing so is what she really, really prefers (deep, deep down) to do. That is to say, if $P$ is indolent enough, even her second-order preferences will be not to exercise. If all of this is true, then promoting the satisfaction of $P$’s second-order preference will differ from promoting $P$’s interests, which surely consists in part in exercising and being physically healthy. The point can again be made without insisting that so-called second-order preferences literally are of a higher order. The basic point then is that one’s well-being can consist in that which one does not in any sense prefer—no matter how deep or carefully one looks.

I do not mean to suggest that there is no one whose well-being, preferences, second-order preferences, and happiness all harmonize. My principal aim has merely been to suggest that these are all distinct concepts. If there has been an additional aim, it has been to use relatively everyday examples in order to show intuitively how likely it is that an average person’s preferences, second-order preferences, interests, and happiness will at least sometimes in fact pull apart. That these concepts can, and at least occasionally do, fail to harmonize – that they can, as it were, point or pull in different directions – is no insignificant fact. At the very least, it means that the conceptual terrain is complex. We have seen that the case for calling the satisfaction of preferences “intrinsically good” is problematic. We can now more clearly see why this is so, and also why even the claim that the satisfaction of individual preferences is instrumentally valuable – and more specifically, valuable in virtue of promoting individual well-being – is contentious.
Section 4: Flawed Accounts of Well-Being

No more mileage can be gotten out of little examples and intuitions pumps. I have suggested that there is a clear conceptual distinction between preference satisfaction and well-being; but more rigor is now required. What we need at this point, I propose, is a full-fledged account of well-being. We need a fundamental understanding of what an interest is, of what a person’s well-being consists in generally, and of what makes it the case (when it is the case) that a change or influence increases a person’s well-being. These are difficult issues, but ones worth exploring. Without a general characterization of well-being, the proponent of CBA is not in a position to assert that the each individual is the best judge of her well-being and that the best way to increase someone’s well-being is by promoting the satisfaction of her preferences. In what follows, I run through, and ultimately reject, two possible general accounts of well-being, namely the teleological account and the so-called objective-list account. In the next section I explain and settle on a third, namely, the full-information or informed-preference account of well-being.

Considerably more plausible than a conceptual identification of well-being and preference satisfaction – though still, I want to say, ultimately unacceptable – is an understanding of the notion of well-being that is rooted in teleology. On a teleological view, something promotes one’s well-being (or at least does not detract from one’s well-being) if it is consistent with, or contributes to, one’s end(s) or purpose(s). Most religions make claims regarding the end or purpose of an individual human life. Perhaps, for example, each person’s purpose is to serve God and/or to be a manifestation, insofar as it is possible, of Godliness. On most such views, we have whatever purpose we have because God has created us and endowed us with the purpose (just as, e.g., a hammer has a purpose –
pounding nails – in virtue of having been made to fulfill this purpose). But not all teleological views make a connection between a human’s purpose and a creative deity. Perhaps each person’s end or purpose is to integrate the various parts of her personality (the intellectual part with the appetitive part, for example), or to contribute to a set of social conditions wherein one’s own, as well all others’, opportunities for creative self-expression are maximized. Perhaps each person’s end or purpose is to do well the things that are distinctively human, or to actualize one’s moral and intellectual potential.

There are at least two interesting questions to consider. The first is: What will a teleological account have to say about the relationship between preferences and well-being? It is not obvious. The standard accounts of humans’ ends or purposes do not seem to be such that each person’s preferences will always and necessarily be consistent with her end(s) or purpose(s). If this is true, then those endorsing a teleological account will not say that the satisfaction of one’s preferences will always promote one’s well-being. Preferences for smoking, gambling, and playing Russian roulette seem like prime examples of those that would detract from rather than contribute to our ends or purposes—unless one’s end or purpose is to engage in risky behavior! Alternatively, someone may argue that each person’s purpose is simply to satisfy her preferences (or something along these lines), and if this argument were made good, then it appears that the economists quoted in Section 2 of this chapter could be teleologists about individual well-being. The prospects for defending such an argument – such a conception of our ends or purposes – seem extremely dim.

This actually leads us naturally to the other interesting question – the more important of the two, frankly – which is: Does the teleological account offer the best way to understand what well-being is? One fact that the previous paragraph helped to illustrate is that it is far
from clear what any human’s (or humans’) purpose or end is. Moreover, it is unclear how one would discover what any human’s purpose or end is and how one would defend a view about these purposes. One could make assertions, but in the end they would seem to be just that; I would not know how I or anyone else could adequately back them up with arguments. These problems plague any teleological account, and they are serious ones. It does seem odd to say that one’s well-being could be promoted by acting contrary to her ends or purposes, so this is a point in favor or the teleological view of well-being. But the account will not ultimately succeed if it cannot be rested on an adequate account of our purpose(s). I myself think that idea of a purpose makes sense in connection with artifacts like hammers and pencils, but am much more suspicious of the attribution of purposes to humans. One worries, too, that the teleological view entails, implausibly, that things like hammers have interests. A hammer pretty clearly has a purpose – it is, at least, to pound nails – but we would not want to say that a hammer thus has interests as well. On what grounds, however, would the teleological account resist the move from hammers’ purposes to hammers’ interests?

Another problem with the teleological account is that we have a better idea of what our interests are than we do of what our purpose is, or that we even have a purpose. This fact would be hard to explain if interests are tied to purposes in the way suggested by the teleological account. For example, if it is supposed that we have what purpose we have in virtue of being endowed with that purpose by some creative deity, then we seem to know what our particular interests are only when we know what our purpose is, which we know only when we know what the creative deity’s plan was for us (when we were created, say). In fact, we seem (on such an account) to know that we have a purpose only if we know that there is a creative deity. This is a shortcoming of the teleological view. Even if we do not
know *exactly* what is in our interests, we are considerably more sure that we have interests, and in at least some cases what our interests are, than we are that there is a creative deity, let alone what that deity’s plan for us was or is. If purposes are not given by a creative deity, then from where do they come? One possible answer is that each person, as it were, gives her purpose to herself. But this kind of view of our purposes – and the view of our well-being that it entails – is surely not what the typical proponent of the teleological view has in mind. Certainly there are contemporary Aristotelians working on accounts of human flourishing that might underpin claims about individual well-being and ultimately make for a kind of teleological account of well-being; but it is hardly a foregone conclusion that these theorists will meet success.

There are also non-teleological attempts to make sense of the notion of well-being. There are what are sometimes called “objective-list” accounts that more or less avoid appeal to the metaphysical spookiness that can be associated with “purposes.” These might also be called “substantive-good” accounts. Martha Nussbaum\(^50\) and John Rawls\(^51\) can be interpreted as giving accounts of this sort. Objectivity and subjectivity are dangerously slippery terms. We should thus proceed cautiously, but not allow the slipperiness to entirely prevent us from using the terms. I follow L. W. Sumner, who says, “We may say that a theory treats welfare as subjective if it makes it depend, at least in part, on some attitude or concern on the part of the welfare subject. More precisely, a subjective theory will map the polarity of attitudes, so that being well-off will depend (in some way or another) on having a favorable attitude toward one’s life (or some of its ingredients), while being badly off will


require being unfavorably disposed towards it.” An account of well-being is “objective,”
then, if it does not treat $P$’s having a favorable attitude towards $X$ as a necessary condition of $X$ contributing to $P$’s well-being. Nussbaum claims that “we had better take our stand squarely in the camp of the substantive good,” and she claims that defects of competing approaches push her “in the direction of a substantive account of certain central abilities and opportunities, as the relevant space within which to make comparisons of quality of life.”

She proposes a list of abilities and opportunities that “will contain a plurality of distinct items,” and the assessment will not “focus on how people feel about their relation to these goods.” For this reason, her account counts as “objective” in Sumner’s sense.

The list that Nussbaum has given includes: not dying prematurely; being able to have good health, to be adequately nourished, and to have adequate shelter; being able to move freely from place to place; having one’s bodily boundaries treated as sovereign; having opportunities for sexual satisfaction; being able to use the senses and to imagine, think, and to do these things in a “‘truly human’ way”; being able to have attachments to things and people outside ourselves, not having one’s emotional development blighted by overwhelming fear and anxiety, and much more. An objective-list account might also, for example, specify that each individual’s well-being consists in having freedom of thought, choice of occupation, access to education, at least a moderate income, and so on.

The two questions that were asked about the teleological account can be repeated here. So, we can first think about how the economists’ claim that an individual’s well-being

---

53 Nussbaum, ibid, p. 166.
54 Ibid, p. 63.
55 Ibid, pp. 77-79. I should note here – looking forward all the way to Chapter 7 – that Nussbaum has more recently worked to extend her account to animals.
is promoted when her preferences are satisfied fares on an objective-good type account of well-being. If something like good health is objectively good and it is true – as it seems to be – that at least some people prefer eating greasy fast food and jumbo-sized sodas at every meal, then an objective-good theorist will clearly disagree with the economists. Nussbaum’s list does mention good health, but what it holds to be objectively good is “being able to have good health” (my emphasis), so her account is thus far not obviously inconsistent with the economists’ claim. (She uses this formulation repeatedly, which is why she calls her approach the “capabilities approach.”) The person who prefers fast food may well be able to have food that is healthier than the fast food is, and in turn she may well be able to have good health. If this is so, then the person’s well-being is perhaps not diminished by having all the fast food she prefers. This actually points us to one of the big problems with Nussbaum’s account, namely the “being able to” phrase. It is clear why she favors it: it avoids the air of paternalism that accompanies so many objective-list accounts. That will not compensate for its deficiencies, however. Two people may be equally able to have good health, even if one of them is currently in very bad health. (The one who is in bad health may be “able to” recover.) On Nussbaum’s account, they thus would appear to be equally well-off (at least with respect to that dimension of well-being that has to do with health). Only one who is not able to be rehabilitated would not be able to have good health and in turn count as less well-off than the currently healthy person. The “being able to” phrase is obviously troublesome.

Setting aside that problem, there is another reason why an objective-list account of well-being might be able to accommodate the economists’ claims. This is because an objective-list theorist might straightforwardly claim a place on the list for something like “having one’s preferences satisfied” or “being able to satisfy one’s preferences.” One can
imagine, however, that this list item might be in tension with other items that would appear on the list. For example, if good health were also on the list, and an individual preferred frequently eating the greasy fast food, it is not clear what judgment this particular objective-list account would issue regarding this individual’s well-being when she eats the fast food and thereby satisfies her preference for it but diminishes her health. Neither “being able to satisfy one’s preferences” nor any subtle variant thereof directly appears on Nussbaum’s list. But Nussbaum herself does seem to be thinking about this issue and to have wanted to make a place for it somewhere in her account. Prior to introducing the actual list, there is a passage where Nussbaum is considering what she calls the “argument from paternalism,” which she says is committed to the claim that “people are the best judges of what is good for them.” (This kind of claim should sound familiar from the economists I have quoted.) Nussbaum says, “This is an important point, and one that any viable cross-cultural proposal should bear firmly in mind… But we can already say that a commitment to respecting people’s choices hardly seems incompatible with the endorsement of universal values,” i.e., with the endorsement of an objective-list account. “Indeed,” she continues, “it appears to endorse explicitly at least one universal value, the value of having the opportunity to think and choose for oneself. Thinking about paternalism gives us a strong reason to respect the variety of ways citizens actually choose to lead their lives in a pluralistic society, and therefore to prefer a form of universalism that is compatible with freedom and choice of the most significant sorts.” (Note that the concern from the previous paragraph about the “being able to” phrase is relevant here again—she has just substituted the phrase “having the opportunity to.”) If there is not some relatively tight connection between preferences and well-being on an

---

56 Which is what she considers her own proposal to be.

57 Ibid, p. 51.
objective-list account, we should wonder whether such an account will ever be able to accommodate the specific substantive claims about an individual’s well-being that we will often have reason for wanting to make (not just in connection with CBA—though there too, certainly).

We turn to the second and more fundamental of the two interesting questions regarding the objective-list account, namely whether it is a compelling account of well-being. Whatever is in the end said about the compatibility or incompatibility between the economists’ views and the objective-list account, this kind of account is, like the teleological account, simply not a satisfying way to understand what well-being is. The biggest problem for those who offer an objective-list account is in trying to explain how and why a certain thing does or does not get onto this kind of list. It is true that many – perhaps all – of the items that appear on Nussbaum’s list are things that most individuals (or the average individual, or whatever) will prefer. But as alluded to in one of the earlier Nussbaum quotes, they do not get there in virtue of being preferred by the people whose good the items are supposed to comprise (or by being preferred by anyone, we can even say). In fact, Nussbaum appears committed to saying that an individual’s quality of life is higher when she possesses a capability on the list than when she lacks it even if she does not avail herself of the capability, even if she is indifferent towards having the capability, and in fact even if she would prefer not to have the capability. The vexing question, then, is: How do these items get onto the list? What is it, really, that makes it true on this view that those things on the list contribute to or constitute well-being? Perhaps goods get onto the list in virtue of fulfilling the agent’s distinct function. This may or may not be Nussbaum’s move. In any case, the
move just takes us back to the teleological account, with which we have already seen significant problems.

While the inability of an objective-list account to provide a compelling explanation of how an item gets onto such a list is sufficient to make such an account untenable, there are additional problems that weaken this kind of account even further. Sumner is critical of objective-list accounts on the grounds that they are unable “to explain adequately what makes an agent’s well-being especially hers.” Sumner says that “we are seeking an analysis of the nature of welfare—and account of what it is for something to make our lives go well—not merely a list of its ingredients or sources. Yet such a list is all that most objectivists give us.” In an important footnote attached to this last statement, he says “a list of virtues is not a theory of virtue, a list of duties is not a theory of duty, and a list of (human) goods is not a theory of welfare.” What is more, the things that are on this list are presumably things that individuals should be interested in pursuing – it would, after all, be odd to say the individual’s well-being consists in things that she has no reason to pursue – but the objective-list account will only be able to give a very shaky explanation of this reason. I connect this up with one of J. L. Mackie’s famous concerns, namely that “an objective good” – which items on the list can be counted as – “would be sought by anyone acquainted with it, not because of any contingent fact that this person, or every person, is so


59 Sumner, ibid, p. 776.

60 Ibid.

61 I intend ‘reason’ here in the internalists’ sense. Internalism about reasons is the view that any claim of the form “Person P has a reason to Φ” (where Φ stands for some verb of action) is true only where P has some motive, i.e., where something that matters to P. will be served or furthered by Φ-ing. See T. M. Scanlon, What We Owe to Each Other, (Cambridge: Harvard Belknap Press, 1998), pp. 363-373.
constituted that he desires this end, but because the end has to-be-pursuedness somehow built into it.” The items on a list of “objective goods” would indeed be, to borrow Mackie’s language, very “queer” things.

Section 5: The Informed-Preference Account of Well-Being

We move back to accounts that take advantage of the intuitive idea that one’s well-being is connected in a fundamental way to the preferences one has, but bear in mind the problems with treating well-being and preference satisfaction as identical. More sophisticated and plausible is what I will call an informed-preference account, which I will articulate and defend in this section of the chapter. This is an account according to which one’s well-being consists in the satisfaction of those preferences one would have if fully informed and rational. While I am not the first to entertain such an account, I try to fill it out in a convincing way and thus to eliminate some of the confusion and controversy that continues to surround it. On my version of the account: (a) one counts as fully informed when one knows relevant concepts, facts, and likelihoods regarding what life would be like in each of the multiple futures open to one, where these things are vividly imagined and considered in a reasonably cool hour; (b) rationality is interpreted as reasoning in according with the principles of formal logic, as well as consistency among one’s fully-informed


63 Some people refer to this general kind of account of well-being as a “full-information account.” Others who have the same basic approach in mind say “informed-desire account.” “Idealized-preference account” is another legitimate name.

64 Some people talk as if everyone basically agrees that an informed-preference (or “informed-desire”) account is the right kind of account of well-being, and that it is just the finest of details of this account that need working out. Would that that were the case. Those objecting to such an account include scholars as different from one another in terms of their theoretical orientation to morality as T. M. Scanlon and Richard Brandt. Make no mistake: that this is the right kind of account has by no means been clearly established.
preferences; and (c) a preference counts as satisfied when what is preferred actually comes about, not when it is merely believed to have done so, although an individual may sometimes prefer not to know what has actually happened. Against various critics, I argue (d) that morally-motivated or other-regarding preferences should not be excluded, and (e) that the exogenous and adaptive nature of preferences does not pose an insuperable problem for the account.

(a) Fully-Informed Preferences:

One of the basic ideas behind an informed-preference account is that preferences and well-being can pull apart when – as is often the case – we operate in conditions of incomplete information. In such conditions, the satisfaction of an individual’s preference does not necessarily increase her well-being; in fact, it may diminish her well-being. Consider the following policy-relevant example: person $P$ prefers drinking relatively cheap well water to paying for and drinking more expensive bottled water, but knows neither that the well water contains considerably more arsenic than the bottled water, nor that arsenic is a potent carcinogen. Again, one view is that, given $P$’s actual preferences, her well-being is better improved by drinking well water than by drinking bottled water (given their relative costs). More plausible, however, is a view that continues to link $P$’s well-being to her preferences, but to the preferences she would have if she had the aforementioned information regarding arsenic. If it is the case that $P$, were she aware of the arsenic, would prefer the bottled water (even given its relative costliness), then it is reasonable to say that the bottled water better contributes to her well-being than does the well water. The informed-preference account does not suggest that the satisfaction of an uninformed preference will in practice never increase one’s well-being; but it does maintain that the more information one has, the more
plausible it is to say that the satisfaction of her preferences promotes her well-being. Well-being does not consist in the satisfaction of uninformed preferences, but rather (at least partly) in the satisfaction of fully-informed ones.⁶⁵

Despite what the arsenic example may suggest, the very notion of full information is not simple. Roughly, in order to have full information, one needs to have a clear picture of what life would be like – for oneself and for others – in each of the multiple futures open to one in accordance with the different choices one could make. But if well-being is to be defined in terms of the satisfaction of fully-informed preferences, and having full information entails knowing one’s level of well-being in the various states over which her preferences range, then the account is circular. Thus, being fully informed must not include knowing directly about one’s well-being. Full information entails knowledge of facts and concepts. It need not, however, include more than those facts and concepts relevant in the particular context at hand. Information is relevant if knowledge of it might reasonably influence one’s preferences. Most of us do not need to know the color of the seventh moon of Neptune in order to be possessed of complete information as the term is intended. In typical cases, though, a variety of facts within the purview of the physical, life, and social sciences will be relevant, as will certain facts about the past.

What facts about the future are required for full information is more complicated. On the one hand, full information is meant to entail more than hopes or reasonable expectations regarding what the various futures would be like. It is meant to preclude the possibility that someone could get what she with full information preferred but be worse off nonetheless because what she got was not all she expected it to be. But full information is not meant to

---

⁶⁵ I say “at least partly” because this is just a preliminary articulation of the informed-preference account of well-being, not the final version. Several modifications are yet to come.
include so much about the future that one always knows which of the various states over which her preferences range actually will obtain. This means that one should not know, e.g., whether she will actually end up drinking the well water or the bottled water. More interestingly, it also means that one need not know whether the arsenic would, if she drank the well water, actually end up giving her cancer, but she must know the relevant likelihoods. Someone with full information will also know how likely it is for various preferences, if satisfied, to make her contented-feeling or happy, and what new preferences would be generated by the satisfaction of her various earlier preferences.

An objection will be raised here. Suppose that $P$ is offered a gamble the odds of which she knows to be in her favor: if a fair coin comes up heads either once or twice in two flips, she gets $50, but if it comes up heads neither time, she pays $50. She does not know what the outcome will be, but she goes for the gamble. Alas, both flips land tails-up. This is a case, the objection goes, where she got what she with full information preferred but is nonetheless worse off. I believe the objection fails. Insofar as she both preferred with full-information to take part in the gamble and was able to do so, she is better off, *ceteris paribus*. The intuition that she is worse off as a result is explained by the fact that she presumably also had an informed preference for winning the gamble, and this preference was frustrated. If the preference for winning was particularly strong, $P$ might indeed, in virtue of getting to take part in the gamble but losing, count as worse off overall; but this is explained precisely by the frustration of this strong (informed) preference.

It is commonly recognized that more must be done to fill out the informed-preference account’s epistemic condition. Sidgwick suggests that “a man’s future good on the whole is what he would now desire and seek on the whole if all the consequences of all the different
lines of conduct open to him were accurately foreseen and adequately realized in imagination at the present time.” Adequate realization in the imagination is an important condition, and can be understood to have two parts. First, in order for $X$ to better contribute to $P$’s well-being than does $Y$, $P$ must not just prefer it in light of an abstract possession of the requisite knowledge of $X$ and $Y$, but prefer it with this knowledge vividly portrayed in her mind. It is one thing to prefer having a cigarette when one knows but is not really thinking about the fact that cigarettes cause cancer; it is another thing to prefer a cigarette even when one vividly imagines what it would be like to have one’s health wrecked by emphysema. Second, $P$ must prefer $X$ over $Y$ in a reasonably “cool hour,” which is to say that $P$ must not be, as it were, in the grip of a passion. One must not, for example, be clouded by rage or raw animal lust. Now, it would clearly be too strong to say that one must be entirely unemotional or dispassionate. Indeed, such a person would probably not even count as having preferences. Moreover, well-being at least sometimes consists in getting/doing that which one passionately prefers, and one’s best insights sometimes come in one’s “aroused” moments. But the basic idea is that one’s emotions must not be so strong as to entirely drown out one’s judgments. Henceforth, I will have all of this in mind when I refer to the condition of full information.

(b) Rational Preferences:

The conditions that a compelling preference-based account of well-being will put on preferences do not stop here. Assume that someone gambles every day, genuinely prefers to do so, and prefers to even as she is fully informed (about, e.g., her tendencies, her financial

---


67 The notion of a “cool hour” is probably best known from “Sermon XI” of Bishop Joseph Butler’s famous *15 Sermons*. 
situation, other ways her family could use this money, the poor odds involved, etc.). Still, it
would be a mistake to conclude too quickly that this means that this individual’s well-being
is enhanced when she gambles. Even someone who has full information can have a
preference the satisfaction of which would diminish rather than enhance her well-being. In
short, this can happen if the preference is irrational.

By making contact with the notion of rationality, the informed-preference account
risks trading one opaque normative notion, namely well-being, for another. In fact, the
understanding of rationality that I recommend here is relatively simple. It is meant to be as
thin as possible. It is just reasoning in accordance with principles of formal logic, and
consistency among one’s preferences. (If the full-information condition is satisfied,
consistency among one’s beliefs will be guaranteed.) One is not irrational simply in virtue of
an inconsistency between an uninformed preference and a fully-informed one. On the other
hand, if the actual preferences of someone who lacks relevant information are inconsistent
with one another, she is irrational. Here, however, it is not just the irrationality but also the
ill-informed nature of these preferences that keeps the individual’s well-being from
consisting in the satisfaction of these preferences. Most notable for the informed-preference
account is irrationality in the form of inconsistency among one’s fully informed preferences.
One’s preference, in order to be such that its satisfaction increases one’s well-being, must be
fully informed and also consistent with all of one’s other fully-informed preferences. This
includes one’s long-term preferences, which it may be more natural to call the long-term
goals one has set for oneself. One implication is that it is possible to act rationally against
one’s well-being. This occurs, e.g., when one pursues the satisfaction of an actual preference
consistent with one’s other actual preferences, but inconsistent with those she would have if given full information.

Appealing to only a thin notion of rationality is contentious. Some will insist, for example, that the preference of an anorexic individual to refrain from eating is necessarily irrational. She is hurting herself, the claim goes, and this is necessarily or categorically irrational; it is irrational regardless of what her other preferences are. Other candidates for categorically irrational preferences include those for terminating one’s own life, as well as those for an outcome that yields less expected value than another that one knows is open to her. On my view, whether an individual with one of these preferences is irrational cannot be established without looking at her other preferences. Suppose the anorexic knows all about the significant dangers anorexia presents to her health, knows that she is already severely underweight, that many people care for her independent of how thin she is, and even that she is thinner than most people consider attractive.\textsuperscript{68} Imagine that still she prefers to refrain from eating. She is irrational only if this preference conflicts with some other informed preference (such as for living a long and healthy life, for enjoying the taste and nourishment of delicious food, and so on). No thicker version of rationality should be invoked.

Suppose that the anorexic’s informed preferences are inconsistent in that she prefers to be healthy but also prefers to be extremely skinny, and that such skinniness is incompatible with good health (and she knows this). Noting this inconsistency does not settle the issue of what her well-being consists in. She could remedy the irrationality by

\textsuperscript{68} Notice that I do not say that she knows that she is so thin as to be unattractive, since I do not want to commit myself to the view that there is a fact of the matter about this. There may be, but I do not need to contentiously suppose so here. How attractive others find someone is clearly inbounds, since it is the kind of question that sociologists, psychologists, neuroscientists, etc., can fairly easily test.
giving up on the preference for thinness, or the one for good health.\textsuperscript{69} Appealing to her second-order preferences – that is, her preferences regarding her preferences – is not the right move, at least unless there is a good explanation of why a reason that compels the move to the second-order will not compel a move to the third-order, and then beyond that even.\textsuperscript{70} Instead, I advocate moving to longer-term and better-considered preferences. The relevant questions here should be: To which of her preferences has she given more, and more careful, consideration? Which has more claim to being counted as a long-term goal? Which is more central to her considered conception of herself and her life plan?\textsuperscript{71} It is possible, however, that even at this level one’s preferences may be inconsistent. If that is the case, then the inconsistency should be pointed out to her, and she should be pressed to resolve it. If she is unwilling or unable, then I have to admit that her well-being is, like a fraction whose denominator is zero, simply undefined. There is nothing for someone who wants to promote her well-being to promote.

One plausible interpretation of Peter Railton’s work on this topic suggests that he would object to my treatment of the rationality condition, though not because he would want a thicker understanding of the notion. Imagine that $P$ prefers to drive rather than fly, even over long distances, because of a fear of flying. This can be interpreted as irrational even on

\textsuperscript{69} This is on the assumption, of course, that giving up either of these would not create other inconsistencies.

\textsuperscript{70} It should also be noted that just as her first-order preferences conflict, so too can her second-order preferences conflict. She could prefer to prefer to be so skinny and also prefer to prefer to have good health.

\textsuperscript{71} This is the same basic move I would make to deal with the problem posed by preferences that change with time. See, for example, James Griffin, \textit{Well-Being: Its Meaning, Measurement, and Moral Importance}, (Oxford, UK: Clarendon Press, 1986), p. 16. Griffin’s example is an individual who “through much of his life… wanted his friends to keep him from vegetating when he retired but, now that he is retired, wants to be left to vegetate.” Being sedentary is not in his interest the moment he develops the preference for being sedentary; but nor should we insist the fact that he long preferred to be active in his retirement means that just because he has reached retirement being active must continue to be in his interest. Being sedentary can be in his interest so long as his preference for being sedentary is based in complete information and is consistent with whatever well-considered, long-term goals he maintains for himself and his life in retirement.
the thin understanding of rationality. Suppose that when pressed, \( P \) reveals that she fears the prospect of dying in a plane after an unimaginably agonizing plunge to the ground. Implicit in this, we can reasonably say, is that she prefers to be safe. In fact, though, one’s chances of being hurt or dying are significantly less in flying than in driving; and we can even imagine that \( P \) knows this, and indeed counts generally as well-informed. So, in effect, she prefers to be safe and prefers a method of travel that, *ceteris paribus*, is less safe than another that is open to her. This is an inconsistency among her preferences, and thus she counts as irrational. Suppose now that the question is whether her well-being consists in flying or driving as part of an upcoming trip. The view I have been articulating suggests that the answer depends on what she would prefer if she were rational. It may seem, offhand, that if she were rational she would prefer to fly, which means that her well-being consists in flying. But Railton’s point is that it would be a mistake to answer this question by erasing her fear, irrational though it is, which is what my account seems to do. Railton says that “an individual’s good consists in what he would want himself to want, or to pursue, were he to contemplate his present situation from a standpoint fully and vividly informed about himself and his circumstances, and entirely free of cognitive error or lapses in instrumental rationality.”\(^{72}\) It is a distinct possibility that the rational, informed (and sympathetic) version of \( P \) would, in virtue of recognizing how traumatic flying would be for the actual \( P \), prefer for the actual \( P \) to drive. Thus, the claim goes, \( P \)’s well-being is better promoted by driving than flying.\(^{73}\)

First, note again that on my view flying is not, in itself, any more rational than driving. This depends on which coheres and which conflicts with one’s other preferences.

\(^{72}\) Railton, “Facts and Values,” *Philosophical Topics* 14 (1986), pp. 5-29 (p. 16).

\(^{73}\) Thanks go to Susan Wolf for helping me to see this as a rationale underlying Railton’s move.
But that point aside, Railton’s move is problematic. If P’s well-being does not consist in flying, then it does consist in travel via less than the safest mode available. This is an odd thing to have to admit, especially given that P does have an informed preference for traveling via the safest mode. It would indeed seem strange for the hypothetical version of P (call her HP) to prefer for the actual (fear-ridden) P to get on the plane, for though flying is in general safer, P would suffer miserably, and may even have a panic attack severe enough to endanger her. But given the other options HP has for what to prefer for P, it also seems crazy for HP to prefer that P drives. A third option is more enticing. Why would HP, if she truly is a sympathetic version of P, not prefer that P prefer to get on the plane? After all, if P herself prefers to get on the plane, then it is either the case that she no longer has the fear, or that though she still has it she wants to confront it. Combine P’s preference to get on the plane with the facts that flying is safer than driving and that P prefers being safe, and it then clearly seems to be the case that flying best promotes P’s well-being (and this in spite of the fact that the actual P still has a phobia regarding flying). The reader should reflect on what irrationalities plague her/him, and then ask: “What would a rational and fully-informed version of me want for my actual self?” The answer will probably be: “I would want myself not to be plagued by the irrationality.” But if that is true, Railton’s move from defining well-being in terms of what one would prefer if informed and rational to defining it in terms of what a rational and informed version of oneself would prefer for one’s actual self adds nothing. My original view would have the same effect, and be more elegant.

(c) On the Satisfaction of Preferences:

The account defended thus far maintains that one’s well-being consists in the satisfaction of one’s fully informed and rational preferences. An additional question for a

---

74 Her panic could become severe enough to frighten another passenger into hurting her, for example.
defender of an informed-preference account of well-being concerns how ‘satisfaction’ is to be understood. When does a preference count as satisfied, and when does one count as frustrated? Is it when the condition that one prefers actually obtains, or when one simply believes it obtains? In James Griffin’s words, the question is whether satisfaction is a state of the world or a state of mind. Griffin quotes Nozick, who asks, “What else can matter to us other than how our lives feel from the inside?” Griffin says that “[Nozick] replies, surely rightly, that we also want to do certain things, to be certain things.” While Griffin talks about desire fulfillment, he means by this essentially what I mean by preference satisfaction; and he maintains that “being fulfilled cannot be understood in a psychological way.” He says, “A desire is ‘fulfilled’ in the sense in which a clause in a contract is fulfilled: namely, what was agreed (desired) comes about.”

I agree with Griffin, but the issue deserves more elaboration than he gives. The elaboration is needed to combat the appearance of unacceptable implications of this understanding of satisfaction (i.e., of being fulfilled). Imagine two cases. In both, P’s rational and fully-informed preference is that her spouse, S, not cheat on her, and in both, S is indeed faithful. In case 1 P rightly believes that S has not cheated, while in case 2 P mistakenly believes that S has. The question is: are 1 and 2 both cases where P’s preference that S be faithful has been satisfied? According to Griffin’s interpretation of satisfaction, they are, and I agree. But it may initially appear that Griffin would thus need to say one of two things, both of which are implausible. One, he might say that because P’s preference has been satisfied in both cases, P is no worse off in case 2 than in case 1. This move just seems implausible; P’s belief, mistaken though it is, should not be irrelevant to her well-being. P’s

75 Griffin, ibid, p. 9.

well-being is diminished by her coming to believe that \( S \) has cheated. Alternatively, Griffin could say that \( P \)’s level of well-being is lower in case 2 than case 1, and could explain this by jettisoning the connection between well-being and preferences. That \( P \) is less well off in 2 has nothing to do with the fact that her (or anyone’s) preferences were not satisfied. From my perspective, this is obviously a dire move. It amounts to abandoning my account of well-being. In fact, neither of these moves is necessary. The fact that \( P \)’s preference that \( S \) be faithful was satisfied does not mean that case 2 does not involve the frustration of some of \( P \)’s other preferences that were in fact satisfied in case 1. We can suppose, reasonably enough, that in addition to having a preference that \( S \) not cheat, \( P \) had a preference for not thinking \( S \) was cheating. Additionally, \( P \) presumably had a preference not to be unhappy that was frustrated when she came to believe that \( S \) cheated. The fact that these two preferences were satisfied in case 1 but not in case 2 suffices to explain why \( P \) is worse off in case 2.

The story will now be modified so as to generate two new cases, 3 and 4. In both of these, \( P \) continues to have a rational and fully-informed preference for \( S \) to be faithful to her, but now \( S \) in fact cheats on \( P \). In case 3, \( P \) rightly believes that \( S \) has cheated, while in case 4 \( P \) mistakenly believes that \( S \) has been faithful. Is \( P \)’s preference frustrated in both? Griffin and I will both say yes. But does that mean that \( P \)’s well-being is identical in these two cases? No. It depends on what her other preferences are. For example, would she prefer to know \( S \) has cheated on her, or to continue believing \( S \) is faithful? Suppose that even if \( P \) knew \( S \)’s behavior towards her will be no different than if \( S \) were not cheating,\(^77\) \( P \) would prefer to know that \( S \) has cheated. If so, she is worse off in case 4, because there her preference for knowing of \( S \)’s cheating has been frustrated. But nothing requires this to be

\(^{77}\) This means, for example, that \( S \) is as affectionate, as interested in hearing about \( P \)’s day, as willing to help around the house, etc., as \( S \) would be if \( S \) were not cheating.
Nozick’s contention is too strong: most of us do prefer “to do certain things, and not just have the experience of doing them,” to “be a certain way, to be a certain sort of person,” and to make “actual contact with... reality,” but not everyone does or must; and the well-being of those who are fully-informed and rational but do not have such preferences can consist in being deluded.\(^{78}\)

The objection may be raised that if one can prefer not to be fully informed – as I have supposed is possible, though surely not required, for \(P\) – then it was a mistake for the account of well-being to have maintained up front that well-being consists in the satisfaction of informed preferences.\(^{79}\) If the view is that fully-informed preferences are more important than uninformed preferences for fixing well-being because people prefer to be fully informed, then there is indeed a problem, for at least two reasons. First, as I have noted, not everyone does (or needs to) have this preference. Second, it would be circular to say fully-informed preferences are important because they are the kinds of preferences people prefer to have. The objection is an important one, and not one to which there is a short compelling response. A full response would involve discussing each competitor to a preference-based account, and showing the severe flaws with each. This would leave us looking at an account that connects well-being and preferences, where we would again confront the considerable shortcomings of an actual-preference account. We want an account that acknowledges

---


\(^{79}\) Thanks to Marc Lange for pressing me to think harder about this issue.
without overstating the important role that information plays in linking preferences and well-being. This is achieved, precisely as I have done, by putting a structural constraint on preferences without constraining their content. One must be fully informed for one’s preferences to count, but the content of one’s fully informed preference may be to lack certain pieces of information.

Two other objections should be briefly considered here. The first is that if satisfaction is understood in something other than a psychological sense, certain of P’s future-oriented preferences – like one for her grandchildren to grow old in good health – could be satisfied or frustrated after P has died. This means, implausibly, that P’s well-being can be affected after she has died. Several responses can be given. Firstly, I see nothing unacceptable about the stipulation that one’s well-being consists in the satisfaction of one’s well-informed and rational preferences, up to but not beyond the time when one dies. Secondly, though, the idea that one’s well-being can be affected after one has died is not so obviously wrong. I do not have a strong view here, but the idea does not strike me as crazy. Even if this is the case, however, one’s well-being would not be liable to change in all the ways it could when one was alive. For example, any preference P has for knowing something will not be one that could be satisfied after she is dead.80

The second objection is not necessarily related to my interpretation of satisfaction, but fits better here than elsewhere, since like the worry about the well-being of the dead, it is based in a concern that preferences range too far. It is that the objects of a person’s informed preferences are likely to include many things that are not related to this person’s well-being, 

---

80 This is true, of course, unless one invokes certain metaphysical claims about the possibility of P existing in another life after this one.
intuitively understood. A paradigm case involves meeting a stranger on a train and forming a preference that she succeeds but then never seeing her again. I must confess: I simply do not feel the force of this objection. If I meet a stranger on a train and form a preference that she succeed in life, my well-being is at least slightly enhanced if she does succeed, even if I never know she succeeds and indeed even if I never hear from or about her again. But if this is not a strong preference of mine, my well-being is not enhanced much. And it is not enhanced as much as it would be if I knew that she succeeds, since I would presumably also have a preference for knowing that she succeeds.

(d) Morally-Motivated and Other-Regarding Preferences:

It is often supposed that a preference-based account of well-being needs to restrict the relevant preferences to those not tied to a sense of moral obligation to prefer/do one thing rather than another. According to J. S. Mill, “Of two pleasures, if there be one to which all or almost all who have experience of both give a decided preference, irrespective of any feeling of moral obligation to prefer it, that is the more desirable pleasure.” David Sobel claims that “the root idea behind this test is that if a person intrinsically prefers X to Y, independent of moral considerations, while fully acquainted with both options and in the ‘cool hour,’ then X is more conducive to the agent’s well-being than Y no matter what other properties X and Y have.”

---

81 This is almost word-for-word from T. M. Scanlon, but I have put the point in terms of preferences rather than desires, and in terms of well-being rather than quality of life. See What We Owe To Each Other (Cambridge, MA: Belknap Press, 1998), esp. p. 115. Griffin is quite concerned by this type of objection.


83 Ibid. I actually understand Mill, in the aforementioned passage, to be proposing a test for distinguishing higher from lower pleasures, which is not the same as giving an account of well-being. Still, some find this to be a reasonable condition to build in to an informed-preference account of well-being. See, e.g., Richard Brandt, A Theory of the Good and the Right. (Amherst, NY: Prometheus Books, 1998), especially Ch. XIII, “Welfare: The Concept, Measurement, and Interpersonal Comparisons.”
I believe it is a mistake to add this condition to an informed-preference account. Suppose $P$ is rational and equally well-acquainted with stealing and not stealing from others: $P$ has done both, knows vividly how each feels, knows the punishments for stealing and the chances of getting caught, knows the harm she has caused to those from whom she has stolen, and so on. Imagine that at this stage in $P$’s life, she prefers not to steal, and this preference is based on a belief that it is morally wrong to steal. Absent this belief, she would (at least sometimes) prefer to steal (meaning that she sometimes enjoys the feeling of doing so, is not worried about getting caught, etc.); but given the belief, she prefers not stealing. Some want to say that on those occasions where $P$ would prefer stealing over not stealing if not for the feeling of moral obligation, her well-being would be better promoted by her stealing. What should be said about this case? Perhaps there is a way of filling in the details so as to make it most plausible to say that $P$ does not actually prefer not stealing. Since preferences are not understood in an entirely behaviorist way, this is possible even if $P$ actually never does steal. But if $P$ is fully informed and rational, then so long as she counts as preferring not to steal – and the fact that she feels morally obligated not to steal does not preclude this – it should be conceded that her well-being consists in not stealing.\footnote{If there are moral facts, one who has full information would then necessarily need to know at least the relevant ones. And if there are such facts, it seems likely that its being wrong to steal from others (except perhaps in certain exceptional circumstances) would be among them. If there are not moral facts, or moral facts are (at least partly) subjective in that they depend (at least partly) on a person’s preferences – perhaps her rational and fully empirically informed preferences – then the mere fact that someone (rationally and with full information) prefers X over Y on the basis of a feeling or judgment that she is morally obligated to do so is sufficient to make doing/getting X more conducive to this person’s well-being than doing/getting Y.}

If we were trying to give a preference-based account of moral rightness, the requirement of independence from a feeling of moral obligation to prefer would be needed. We would want to say that if what is preferred is preferred due to moral considerations, i.e., feelings of moral obligation, then the fact that it is preferred (even where the other conditions
are satisfied) is not a reliable guide to its being morally right. But I am aiming for an account of individual well-being, not moral rightness, and so appealing to preferences that are not independent of a feeling of moral obligation is not circular. If this condition is included in an account of well-being, we flatly rule out the possibility that P’s well-being consists at least partly in her sense of herself as a morally decent person, and in doing what she believes to be morally obligatory. To do so would be a mistake. To require that preferences be independent of feelings of moral obligation is to build too much into the informed-preference account of well-being.

This is the same basic response that should be given to those who build into an informed-preference account the requirement that preferences be self-interested. But here my response is even stronger. A self-interested preference seems to be based only on considerations of one’s own well-being. But if that is so, those who define the notion of well-being in terms of satisfaction of self-interested preferences invoke in their definition of well-being the very notion that they are trying to define. Nonetheless, there is a reasonable concern that can underpin the move to restrict preferences only to self-interested ones. The worry is that if the preferences whose satisfaction well-being consists in are not so restricted, it will be logically impossible for someone who is rational and fully-informed to prefer to act in a way that diminishes her well-being. This seems to render altruism impossible, since altruism is reasonably understood as voluntary behavior sacrificing at least some of one’s own well-being for the sake of the well-being of another.

My account does indeed deny that one can rationally and with full information prefer to sacrifice (even a little of) one’s well-being. This does not mean, however, that it denies

---

the possibility of altruism. Why not? On the view I have urged, well-being is not identical to any actual mental or psychological state. One’s well-being can be increased without one actually feeling better, and without one actually counting oneself as better off. It is possible, in fact, for one to count as generally well-off without ever experiencing any sensation of satisfaction. This might be the case if none of one’s actual preferences is satisfied, but what one actually prefers is always opposite of what one would prefer if one were rational and fully informed. Well-being is distinct from happiness, which, whatever it is exactly, surely must involve some actual psychological good feeling. An individual can feel great, sincerely claim to be contented, frequently and genuinely laugh, etc. – in short, be happy – but unknowingly have a terminal disease (something she surely would rationally and with full information prefer not to have). Though the disease-stricken individual will presumably eventually be unhappy, well-being is also distinct from long-term happiness. I suspect that many people are constituted such that the same things that promote their well-being will promote their long-term happiness; but nothing guarantees that one could not rationally and with full information prefer to commit oneself over the course of one’s life to projects that one deems to be worthwhile but that will not make one happy.

Notice that preferring to sacrifice one’s happiness for the sake of another – either on a certain occasion or over the course of one’s entire life – is a fine description of altruism. It is because my account permits such preferences that the account allows that rational and fully informed people can be altruistic. Once the possibility of altruism is saved, I see no grounds for taking the fact that my account precludes the possibility of a rational and fully-informed preference for sacrificing one’s well-being to be a reason for objecting to the account.

(e) The Exogenous and Adaptive Nature of Preferences:
An additional objection to an informed-preference account is that the exogenous and adaptive nature of preferences makes them ill-suited to play such a key role in an account of well-being. Why should we think that one’s well-being consists in what one would rationally and with full information prefer when it is the case that one’s preferences are shaped, sometimes very significantly, by a whole variety of external forces, including advertising, social norms and pressures, one’s upbringing, and possibly even factors like a desire to impress others (including others who are cruel, vain, excessively materialistic, etc.)? Furthermore, we know that people’s preferences adapt to their circumstances. If, for example, you are told repeatedly and from an early age that you are not worthy of, and have no chance to possess, political power, then you may well come to prefer not to have political power. As Rousseau famously says, “Slaves, in their bondage, lose everything, even the desire to be free. They love their servitude.” This adaptation of preferences is intelligible – it is a critical coping mechanism – but surely the well-being of such an individual does not consist in being a slave.

These are serious challenges to an informed preference account of well-being, though not so powerful as to prove that this sort of account must be abandoned. At the risk of trivializing the issue, consider my extreme love for Bob Dylan’s music. It is a bit embarrassing to admit to having collected and listened to recordings of over 1000 complete concerts of his. A wide variety of factors have worked to influence my preference for listening to Dylan. A long (and true) story could be told about how I came to have this

---


preference: it might mention what my parents listened to when I was young, the musicians to whom my friends exposed me, the English classes I took that fostered in me an appreciation for kinds of poetry to which Dylan’s lyrics are akin, and so on. The story would show that my preference for listening to Dylan is exogenous, but it would not make it any less plausible to say that my well-being is increased when my preference for listening to Dylan is satisfied. To the contrary, telling this story would actually help explain why my well-being is increased when I get to listen to Dylan. My wife was not a big Dylan fan when she met me, but spending time with me meant spending time listening to Dylan. She could have avoided Dylan by spending less time with me, or she could have spent time with me but let the music make her miserable. What actually happened was that as I continued to bombard her with Dylan, her preferences adapted such that she too came to enjoy and prefer this music. That her preferences adapted does not prevent us from saying that her well-being is now enhanced when her preference for listening to Dylan is satisfied.

These stories show that it would be a mistake to require that a preference not be exogenous or adapted in order for an individual’s well-being to consist in the satisfaction of that preference. Still, serious concerns remain. What should be said, for example, about the person whose preferences were significantly influenced by her parents vigorously pushing her to play the piano, or be a gymnast, or a beauty queen? What about the well-being of an individual who through abuse has developed extremely low self-esteem and as a result prefers to stay in a relationship where she is not respected, or worse? What about a person who develops certain preferences through brainwashing or hypnosis? One way of handling these cases is simply to add to the informed-preference account the requirement that preferences not arise from hard pushing by parents, abuse, low self-esteem, brainwashing, or

---

88 It might even plausibly be said that these factors have entirely created my preference for listening to Dylan.
hypnosis. But simply compiling a list of disqualifying influences on preferences is unacceptably *ad hoc*. A more elegant move would be to identify what it is that these putatively problematic influences have in common, and to explain why this common element is problematic. One such attempt would be to say that they are cases where the preferences are not autonomous. The worry is that as autonomy is most intuitively understood, neither my own nor my wife’s preferences for Dylan are autonomous. Indeed, ‘autonomous’ and ‘not exogenous’ are essentially synonyms. I see no good way of filling out autonomy – or any other condition, for that matter – so as to exclude the preferences of the brainwashed, hypnotized, and abused individuals, but to include my own and my wife’s preferences to listen to Dylan.

This is fine. It is not the exogenous and adaptive nature of the preferences involved in the problematic cases that makes the cases problematic, but rather two other features of the cases. The first is that we are unsure whether the individuals in question – those who were pushed extremely hard by others, or were abused, brainwashed, hypnotized, etc. – really count as having the preferences that their behavior seems to suggest they have. For a preference to be such that its satisfaction increases *P*’s well-being, it must genuinely count as *P*’s preference, not the parents’, the abuser’s, brainwasher’s, hypnotist’s, or anyone else’s. And it must of course also be a preference. It must be more than a mindless pull, must be constituted by particular beliefs, and must involve an attraction that one endorses. If we were sure the individuals in the problem cases really had preferences, the grounds for concern would be diminished.

The second feature of the problem cases that makes them troublesome is that the individuals involved seem not to have, as it were, a clear view of the requisite information.
A particularly important resource of the informed-preference account is that someone who satisfies its conditions will have full information about her reasons for preferring what she does (where these reasons could be understood in psychological as well as causal terms). Insight into one’s reasons for preferring $X$ to $Y$ may alter one’s preference for $X$ over $Y$. If the abused individual fulfills the informed-preference account’s conditions, she will know (among other things) that she was abused, that she now has low self-esteem, that and how the abuse contributes to her low self-esteem, and that and how her low self-esteem contributes to her staying in this relationship. It is possible, though not guaranteed, that were she to know this she would prefer not to stay in the relationship. If so, then her well-being is diminished by staying in the relationship. The hypnotized person will know that she was hypnotized, will know what was said to her under hypnosis, and will know how the hypnosis modified her preferences. I believe that it is the fact that these things are typically unknown that makes it so uncomfortable to have preferences generated under hypnosis, through past abuse, etc., play a constitutive role in an account of well-being.\textsuperscript{89}

More could be said by way of elaboration and defense of the informed-preference account of well-being, but if what was argued for in the preceding discussion was reasonable, it should suffice to establish this account as offering a compelling view of the nature of well-being. The account is not a naively subjectivist one, but it is also not implausibly objectivist. I count it as a partial- or quasi-subjectivist view of well-being. ‘Satisfaction’ is given an objectivist construal: that which one prefers can obtain even when one does not know it has

\textsuperscript{89} There are reasonable reservations about the informed-preference account based in a concern that it does not seem to be a plausible account of the well-being of children. The problem, I think, is not with the account, but with the fact that it is so hard to even imagine children having and processing all the information the account requires, and formulating an internally consistent set of preferences in response to it. But in any case, I will be happy if it is granted to me that my account is plausible with respect to adults. I am not entirely averse to the idea of giving a different type of account of well-being for children and animals, (and even possibly plants and ecosystems).
done so. The facts that one must know in order to count as well informed are objective ones: for example, that arsenic is a carcinogen is true regardless of what anyone happens to think or of how anyone happens to feel about this matter. The demand that preferences be rational also has an objectivist strain: the fact, for example, that a preference for $X$ and a preference for not-$X$ are not consistent with one another is true regardless of anyone’s feelings or preferences regarding this pair of preferences. In spite of all of this, we cannot get away from preferences. We cannot avoid appeal to the fact that a thing is, or would in certain conditions be, preferred by a person in our explanation of why it is that this thing contributes to this person’s well-being. If we cannot point to some preference – at least some idealized preference – of this person that counts as satisfied, there is no compelling reason for thinking that this person’s well-being has been enhanced. This is the irreducibly subjectivist element of the account. Taken together, this all makes for an account that travels just the right middle course between the implausible ethereality of pure objectivism and the unacceptable viscerality of pure subjectivism.
CHAPTER 3: THE ROLE OF INFORMED, RATIONAL PREFERENCES IN PUBLIC POLICY

Section 1: Introduction

It is time to take stock. I began with the question: With respect to CBA, what counts as a benefit, what counts as a cost, and why? It was observed that the claim made by proponents of CBA is that with respect to any potential governmental regulation or public investment project, the preferences of individuals who would be affected by the policy decision ought to be taken into account. Individual preferences for or against the policy, it is claimed, make it the case that the potential project will proceed or that it will not. A benefit (for an individual—and again, for my purposes, it is always benefit and cost for an individual that are relevant, since I am leaving off to the side the question about total, i.e., aggregated, costs and benefits are figured) is obtained when (that individual’s) preference is satisfied, and costs are incurred when such a preference is frustrated. If, for example, the state is considering spending money to clean up a polluted river, and this river once provided – and if clean could once more provide – a major form of recreation to those around it, proponents of CBA say that the issue of whether this project should go ahead should be settled by appeal to the preferences of the people affected. The cost-benefit analysts’ task, on this view, is to work to measure and record the preferences of those people who are affected by the condition of the river. But the question arose why individual preferences should play this role in public policy. Are supporters of CBA right to count the satisfaction of a preference as
a benefit, and if so, why? (I put off the question of whether preferences can all be expressed monetarily, and will continue to do so, though it will be addressed eventually.)

I focused in on one key assertion made (in effect, at least) by proponents of CBA, namely that the satisfaction of individual preferences is instrumentally valuable. The state, many advocates of CBA claim, ought to promote the satisfaction of individual preferences because doing so has good consequences. The promotion of well-being, we are told, is the ultimate aim of the state through its public policy decisions, and promoting the satisfaction of individual preferences is asserted to be the best way to increase individual well-being (and overall social welfare by extension). Each person, we were told, is the best judge of her own well-being. The idea, then, is that benefits (to an individual) are obtained when (that individual’s) well-being is enhanced, and since an individual’s well-being is enhanced when her preferences are satisfied, benefits are in turn identified with the satisfaction of preferences. After outlining all this, I then argued that the connection between the satisfaction of an individual’s preferences and the increase of that individual’s well-being is more tenuous than CBA’s proponents admit: merely a little reflection suffices to show that satisfaction of one’s preferences and improvements in one’s well-being are conceptually distinct. Several ways of understanding the notion of well-being were then reviewed. These were teleological, objective-list, and preference-based accounts. I worked at length to characterize one preference-based account – an idealized preference account that was labeled the informed-preference account – in a way that allows it to handle the most important objections that might be leveled against it, and concluded that it is indeed a compelling account of individual well-being. The informed-preference account holds on to the intuition that preferences and well-being are related to one another, but maintains that the satisfaction
of an individual’s preference does not necessarily increase that individual’s well-being. Instead, this account holds that an individual’s well-being consists in the satisfaction of those preferences she would have if she were fully informed and rational.

What is called for now is a discussion of the implications of this account of well-being for our understanding of what counts as a benefit and what as a cost. Given what we now know about the nature of well-being, is it true – as the standard view within economics maintains – that each individual is the best judge of her own well-being? Is the best way to promote a person’s well-being by promoting the satisfaction of her preferences? Even granting that the cost-benefit analyst’s methods for measuring people’s preferences – namely the revealed- and stated-preference methods – do so accurately, is there reason to think that they offer the analyst the best possible view of what will promote individual well-being? If a person’s preference is for the river not to be cleaned up, should the clean up of the river count as a cost for this person, even though, if this person is ill-informed or irrational, the satisfaction of this preference may actually diminish her well-being? And beyond this, we need to ask whether it is actually even the promotion of well-being at which the state, through its policies (and especially its regulatory and investment decisions), ought ultimately to be aiming. If it is not, what should be the ultimate goal of public policy? What really should count as a cost and what as a benefit? (To reiterate from Chapter 1, I will not be addressing in this project the issue of whether gains by one person can justify a (smaller) loss for another person, or whether/how individual well-beings can be aggregated into “social welfare.” My initial question is about what ought to count as a benefit and what ought to count as a cost, and a later question will be about whether costs and benefits can all be expressed monetarily.)
Proponents of CBA worry, I believe, that a move beyond the purely subjectivist conception of well-being and to the more nuanced account that I have defended opens the door to paternalism (at best, and at worst, even tyranny) in government. The state can now simply tell people what promotes their well-being and what does not. It can just tell people what counts as a benefit for them and what counts as a cost. In this chapter, I argue that the move does not do all of this. Several interpretations of the claim that each person is the best judge of her own well-being are explored. I show that this claim is not true in any of the ways in which those who defend the claim believe it to be, and thus that this claim will not entirely suffice to show that the state – even conceding that its proper ultimate aim is to promote well-being – should identify benefits to an individual with the satisfaction of that individual’s actual preferences. Now, there are many cases in which the state will not have a better sense than the individual herself of what this individual would prefer if she were fully informed and rational; and even where it does, the state will in most cases not have a perfect sense of this. These facts guard against too quick an inference from an account of well-being (like mine) that is not purely subjectivist to an endorsement of paternalism. They make reasonable the claim that someone who believes that the state’s ultimate aim in making policy decisions is to promote well-being should identify benefits with preference satisfaction. But they do not definitively vindicate that claim.

It is possible (at least in theory, and will probably sometimes be the case in practice) that the state will have an accurate sense of what will promote an individual’s well-being, while the individual’s own sense of this is quite muddled. This means, given the nature of well-being, that the state might be in a position to best promote a person’s well-being by frustrating the person’s preferences. If a person is very ill-informed and/or extremely
irrational, the state could at least in theory promote the this person’s well-being by frustrating many of her preferences. I maintain that this is not always permissible even if it would promote people’s well-being, and that this shows that it is a mistake to think that promoting well-being is the ultimate aim of public policy decisions. Instead, the state, in its policy decisions, must balance aims that sometimes compete with one another. Promoting preference satisfaction will sometimes promote well-being, but frustrating preferences might sometimes be clearly demanded by the reasonable aim of promoting well-being. Even in this latter kind of case, however, respect for autonomy and a recognition of the importance of the experience of satisfaction recommends satisfying even welfare-reducing preferences. I maintain that the state can and should endeavor to achieve all of these aims by helping people to have informed and rational preferences and then aiming to satisfy these. This means, in other words, that as far as possible, benefits to an individual should be identified with the actual satisfaction of the individual’s actual well-informed and rational preferences rather than with the satisfaction of ill-informed or irrational preferences, or with that which would satisfy the person’s preferences were the person well-informed and rational. However, to the extent that it is not possible to do so, i.e., to the extent that the preferences an individual ends up with still are or seem to be ones that pull apart from her well-being, the state must balance identifying benefit to an individual with the promotion of the individual’s well-being (over and against the satisfaction of her actual preferences) and with the satisfaction of her actual preferences (over and against her well-being).
Section 2: Each Individual Is the Best Judge of Her Own Well-Being: Four Interpretations

All the economists quoted in the previous chapter (specifically Section 2) maintain that each individual is the best judge of her own well-being. It apparently follows from this that if we want to know what best promotes an individual’s well-being, we should ask that individual what she prefers to get, do, pursue, have happen, etc. Is it true that each individual is the best judge of her own well-being? An important part of assessing the claim is to disambiguate it. It could mean at least four different things. In this section, I discuss each of these interpretations. I ask of each whether the claim, so interpreted, is true. In the next section, then, I move to assessing the implications of the claim’s truth or falsity for the theory and practice of CBA. The four interpretations, in short, are as follows: (1) each person $P$ is the best conceivable judge of her own well-being; (2) each person $P$ is a better judge than any other actual (i.e., non-hypothetical, non-idealized, non-counterfactually-described) person is of $P$’s own well-being; (3) each person $P$ is always a better judge than the state is of $P$’s well-being; and (4) each person $P$ tends to have a better sense of what promotes $P$’s well-being than anyone else does of what promotes $P$’s well-being, and thus no one is on the whole a better judge of $P$’s well-being that $P$ herself is.

Let us begin with the first interpretation, according to which the claim is that each person $P$ is the best conceivable judge of her own well-being. This means that no better judge could even be imagined. On this interpretation, the economists’ claim is plainly false, at least where $P$ is not actually perfectly informed and rational, and where we are talking about $P$ as she actually is, not an idealized version of $P$. Actual people are not (or at least not always, and probably even not often) perfectly informed and rational. This is an empirical claim, but I believe it to be one that will be readily conceded by anyone who understands the
informed-preference account of well-being and who thus knows how strong these conditions are. So long as a person $P$ is not in the relevant respects fully informed or entirely rational, we can at least imagine a better judge of $P$’s well-being, namely an entirely rational and fully informed version of $P$. Certainly, we cannot imagine a better judge of $P$’s well-being than the fully-informed and entirely rational version of $P$ would be. This is strictly entailed by the informed-preference account of well-being. This is an important point, and one to which I will return. But it is important in a subtle way. In any case, it is plainly not the claim the economists have in mind when they maintain that each person is the best judge of her own well-being.

A second possible interpretation deserves much more attention. Here, the claim is that each person $P$ is a better judge than any other actual (i.e., non-hypothetical, non-idealized) person is of $P$’s own well-being. Interpreted in this way, the claim is interesting and important. Nothing I have said unequivocally refutes it. It is by no means obvious what substantive judgments regarding $P$’s well-being will be issued by the informed-preference account. Actually, even if some other person is well-informed and rational, nothing ensures that this person will know what will promote $P$’s well-being. Now certainly two people who have exactly the same mental/psychological sets and who are in the exact same external circumstances will have all the same preferences, and as a result their well-being will consist in the same thing. But it is possible for two people both to be rational and fully informed but nonetheless have different preferences and thus have their respective well-beings consist in different things; and it is possible for one fully-informed and rational person not to know what another person $P$ would prefer if $P$ were fully informed and rational. What is more – and this, while related to the previous point, is not identical to it – it will often be difficult for
a person $Q$ to know, in any particular case where $P$ specifies or evinces a preference, whether $P$ and $P$’s preference are such as to satisfy the conditions laid out by the informed-preference account. If $Q$ observes $P$ specifying or demonstrating a preference, it will be difficult for $Q$ to know whether the account’s conditions have been satisfied, and thus difficult for $Q$ to make judgments regarding the extent to which $P$’s preference lines up with $P$’s well-being.

To take a rather silly example, two people reasoning equally well and both possessed of full information regarding The Who’s album “Who’s Next” could each fail to know that the other is (in the relevant respects) informed and rational. Each could thus fail to know whether the other’s preference for or against listening to this album would, if the preference were satisfied, enhance the well-being of the person who had the preference. They could also reasonably disagree on how listening to the album would affect their respective well-beings. I take myself (of course!) to be, in this context, informed and rational, and thus given my preferences, listening to the album increases my well-being markedly; but I know people who would say that listening to it would actually diminish their well-being (and much as I might be tempted to, I cannot call them irrational just for failing to enjoy the album)! Moreover, my being fully informed and rational does not give me clear \textit{a priori} insight into what this other person’s informed and rational preference regarding the album would be. Similarly, two rational and fully-informed people could reasonably disagree on how a reduction of arsenic from 50 to 25 parts-per-billion in their drinking water would affect their respective well-beings (even where the cost would be the same for the two people); and it is possible that, prior to each expressing/stating her preference, neither will know (precisely, or perhaps even roughly) what the other’s informed and rational preference will be.
This all gives some plausibility to the claim that each person $P$ is a better judge than any other actual person is of $P$’s well-being. This is important—enough so as to compel me to return to this point later. Even so, the claim is certainly not definitively vindicated. What has been shown is that people who are uninformed or plagued by irrationality are not perfect judges of their well-being. I have, moreover, made the empirical claim that in many cases people are to a significant extent uninformed and plagued by certain inconsistencies among their preferences (though this was not meant as a suggestion that there are not even particular occasions where anyone counts as informed and rational in the relevant respects). I have further suggested that it is not easy for another person to know what $P$’s well-being substantively consists in. But the informed-preference account is one on which it is also difficult for $P$ – if she lacks full information or is irrational – to know what it is in which her own well-being substantively consists. It is even difficult for $P$ herself to know whether the account’s conditions have been satisfied, i.e., to know whether or not she herself counts as fully informed and rational. This opens the door to at least the possibility that some other actual person $Q$ is a better judge than $P$ of $P$’s well-being. $Q$ might know that, and at least a bit about how, $P$ is misinformed and/or irrational, might have better knowledge of this than $P$ herself, and might have better insight than $P$ into what $P$ would prefer if $P$ were better informed and rational. $Q$ might be better informed than $P$ and not plagued by $P$’s inconsistencies; but so long as $P$ has no idea that $P$ is ill-informed or plagued by inconsistencies, $Q$ does not need to count as perfectly well-informed and rational, and indeed does not even need to count as on the whole better informed and more rational than $P$, in order to have (at least on occasion) better insight than $P$ into $P$’s well-being.
One of P’s best friends might know P well enough to know that if P were aware of (say) the amount of arsenic in the well water P currently prefers to drink, P would instead prefer to drink the bottled water, even (say) if it would cost P an additional $0.05 per gallon. This friend could, for example, have knowledge of how much P spends on products at the health-food market sufficient for counting as having justification for the true belief that P would prefer the bottled water at $0.05 more per gallon. This seems plausible. In fact, if the bottled water were no more expensive than the well water, I suspect that nearly anyone who was aware of the arsenic in the well water would be in a position to know that P would, if P herself knew of the arsenic, prefer to drink the bottled water. Most people, if they had information that their spouse or best friend did not have, would express confidence in their ability to judge, in at least certain cases, what their spouse/best friend would prefer if the spouse/best friend had this extra information. Imagine that P and her spouse S have been happily married for 50 years. They know each other very well. S goes to the grocery store to do the weekly shopping and has been told by P what P prefers for S to pick up. Grapes are not on the list P supplies but turn out unexpectedly to be on sale. It is entirely realistic to imagine that S knows that while P’s expressed preference was for S not to buy grapes, P would have preferred that S buy them if P would have known that they would be on sale. This is a case where S is a better judge than P of what will promote P’s well-being.

Clearly for many of us, on many occasions, it will be difficult to issue justified, true judgments regarding what someone else’s well-being (substantively) consists in. It will be difficult for us to know better than the person herself what will increase her well-being. This difficulty should not be underestimated. But it should not be overestimated either. Several more scenarios help pinpoint the extent of the epistemic obstacle. Imagine that my wife
prefers to eat greasy fast food for lunch every day. She enjoys the taste of the greasy food more than the taste of healthier, leaner foods. Suppose that – for whatever reason – she does not know that her cholesterol numbers total 330, and even if she did know this, would not know that this is considered extremely high for someone her age. The question now is: If I am interested in promoting her well-being, ought I to set myself (as far as possible) to satisfy this preference of hers for fast food? When she and I are out on the town together and want to get lunch, should I offer to take her for the fast food she prefers, or should I take her to a restaurant where the food is healthier? When I make dinner for her, should I make the greasy but unhealthy kind of food that I know she would prefer, or something more nutritious? It is not patently obvious how to answer these questions, but it is surely not crazy for me to judge that if my aim is indeed to promote her well-being, I should not take her for, or prepare for her, the greasy food that she prefers. My judgment here is that if she were aware of how high her numbers were and how this jeopardizes her health, she would not continue to prefer the fast food.

Consider two additional related cases. In the first, my wife has the same penchant for fast food, but here she is fully aware of how high her numbers are and of the threats this poses for her health; but she shrugs this off. She continues to prefer going out for fast food. Assuming my aim is still to promote her well-being, what weight should I give to this preference? In the second case, imagine that this person with a penchant for fast food but high cholesterol (and who knows it, and knows how dangerous this is) is not my spouse, but rather my neighbor (who counts as an acquaintance of mine, but by no means a dear friend). The same question is posed, namely, whether (on the supposition that I want to promote his
well-being) I ought to aim to satisfy his preference for fast food. Am I more justified in one case than in another to aim at actually frustrating the individual’s preferences?

In this second case, I still have little doubt about the fact that the fast food that my wife prefers will detract from rather than contribute to her well-being – at least if she has more than a little of it. It is not arrogance or “mere paternalism” that leads me to say so. I am slightly less certain of this when she prefers the fast food even knowing and understanding her high numbers as I am when she is altogether ignorant of her numbers, but I am still quite confident. I also am fairly confident that (more than a little of) this food would detract from my neighbor’s well-being. If they both have a fully informed preference for this food, it does not seem like a huge stretch to presume that they are irrational, and to say in turn that I am a better judge in this case of what will promote their well-being than either of them is. But this is clearly something I have a more difficult time knowing regarding my neighbor than I do regarding my wife. I do know my wife’s cholesterol numbers, and they were easy for me to learn. In reality, I do not have a single neighbor whose numbers I actually know. This is not something my neighbors tend to share with me. But suppose that I do know a particular neighbor’s numbers, and that they are indeed 330. What then could make me less sure about whether the fast food that my neighbor prefers detracts from his well-being than I am that it detracts from my wife’s? Presumably it has at least partly to do with the fact that nothing guarantees that eating healthily – doing so on this one particular occasion, or even always – makes a greater contribution to the well-being of the person with high cholesterol than would eating unhealthily. My wife and my neighbor are both going to get considerable enjoyment from a tasty, greasy meal. Nothing guarantees that pristine health better contributes to one’s well-being than enjoyment from a good-tasting though
unhealthy meal does. Nothing ensures that each person’s well-being is enhanced by a long life. Whether it does depends on whether it would be preferred by those who are rational and fully informed. Perhaps long life is only something that a rational and fully information person would prefer if she knows that her life will be filled with enriching activities and relationships.

I believe I know that my wife’s fully-informed and rational preference would be for a long life – or at least a longer one than she will live if she keeps eating the fast food. If she truly, vividly recognized the incompatibility of her preference for fast food and her preferences for a long life, seeing her children grow old, traveling the world after we have retired, etc., and if she were prodded to eliminate the inconsistency among these preferences, she would give up the preference for fast food, or would at least prefer to give up that preference. I know that a longer life is in her interests, and life-shortening fast food is not, and when I make dinner for her or take her out to eat, I do not need to satisfy her preference for fast food in order to promote her well-being. With my neighbor, on the other hand, I am less sure that a longer life really is in his interests. Perhaps his informed preference for fast food does not conflict with any of his other informed preferences. Perhaps his life is and will continue to be sufficiently devoid of other pleasures that his fully-informed and rational preference would be to pursue (on occasion, or very regularly) the great pleasure that eating greasy food represents for him. I do not know for sure, and (at least partly) because of this I am less justified in frustrating his preferences than I am in frustrating those of my wife. Still, to concede this is not to admit that each person $P$ is always a better judge than any other actual person of $P$’s well-being. People are not so terribly mysterious that we can never confidently say of any person what she would prefer if she were fully informed and rational.
Sometimes, and with respect to certain people, acting in accordance with our reasonable estimations of what a person P would prefer if fully informed and rational preferences will better promote P’s well-being than promoting the satisfaction of her actual preferences would.

My claims thus far have been that if – as is often the case – P is not perfectly informed and rational, then P is not the best imaginable judge of her well-being, and P is not always a better judge than every other actual person is of P’s well-being. It may be objected here that I have made an illicit move by talking about the ability of one’s spouse, best friend, and even neighbor to make judgments regarding one’s well-being. We are interested, the objection goes, in the question of how the state should go about promoting individuals’ well-being, and the fact that one can occasionally be a better judge than one’s spouse of one’s spouse’s well-being does not mean that the state is wrong to promote people’s well-being by promoting the satisfaction of their preferences. The objector is right that my expressed goal is ultimately to focus on policy-relevant issues. That P’s spouse or best friend occasionally knows better than P herself what will promote P’s well-being is not going to tell us all that we need to know in order properly to appraise CBA. But this kind of discussion is not out of bounds. The question of how to promote well-being, and more specifically of whether to do so by promoting people’s actual preferences, is not just one that public policy analysts face, but that all of us face in each of the different relationships we have. I am in a position to affect my spouse through my actions (and, in fact, even through my inaction). I am also in a position to affect my siblings, parents, friends, neighbors, colleagues, students, etc., through my actions and inaction. In fact, I am in a position to affect many people who are complete strangers to me. These different relationships are each worth thinking about. The question
of what role, in these different cases, these people’s preferences should have for me if I am interested in promoting their well-being is interesting in its own right. But in addition, and more importantly for my purposes, working to answer this question will provide insight into the citizen/state relationship, and more specifically, will inform the issue of whether/when/why policy analysts should aim to satisfy individual preferences, which is my primary concern.

In any case, this objection leads us to a third possible interpretation of the economists’ claim that each person is the best judge of her own well-being. Perhaps this is not supposed to mean that each person $P$ is always a better judge than every other person of $P$’s well-being, but that each person $P$ is always a better judge than *the state is* of $P$’s well-being, and for this reason, if the state wants to promote an individual’s well-being, it should promote the satisfaction of her preferences. This claim too is false. In spite of the genuine obstacles that exist, the state may sometimes know better than $P$ does what will promote $P$’s interests. Cost-benefit analysts are often possessed of information that ordinary citizens do not have. There are many examples, and I will run through a number of them in some detail in later parts of this project; but a simple one on which I have already touched can illustrate the point. Suppose that $P$ believes that the water she drinks from her private well is extremely clean. She is utterly unaware of the fact that it is actually contaminated with high amounts of arsenic, and even unaware that arsenic is dangerous to ingest. Moreover, the water from her well tastes great to her. If an analyst simply asks $P$ whether she would prefer to have bottled water, she will say no, and she will say no even if this the bottled water would not entail additional expenses for her. As far as she knows there are no problems with the water she has. But the state – and more specifically the analysts it employs – may know
better. If a CBA is being conducted regarding an issue like rural water provision, it would be very common for the state to know about the levels of arsenic in water coming from area wells. (What the state knows in such a case will probably exceed what the majority of area residents know.) In this case, there is essentially no question that if P had the information regarding the arsenic and clean water that is no more expensive but tastes the same were available, P would prefer the clean to the contaminated water. The water that is cleaner and equally good tasting but no more expensive would better promote her well-being; but one who is striving just to satisfy her actual preferences will not even aim to give her this cleaner water.

Perhaps the economists meant not to deny that others, including even the state, can sometimes be a better judge than P of P’s well-being. A fourth (and, for my purposes, final) interpretation of the economists’ claim that each person is the best judge of her well-being is that though at least certain others, the state included, may now and again know better that P what will promote P’s well-being, no one is on the whole better able to know this than P is. Each individual tends to have a better sense of what is in her interests than does anyone else, and thus each person’s preferences tend to promote her interests. Thus, the claim is that the best general approach to promoting a person’s well-being is to promote the satisfaction of her preferences. This can work in tandem with the third interpretation: it is especially true, the claim goes, that each person P tends to have a better sense than the state tends to have of what contributes to P’s welfare. The state’s promotion of the satisfaction of citizens’ preferences does not guarantee the promotion of their interests in each and every case, but is nonetheless the most effective way the state has for promoting well-being, in general and over time.
This interpretation may (finally) be one that makes the claim true. Nonetheless, proponents of CBA should not be content. The person who makes this claim faces a serious problem akin to the one the rule-utilitarian faces in opposing act-utilitarianism. Both act-utilitarians (AU) and rule-utilitarians (RU) cash out moral rightness in terms of the maximization of goodness. But they differ with respect to how they answer the following question: What is it that is right in virtue of having the best consequences? The AU says that individual acts are right in virtue of having the best consequences. The RU, on the other hand, says (roughly) that moral rules are right in virtue of having the best consequences, and individual acts then are right when they are prescribed by the right rules. Rule-utilitarianism has the virtue of accommodating many more strongly held intuitions than does act-utilitarianism, but it faces a powerful objection from the AU. In some cases, the AU will claim, the rule the following of which promotes the best consequences will actually prescribe actions the performance of which produce the best consequences. But, says the AU, this is not true in every case. If, for example, everyone were without exception to follow the rule “Do not lie, ever,” better consequences would probably be produced than were everyone without exception to follow the rule “Tell the truth except when lying is convenient.” But this does not mean that there will never be a single occasion where an individual lie will have better consequences than would telling the truth. The AU claims that the RU is wrong to follow rules in every single case, including those specific cases where not following the rules would actually produce the best consequences.

---

90 There are of course numerous versions of rule-utilitarianism. A classic article on this topic is Richard Brandt’s “Some Merits of One Form of Rule-Utilitarianism,” University of Colorado Studies (1967). This article is reprinted in 20th Century Ethical Theory, Steven M. Cahn and Joram G. Haber, (eds.), (Upper Saddle River, NJ: Prentice-Hall, 1995), pp. 418-436. Brandt says that “rule-utilitarianisms may be divided into two main groups, according as the rightness of a particular act is made a function of ideal rules in some sense, or of the actual and recognized rules of society.”
This objection was famously offered by J. J. C. Smart. Smart asks his readers to imagine that the behavior of the sun, moon, and planets is almost as it is now, but that once in a great while this behavior is irregular in a way that made a nautical almanac unreliable on these occasions. Thus, the almanac would in effect give us “rules of the form ‘in 99% of cases where the observations are such and such you can [use the almanac to] deduce that your position is so and so.’” But suppose moreover, Smart says, that there are methods that, though laborious, enable direct calculation from the original astronomical data that would give us our correct position in 100% of the cases. Smart notes that “seafarers might use the almanac because they never had time for the long calculation and they were content with a 99% chance of success in calculating their positions. Would it not be absurd, however, if they did make the direct calculation, and finding that it disagreed with the almanac calculation, nevertheless they ignored it and stuck to the almanac conclusion?”

The parallel with rule-utilitarianism should be clear. Smart accuses the rule-utilitarians of doing something precisely akin to insisting on sticking to the almanac conclusion. He writes:

Suppose that there is a rule R and that in 99% of the cases the best possible results are obtained by acting in accordance with R. Then clearly R is a useful rule of thumb; if we have not time or are not impartial enough to assess the consequences of an action it is an extremely good bet that the thing to do is to act in accordance with R. But is it not monstrous to suppose that if we have worked out the consequences and if we have perfect faith in the impartiality of our calculations, and if we know that in this instance to break R will have better results than to keep it, we should nevertheless obey the rule? Is it not to erect R into a sort of idol if we keep it when breaking it will prevent, say, some avoidable misery? Is not this a form or superstitious rule-worship (easily explicable psychologically) and not the rational thought of a philosopher?

91 “Extreme and Restricted Utilitarianism,” Philosophical Quarterly 6 (1956). This essay is reprinted in 20th Century Ethical Theory, Steven M. Cahn and Joram G. Haber, (eds.), (Upper Saddle River, NJ: Prentice-Hall, 1995), pp. 307-313; and it is to this reprinting that my citations will point.

92 Ibid, pp. 311-312.

93 Ibid, p. 311.
While contemporary rule-utilitarians have worked to develop responses to this criticism, the criticism continues to impress many people.\textsuperscript{94} At the very least it provides a deep challenge to rule-utilitarianism. The same basic problem plagues those who would endorse this fourth interpretation of the economists’ claim and appeal to the model akin to rule-utilitarianism to explain why the state ought to promote preference satisfaction. It may be true that the state \textit{often} has trouble knowing what is in each individual citizen’s interest, and it may be true that \textit{always} promoting the satisfaction of actual preferences is a better way to promote well-being than \textit{never} promoting the satisfaction of actual preferences. But it is also clearly true that a citizen will at least occasionally \textit{not} have better knowledge of what is in her interest than the state will. Indeed, Smart’s point has special bite when taken as a criticism of the fourth interpretation since it is reasonable to suppose that the frequency with which the individual knows better than the state what will promote the individual’s well-being is significantly less than 99\% (even if it is, say, still well over 50\%—and what the percentage \textit{actually} is, I obviously do not claim to know). If the state did in some particular case know better than \textit{P} what is in \textit{P}’s interest, then the fact that the \textit{P tends} to know better than the state does what will promote her well-being provides no justification \textit{in this case} for the state to promote the citizen’s preference over and against her well-being. For this reason, this fourth interpretation fails as one that would make the economists’ claim convincing.

\textsuperscript{94} For a recent and sophisticated defense of rule-utilitarianism against this and other objections, see Brad Hooker, \textit{Ideal Code, Real World: A Rule-Consequentialist Theory of Morality}, (New York: Oxford University Press, 2000).
Section 3: The Role of Individual Preferences When Benefits are Ultimately Identified with Gains in Well-Being

What are the implications of the preceding discussion for the theory and practice of CBA? The move from the economists’ purely – I would say “naïvely” – subjectivist account of well-being to my informed-preference account is such as to make the first three interpretations of the claim that each individual is the best judge of her well-being ones according to which the claim is false, and the fourth interpretation, though it made the claim true, did not suffice to establish that each individual will always know better than the state what will promote the individual well-being. Does this thus open the door to paternalism? Does it license the state on welfare grounds regularly to override an individual’s preferences in making an assessment of what counts as a cost or benefit for that individual? For several important reasons, I believe that suggestion is at best an oversimplification. The first of the reasons is grounded in epistemic considerations. Later, I will discuss another reason, which is based in the idea that promoting well-being is not even the only purpose of public policy decisions.

Let us begin with the epistemic considerations, which is really what the previous section of this chapter was devoted to beginning to explore. Suppose that the textbook line by the economists is the right one: the purpose of public policy decisions is to promote well-being (and leave aside questions as to how individuals’ well-being get aggregated into social welfare). What considerations suggest that the state should simply count the satisfaction of an individual’s actual preferences as benefits to that individual? It was noted that if cost-benefit analysts committed to promoting individual well-being had to choose between always promoting the satisfaction of individuals’ actual preferences and always entirely ignoring these, they may well fare worse by choosing the latter. Again, though, these are patently not
the only options. The state does not need to commit itself now and forever to either of these courses. Another option, obviously, is for the state sometimes to promote the satisfaction of actual preferences while at other times to ignore actual preferences. Specifically, it might ignore actual preferences on those occasions where it is confident that an individual’s actual preferences are not the ones she would have if she were fully informed and rational. So, the state may on some occasion believe that the satisfaction of these preferences will detract from the person’s well-being. And indeed the satisfaction of these preferences might do that. But does this belief constitute knowledge? The state might believe that a person is not informed and rational, but not be sure she is not. Another possibility is that the state might know she is not, but nonetheless have no idea what the person’s informed and rational preference would be. In either of these cases, the state might well do less to promote the individual’s well-being by identifying benefits with that which, in the state’s best estimation, she would prefer if informed and rational than it would by simply counting as benefits (to her) the satisfaction of the preferences she actually has. This is just to say, in other words, that the state, if its aim is to promote an individual’s well-being, will frequently have reasons to identify benefits for that individual with the satisfaction of her actual preferences.

Perhaps the state could commit itself to contravening actual preferences on all and only those occasions on which it is extremely confident it has a better sense of what would promote a person’s well-being than the person herself does. But here again the grounds for welfare-based resistance to the satisfaction of the individual’s preferences – paternalistic intervention, in other words – are a bit shaky, for as the preceding discussion made clear, even on these occasions, the state may lack a perfect sense of what would promote this person’s well-being. Policy analysts might know that P would, if she were well-informed
and rational, prefer bottled-water to well-water, *ceteris paribus*. But precisely how strong would her preference for avoiding arsenic be? Would *P*, if well informed and rational, avoid arsenic altogether? Would she pay $0.50 more per gallon for water with 25 ppb than for water with 50 ppb? Would she pay $0.05 more per gallon for that reduction? Would she pay $0.01 more per gallon for water with 9 ppb than for water with 10 ppb? It is hard to know. I frankly do not know how my own wife would answer these questions; and the state is in an even worse position – significantly worse, in fact – to know how each of its many and varied citizens would do so. This is, however, the kind of thing that policy analysts are interested in knowing. There are many other examples illustrating this same issue. This provides some support for actually soliciting preferences—be it by observing people or by putting questions to them. Judging on the basis of something other than an individual’s actual preferences *precisely* how a change would affect that person’s well-being (including precisely how much it would increase or diminish her well-being) is going to be quite difficult. (In fact, even if policy analysts were not trying to reduce everything to willingness to pay, there will nonetheless be specificities regarding individual preferences that the state will have a very difficult time knowing without actually putting specific questions to people.)

In spite of all of this, the welfare-based case for identifying benefits (to an individual) with the promotion of (this individual’s) preferences is hardly an open and closed one. The aforementioned considerations provide some reason on the grounds of well-being for identifying benefits with preference satisfaction. But they do not mean that these should *always* be identified. As the previous section of the chapter indicated, there will be occasions where the state has a better sense of what (in substantive terms) will promote an individual’s well-being than the individual herself does. It is true that it will be difficult to know
precisely how strong her informed and rational preference would be, but that issue is apt to distract us from the important point here (though as the next paragraph starts to illustrate, that issue does have relevance to the broad discussion). Both the state’s estimate and the individual’s own preferences may provide an imperfect sense of what will promote the individual’s well-being, but if these are our two choices, it may at least sometimes be better to related benefits and costs to what the state estimates the change would be on the person’s well-being than to what the individual’s own preferences indicate.

As it turns out, those are in fact not our only two choices. A third option exists, and is better. In short, it is possible – at least in many cases – for the state to supply people with information and help them understand and process it. Of course the state will not always have good information, and it will perhaps never have information that suffices to count as “full” in the sense strictly demanded by the informed-preference account of well-being considered in its most philosophically abstract form. But the state will often have information – especially information relevant to the investment or regulatory decision at hand – that affected individuals lack. When the state provides individuals with this information, helps them understand, process, and develop consistent preferences in response to it, and then solicits preferences from these people, it ensures that it is in the position where, by satisfying an individual’s preferences, it is promoting the individual’s well-being. If proponents of CBA are genuinely committed to promoting individuals’ well-being, they should maintain that it is after people are given this information and assistance in processing it that analysts should record people’s preferences and ipso facto make an assessment of costs and benefits. The analyst who claims to be striving to promote well-being is either exaggerating her insightfulness into what others would prefer when fully informed and rational, or ignorant of
the epistemic challenges she faces, if she regularly makes specific, substantive judgments about an individual’s well-being without attending at all to the individual’s preferences. But this same analyst is either disingenuous in her claim to be promoting well-being or badly misconceiving the relationship between preferences and well-being if she attends to individual preferences without first aiming to ensure that individuals are both rational and fully informed and then supplying individuals with such information and encouraging consistency among these preferences.

This proposal is obviously much easier to state than it would be to implement in practice. I certainly do not want to gainsay the challenges the proposal faces; and for this reason, I will devote more time to it later in this project (especially in the next chapter). For now, note simply that the proposal has implications for both the revealed- and the stated-preference methods. Those who use the revealed-preference method should acknowledge that if there is reason to think that an individual being observed is ill-informed or irrational – both of which are, I believe, frequently the case for individuals acting in the market – there are serious risks that her market behavior (from which analysts infer her preferences) will provide analysts with misleading information about her well-being. If it is going to be standard practice for analysts to read off a person’s preferences from her behavior in the market, then there should also be a push to make sure that relevant information is available to each person who is making a purchase. If relevant information exists but is not available to the individual, or the information is available to the individual but she clearly has not accessed it, analysts should be wary of using the revealed-preference method. (Again, though, it should be noted that an individual may have an informed, rational preference for not having full information.) Furthermore, those who design and administer CV surveys
should be concerned to ensure that each respondent is well-informed before responding, and should be prepared to supply a respondent with information when she is ill-informed. Analysts should ensure that the respondent understands the information, should help make this information vivid, and should as far as possible help ensure that the respondent’s informed preferences square with one another. Supplying respondents with information will not include telling them what their moral obligations are, but because feelings of moral obligation are not excluded from the account of well-being I defended, the analyst should encourage individuals to give serious thought to what moral obligations they take themselves to have before these individuals register their preferences. In at least certain cases, if the analyst refrains from doing any of these things, there is serious risk that the assessment of costs and benefits generated by a CV will tell us little to nothing – and in any case less than it might – about the well-being of the respondents.

Section 4: Why Promoting Well-Being Is Not the Ultimate Aim of Public Policy Decisions

In the previous section, I asked whether the move from a naïve to a more sophisticated preference-based account of well-being licenses the state on welfare grounds regularly to override an individual’s preferences in making an assessment of what counts as a cost or benefit for that individual. Drawing on points established in the previous section, I then talked through some epistemic considerations. I argued that on the supposition that the state’s ultimate aim in making public policy decisions is to promote well-being, some (actual) individual preferences will be such that their satisfaction counts as a benefit to that individual, while others will be such that their frustration is what should count as a benefit to that individual. Moreover, I maintained that the ideal scenario is one in which the state
provides people with information and aids in the understanding and processing of this information. Once this takes place, then a person’s preferences and well-being will not diverge, and the satisfaction of actual preferences – now actual informed and rational ones – can be identified with policy-level benefits. (And if there is reason to believe that a person does not actually become sufficiently informed and rational through this, then we are back to the position where, on grounds of well-being, the satisfaction of [actual] individual preferences will sometimes count as a benefit to that individual but at other times count as a cost.)

All of this, however, was just based on what was basically a supposition, namely that the state’s ultimate policy aim is to promote well-being. The claim that this ought to be the state’s ultimate policy aim was not so much argued for as it was just asserted. Upon more careful analysis, it turns ought to be a rather indefensible claim. I now want to explore considerations that challenge the idea that promoting well-being is the only purpose of public policy decisions. As noted earlier, one question that could be pressed is: Why think that social welfare is properly conceived of as an aggregate of individual well-beings (where, say, social welfare is promoted when the sum of individual well-beings is made greater)? This is a fine question, but it is not the one on which I want to focus. I want to ask: Even if it is granted that social welfare is properly conceived of as some kind of aggregate of individual well-beings, why should one think that the state, through its policy (especially regulatory and investment) decisions, ought ultimately to be aiming to promote individual well-being? In other words, why accept the textbook line that benefits (to an individual) ultimately have to do with improvements in (that individual’s) well-being? I will argue that that line exaggerates the importance of the promotion of well-being. No one should doubt that this is
among the proper aims of public policy decisions. But it is not the only one. This means that there is a connection between increases to (an individual’s) well-being and benefit (to that individual) – whether or not this is done by satisfying the individual’s actual preferences – but that whether a change counts as a cost or benefit is not explained entirely by its affects on well-being.

Suppose that the state has perfect knowledge of what will promote and what will detract from the well-being of a particular ill-informed or irrational person, including knowledge of precisely how much and in what ways the changes that various possible policies would introduce would affect the individual’s well-being. (Again, in actuality, the state will almost certainly not have such knowledge, but we gain insight by imagining that it does.) I believe that once we do this, we can see more plainly that it is a mistake to think of the promotion of well-being as the proper ultimate aim of public policy decisions. Instead, we should understand the state as having multiple aims or purposes in making public policy decisions. In short, it should be aiming not just to promote well-being, but also to respect individual autonomy and to help individuals actually experience satisfaction. A commitment to promoting well-being will at least sometimes recommend counting the frustration of an individual’s actual preferences as a benefit to that person. A commitment to respecting autonomy will not recommend this. Nor, obviously, will a commitment to helping individuals to actually experience satisfaction. What this means is that these various purposes the state has in making policy decisions might at least sometimes conflict with one another. This recommends a pluralistic conception of benefits (and costs too, of course). But the recommendation that the state help as far as possible to create actual well-informed and rational preferences and then promote the satisfaction of these has even more going for it
than was discussed above. It makes genuinely viable the prospect of the state achieving all of its aims at once.

It may again be instructive to consider other relationships along with the one between citizens and their state. Recall that I expressed less confidence in the judgment that eating fast food is contrary to the interests of my fast-food-loving neighbor than I did in the judgment that eating such food is contrary to my fast-food-loving wife’s well-being. Importantly though, even if I were as confident in the case of my neighbor as I am in the case of my wife – and indeed even if I were perfectly confident in both cases – I would take myself to be less justified in promoting my neighbor’s well-being over and against his preferences than I am in doing so with respect to my wife. I suspect I am not unique in this respect. Where epistemic issues are ruled out, what could be the explanation for this? And relatedly, will it be an explanation that also suggests that if the state certainly knew that a citizen’s preferences and well-being diverge and knew in addition precisely what would promote that citizen’s well-being, it would somehow be inappropriate for the state to promote her well-being over and against the satisfaction of her preferences?

It is difficult precisely to pinpoint, but it surely has something to do with differences in the nature of these two relationships. The intuition is roughly that the closer, the more personal, the relationship between promoter and promotee, the more justified the promoter is in promoting the promotee’s well-being rather than preferences when these pull apart. My neighbor and I do not have an adversarial relationship. Far from it; I quite like him. But my relationship with him is less close and less personal than the one I have with my wife. One need not see the state as having an adversarial relationship with its citizens, need not dislike government, and need not be an anarchist or even a libertarian, in order to say that the
citizen/state relationship is even less close and less personal than the one I have with my neighbor. If this is right, it will not seem unreasonable to say that when one’s well-being and her preferences diverge, the state’s case for promoting the former over and against the latter – again assuming that the state could know that what one prefers is contrary to one’s well-being, and could know (substantively) precisely what one’s well-being in this case consists in – will be contentious.

This is a gesture, but something more substantive is required in order to make good on the basic claim regarding the difference in appropriateness between aiming to promote my wife’s well-being and aiming to promote my neighbor’s. If it is less acceptable, even apart from epistemic considerations, to promote my neighbor’s well-being over and against his preferences than it is to do so for my wife, this could perhaps be explained by the fact that I have taken an explicit and public vow to care for my wife, but have not taken one to care for my neighbor. The move, in turn, would be to point out that neither the state nor its elected or appointed decision-makers have taken any such vow to care for citizens. There are several problems with this suggestion. Firstly, why think that the “caring” thing to do is to promote a person’s well-being over and against her preferences? Is it entirely obvious that the “caring” thing to do is to push my wife, against her own preferences, to eat healthily? And more importantly, why is a vow of any sort necessary in order to promote someone’s well-being over and against her preferences? There is something to the idea, but it is tough to make completely good on it. A second attempt at an explanation turns on the claim that my wife has, as it were, tacitly authorized me to make decisions regarding her based on what I know to be her interests. My neighbor has not ever authorized me – even tacitly – to make decisions based on his interests. The related point then pressed would be that citizens (at
least typically) have not authorized the state to do so. Again, the suggestion is reasonable, but not entirely without problems. For one, it is not clear that even my wife has ever authorized this. But even if she has, this explanation is insufficient. Importantly, why I need authorization to make decisions regarding another individual based on that individual’s well-being, but not to make them based on that individual’s actual preferences, stands still in need of explanation. In fact, this explanation is dangerously similar to the one mentioned earlier that Sagoff rightly warned directly invited the regress, namely that individual preferences ought to be satisfied because that is what individuals prefer.

Consider a different pair of scenarios. First, imagine a case in which my teenage son prefers to hang out with a man of questionable character – one who sells illegal drugs, let us say. Then consider another case in which it is not my son but my colleague who prefers to socialize with the drug dealer. Suppose I know that though both prefer doing so, it is in neither my son’s nor my colleague’s interests to be spending time with this unsavory character. (This is contentious, but grant it for the sake of this discussion.) Still, is the well-being of each of these people what I should be aiming to promote? My intuition is that the aim of promoting the well-being of my colleague is not as obviously justifiable as the aim of promoting my son’s well-being. Is there any way of making sense of this intuition, and will it also be such as to suggest that the state should not be aiming simply to promote well-being, even if it could be sure that it knew what it is (substantively) in which the well-being of each of its citizens consists?

Perhaps the justification vis-à-vis my son is that he is someone for whom I am directly responsible, perhaps quasi-contractually in virtue of having “commissioned” his existence, or perhaps somehow in virtue of the brute physical fact that I am his father, while I
do not have the same kind of responsibility for my colleague. The subsequent move is to claim that the state clearly does not have the same kind of responsibility for its citizens. Again, there is an intuitiveness, but there are worries too. Why exactly does the fact that I did not commission my colleague’s existence relieve me of responsibility for her well-being?

If the brute physical fact that one is my son somehow makes me responsible for him, why does the brute physical fact that I am one’s colleague not make me responsible for her? And perhaps most fundamentally, why does living up to my responsibility for my son require the promotion of his interests rather than his (uninformed or irrational) preferences when the two diverge? Perhaps the explanation for why it is more appropriate for me to make decisions based on my son’s interests than it is for me to do so for my colleague goes to the fact that my son and I are in some sense one: we are one family. (This could also be used as an explanation in the wife/neighbor case.) The claim, then, will be that the state and its citizens are not one in the same sense, and so the state may not promote citizens interests over and against their preferences. But why are my colleague and I, or my neighbor and I, not also in some important sense one: (at least part of) one workplace, or one neighborhood? (And why are the state and I not in some meaningful sense one: one polis?)

There are two additional lines of reasoning that are even more persuasive, and the upshot of each is to challenge the textbook line that the aim of the state in making public policy decisions is to promote individual (and in turn social) well-being. What these amount to are reasons for counting the satisfaction of preferences – even ones that are clearly ill informed or irrational and thus welfare reducing – as benefits. One feature of the claim that each individual is the best judge of her own well-being, however it is interpreted, is that if it is true, it makes the claim that the state ought to promote citizens’ preferences at best
contingently true, not true as a matter of principle. After all, despite a fundamental conceptual connection between a person’s preferences and her well-being, it is, in an important sense, at best a contingent truth that the best way the state has of promoting each individual’s well-being is by promoting his/her actual preferences. The proponents of CBA who focus only on epistemic justifications for preference satisfaction should be uncomfortable that the case for promoting preference-satisfaction rest on this contingent connection. As has been made clear, if actual people are ill-informed or irrational, and the state did somehow know precisely what would promote people’s well-being, it could promote well-being without attending to people’s actual preferences. The first and more important of the two lines of reasoning notes that this would be inconsistent with properly respecting individuals, particularly their autonomy.

Earlier I said that Sagoff considered the possibility that the satisfaction of a preference is good in itself and the possibility that it is good for its consequences. He claims that it is not good in either of these senses. He seems, moreover, to have believed that these two possibilities are exhaustive. They are not. One possibility that Sagoff does not consider is that the satisfaction of preferences is good as a necessary constituent part of another thing that is uncontroversially good. I propose that the promotion of the satisfaction of preferences – and I do not just mean idealized (informed and rational) ones – is good as a necessary part of freedom or autonomy. (Some may be more inclined to say that the satisfaction of these preferences is good as a means to respecting individual autonomy, i.e., that promoting the satisfaction of these preferences is good for its consequences, namely that doing so has the effect of being respectful of individual autonomy. I will not insist that the point not be put in those terms, though to me it seems more natural to talk about the promotion of preference
satisfaction as a part of respect for autonomy rather than as a means to this.) I will give as
detailed an analysis of autonomy or its relationship to preferences as I did for well-being.
Perhaps there will be an occasion for such an analysis, but this is not such an occasion.
Nonetheless, I certainly must clarify what I have in mind by autonomy and what I take its
relationship with preferences to be. It would be too strong to say that one is entirely unfree
any time one of her actual preferences is frustrated, and even too strong to say that the
individual’s freedom is somewhat impinged upon any time one of her actual preferences is
frustrated. Nor does one lack freedom, or is one even less free ceteris paribus, simply in
virtue of the state refraining, on a particular occasion, from attending to one’s preference.
However – and the point can be put cautiously at first – there is clearly a tension between
respecting an individual’s autonomy and ignoring the individual’s preferences, even if the
reason the preference was ignored was because its satisfaction would not have promoted, and
indeed would have detracted from, the individual’s well-being.

One of morality’s central demands is for respect for individual autonomy. This claim
should be uncontroversial enough as to pass without argument. But how important it is again
is context dependent. Frankly, part of what it is to get married is to renounce some of one’s
freedom. I am accountable to my wife for my actions, and when I got married, my wife
 gained some say over how I conduct my life. She did not gain complete authority over me,
and thankfully does not want that anyway, but her authority to regulate how I conduct my life
did increase when we got married. And my say over how she conducts her life increased too.
What I gained, among other things, is the right to try to influence her to eat more healthily,
even when doing so requires frustrating her preferences. I infringe upon her liberty when I
do so, but this is not always inappropriate in a spousal relationship. It is, however, typically
inappropriate in a neighborly relationship. Likewise, I limit my son’s freedom by frustrating his preference to hang out with the drug dealer, but this is permissible. He is relatively young and his freedom is limited in a number of respects – he cannot drive, drink alcohol, vote, see certain movies without an adult, etc. – and for good reason. A parent who never restricts his/her child’s autonomy is probably not a good parent (unless the child is an exceptionally unusual one, with the child’s preferences and interests always lining up).

There is a sharp contrast between these relationships and the one between citizens and the state. Freedom is one of the chief political virtues. A good government is one that does not regularly or severely impinge on its citizens’ freedom. Again, the truth of these claims regarding the value of freedom is intuitive. Because it is not among my aims in this section – or anywhere in this project – to defend political liberalism, these claims will not be argued for. I am content here to address myself just to those people who find political liberalism tenable (and I take this to be the large majority of Americans, including the large majority of those who are “illiberal” in the sense that they are not “Democrats”). Those who take it to be important for the state to respect the freedom of its citizens should also take there to be some reason for the state to aim at satisfying individual preferences. Considerations of autonomy thus provide some reason for identifying benefits (to an individual) with the satisfaction of (that individual’s) preferences. The preferences to which these claims pertain are not just those that would be had under appropriately idealized conditions. In other words, there is some reason for aiming to satisfy preferences and identify benefits with preference satisfaction even if the satisfaction of an individual’s preference would make her worse off.\footnote{One fact has been mentioned before, and will be elaborated on in the following pages, but should be reiterated here: My view is ultimately that the state provide people with information and help them reason 05

One does not count as truly autonomous if she is forbidden by others from harming herself.
It is commonly said that the state respects people’s autonomy by not infringing on their speech or their worship, by not restricting the press, by not regulating how people assemble, by not preventing people from traveling or (somewhat more controversially) from owning firearms. It is also recognized that in order to respect people’s autonomy, the state must allow people the opportunity to vote in the election of public officials and on certain referendums. This is all reasonable. But the notion arises in the context of analysis of proposals for public investments and regulations – which is to say that it arises in the context of CBA – as well. Part of what is to be free as a citizen within a state is to have one’s preferences – one’s actual preferences, not just those one would have under certain ideal conditions – taken into account when the state makes decisions that will affect one, which will include occasions on which CBA is employed. My claim is not that people must have input on every policy decision. We elect representatives and authorize them to make policy decisions on our behalf. These representatives even appoint individuals to make certain policy decisions, and that is legitimate. Elected representatives and their appointees make many decisions, and doing so becomes their job. They are able to devote considerable time to these decisions while the rest of us devote ourselves to other jobs, educations, families, etc. A political system that allows us to elect representatives who will make many decisions on our behalf is certainly not one that can be charged with a wholesale failure to respect individual autonomy. The point I have been arguing for here is simply that there is opportunity in the context of the analysis of public policies for the state to demonstrate its respect for individual autonomy: perhaps unexpectedly, the opportunity exists even in the

through this, and then aim to satisfy these preferences. In this way, the state respects people’s autonomy by taking their preferences seriously, and does so in a way that promotes rather than detracts from individual well-being.
context of the technical policy questions that tend to be addressed by CBA. Cost-benefit analysts (and the elected officials who request the CBA) want to solicit preferences from individuals who stand to be affected by a decision. I believe there are good reasons to do so, and one of the reasons is that insufficient respect will be shown for individual autonomy if public officials and the policy analysts they employ regularly issue judgments about what will promote individuals’ well-being in opposition to, or without taking due interest in, the preferences that ordinary individuals assert and evince.

Now, it is true that there are many conceptions of autonomy or liberty, and moreover great divergence between certain conceptions. This was powerfully illustrated in Isaiah Berlin’s famous “Two Concepts of Liberty.”\(^\text{96}\) That considerations of autonomy provide some reason for recommending the satisfaction of an individual’s preferences even where this will diminish that person’s well-being will not be obviously true on the positive conception of freedom that Berlin discusses. But Berlin’s discussion is complicated; how the version of “positive liberty” that he presents connects up with preference satisfaction is not obvious. I said I did not want to give a full-scale analysis of the notion of autonomy here, but that is roughly what a careful discussion of Berlin would amount to. Thus, I propose to leave for another time the exploration of the implications of a positive conception of liberty for the question of what counts as a cost and what counts as a benefit.

Cass Sunstein discusses the connection between preference satisfaction and autonomy, and his discussion is well worth pausing to explore here. He does not seem to be working with a positive conception of liberty/autonomy, but does maintain that “the notion of autonomy should refer… to decisions reached with a full and vivid awareness of available opportunities, with reference to all relevant information, and without illegitimate or excessive

constraints on the process of preference formation. When these conditions are not met, decisions should be described as unfree or nonautonomous; for this reason it is most difficult to identify autonomy with preference satisfaction.”97 I myself have not identified autonomy and preference satisfaction, but rather have said that taking someone’s preferences seriously is an important part of respecting her autonomy. I have said that considerations of autonomy provide a reason for identifying benefits (to an individual) with the satisfaction of (that individual’s actual) preferences. Still, this certainly makes contact with Sunstein’s discussion. He would seemingly maintain that taking someone’s preferences seriously is a requirement of respecting her autonomy only if the aforementioned conditions he specifies – viz., the person expressing a preference is fully and vividly aware of available opportunities, has all relevant information, etc. – are met. If he is right, then just as considerations of well-being leave us with a question as to whether to count the satisfaction of individual preferences as a benefit (when these seem to be ill informed or irrational), so too can considerations of autonomy leave us with this question.

Sunstein is mistaken, or at least his point is one that it is easy to overstate. It is not the case that one’s decisions are unfree in every case where they are made without one having full and vivid awareness of all available opportunities. Imagine that P decides to drive from her home in North Carolina to her parents’ home in Ohio. P does so after exploring flying and taking the train as options for making this trip. Suppose that there is, unbeknownst to P, also a bus that could take her from the one place to the other. P did not have a full and vivid awareness of all available opportunities for travel between these two places, but contrary to what Sunstein seems committed to maintaining, it would surely be a mistake to say that her decision to drive thereby counts as in some way unfree. Now it is of

course true that if $P$ does not know that she has any other options beside driving, then her decision to drive is not free. $P$ may still be free not to go at all, but if she has to go and does not know of any way but driving to get there, then her decision to drive is not free. (In fact, if $P$ has to go and does not know of any way but driving to get there, it may be infelicitous to say that $P$ makes a decision – free or otherwise – to drive. But nothing else that I have to say rides on this last claim.) Sunstein’s claim, however, was much stronger than the claim I identified above as true.

Consider again the case where $P$ knows that driving, flying, and taking the train are all options but fails to realize that the bus is also an option. Now suppose you are in a position easily (i.e., without cost to you) to promote or frustrate $P$’s preference for driving. You could, for example, assist $P$ or hinder $P$ in readying her map and getting her car’s oil changed, tires filled and rotated, etc. Contrary to what Sunstein suggests, if you are entirely committed to respecting $P$’s autonomy, it would be a mistake to interfere with her ability to drive from here to there and then attempt to justify this on the grounds that her decision to drive was itself not an autonomous one and that respect for $P$’s autonomy thus permits or demands such interference. Importantly, however, a more reasonable point is that if you know that $P$ has to get to Ohio from North Carolina, and you know both that the bus is a way of doing so and that $P$ is unaware of this option, you will lack credibility if you claim to be fully committed to respecting $P$’s autonomy but – regardless of whatever else you do – you do not tell her about this option. $P$ is not unfree simply in virtue of not having full information regarding all her options for getting to Ohio, but if you help satisfy $P$’s preference for driving without informing her that the bus is an option, then it will be odd for you to point to the importance of respecting $P$’s autonomy as your reason for helping to
satisfy her preference. If you were genuinely committed to respecting \( P^{\prime} \)’s autonomy, you would have told her that the bus was an option she had not yet considered.

This all squares very nicely with the suggestion I made in the previous section of this chapter. I have maintained that rather than aim to satisfy whatever preferences individuals happen to have – which will often be ill-informed or irrational ones – the state should work to supply citizens with information and help them understand and process it, thereby creating informed and rational preferences, and then to aim to satisfy these. When it was initially presented, this recommendation was motivated largely by epistemic concerns: it seemed on the whole to be the best way of discovering precisely what would best promote individual well-being. This recommendation now has additional support—namely support from considerations of respect for autonomy. Even if the state could know (on some occasion) that an individual’s preferences are ill-informed or irrational and could further know precisely what this individual’s fully-informed and rational preferences would be, there is reason for the state to take my middle course between promoting the satisfaction of this person’s actual (ill-informed or irrational) preferences and contravening these preferences and promoting what the state knows to be conducive to the individual’s well-being. Simply disregarding a person’s preferences, even where these preferences are ill-informed or irrational, is not consistent with due respect for this person’s autonomy. But the same is true of promoting the satisfaction of a person’s preferences when you know or suspect that these preferences are ill-informed or irrational and you are in a position to help the person become better informed and more rational. While lacking complete information is (contrary to Sunstein’s claim) not strictly incompatible with autonomy, providing someone with information will not infringe on that person’s autonomy. Taking someone’s preferences
seriously, even if they are uninformed and irrational, is an important part of respecting that person’s autonomy; but if you have information that this person does not have, then first giving her this information, helping her to vividly perceive it, understand and organize it, and deliberate about it, and then taking her now-informed preferences seriously, is clearly in such a case demanded by respect for that person’s autonomy. Indeed, this is the ideal way to respect someone’s autonomy. This is true when I interact with my neighbor who loves greasy fast food or my colleague who spends time with a drug dealer, as well as when the state makes an assessment of the costs and benefits of a proposed policy.

There is another part of the aforementioned quotation from Sunstein that has not yet received sufficient attention, and this is the claim that “the notion of autonomy should refer… to decisions reached… without illegitimate or excessive constraints on the process of preference formation.” What should be said about this? It is a bit more reasonable than his demand for full awareness of one’s options, but if not handled delicately, it too will be overstated. In short, I think it should be handled essentially like that other demand was handled. Just as preferences were not excluded from my preference-based account of well-being just in virtue of being exogenous, so too should we resist restricting those preferences attention to which is demanded by respect for autonomy to those that are formed “without illegitimate or excessive constraints.” Every one of my preferences is formed under certain constraints. What I prefer is not determined by some radically free will. Consider that all sorts of constraints operated in the process in which my preference for listening to Dylan’s music was formed. Nonetheless, it seems clear that my decision to listen to Dylan should count as autonomous. This must mean, if Sunstein is right, that these constraints were not illegitimate or excessive. But I have no idea why they were not, or what counts as an
illegitimate or excessive constraint. I suspect that efforts to specify what kinds of constraints count as illegitimate or excessive will fail; they will at best amount to an ad hoc list of disqualifying constraints. In at least a great many cases, decisions based on preferences the formation of which was constrained in all kinds of ways should count as autonomous. However, it may be that certain constraints are such as to make it the case that the decision that is made is one that is somehow compelled, not preferred. Again though, the recommendation that individuals be provided with and helped to process information is relevant. If an individual is fully informed and rational, she will know all about what did and did not constrain the process by which her preferences were formed, and in this way the notion of full information does all the work that needs doing. Whatever preferences she has in light of this knowledge are ones that respect for her autonomy demands be taken seriously. The more one can be led to see what has influenced the formation of one’s preferences, the better one can evaluate these preferences and decide which are worthy preferences to have. Once ones sees this and does one’s evaluating, the preferences one ends up with cannot be ignored by anyone who claims to be fully committed to respecting individual autonomy.

There is a second (and more straightforward) line of reasoning that tells against ultimately identifying the claim that if the state could overcome the relevant epistemic

---

98 One may ask here whether considerations of well-being and considerations of autonomy are not actually related even more closely than I have suggested. Is it possible genuinely to promote someone’s well-being without respecting that person’s autonomy? Can one really respect someone’s autonomy without promoting that person’s well-being? This is a complicated issue. The important point for my purposes is simply that there is no conceptually necessary link here. If there is a connection, it is a substantial rather than formal one. It is a connection that exists in virtue of what a person prefers—though it may be that we are all generally psychologically put together so as to prefer to be treated (at least in certain contexts and up to a certain point) as autonomous individuals. Moreover, one might wonder whether my suggestion is not that having one’s autonomy respected is good for one independent of one’s preferences. The actual suggestion is that respect for autonomy is good independent of one’s preferences, but not necessarily good for one independent of one’s consequences. That is, the goodness of respect for one’s autonomy does not consist (solely) in the contribution it makes to an individual’s well-being. Precisely why it is that respect for autonomy is good is here left generally unanalyzed.
problems, then benefits should be associated simply with the promotion of one’s well-being rather than the satisfaction of her preferences when the two diverge. This reason became intelligible when the nature of well-being came into focus in the previous chapter. In short, it is that one might – again, at least in theory – promote a person’s well-being without helping that person to experience any satisfaction. However, helping people to experience satisfaction should be among the ultimate aims of the state through its public policy decisions. Were the claim that this is the only appropriate ultimate policy aim of the state, it would obviously be very controversial. Indeed, as has in effect been argued, it would be false. But where the claim is that this is one of several appropriate ultimate aims of the state through its public policy decisions, it seems quite reasonable. I still do not quite know how to argue for that claim. This is that sort of ground-level kind of consideration where arguments do not work. It is one of those claims that serves as a premise in other arguments, but cannot itself be argue for. My hope is that once it is clear what this means and that it is not entailed by the promotion of well-being, it will seem obvious that there are reasons – including reasons that hold in the context of the citizen/state relationship – for helping people to experience satisfaction. I do not call this more important than promoting well-being or respecting autonomy. Nor do I maintain that it is less important. Alas, I do not have a view on the relative importance of these various aims. Nonetheless, the experience of satisfaction seems undeniably to be of some importance, even if the promotion of experiences of satisfaction brings about no further good consequences.

According to the informed-preference account – of whose plausibility I am convinced – if an individual is not actually fully informed or rational, then the satisfaction of her actual preference may actually diminish her well-being. Given this understanding of well-being,
there exist two ways in which one might promote a person’s well-being without helping her to experience satisfaction. First, let us suppose that some person \( P \) is indeed ill-informed, and that the state knows this, is possessed of the information that \( P \) is lacking, knows what \( P \) would prefer if she had this information, and knows that this is precisely the opposite of what \( P \) currently (in her ill-informed state) prefers. The critical point is that in this case the state could promote \( P \)’s well-being without the actual \( P \) receiving any experience of satisfaction. \( P \) would be better off if she got what she would prefer if she were informed and rational, but when this is not what \( P \) actually prefers, giving her what she would under the counterfactual conditions prefer is not going to suffice for the actual \( P \) to feel satisfied. Indeed, if all of her actual preferences over the course of her entire life are exactly the opposite of what they would be if she were fully informed and rational, and the state could know this, it could promote her well-being over the course of her entire life without her ever experiencing any sensation of satisfaction. Granted, it is quite unlikely that anyone actually matches this extreme description. Most people are such that at least some of their preferences are exactly those that they would have if they were fully informed and rational, and most people at least occasionally, and within certain contexts, actually count as fully informed and rational. Regardless, the theoretical point stands.

Second, we must recall just how the notion of “satisfaction” being worked with by the informed-preference account of well-being is to be understood. On this account, satisfaction is a state of the world, not a state of mind, so it will be possible, in at least some cases, for the preference of a person \( P \) to be satisfied without \( P \) actually experiencing satisfaction. (To go back to the standard example, \( P \)’s preference may be for a person \( Q \) that \( P \) meets on the train to succeed in life – by accomplishing task \( T \), say – and if \( Q \) does accomplish \( T \), then \( P \)’s
preference has been satisfied. This is true even where $P$ does not know what happens to $Q$, and so where $P$, upon $Q$’s accomplishing $T$, is no different psychologically from how $P$ would have been had $Q$ not accomplished $T$. So, a policy that helped $Q$ to accomplish $T$ would promote $P$’s well-being, but if $P$ does not know that $Q$ accomplishes $T$, then $P$ experiences no satisfaction. If, somehow, all of that which $P$ prefers were to obtain without $P$ knowing it – a possibility that is generally left open by the informed-preference account – then $P$’s well-being could be promoted without $P$ experiencing any satisfaction. Of course, among the things one prefers might be to have experiences of satisfaction, and this preference could surely not be satisfied without one experiencing satisfaction! Nonetheless, again, the general theoretical point stands. One question is whether one who satisfies the conditions set out by the informed preference will necessarily know whether that which she prefers has actually obtained. This need not be the case. For example, having a fully-informed preference about that person one meets on the train involves knowing what this person is really like before one forms a preference that (say) this person succeeds in life; but one may count as fully informed without knowing whether this person actually will succeed in life.99

The final element in this particular argument is almost too obvious to merit stating. It is simply that the fact that promoting experiences of satisfaction is among the state’s ultimate aims provides a reason – another reason, as it turns out – for counting the satisfaction of

99 Why do I focus on the experience of satisfaction rather than the experience of happiness? Perhaps one should not insist on a distinction between these ideas. In fact, while I think these ideas are not far apart conceptually, it does seem to me that there is a difference, and that the former is the one on which it is worth focusing. Satisfaction is a richer concept than happiness. Satisfaction can consist in more than – indeed in something other than – happiness, and if one find satisfaction in a lifestyle that does not exactly make her happy, there seems here nothing remiss. I can imagine the brutally hard work and stress of being President of the United States to involve (at least sometimes) experiences of satisfaction, but not all of these – perhaps very few of them, in fact – would, I imagine, involve experiencing happiness. This is just a gesture at an answer to the original question. It is a complex one, and frankly one I would like on this occasion simply to brush past.
preferences as a benefit. (The fact that the respecting of individual autonomy is among these aims does the same. The fact that well-being is also among these aims sometimes provides another reason for doing so; though this fact – given the nature of well-being and some reasonable epistemological claims – sometimes provides reasons for counting the satisfaction of preferences as a cost.) The recommendation that I have been pressing, namely that the state actively provide information, assist in its processing, etc., and then help promote the satisfaction of the resulting (now informed and rational) preferences has less going for it in connection with this third aim. Importantly, though, my recommendation does not in any way compete with this third aim. Purely from the point of view of the experience of satisfaction, it is no better, but it is also surely no worse, that these are informed and rational preferences.

Summarization is in order. After breaking CBA into a theory of value and a theory of right action, I set out to explore the former, and more specifically, the part of CBA’s theory of value that specifies what is to count as a cost and what is to count as a benefit. The textbook line, according to which benefits should be identified with preference satisfaction because preference satisfaction increases well-being and promoting well-being is the ultimate aim of public policy decisions, was argued to be mistaken, or at least significantly oversimplified. Given the appropriate conception of well-being, preference satisfaction does not always enhance well-being. The related claim that each individual is the best judge of her own well-being is either not true or not sufficient to justify – even supposing that the ultimate aim of policy decisions is indeed to promote well-being – the idea that the satisfaction of a preference ought always to count as a benefit. Sometimes it should, but sometimes this will diminish an individual’s well-being and for this reason count (on grounds
of well-being) as a cost. It was argued, moreover, that it is actually a mistake to think that the promotion of well-being is truly the only appropriate ultimate aim of the state through its policy decisions. At least in theory, well-being can be promoted in a way that is not duly respectful of individual autonomy and that does not lead to experiences of satisfaction, but both of these things are important. With the promotion of well-being, respect for individual autonomy and the promotion of experiences of satisfaction are also ultimate aims of public policy decisions. These three aims might sometimes compete with each other. With respect to some policy, there may be a preference the satisfaction of which should, on grounds of promoting well-being, clearly count as a cost, but on grounds of respecting autonomy and promoting experiences of satisfaction, count as a benefit. (And if it is a merely idealized, hypothetical preference, then the opposite may be the case.) In such cases, the state must strike a balance between these aims.

In practical terms, nearly every real-world policy case will be one in which there will be reason to doubt that a person’s preferences regarding a proposed policy change tell us precisely how that change truly will affect her well-being. The state should not simply estimate what this effect will be. Nor should it simply make its assessment based on the individual’s actual preferences. My fundamental point is that there is a way in which the state can at least approach achieving all of its aims – at least all those in has in making policy decisions – at once. Everything recommends counting as a benefit the satisfaction of preferences that are actually well informed and rational. The contrast is with counting as a benefit the satisfaction of preferences that, though informed and rational, are merely hypothetical (because not had by actual people), and counting as a benefit the satisfaction of preferences that, though actual (because had by actual people), are not both informed and
rational. There should thus be a major push by the state to create actual preferences that are informed and rational. This is the ideal. It is what policy analysts should strive for when they set about figuring out what the costs and benefits of a proposed policy would be. However, if it becomes clear that an individual is not going to be able to become sufficiently well-informed or rational, then the state is back to striking a balance, somehow, between those three aims. When this happens, I cannot say precisely how these should be balanced—I certainly have no algorithm. The three aims may or may not all be of equal importance. Reflections on that question will have to be postponed for another occasion. For now, the idea is just that policy makers must conscientiously recognize that no single aim is of absolute importance, but none is unimportant either. With this, they must muddle through.
CHAPTER 4: PROVIDING INFORMATION AND ASSISTING IN ITS PROCESSING: THEORY AND PRACTICE

Section 1: Introduction

The previous chapter discussed in fairly general and abstract terms what the proper ultimate aims of public policy decisions are. The fundamental recommendation I made was that there are reasons for the state to actively provide individuals with information and help them to understand, process, and reason consistently about this information. When this happens, benefits (to an individual) can rather uncontroversially be identified with the satisfaction of an individual’s preferences—that is, her now informed and rational preferences. In this chapter I work to more fully fill out my view, which still faces many important questions. I begin to address them by considering one rather ordinary decision that might be addressed by way of CBA, namely one where a small city is considering whether to buy a wooded tract of land and preserve it as a place for quiet reflection and leisurely outdoor recreation. Through this case I shed light on the related general issues of whose preferences should be counted and what information individuals should have before their preferences are gauged. In connection with this same case, I then consider two issues that have occupied an important place in the literature, namely transformative value and existence value. I explain what each is and argue that CBA can and should accommodate both.

I then address in more detail what it means to provide full information, especially in cases where there is disagreement over the facts or where it is agreed that our knowledge is
limited. The debate over global warming is used as the example. I endorse the model of journalistic neutrality regarding the presentation of information in controversial policy cases, though warn against the possibility of creating an impression of greater disagreement about the facts among experts than really exists. I then consider the worry that individuals will not be able to hold on to and process all the information that my suggestion seems to require them to have. Here too, I argue, the state has a role. In the end of the chapter, I consider and respond to a number of other important questions, including the following: Does not my suggestion make the analyst’s job impractically big? What if we become convinced that people are not getting accurate information, or are not correctly processing the information they get? Is it not true that many people just want elected officials to make a decision? Is it really my view that individuals’ preferences must be sought out for each and every policy decision?

**Section 2: One Basic, Concrete Example**

In order to flesh out my view more fully, I want to begin by looking at a fictional but plausible case in which CBA confronts conditions of incomplete information. The case is relatively unexceptional. It does not require imagining the possibility of a project that would flood over the Grand Canyon, make bald eagles go extinct, poison thousands and thousands of people for many generations to come, or anything such as this. But I like the case precisely because it is commonplace; it is the kind of scenario that confronts real people in real towns and cities across the county. As should hopefully become clear as I proceed in the chapter, I do not pick this example in order to skirt difficult questions that (one might think) arise in connection with more “dramatic” policy questions.
Imagine a relatively small city that is considering buying a 100-acre wooded tract of land within that city’s borders. The tract has long been (and is still) privately owned, and the city has grown around it. It thus makes for very desirable real estate, and also conveniently located for people to access it for recreational purposes. Indeed, city residents have long been permitted by the owner to use it for hiking, mountain biking, quiet reflection, and the like. If the city were to make the purchase, the city would maintain the tract as a place for this kind of recreational and leisurely purposes, but if the land is not sold to the city, the owner has said – and is sincere, let us say – that she will sell it to a developer who will clear much of it and build a large neighborhood full of houses. Whether the city ought to purchase the land is among the kinds of governmental decisions for which economists and many policy analysts are likely to recommend the use of CBA. The analyst would work to measure people’s preferences, and more specifically how much people would be willing to pay to preserve the land in its current undeveloped state. The basic question for the cost-benefit analyst is whether city residents’ collective WTP exceeds the purchase price.\(^{100}\) This is the kind of case that lends itself rather naturally to contingent valuation. Analysts will contact people and ask them a number of questions intended to tell the analyst about people’s sincere WTP. Let us imagine, plausibly enough, that hundreds of people use the land for recreational activities, and are aware that it is going to be sold either to the city (which would preserve the land) or to a developer. Others, we can imagine, use the tract but do not follow the news well enough to know that it is up for sale. (Indeed, they might believe, e.g., that the city already owns it.) Many of the city’s other residents who do not use the land are aware it is up for

\(^{100}\) There may be some other factors to think about, including whether there will be maintenance costs into the future, and if there would be, questions will arise as to how/whether to discount these in order to put them in terms of present value. The statement above was not meant to be perfectly precise.
sale. But there are also many people in the city who know basically nothing whatsoever about the tract of land.

This is the basic scenario. It will be modified in a number of different ways in order to facilitate the posing of certain kinds of questions and the illustrating of particular points. Generally, though, I want to focus on two important questions, and they are related. One is: who exactly ought to be approached and surveyed? (In a typical CBA, the method is not usually to survey everyone, because this would simply involve too much work. Instead, a random sample is commonly generated. The use of random samples is not a concern I want to press here. Instead, the question is, in essence, about who should be included in the group from which the random sample is drawn.) Another question is: what information should be supplied to respondents before their preferences are solicited? Alas, both questions are more complicated than they may initially appear to be.

Presumably all will agree that everyone who uses the tract of land should be included in the group from which the random sample of respondents is drawn, and this includes both those who do and those who do not know about the decision the city faces as to whether to buy the tract. In order to count as being in a reasonable position to specify their preferences for continued access to the undeveloped land, those who currently use the land should know, e.g., where other opportunities for similar recreational activities can be found. If someone enjoys mountain biking on this land, but mistakenly believes that there is another publicly accessible tract of land not far down the road from this one, when in fact the only opportunity for similar recreation is an hour’s drive away, the preference she expresses will probably be different from the one she would have expressed had she known that the nearest similar opportunity was so inconvenient. If we suppose that she would on the whole prefer for the
land not to be developed, we can reasonably suppose in turn that she would prefer this more strongly if, instead of believing a similar place to mountain bike were down the street, she rightly knew that the closest such place were an hour away. If she does not know this, the preference she specifies is suspect, which is to say that we cannot be confident that it tells the analyst what the analyst reasonably wants to know, namely, how bad this person would take the development of this land to be. The preference she specifies in this case does not accurately inform us about how her welfare would be affected by the development. But analysts should not simply say without contacting her what her informed and rational preference would be. Doing this underestimates the difficulty of knowing precisely what that preference would be, and also fails to respect her autonomy. Supplying her with better information and helping her process it does not overstep the appropriate bounds of the citizen/state relationship, and will help ensure that her preference, when satisfied, increases her well-being.

What else should these people be informed about? One important issue is whether there will be any effects on their health that would result from the development. Perhaps there are; but if so, there is no reason to think that these would only plague those who know about and currently use the tract. Imagine – hopefully not too fancifully – that developing the land would unavoidably spill harmful pollutants into a nearby creek, which would then flow downstream and be spilled into a swimming hole used by numerous residents of the city. There would be deleterious effects on the health of nearly all of those who swim in this hole, though (let us imagine) those affected would not be able to trace these effects back to the development of the tract, and indeed would not even know that their swimming water had been polluted. Suppose that some of the people who swim in this hole are people who also
use the tract of land for recreation; they are general enthusiasts for outdoor recreation. But at least some of them, let us imagine, would be among those who know nothing about the tract or its development. (An alternate story, for those who do not like this one, is that property values around the tract would decrease if the tract were developed. This is another effect that would make it such that the development of the tract would affect more than those people who use the tract for recreational purposes. This alternate story is one the reader will have to follow through with on her or his own.) It is possible, of course, that the spilling of the pollution is the kind of result no one at all could possibly anticipate. Perhaps the best urban planners and environmental scientists were entirely convinced that this would not happen. Sometimes things occur that could never have been anticipated. In the case where no one even has a suspicion this might happen, there is of course no obligation for anyone to share with the city’s residents the fact that developing the land would result in the spill of pollutants into the water. But suppose instead that while the average citizen does not know that development will pollute the water in this way, the projects’ planners do, and government officials do too. In fact, they are confident that they know what kind and amount of pollutant would go into the water, as well as what kind and amount of sickness this would cause to the swimmers.

In this case, there may be disagreement about who should even be counted as having preferences regarding the possibility that the city would buy the land. Some economists will claim that those people who do not know of the tract or the possibility of its development do not have any preferences regarding the tract or its future. In turn, the economist will maintain that a CBA does not need to concern itself with these people. CBA aims to measure people’s preferences for particular outcomes in order to determine whether a proposed
change ought to be brought about, but because these people do not have preferences regarding the tract, they provide the analyst with nothing to measure. It may even be said that it is not even clear that those who swim in the would-be-polluted water – at least those in this group who do not know about the tract – have preferences regarding the tract. After all, they do not know that it will pollute their swimming hole, and as I have told the story, will not know this even after that pollution has been generated.

This view is at best overly simple. We can reasonably say that none of those people ignorant of the tract and the possibility of its development has the issue on her or his mind. None of them will, e.g., initiate a conversation about the tract. We can say of these people that there is nothing that they actually \textit{will} give up, e.g., to keep the tract from being developed (at least if they remain ignorant of it). In this sense they are without preferences regarding the tract. But we cannot say of them that there is nothing that they \textit{would} give up. Those who swim \textit{would} perhaps give up something \textit{if} they knew the development of the land would pollute the water they swim in and in turn diminish their health. They surely all have preferences regarding their health, and whether or not the tract is developed has implications for the health of at least some of them, namely the ones who swim in the water that would be polluted; so by extension it seems reasonable to say they have preferences regarding the tract and its future. Even if I am wrong about this, we should certainly say that if the land is developed and the water is polluted, then it will not be only those citizens who use the tract, and not even only those aware of its existence, but also at least some people who are entirely ignorant of its existence, who will be affected by its development. Even if they do not count as having preferences now, they would, if they were aware that they were affected by the development, certainly count as having preferences regarding how they are affected (e.g., the
preference for not having their water polluted, for not having their health damaged by this pollution, etc.). For this reason, there is strong reason for the analyst to work to include all those who would be affected by the development and pollution in the group from which the random sample of respondents is drawn.101

Not only must these people be counted, but moreover, the information that the state has regarding the pollution- and health-related repercussions should be shared with them. Though none of the city’s citizens knows about the pollution to the water that the development of the land would cause, this is clearly an instance for the swimmers at the very least where what they do not know can hurt them. It is a case where there is good reason to think that what such an individual actually prefers will not give us a good sense of how her well-being would be influenced. This does not mean that the state should, as it were, on its own make an assessment of the costs and benefits of the purchase by simply anticipating what the people’s preferences would be. This is partly because it is so unclear what the effects, on the whole, of this project on people’s well-being would be. Presumably the polluted water would make swimmers worse off, ceteris paribus. But whether every one of them is on the whole worse off as a result of the project is not clear. It may well be that at least certain people – those, say, who do not and would not use the tract, who would be likely to buy one of the news houses if the tract were developed, and who do not go to the swimming hole often – would, if fully informed and rational, prefer for the city not to buy the land and for the development to go forward. Even if the city knew who would on the whole

101 Note that not everyone who is currently unaware of the existence and possible development would have develop preferences regarding the tract if they knew of its existence and possible development. Some people, even if they knew of all this, might be completely indifferent to it all. Others, though, would not be indifferent, and at least where these people satisfy certain other conditions (which I will discuss later), what they would prefer if informed and rational – and more precisely, what they do prefer when they are informed and rational – is relevant to an assessment of the costs and benefits of purchasing the tract.
be made better off by its purchasing the land and who would on the whole be made worse off, it would still presumably have a very difficult time knowing how much better/worse off each would be made. (Much more will be said later about the difficult issue of judging the extent of an effect, i.e., how much better or worse off one is made.)

This provides a good reason for the state to seek to gauge people’s preferences. As I have said, though, even if the state could know all of this, there would be good reason for it to avoid determining the costs and benefits of the project in a way that did not directly gauge people’s preferences. There are, as I have argued, considerations of respect for autonomy and consideration of the value of experiences of satisfaction to consider. However, this does not mean the state should simply aim to satisfy the ill-informed preferences that the citizens actually have. Instead, the health effects associated with development should be shared with those who would incur them. The analyst should work to provide those affected by the pollution with information that is as complete and comprehensible as possible. The preferences that ought to be measured and to which policy decisions ought to respond are these well-informed preferences. Providing this information and then making policy that is genuinely based on and responsive to these informed preferences is a way of helping to ensure that the state, in satisfying people’s preferences, would be promoting their well-being. But my basic proposal also helps ensure that individual well-being is promoted in a way that is ideally respectful of individual autonomy in the context of the citizen/state relationship. (Again, this proposal does nothing to facilitate actual experiences of satisfaction, but it does nothing to undermine these either.)

A general objection will be raised here to the proposal that the state provide people with information and help them reason through it before registering their preferences. Some
will claim at this point that the purpose of a CBA is to measure, not influence or create, preferences. They will note that the very act of using contingent valuation to solicit and measure people’s preferences regarding a project risks influencing preferences or creating preferences that did not exist, and will claim that to supply people with complete as possible information (as I have called for, in fact preceding the use of both the revealed- and stated-preference methods) is going to exacerbate this problem. There are several things to say in response to this objection. The first is that if CBA were guilty on this count, it would not be uniquely guilty. There are plenty of instances of measurement tools modifying via the act of measurement the very thing that the tool was aiming to measure. Indeed – and I hope the reference here is not thought to be excessive – as Werner Heisenberg’s famous Uncertainty Principle explains, it is impossible to observe a thing without the act of observation modifying the thing. To measure people’s preferences will indeed at least sometimes modify preferences. To consciously work to supply respondents with information (about, e.g., the health consequences of a project) before measuring their preferences will indeed further risk influencing the preferences. Sometimes it will clearly do so. For example, people’s preferences for the aforementioned tract of land to remain undeveloped would presumably be stronger, ceteris paribus, when they know that development would negatively affect their health than it would be if they believed there would be negative health effects. Sometimes it may altogether create preferences. Such information must be supplied nonetheless.

It has been argued that promoting well-being is one of the aims of public policy, and it should be influencing and even creating preferences does not compete or interfere with the aim of promoting people’s well-being, and may often even be demanded by it. Consider that a scientist who wants to measure the behavior of an object, but who in the process
unintentionally (though perhaps unavoidably) influences the thing so that it behaves differently has problems. But this is not the parallel of the cost-benefit analyst who aims to promote well-being and supplies people with information before measuring people’s preferences. Rather, the parallel to the analyst is a scientist who wants to measure a substance’s behavior at (say) an extremely high temperature, and so who begins by subjecting the substance to heat. This scientist has no problem. She is doing what she needs to do in order to measure that in which she is interested. It has been shown that an individual’s well-being consists in the satisfaction of those preferences she would have if she were fully informed and rational. Thus, while supplying someone with information likely will influence and modify her actual preferences, it will not alter her well-being. Rather, it will simply increase the likelihood that the preferences she expresses are ones the satisfaction of which would actually promote her well-being. Cost-benefit analysts should provide people with information and help them process this (the analogue of heating the substance) and then measure their well-informed, rational preferences (the analogue of measuring the substance at the high temperature). It is clear, therefore, that the mere fact that it would influence preferences does nothing to show that the analyst is wrong to inform the people whose preferences are being sought for measurement about the negative effects on their health associated with development. In fact, the analyst’s claim to want ultimately to promote people’s well-being patently lacks credibility if the analyst measures preferences without ensuring that information that can be reasonably be expected to influence people’s preference is not shared where possible. I will not repeat the argument that influencing people’s actual preferences does not conflict with – and is in fact demanded by – respect for autonomy, for this should be apparent, at least where the influence is just in the form of the
provision of information and the provision of assistance in understanding and processing that information. Similarly, it patently does not interfere with the aim of promoting actual experiences of satisfaction, though (as I have conceded) it is not demanded by this aim.

Section 3: Transformative Value and Existence Value

If development were to go ahead, there will be many people who would not feel any different from how they felt before the development. This will include people who did not use the tract at all, including those who did not even know the tract existed. But this – even on the supposition that these are people who do not swim in the water that would be polluted and whose health would thus not be threatened by the development-produced pollution – does not mean that the development would not negatively affect their well-being. It does not mean that the cost-benefit analyst need not take account of them. If the land were developed, these people would lose the opportunity to discover the tract and the recreational opportunities therein. This point connects up with one made by Bryan Norton\textsuperscript{102} and Andrew Brennan.\textsuperscript{103} Norton tells a story about a teenager who is required to attend a performance of classical music. The teenager is not at all excited to go; she has always preferred pop music, which is essentially the only kind of music to which she has ever listened. Much to her surprise – perhaps to everyone’s surprise – she has a great experience at the classical music concert. She goes on to develop a sincere interest in, and eventually a deep love of, classical music. Brennan says that this example shows that “attending the concert has transformed her


considered preferences.” We can imagine that had she not had the opportunity to go to that classical concert, she would not have become this great lover of classical music.

Similarly, we can imagine someone who has spent no time in “undeveloped nature.” This person might understandably have no appreciation for the kinds of recreational opportunities it provides. She is the kind of person who would not know that the tract even exists. Still, this does not mean that they would not, under certain conditions – including, perhaps, if she were able to spend time out in nature (exploring this tract, etc.), especially with someone knowledgeable and experienced – come to prefer for it to remain undeveloped. One implication for CBA is simply that those who know nothing about the tract should not be excluded from responding to a survey about the tract, since they could, if it were not developed, at some point learn of its existence and then come to use and love it. This is not purported to be an overriding reason for not developing the land, because these people could, on the other hand, learn of the tract’s existence and form the view that it is an utter waste of valuable real estate space. The point is simply that these people’s views matter. The fact that they do not currently know about the issue is not a reason to exclude them from the survey. This is an issue that their government has taken up. If the government is seeking guidance on the issue, it should not simply consult those who already know about the tract. The city should seek guidance from all of its citizens. Perhaps this does not mean that they should consult every single citizen, but each citizen should be included in the group from which the sample is drawn.

Some worry CBA is unable to take into account the transformative value of new experiences. Brennan says that “the impact of transformative value is completely unpredictable, and the degree of impact (when it occurs) will vary massively from person to

\[104 \text{ Ibid.}\]
person.” This is an understandable worry; but it is too strong to say that CBA is unable to take transformative value into account. Perhaps we should not force people to go spend time on the tract of land in the way that the teenager was forced to go to the classical concert (though perhaps we should. I will say more about this later). But at the very least, providing complete as possible information about the tract to the townspeople would increase the chances that preferences liable to transformation will actually be transformed prior to being measured and recorded by the analyst. It is also unfair to single out CBA for failing to completely take transformative value into account. Many experiences will transform people’s considered preferences, and many projects will limit the ways in which people’s preferences can be transformed in the future; but this is no less true of policies that are not justified via a CBA than it is of those that are supported by a CBA. It is, for example, no more severe a problem for CBA than it is for voting, which is another mechanism for public decisions.105

There are other difficult questions regarding what the best possible CBA regarding this tract of land will look like. What other effects would there be to townspeople if this land were cleared and homes were built? What, if any, ecological services are being provided by this tract that would no longer be provided (at least in the same way) if the tract were developed? Imagine that I know about and am accustomed to using the tract, that I know it may be developed, know that development would not have any deleterious health effects, and know that similar opportunities for recreation are an hour’s drive away. Imagine, in addition, that I believe that if the land were developed, the animals who live there would all be able to relocate to another wooded area, but that it is in fact the case that most of them will die.

---

105 See Chapter 5 for an extensive elaboration of similarities and differences between CBA and voting.
Finally, imagine that I care a great deal about animals; I do not want them to be killed, and would be willing to pay something to help prevent them from being killed.

If the state knows that the animals will die, must it share this information with me before it measures my preferences? The answer is yes. One possibility is that the death of the animals has the same basic kinds of effects on area residents as the aforementioned pollution would on the swimmers. If I am one of the swimmers, the pollution affects my well-being independent of whether or not I am aware of the pollution. Does the death of animals that I care about affect my well-being even if I am unaware of the deaths? Often yes, though it will be less direct than in the swimming case. In many situations, animals provide services that have direct physical consequences on me. For example – and here I am oversimplifying, though seemingly not in a pernicious way – birds in a flock eat from plants and later, in excreting seeds from the plants in another location, help the plants propagate. The excreted seeds grow into adult plants that produce fruit that I can eat, trap CO₂ and release O₂ that I breathe, and prevent soil from washing into and contaminating water that I drink. If animals providing services to me die and the services go unperformed, I am worse off whether or not I know about their death. Consequently, if the state has knowledge that the animals will die, it should share this information with me.

Let us modify the case again to delve into deeper issues. Suppose that the animals who are expected to be killed do not perform the aforementioned kind of services, but that I still care about them and how they fare. Would there still be reason for the state to share with me (and others) the information that the animals are expected to die? Here I clearly have preferences regarding the animals and their futures. These are other-regarding preferences. My concern for them is not rooted in a concern about how the way they fare affects the way I
fare. It is the same kind of concern that non-swimmers might have for swimmers who would be affected by the pollution. Perhaps, for example, those who regularly swim in the swimming hole are my employees; I then have self-interested reasons for wanting them to be healthy. But it is also possible that they are simply people about whom I care. I could care about them without even knowing them personally. I simply know that they enjoy swimming in the hole, and if I knew that it would be polluted, I would be willing to pay something to ensure that the land was not developed. These preferences – be they for the welfare of the animals or that of the swimmers – should be among those captured by the CBA. Information that would affect these preferences – that the animals will in fact be killed, or that the swimming hole will in fact be polluted – is relevant, and should be provided.¹⁰⁶

On what grounds could anyone justify withholding such information? Surely those who maintain that the purpose of CBA is to promote individual and societal well-being could not do so. The informed-preference account surely could endorse the judgment that I am worse off when animals about whom I care have been killed – that is, worse off independent of any services the animals perform or provide (or would have) to my benefit – even if I do not know that anything bad has happened to them. If I had full-information, I would prefer for the animals about whom I care not to die (ceteris paribus, at least). If I would, were I fully informed, have preferences that the animals not die (ceteris paribus), then the death of

¹⁰⁶ One may be tempted to say at this point that the real problem with the cases I have introduced involving the death of animals about whom the respondent cares is that animals and swimmers have a welfare that matters independently of whether or what people are willing to pay for them. I believe, and will later argue, that it is true that animals’ welfare and preferences matter independently of what people are willing to pay. Animals are not mere resources. But there is an additional point that I have been after here. The point could have been made via an example that had nothing to do with animals. The reader may satisfy her- or himself by revising the preceding example such that it is a totally inanimate object that the development would damage, where this inanimate object is still cared about by those whose WTP is sought, and still is a consciousness-independent welfare influence.
the animals affects my well-being, and does so even if I do not actually know about the death. Another way of putting the point is to say that how much I am willing to pay will differ according to (among other things) whether or not I know about the animals’ death. Surely the analyst wants to measure my preferences regarding what actually will happen, or at least regarding what the most educated extant prognoses suggest actually will happen, not my preferences regarding whatever it is I happen to believe will happen. By providing me with information – in this case, information about the expected fate of these animals – and then working to satisfy my informed preference, the state is able to promote my well-being, and does so in a manner consistent with respecting my autonomy within a citizen/state relationship.

The general response is the same regarding the person who cares about the health of the swimmers, and does so independent of any services that these swimmers may generally provide for her. Some people – though not all – are such that their well-being is connected to things that do not benefit them, and certainly not in any direct, physical sense. Surely a parent would deserve to have her WTP for the tract gauged if her child was swimming in the swimming hole that would be polluted by development of the tract. This would be true even if the parent did not actually know that the development would cause such pollution. This parent would also need to be provided information before her WTP was gauged. Presumably no stranger will care about this child’s health as much as the child’s own parent will. But there are probably many people who are not the child’s parent, and who do not do any swimming in the swimming hole, who would be willing to pay something to keep the child’s swimming water from being polluted. These people should be provided with the information about the pollution.
This discussion is important in no small part because of how it bumps up against a large and difficult issue, namely the role that so-called “existence values” or “non-use preferences” should play in CBA. Existence value is another one of those terms whose meaning is often not made sufficiently clear by those who use the term. The basic issue is that we are capable of having preferences for an enormous quantity and variety of things, and that these things are sometimes very far away from us, and things that we will not ever, and will not even ever want to, consume, use, or even encounter in person. For the lack of a better term, I will call these non-use preferences, though technically this is a bit of a misnomer. To see why it is a misnomer, suppose that I am a Parisian who prefers going every day to the Louvre to see the Mona Lisa. It is odd to say that I use the Mona Lisa. I certainly do not use it like I use a kitchen knife when I am cutting a carrot, or like I use that carrot after I have cut it. But I do experience it first hand, and on a daily basis, and I do derive pleasure from it. Is this a case of me demonstrating a non-use preference? As I intend the term, this is not such a case. Strictly speaking, my preference here is not one that should be counted as a non-use preference. It is, in this case, not merely the existence of the Mona Lisa that matters to me, but also my ability to see it regularly. The kind of case that will really illustrate non-use preferences is one where I prefer something and where the only benefits I enjoy come merely from knowing that the thing exists (or that it flourishes, or whatever). Such preferences could (and do, I am sure) exist for the Mona Lisa, but the best illustration of a person who has such preferences is someone who has preferences for the Mona Lisa – wants it to be preserved, for example – but who never even plans (or hopes) to see it.
To say that something has existence value is to say that a person has a non-use preference for it. A thing has existence value to someone if that person cares about how the thing fares though he or she does not use it, does not have first hand experiences with it, and does not engage in any behavior vis-à-vis the thing. What is so interesting about existence value is that if a thing “has” such value, there is essentially no way to prevent people from enjoying the good. This is what Koff has in mind when he says that “non-use or existence values are nothing more than values for a very pure public good.” But note that many things that are not conventionally thought of as pure public goods can have existence value. A grove of ancient Redwoods on my private property is not typically thought of as a pure public good, because I can exclude people from seeing them. But again, what I cannot do is keep people from caring about them, and from deriving some kind of psychological gain from the mere fact that they exist. (Another way of putting this would be to say that all that I could do to keep them from deriving benefits from the Redwoods is to interfere with their awareness that the Redwoods exist.) This is what it means to say that the Redwoods – or whatever one might want to say it about – have existence value. The person who has non-use preferences for the good – the Redwoods on that piece of private property, say – does not see, hear, touch, see, or (tangibly) feel the good, or even plan or hope to do so; but what happens to the good can affect whether that person’s preferences regarding it are satisfied, and thus can truly affect that person’s well-being.

How does this bear on CBA? Notice that a person who has non-use preferences for a good that is technically another person’s property can derive benefits from that good – more specifically, from its continued existence – without ever the owner ever knowing, let alone being able to collect some compensation for this. The market will thus surely not capture
this value. The good’s “price,” were the owner to take it to market, would perhaps reflect the “standard” (use-based) preferences people have regarding the good, but would sure not really reflect the non-use preferences people have regarding it. This has something to do with why CBA exists in the first place: economists recognized that a good’s price may diverge from its “true value” and allows for this good to be shadow priced. But this is not the only implication that the existence of such preferences has for CBA. The existence of non-use preferences makes even more difficult the question of precisely which preferences CBA ought to take into account. More to the point: there is disagreement as to whether CBA ought even to take existence value into consideration, which is to say that there is disagreement as to whether what I have called non-use preferences ought to be counted in a CBA.

Why is this? We can imagine someone who does not live near the Arctic National Wildlife Refuge (ANWR), does not plan or want to visit, does not think that an undrilled ANWR contributes to his health better than it would if drilled, and does not even believe that the reason to refrain from drilling ANWR now is so that we have the option to exploit its resources in the future (when, say, accessing them might be more vital), but who nonetheless prefers that ANWR not be drilled for oil. This person simply derives satisfaction from knowing that this huge piece of land exists in its relatively pristine state, and prefers for ANWR to remain that way. We can even say that he is willing to pay to maintain ANWR in an untouched state. This is a non-use preference. But should this preference be included in a cost-benefit analysis? Well, the fact is that there are people all over the United States, and indeed all over the world, who have these kinds of non-use preferences regarding ANWR. The worry is that a CBA that has to measure ANWR’s existence value, i.e., that has to take
account of the preferences, including the non-use ones, of everyone who has a preference regarding ANWR, that has to take into account the effect on everyone who claims his or her well-being is affected by whether ANWR is left untouched, will simply be an unfeasible CBA.

As Rosenthal and Nelson note:

The range of possible existence values may well be limitless and certainly extends far beyond the current application of the concept. Because existence value refers to any nonuse-related change in the state of the world that affects utility, anything that shapes perceptions of the world becomes potentially eligible for estimation as an existence value. This vastly increases the possible scope of economic analysis and the potential burden of analytical work. If no practical way of limiting the range of existence values can be developed—and none has been established to our knowledge—then there may be no closure to the analytical efforts required of economists and the very concept becomes infeasible.  

But as Koff says in response, Rosenthal and Nelson “are not suggesting that existence values are an illusion. They would surely agree that these values are real; that they are held by some people; and that, if the objects giving rise to these values are diminished, the well-being of these people is similarly diminished.”

“Is it outside the range of economics to inform these decisions about the size and distribution of these values?” Koff asks. “I think not,” he says, “and moreover, I believe we would be derelict in our duty if we turned our backs on these questions. The fact that it would increase the work load of benefit-cost analysts seems a weak reed upon which to base a rejection of non-use value.” In short, it seems more than a little ad hoc, after having identified the policy-level benefits with the satisfaction of

---


preferences – fully-informed and rational ones, or otherwise – to exclude non-use preferences from consideration.

Proponents of CBA thus have the difficult task of specifying which preferences – whose, and what kinds – ought to be considered by the analysis. If they include existence values, the analyst’s workload is potentially enormous; but there are no good theoretical grounds for excluding them. My suggestion that the state is obligated to provide people with as complete as possible information does not make this task any easier. In fact, it seems to compound the problem. If even I, all the way over here in North Carolina, can have preferences – non-use preferences even – regarding ANWR, then the state’s task of disseminating information is a big one indeed. And when more information is disseminated, more people will come to have non-use preferences regarding what goes on in ANWR. This of course makes the analyst’s job of measuring preferences even more difficult than it would have been when many people did not even know anything about, and so did not have any preferences at all regarding, ANWR.

This is indeed a problem, but it is not an insuperable one. I do not maintain that the particular government proposing the policy – be it the federal government, or that of a state, or county, or city, etc. – is responsible for providing everyone in the world with as much information as possible. There are important issues of standing that must be addressed. Properly addressing this issue moderates the concerns articulated in the preceding few paragraphs. If the town of Chapel Hill, North Carolina is doing a CBA in order to determine whether it ought to purchase a tract of undeveloped land and establish this land as a park, it is probably not obligated to take into account the non-use preference that someone in, e.g., California or Texas has for green spaces to be preserved (or developed). The town of Chapel
Hill is under no obligation to provide people all over the world with information about this prospective purchase, even though some people very far away from Chapel Hill would have genuine preferences regarding this tract if they were informed about its existence and the town’s decision. There may be people in far-off places whose well-being is genuinely affected by their mere knowledge that (say) a currently-undeveloped tract was preserved. But people in far-flung places do not have standing in Chapel Hill policy issues. Consider that no one would say that if the town of Chapel Hill were to put the issue of purchasing this tract up for a vote, people all over the world should be able to participate in the vote. Both CBA and voting restrict, and quite reasonably, the number of people who are permitted to participate in the project.110

This is not to say, of course, that there are never issues that are truly national, and in which all people from a wide range of locations should be permitted to have input. ANWR is importantly different from the tract of land in Chapel Hill because ANWR is national land. The Refuge was established by federal laws and using federal money. If the federal government is conducting a CBA on whether to drill in ANWR, it is obligated to take into account the preferences of people all over the country. And if it does so, it is obligated to provide people with complete as possible empirical information regarding the Refuge, though doing so will influence the preferences that people have, and will in fact create preferences in at least some people, including those who before did not even know that ANWR existed. It is obligated to do this in spite of the fact that it will require work to get information out across the country and to measure the preferences that people all over the country have. It is not the case that people all over the world have standing regarding Chapel Hill’s decision,

110 Again, I do much more in Chapter 5 to compare and contrast voting and CBA. Nonetheless, the comparison here is obviously relevant, and indeed too important to postpone until the next chapter.
but people all over the country do have standing regarding the federal government’s decision of whether to permit drilling in ANWR. And if the preferences regarding ANWR that are measured are to be authoritative, are to be taken as legitimately establishing what the policy vis-à-vis drilling in ANWR ought to be, then these preferences must be as well-informed as possible. Addressing the issue of which people exactly have standing on which issue, and why, is beyond my purposes here. What I want to make clear is that the mere fact that people all over the world can have preferences regarding what goes on even in some small, isolated place does not mean that the state is obligated to provide complete information to people all over the world. There are jurisdictional considerations of which to be mindful. I engage with these issues again in the next chapter, and again in Chapter 7. And in the last section of this chapter, I try to respond to concerns about the practicability of my suggestion regarding the importance of ensuring that people provided with, and helped in understanding and processing, relevant information.

Section 4: More on Providing Information and Helping to Ensure Rationality

Having moved away from the case of the city’s decision regarding the 100-acre tract, I want to continue to talk here in more general terms. I have expressed the view that, in general, the less of the relevant information an individual has, the more tendentious the case for identifying policy-level benefits with the satisfaction of this person’s preferences. It is the preferences that people would have if rational and fully informed about empirical facts that CBA should be aiming to measure. Analysts will not be able to supply every piece of information in the world, or even every piece of information regarding a project, and indeed there would be no need for them to do so. This is almost too obvious to merit saying.
Analysts typically know, with respect to a project, what information is important. If a piece of information would dramatically change someone’s specification, then the information is dramatically important; if it would only mildly change someone’s specification, then it is only mildly important.

Still, it will be objected at this point that many of the policy problems that face us, and many for which politicians are most apt to turn to cost-benefit analysts for help, are ones where we face tremendous uncertainty regarding the empirical facts – much greater uncertainty than we face regarding facts surrounding the small tract of land described above (which even in that case may have involved substantial uncertainty). Take, as an example, the question of what the proper policy responses are to threats of global warming. What is the real extent of global warming? How do we distinguish anthropogenic global warming from “natural” rises and dips in the Earth’s temperature? How quickly is such warming proceeding? What can slow it? What will be the consequences if the earth does warm up? How gradually or suddenly will these consequences be manifest? What would the effects on the economy be of making changes now to lower emissions of, e.g., methane, carbon dioxide, nitrous oxide, hydrofluorocarbons, and other such greenhouse gases? What contribution does deforestation make to global warming? Would saving old growth forests do more than planting new forests to sequester atmospheric carbon? These are all important questions, and we do not have precise answers to any of them. Indeed, to at least some of them, our answers are far from precise. There is little disagreement – at least within the scientific community – over the claim that global warming is occurring, but beyond that, there is a great deal of uncertainty.
The same point could be made about many other policy issues. With regard to a dam on the Colorado River: How quickly will silting develop behind the dam? How will fish in the river fare as a result? How will those downstream from the dam who use the river’s water for irrigation be affected? With regard to allowing snowmobiles in Yellowstone: How will animals in the park respond? What will the effects be on the health of the ecosystem? To what extent will the presence of the snowmobiles in the park disrupt other visitors to the park? With regard to reintroducing grey wolves to federal lands in Idaho, Montana, and Wyoming: How quickly will the wolves repopulate? To what extent will they prey on ranchers’ cows and sheep? What effect will the wolves have on controlling the elk population, and how will this affect the health of the broader ecosystem in the region? With regard to certain construction projects in a coastal region that would guard against damage in the event of a massive hurricane: What are the chances that such a hurricane would hit? What damage would it be likely to do if it did hit? How many people would be hurt or killed in the process? What might the extent of property damage be? To what extent would barrier islands, levees, pumps, and the like, make a difference? These are just a few examples, and in each case the answers to the question are difficult to come by.

The general charge, then, is that by focusing on such a simple example as whether or not a small city should develop or preserve a relatively small tract of land, I severely underestimate the extent to which there is a lack of information in so-called “big-ticket” policy problems. I certainly do not want to underestimate this. Nor do I want to come across as terribly naïve about the existence of all of these unanswered questions, or as trying to obscure or underestimate the incompleteness of the information possessed not just by policy

111 This final series of questions is obviously motivated by the recent Hurricane Katrina. The decision to not build the levees in New Orleans higher and stronger was surely based on an assessment of costs and benefits. One wonders what exactly that CBA looked like, and how it will be appraised in hindsight.
make, but even by our best scientists. Still, phrases like “full information,” “well informed,” and so on, are not meaningless in connection with policy problems that involve so much uncertainty. As alluded to above, I do not maintain that we can always know what the consequences of a project will be. Sometimes we certainly can, and other times those with intimate knowledge of the project and technical expertise can reasonably anticipate what the consequences will be. In this case, the consequences that are most reasonably expected, and the basic degree of confidence with which they are expected should be shared. Other times still, the consequences will be ones that no one has or could have foreseen. Here I have said that there is obviously no expectation that these consequences should be shared with people before the fact. But if it seems likely that the consequences can be foreseen by those with intimate knowledge and technical expertise and these consequences are not shared with a person specifying a preference, the specification is not as informative as it could be.

In cases of scientific uncertainty, people’s informed preferences can still be linked with costs and benefits, and will count as informed (with respect to a certain issue) when those whose preferences are gauged are made aware of all of the scientific uncertainty regarding that issue. Before their preferences are registered, people should be made aware of the elements of uncertainty at the empirical level. They should be informed about where disagreement exists, whether one side is a minority position, and who comprises the camps on the different sides of the issue. In cases like those I referred to above, people should hear about what the best-case and worse-case scenarios are. It is difficult to say in general terms exactly what this would look like or require in practice; but we certainly recognize someone who is well-informed on the global warming debate when we meet her. Such people do
exist, though they are not people who know with certainty answers to all of the questions that were posed in the preceding paragraph.

Current debates in journalism about what exactly journalistic balance requires are relevant. Journalists, even in short articles, aim to find and present the most dominant widespread positions and then tell both sides of the story. According to media scholar Robert Entman, “Balance aims for neutrality. It requires that reporters present the views of legitimate spokespersons of the conflicting sides in any significant dispute, and provide both sides with roughly equal attention.” However, neutrality is a deceptively complex notion. Moreover, Boykoff and Boykoff note that “balanced coverage does not, however, always mean accurate coverage. In terms of the global warming story, ‘balance’ may allow skeptics—many of them funded by carbon-based industry interests—to be frequently consulted and quoted in news reports on climate change.” They cite Ross Gelbspan, who, “drawing from his 31-year career as a reporter and editor, charges… that a failed application of the ethical standard of balanced reporting on issues of fact has contributed to inadequate U.S. press coverage of global warming.” Gelbspan writes:

The professional canon of journalistic fairness requires reporters who write about a controversy to present competing points of view. When the issue is of a political or social nature, fairness—presenting the most compelling arguments of both sides with equal weight—is a fundamental check on biased reporting. But this canon causes problems when it is applied to issues of science. It seems to demand that journalists present competing points of view on a scientific question as though they had equal scientific weight, when actually they do not.\textsuperscript{113}


\textsuperscript{113} Ibid.
Boykoff and Boykoff “analyzed articles about human contributions to global warming that appeared between 1988 and 2002 in the U.S. prestige press: the New York Times, Washington Post, Los Angeles Times and Wall Street Journal… From a total of 3,543 articles, we examined a random sample of 636 articles. Our results showed that the majority of these stories were, in fact, structured on the journalistic norm of balanced reporting, giving the impression that the scientific community was embroiled in a rip-roaring debate on whether or not humans were contributing to global warming.” Among other things, they “found that 53 percent of the articles gave roughly equal attention to the views that humans contribute to global warming and that climate change is exclusively the result of natural fluctuations.” But within the scientific community, there is much more consensus. Given the extent of scientific agreement on the issue, it is a mistake to give so much attention to a position that is held by such a small number of people. It is not wrong to tell people that there is a view that the climate change we are experiencing is just a natural fluctuation, but it is impermissible to do so without mentioning how few scientists agree with this view. Journalists, it seems, are now catching on. The cover story for the April 9, 2001 edition of Time Magazine was about global warming. Almost 5 years later, in the April 3, 2006 edition, global warming is again the cover story. The newer story is significantly longer, and yet less attention is given to the views of the skeptics. In fact, the new story gives no coverage at all to the skeptical position; rather, global warming is talked about as an inevitability, and one that will affect us sooner rather than later.

Clearly, the issue of what constitutes full information in connection with complex policy problems is not a simple one. In any case, the fundamental point is that full empirical

---

information is a notion that makes sense even where there is considerable scientific uncertainty, and what it requires is that people have a good understanding of the scientific uncertainty. There is, however, one kind of case that is particularly difficult for my argument to handle. Recall the scenario, considered within the discussion of the informed-preference account of well-being (Chapter 2, Section 5), where a person \( P \) prefers for her spouse \( S \) to be faithful. In two of the cases I imagined, \( S \) was not faithful. I argued there that it is possible, though not required, that \( P \)'s informed and rational preference could be not to know that \( S \) has cheated. In other words, there may be cases where an individual’s informed and rational preference is to lack full information. Scenarios such as this are ones that it is difficult for the claim that I am here defending, \( \text{viz.} \), that the state should help provide people with complete as possible information, to handle. If my well-being consists in not knowing \( X \) – which is to say that my fully-informed preference is not to know \( X \) – then when the state provides me with full information, this may include information that or about \( X \). In this case, the very act of providing information frustrates my informed preference. Thus the state is not promoting my well-being, is not helping me to experience feelings of satisfaction, and is not clearly respecting my autonomy.

Thankfully, I do not think there are many real-world policy-relevant examples where this dilemma arises. There are very few policy-relevant cases where my fully-informed preference would be not to have full information. In fact, I can not think of a single realistic policy-relevant example. This is probably because most cases where one might have a preference for lacking information are when one is, as it were, resigned to one’s fate. If something bad is going to befall one, and one could not do anything at all to change it, one may well simply prefer not to know about that impending bad thing. Policy-relevant cases
are not like this: the very fact that policy analysts are addressing the cases makes them ones where one’s preference has the chance of being influential, and so one would typically prefer to be well informed. At the theoretical level, I guess I would say that if, but only if, the state knew that my rational and fully-informed preference would be not to have full information, it would be excused from working to provide me with such information. At the practical level, it would of course be essentially impossible for the state to know this. But as I have said, at the practical level there are few, if any, instances where this situation is likely to arise. Perhaps I should leave open the possibility that one might simply prefer to have no part whatsoever in public policy decisions. This person would prefer not to be provided with any policy-relevant information. If that is her preference, she should not be forced to have her preference regarding the policy registered. But in a case like this, it may actually be easy enough to avoid giving this person such information: if this person is approached with a survey, the analyst will presumably learn right away that this person prefers to have no part in public policy decisions.

Of course, the question is not just how to conceive of and help bring about informed preferences, but also about how to conceive of and help bring it about that these are rational. Again, when I say rational, I do not mean that the state should help individuals remove all affect or emotion from decisions. Preferences are not pure affect, but they are also not pure reason; they have an essential affective component. Moreover, a preference can also be rational in my sense while being different from what the “average person” would embrace, and indeed even while being what this average person might find to be extremely strange. What is required for preferences to be rational is simply that they not be inconsistent with one another. Importantly, this includes consistent with one’s long-term preferences, which it
may be more natural to call one’s long-term goals, or even one’s life plan or ultimate ends. Such talk is permissible – it is not part of a teleological theory the likes of which I expressed deep skepticism about back in Chapter 2 – if the ends are ones that each (well-informed) individual sets for herself. The state should stay neutral, as far as possible, regarding what ends people adopt. However, this does not mean that it is wrong to challenge people to think deeply about these ultimate goals. These are not so sacrosanct that people cannot also be encouraged to reevaluate the life plans and goals they adopt and pursue. If a person reflects carefully and deliberately about her long-term goals, takes them to be ones worth pursuing, and develops well-informed preferences with respect to a policy that are consistent with these long-term goals (and all her other preferences too), then there is no good basis on which the state, or any analyst, may call these preferences irrational.

Let us consider another example. Imagine that a relatively small town in Africa is trying to decide whether to invest in a sanitation system that would allow for sewage connections to, and water-sealed toilets inside of, the town residents’ homes. Suppose that currently there are a few central locations for going to the bathroom. Lines sometimes develop at these locations. People are not always willing to wait in line – sometimes are not even willing to walk to one of the centralized locations – and so will sometimes relieve themselves in the bushes, the creek, etc. Whittington et al note that “sanitation planning in developing countries has all the characteristics of what policy analysts have termed ‘wicked problems.’” There are several reasons why improving sanitation service poses such a complicated policy and planning problem. First, the costs of conventional water-borne sewerage solutions (e.g., on-site facilities such as water-sealed toilets, sewerage networks, and wastewater treatment) are expensive—on the order of US$25-35 per household per
month (Lauria et al, 1995). This is equal to the total monthly income of many poor households in urban areas of some developing countries.” They write that an additional “challenge to effective sanitation planning is the lack of public awareness of the benefits of sanitation services.” (Presumably they do not use the term “benefits” here to mean “policy-level benefit”; their suggestion is not that sanitation services are a policy-level benefit regardless of what people think or feel about them, but rather that sanitation services provide health benefits.) I read this as Whittington et al acknowledging one thing that I have been arguing, namely that incomplete information presents an obstacle, though the article does not go on clearly to suggest, as I have, that people should be given complete as possible information before these people articulate their WTP. They should, e.g., be well-informed regarding the health benefits of keeping excrement out of local waterways (which would obviously be easier to do if each residence had its own working indoor toilet).

Imagine now that people register a relatively low WTP for town-wide in-home sanitation services, and that their WTP is much higher for cable television. They do not currently have either of these things, and by their own account would much rather that cable TV be brought into every home than that there be functional water-sealed toilets in each home. Whittington et al note that being confronted with situations such as this “has often led public health specialists, planners, and engineers to rely solely on their expert opinion and to ignore presumably uninformed wishes of households.” In response to an expressed preference for cable TV over sanitation services, water and sanitation engineers may feel moved to somehow make people take the sanitation system. “Such professional arrogance,”


117 Ibid.
says Whittington, “has resulted in many spectacular planning failures.”

This is an important point, but then surely so too has simply promoting the satisfaction of whatever preferences people happen to have, without analysts working to ensure that these preferences are well informed. What I really want to note here is that the issue of consistency among preferences is left entirely out of Whittington’s discussion. My guess is that many people – many Americans, at least, and this will include many who really enjoy watching television – will not just presume the preference for television over a water-sealed toilet that would contribute to health-promoting sanitation conditions to be uninformed, but will also be tempted to call the preference an irrational one. I will admit that to me, the preference for TV over the toilet seems extremely bizarre. However, simply dismissing this preference as irrational would be a mistake. Importantly, though, so too would simply promoting the satisfaction of this preference without analysts working to ensure that this preference is consistent with the individual’s other preferences.

No one should obstinately and categorically refuse to label the satisfaction of an individual’s preference for TV over toilets as a benefit simply in virtue of the belief that this preference is irrational. Instead, the state (in this case, the town) and the analysts it hires should work not just to give people information, but also to ensure that people have processed the information and understand, e.g., the relative health implications of getting TV and getting toilets (among other empirical facts). Indeed, more than this should be done: the analyst should press people not just to say what they prefer in this single situation, but also to think about what their long-term preferences, i.e., their broad goals, are, and even pressed to think about whether these goals are ones that they are deeply committed to embracing. The analyst should the press the individual to consider whether what she prefers with respect to

---

118 Ibid.
the case at had is genuinely consistent with those long-term goals that she on reflection embraces. If this person then truly believes that having a TV is a better means than having a water-sealed indoor toilet would be to the end that she firmly and on reflection embraces, then it is the satisfaction of the preference for the TV that should count as a benefit for that person. My recommendation here will surely have implications for both the stated- and revealed-preference methods. I leave it to others – to technicians – to decide exactly how to achieve this in practice. My point has principally been to suggest that the theoretical considerations that recommend this task to the technicians are solid. I will say just that I suspect that the need to check for and encourage consistency among an individual’s preferences will make considerable difficulties for the revealed-preference method, which just adds to my skepticism regarding that way of figuring costs and benefits. On the other hand, it would not be an impossibly difficult suggestion for the stated-preference method to accommodate. More on this will be said in the next suggestion.

Another worry should be entertained here. Most people’s preferences change over time. Some have been concerned that this makes a problem for CBA. Because I have been addressing the rationality condition, and because changing preferences raise the specter of inconsistency, it seems fitting to discuss this concern here. The question is: Does it count as a gain if an outcome that I once preferred but no longer do obtains? James Griffin writes:

If our desires never changed with time, then each of us would have a single preference order, by reference to which what most fulfilled his desires over the course of his life could be calculated. However, life is not so simple; preferences change, and not always in a way that allows us to totally discount earlier ones. Suppose that for much of his life a person wants his friends to keep him from vegetating when he retired but, now that he is retired, he wants to be left alone to vegetate. Is there any
This is an interesting theoretical question for anyone who is interested in making an important place for people’s preferences in politics. And for CBA, it is not just an abstract issue. Imagine that a town government is considering building a recreational center for senior citizens. What weight should be given to the lifetime of preferences this man had for being active in his retirement, being as though now that he is actually in his retirement he simply wants to vegetate? Or imagine that I desire that a road in my town be widened; I sincerely specify a positive sum of money that I am willing to pay for the project. A thorough CBA is then conducted, the project is approved, and the road widening begins; but prior to the project’s completion I have a change of heart – for whatever reason – and now desire that the road not be wider than it originally was. Which of these preferences will the cost-benefit proponent count as having normative weight?

The response is two-fold. First, my suggestion that the state provide people with, and help them think as clearheadedly as possible about, information relevant to the policy question is actually a way of avoiding having to ask the aforementioned questions in the first place. The better people organize and process information, the more carefully they reason, the less likely this sticky problem is to present itself. The more someone is encouraged to think carefully about her final ends and the effectiveness of various means to achieving these ends, the less likely it will be that the preference she articulates will differ from those she has long had. The old man may be led by reflection to see that his preference for vegetating is not actually consistent with the plans that deep down he still does have for himself. For

example, perhaps a kind of purposelessness that he felt upon retiring kept him from seeing this clearly on his own. But there is, of course, no guarantee that his preference to vegetate is not a reasonable preference for him. Perhaps he is able to see something now that he has reached retirement that he was not able to see throughout the earlier parts of his life in which he expressed a preference to be active in his retirement. Perhaps this means that he has information now that he did not have earlier in his life, or perhaps it means that there is something about actually being retired that quite reasonably leads him to process the information he has always had in a new way. This possibility cannot be ruled out on principle. In any case, I am comfortable saying that if his sincere, well-informed, well-considered (i.e., rational) preference is to be able to vegetate, the state should count the satisfaction of his preference for vegetating as a benefit, even in spite of his previous long-standing preference for an active retirement.

Even if all of the aforementioned steps are taken and an individual’s preferences nonetheless change, I see no reason why this should concern proponents of CBA. The standard CBA is one conducted *ex ante*, that is, before the project is begun. The results of the CBA make it the case that the project goes forward or that it does not. *Ex ante*, a decision must be made. What counts as a cost or as a benefit is relative to what each individual prefers at the time the decision must be made. But CBA’s can also be carried out *ex post*, i.e., after the fact. If an individual’s preferences change after the project is completed, then the *ex post* analysis will take account of the new preference. With respect to this person, what counts as a benefit relative to this analysis would then be different from what counted as a benefit relative to the *ex ante* analysis. Depending on what other people’s preferences are and what theory of action (what view regarding aggregation, etc.) is settled on, the *ex
post analysis may recommend turning around and undoing the work that was just done, or initiating right away some altogether different project. In fact, for some massive public investment projects for which much time passes between the beginning and end of the project, a new CBA may be conducted somewhere between these two points in time. The same goes for a regulation: even after a regulation is put in place, a later CBA can be conducted to assess the costs and benefits of keeping that regulation in place (or instituting a new one).

**Section 5: Responses to Concerns About Practicability**

There are reasonable concerns about the practicability of my suggestions, and I aim in the last section of this chapter to assuage some of these. Certainly, it is not easy to conduct a CBA. They cost money, and the more expensive the CBA is, the farther the project’s benefits will need to outweigh its costs in order for the project to pass muster – assuming that the cost of doing the CBA is counted as a cost of the project, and that the project will be approved only in the event that its benefits outweigh its costs. Working to ensure that people are well-informed and rational will may be a very cumbersome task for the analyst—more cumbersome, in any case, than conducting a CBA without worrying at all about the informedness or rationality of those individuals whose preferences are being gauged. This does not change the fact that the less of this information the individual – even the sincere one – specifying a preference has, the more grains of salt, as it were, the specification should be taken with, the less authoritative the specification should be considered to be, and the less reason we have for counting the satisfaction of the preference as a policy level benefit. But it does raise a number of important questions.
Though I do not take myself yet to have said anything directly about what at the outset of the project was dubbed Claim 3 – that is, though nothing I have said yet has directly supported or opposed the claim that all preferences can accurately be expressed in terms of WTP – there is a point regarding money that is clearly relevant. Providing people with information and helping them to understand and process it before making an assessment of costs and benefits is going to require time and money. The question then arises: How much time and money is it worth devoting to this purpose? I do not have an answer. Clearly, there will be a point at which too much money and time are spent on this task. One proposal for figuring out how much time and money to spend on the task is to perform a cost-benefit analysis. But then we need to know what to count as a cost and what to count as a benefit. It was wanting to know precisely this that led in the first place to the suggestion that people be provided with information and helped to understand and process it. It is clear, then, that this proposal leads to a regress. The question of how much time and money to devote to informing people must not be answered by CBA. However, I am optimistic that cost-benefit analysts will be able, without too much difficulty, to help people become well informed and rational, and for this reason I think analysts will generally be able to satisfy my proposal without devoting to the task an excessive amount of time or money. I will try to explain some of the thoughts I have that ground my optimism.

To be sure, the requirement that the analyst reasonably ensures that individuals be well informed before registering their preferences is not fulfilled just by generally ensuring that the relevant information is available somehow, somewhere, at least to those willing to comb through every book, booklet, textbook, guidebook, handbook, brochure, catalog, circular, dictionary, encyclopedia, flier, folder, folio, journal, leaflet, magazine, manual,
pamphlet, program, tome, tract, and treatise. There must be a commitment to making relevant information readily accessible to people. Taking this commitment seriously may well add significantly to the analyst’s task. It obviously takes time and energy to compile this information and put it into a form in which it can be apprehended as easily as possible. It takes time and energy for the individuals administering contingent valuation surveys to share this information with respondents. For real analysts to fill this role well, a great deal of skill would be required, which means that a great deal of costly and time-consuming training would be required, and the analysts’ work out in the field may well also be quite time consuming and expensive. This need not mean that there is no place for CBA, or for CV, but may be enough to suggest that some changes are in order regarding how they are done.

For one, public debate can and should play a bigger role in the policy process, even if it going to still terminate in surveys asking people to specify their preferences. Public debate is one of the ways of helping people to gain complete as possible information. It can also help people organize and process the empirical information. It can reveal to someone that her preferences are inconsistent with one another. Public debate has the ability to lead people to think harder about their priorities, the life plans, their ultimate goals. It is healthy to subject an individual’s preferences to public debate, to challenge her to defend her preferences against competing views regarding what kind of preferences, and indeed what kind of life plans, are most worthy, and to embrace her goals, plans, and particular preferences only after careful consideration. These are all incredibly important functions of public debate. As I have said, an individual’s preferences – even those that are well informed and consistent with the individual’s other preferences – are not so sacred that it is wrong to put the individual into a context in which the topic regarding which she has preferences is
publicly debated and scrutinized. Not are preferences so fundamental or primitive that it is
*purposeless* to do so.

Consider first an issue that (so far as I know) has no policy implications whatsoever. I start with it because despite its simplicity it is instructive. Suppose that a friend of mine prefers listening to a wide variety of other British Invasion bands over listening to The Who. I find this hard to believe! I engage him in discussion to see what he really knows about The Who. Does he know the band’s history, their influences, what the many songs they have recorded sounds like? Even if I am convinced that he does, I could still say something like this: “You, friend, should nonetheless rethink your preference. Listen more carefully to the drumming of The Who’s Keith Moon. Surely upon doing so you would acknowledge that he was more skilled drummer, and brought much more energy to his playing, than Ringo Starr (from The Beatles) or Charlie Watts (of The Rolling Stones).” I could even present him with reviews written about Moon by various people who my friend would concede are authorities on rock and roll drumming. I might continue: “You surely know that The Who pioneered the rock opera, but I think you discount how truly special a fact about them this is. The rock opera gave rock music a kind of substance that it had never before and allowed for a fuller development of a theme over the course of an album. What did the bands that you prefer over The Who do to match this? Is it not also noteworthy that The Who’s inversion of the typical structure of having the bassist lay down the rhythm and the guitarist play the melody makes them different in a neat way. That surely counts as one of their great virtues, and one that should earn them a more favorable place in your catalog of great bands.”

The same thing goes when the preference is one regarding a policy issue. Public debate can help someone gain, organize, and reason well about information pertinent to a
policy decision. Even once an individual has a basic grasp of the information, it is good for her to hear about how and why these facts matter to others, what preference other people form in response to the empirical information, and why. Subjecting a policy issue – and in a sense then, people’s preferences regarding that policy proposal – to public debate is a way of ensuring as far as possible that each person is comfortable upon reflection with the preference she expresses. This is all work that the analyst could do out in the field as she approaches individuals one by one with the survey device. But the analyst’s task is much less cumbersome when the registration of a preference is preceded by public debate. CBA suffers, and my general recommendations are much more unwieldy, to the extent that an analysis attempts to bypass public debate on the way to making policy recommendations. Certainly, CBA is not *incompatible* with public debate. It is possible for there to be debate – extensive debate even – on an issue, and only *after this* for preferences to be solicited via the revealed and/or stated preference method(s). But this is typically not how it works. CBA is typically used as a way to get a policy decision without having to go through the “messy,” time consuming process of public debate. Sometimes this is understandable. There is not enough time, space (be it in newspapers, television programming, town halls, etc.), or energy for public debate on every policy issue. Still, proponents of CBA make a mistake if they fail to acknowledge the value of debate. For this reason, the more important the policy issue, the more important it should be to have public debate, and the more skeptical we should be of policy recommendations that are made in the absence of such debate. Oftentimes, the real blame here rests with the elected officials who solicit CBAs in the first place.

Experts have an important role to play in uncovering the information that is passed on to citizens, in helping to convey this information to people, and in helping people to
understand and process this information. But their role is not one that renders the preferences of lay people, of non-experts, unimportant. Ideally, the experts are not to say what is in people’s interests, and are not to by themselves decide policy. One may well wonder why, if the information the experts have is so important, we should not just let them make the decision? Again, doing so is not consistent with respecting people’s freedom, and moreover, the experts are simply not always, despite their access to such good empirical information, in a position to say what will promote people’s well-being, since, as I explained earlier, the informed-preference account of well-being is not a convergence account. Two people reasoning equally well and with the same complete information will not necessarily respond to the information with the same preferences. The experts are (essentially by definition) possessed of superior empirical information, but the normative information – how good or bad certain outcomes would be – must be supplied by the individuals affected by the policy.

The internet also has a significant role to play. Obviously not everyone has internet access or knowledge of how to use it. Moreover, simply tucking a piece of information in some obscure “corner” of some obscure website is not going to suffice to satisfy the demands of the provision of information. But it possible to house relevant information on well-publicized websites. Consider, for example, the following information provided by the U.S. Environmental Protection Agency:

In 1984 a deadly cloud of methyl isocyanate killed thousands of people in Bhopal, India. Shortly thereafter, there was a serious chemical release at a sister plant in West Virginia. These incidents underscored demands by industrial workers and communities in several states for information on hazardous materials. Public interest and environmental organizations around the country accelerated demands for information on toxic chemicals being released “beyond the fence line” – outside of the facility. Against this background, the Emergency Planning and Community Right-to-Know Act (EPCRA) was enacted in 1986.
EPCRA’s primary purpose is to inform communities and citizens of chemical hazards in their areas. Sections 311 and 312 of EPCRA require businesses to report the locations and quantities of chemicals stored on-site to state and local governments in order to help communities prepare to respond to chemical spills and similar emergencies. EPCRA Section 313 requires EPA and the States to annually collect data on releases and transfers of certain toxic chemicals from industrial facilities, and make the data available to the public in the Toxics Release Inventory (TRI). In 1990 Congress passed the Pollution Prevention Act which required that additional data on waste management and source reduction activities be reported under TRI. The goal of TRI is to empower citizens, through information, to hold companies and local governments accountable in terms of how toxic chemicals are managed.\(^\text{120}\)

The TRI is available online (http://www.scorecard.org) and is indeed a fantastic resource for those seeking the kinds of information it makes available. It provides information that is relevant – though not all such information – for certain regulatory decisions. The question for policy technicians is how to ensure that all those individuals whose preferences are being gauged are made aware of this information.

This leads to what is perhaps my most important response to the concern about practicability. I can imagine a kind of approach to making public investment and regulatory decisions that blends elements of traditional CBA with elements of the legal system’s use of a jury. The idea is that a group of individuals – not just specialists, policy wonks, scientists, etc., but the kinds of ordinary people who make up a standard jury – would come together to consider a policy issue in great detail. These individuals would be selected at random. Perhaps there would then be some more nuanced selection process from among these individuals. Perhaps participating when one is called would be touted as one’s “civil duty” but not legally required. Perhaps one would – absent special circumstances – be legally required to participate like when one is currently called for jury duty. I leave these questions unanswered. In any case, it would be with these individuals’ preferences that the costs and

\(^{120}\) See http://www.epa.gov/tri/whatis.htm.
benefits of the proposed policy would be identified. But of course it would not simply be the preferences that these people enter the process with that would matter. Rather, it would be those they express after they have been exposed to and deliberated over a robust amount of information relevant to the policy. Consistent with my desire throughout this project to stay neutral on CBA’s theory of right action, I do not have a view on whether the members on this kind of “jury” should be required to come to a unanimous decision (like a normal jury must) or whether each individual should, after all the discussion has taken place, simply specify whether she counts the policy as a cost or benefit (and to what extent), with these costs and benefits then being added up in some particular way.

The fact that it is only a relatively small group of people whose preferences will determine the costs and benefits of a policy that will affect many people, including people not on the jury, should not cause concern. Contingent valuation studies typically approach only a random sample of people within the target population. There are methods of doing this that generate reliable results. Of course, it is possible that someone who would have had a very extreme preference – and incredibly high WTP, say, or an infinite WTA (on the supposition that these measures are employed within this jury) – and then the policy decision is affected by whether or not this person is included. Suppose this person’s incredibly strong (informed and rational) preference favors the policy being considered. She could throw off the generalization. If she is on the jury, will the total benefits of the policy be exaggerated? If she is not on the jury, will the total benefits of the policy be underestimated. This is a complicated question, but the problem that underlies the question is no more severe for the jury model of decision making that I have proposed than for the standard model of
approaching a randomly selected group of people with a survey and then generalizing from these findings.

There is, however, much to recommend the jury model. Again, for the average policy, the empirical information will be quite extensive and complex. With respect to the average policy, I cannot imagine that those administering CV surveys will be able to get each respondent to the point where she counts appropriately well informed and rational. It is in this context, but perhaps only in this context, that ordinary citizens can be sufficiently well-informed about a policy for their preferences to tell us all that we want them to tell. Suppose, however, that we become convinced that even this method – more generally, that no method – does a satisfactory job of supplying people even relatively accurate and complete policy-relevant information and helping them develop reasonably rational preferences in response to this information. Again, I am confident that this problem can be avoided; but if it does happen that people are clearly ill-informed or not reasoning well, the proper response is not to just give up on informing people and to proceed, via the revealed-preference or the stated-preference methods, recording people’s actual preferences altogether ignoring the fact that they are uninformed or irrational. These uninformed or irrational preferences should in this case be recorded and given some weight in the policy decision, but input could at this point also come directly from experts. The view is not the experts always know better. Nonetheless, in cases where people are irredeemably ill-informed or irrational, the next best alternative to creating actual informed and rational preferences is to give a select group of experts a significant role in crafting the policy. Ultimately, decision makers would then have to balance these inputs, and more generally, the various aims that I discussed in so much detail in the previous chapter.
Two points mentioned earlier bear repeating before the chapter closes. First, it is possible that a person simply does not want to be so involved in the policy process. She just wants the state to take care of her, that is, to make decisions on her behalf and with what the state takes to be her best interests in mind. I do not want to deny that there are people of this sort, or to insist that it is impermissible for them to want this. If someone has an informed and rational preference for being disengaged from the policy process, then I would certainly not insist that this preference nonetheless be required to be involved. Second, I have never maintained that people need to express their preferences directly on each and every policy issue. This would be ideal in a world where there were fewer policy decisions, or where people were better informed or could more easily handle significant amounts of information. But in the world we live in, it would be silly to take the position that this needs to happen. There are other ways for people to express preferences, and principal among them is by electing officials who will (at least in theory!) devote time to becoming informed on policy issues and then making decisions. This is to say, in other words, that sometimes the well-informed and rational preference that we are interested in is the one that people register at the poll for representatives who will make policy decisions on their behalf. (The entire next chapter is devoted to the issue of voting as a means of registering individual preferences.) This is not inconsistent with anything I have argued thus far. The aim should be to create a well-informed and rational group of voters who will elect decision makers. These decision makers are then at liberty to make policy choices based on their impression of people’s well-being. In this case, public officials should continue to inform people as far as possible about the policy process, because this will ensure that when they make decisions they will be satisfying people’s actual informed preferences, and will be respecting people’s autonomy.
None of this, though, diminishes the point that where analysts are committed (directed by politicians, etc.) to recording people’s preferences regarding a particular policy decision (via some means other than voting), there must be an accompanying commitment to provide people with policy-relevant information and help them process this.
CHAPTER 5: WHY CITIZEN PREFERENCES AND SOCIAL VALUES DO NOT THREATEN COST-BENEFIT ANALYSIS

Section 1: Introduction

The previous three chapters were on the whole given to exploring the question of what role individual preferences ought to play in public policy, and more specifically whether there is good reason for counting the satisfaction of individual preferences and a benefit and their frustration as a cost. This chapter continues to probe the same basic topic, but it explores decidedly different kinds of arguments. If one is going to challenge the role in CBA that economists standardly assign to individual preferences, it seems that one ought to have in mind a candidate for what should stand in place of individual preferences. If benefits are not to be identified with the satisfaction of preferences, then what is it with which they should be identified? One obvious proposal, we have already seen, is to identify benefits with gains in well-being, which may pull apart from the satisfaction of individual preferences. But this is not the only proposal, and at least certain competing proposals have had significant influence. I this chapter, I consider and reject several of these competitors.

One such proposal may not so much deny that CBA is right to give a role to preferences as it insists that there are, as it were, multiple kinds of preferences and that CBA attends to the wrong kind. Roughly, the idea is that people take on different roles at different times or in different contexts. We are sometimes consumers, and have consumer preferences that correspond to this role; while we are at other times citizens with corresponding citizen
preferences. Others say – seemingly along quite similar lines – that the difference is between consumer preferences and citizen values. It is surely interesting to note that while CBA is frequently disparaged – at least in certain circles – voting is widely esteemed, and is indeed counted as a fundamental right. Those who argue that CBA inappropriately treats people just as consumers typically claim that voting is superior to CBA as a mechanism for public decisions because voting engages people as citizens.

The insistence by CBA’s opponents on a consumer/citizen distinction is, I believe, motivated by a sense that the use of “willingness to pay” (WTP) is problematic. Nonetheless, the consumer/citizen distinction is not needed in order for that problem to be diagnosed or treated. In fact, the distinction is part of a misdiagnosis and serves only to obscure the problem. My view is that there are reasonable grounds for concern with CBA’s commitment to the use of WTP, and I explore these later. Importantly, though, these concerns can be pressed without impugning the claim that individual preferences – especially, as I have argued, informed and rational ones – play in CBA. The examples that appear to illustrate the existence of these two different kinds of preferences, or of citizen preferences and social values, can be explained without any appeal to a distinction between citizens and consumers or between preferences and values. More specifically, they can be explained by appeal to the idea of the Prisoners’ Dilemma. The preferences people revealed in the market frequently are different from those they reveal or express in political contexts, but this is because collective-action problems frequently constrain how it makes sense for people to act in the market. Importantly, though, it does not mean that there are consumer

---

121 A brief anecdote: I recently heard a well-known economist say that the ancient Greek city-state with its direct participatory democracy was the ideal mechanism for public decision making. It is, the economist suggested, just for logistical reasons – which, granted, are not trivial, but somehow feel to me cheaper than good theoretical reasons – that we turn so often to CBA. He was not quite disparaging CBA, but to my surprise was also not exactly ardently denying that it is a second-class decision-making tool.
preferences and citizen ones, and one of the strengths of CBA is that it creates the possibility to inquire as to what people would prefer if they were not constrained by the collective-action problem. Beyond situations in which there are collective-action problems (and to some extent even *in* such situations), I argue that the tendency a person has to spend money in self-interested rather than community-minded ways is not something for which CBA should be faulted. In connection with this, I argue that voting is actually committed, fundamentally, to doing just what I take CBA (at its best) to be at least aiming to do, namely registering and making public decisions in response to individual preferences. Difficulties that plague CBA in registering non-use preferences and the interests of future generations plague voting to no less an extent. This is not to deny that there are not important differences between voting and CBA as it is traditionally taught, understood, and practiced. Where they differ, it is sometimes in ways that favor CBA. For example, unlike voting, CBA aims to capture not just the direction but also the strength of a preference. Where the differences favor voting, the real problem is with CBA’s commitment to expressing all preferences monetarily, not with its commitment to identifying benefits with the satisfaction of preferences (again, at least informed and rational ones).

An additional (and indeed final) proposal to be considered in this chapter is that CBA is flawed because it bases public policy decisions on mere individual preferences rather than on social values. (Despite the similarity in rings, this is actually a different criticism than the earlier one.) The charge is that there are principled commitments of society, and that these can conflict with what individual preferences would recommend. This is, I believe, a reasonable objection only to the most extreme forms of CBA. Few proponents of CBA would deny that there may be certain fundamental restrictions on what the state may do –
restrictions that may, for example, be enshrined in a constitutional document (state or federal) – or that public policy decisions cannot contravene these restrictions. These restrictions may mean that, in certain cases, one might have a preference the satisfaction of which does not count as a benefit; but this is because that which is preferred is constitutionally forbidden and no reasonable advocate of CBA would even recommend that such an analysis be conducted for a policy that would clearly fly in the face of one of these restrictions. This is no threat to CBA. Moreover, I will argue that even in these cases, individual preferences are relevant in a number of ways, including in reshaping that constitutional document.

Section 2: Citizens Versus Consumers

It is clear from earlier parts of this project that Mark Sagoff is critical of the view standardly endorsed by economists and other proponents of CBA regarding the role that individual preferences should play in policy analysis. One of Sagoff’s influential arguments – one, in fact, that has thus far in this project not been referred to – goes as follows:

The question arises… whether what we want for ourselves individually as consumers is consistent with the goals we would set for ourselves collectively as citizens. Would I vote for the sort of things I shop for? Are my preferences as a consumer consistent with my judgments as a citizen?

They are not. I am schizophrenic. Last year, I fixed a couple of tickets and was happy to do so since I saved fifty dollars. Yet, at election time, I helped to vote the corrupt judge out of office. I speed on the highway; yet I want the police to enforce laws against speeding. I used to buy mixers in returnable bottles—but who can bother to return them? I buy only disposable now, but, to soothe my conscience, I urge my state senator to outlaw one-way containers. I love my car; I hate the bus. Yet I vote for candidates who promise to tax gasoline to pay for public transportation. I send my dues to the Sierra Club to protect areas in Alaska I shall never visit… And of course, I applaud the Endangered Species Act, although I have no earthly use for the Colorado Squawfish or the Indiana bat. I support almost any political cause that I think will defeat my consumer interests. This is because I have contempt for –
although I act upon – those interests. I have an ‘ecology now’ sticker on a car that drips oil everywhere it’s parked. 

Despite some shifty wording – Sagoff sometimes distinguishes between two types of preferences, namely consumer preferences and citizen preferences, while at other times the distinction appears to be between consumer preferences and citizen values – the general point is sufficiently clear. CBA is purported to capture something fundamentally different from that on which public decisions ought to be based (and, as will be discussed at more length later, from what is registered via the vote); and thus proponents of CBA are confused in simply recommending individual preferences as the proper basis for public policy decisions. The charge is that through its use of WTP, CBA generates policy decisions from each individual’s allocation of her income as a consumer, which is to say, her actions “as a private individual determined by self-interest and… personal wants.” However, “the individual voter dealing with political issues has a frame of reference quite distinct” from this: she acts as a citizen, which is to say, “as a political being guided by [her] image of a good society.”

For support, Sagoff cites Stephen Marglin, who says:

> The preferences that govern one’s unilateral market actions no longer govern his actions when the form of the reference is shifted from the market to the political arena. The Economic Man and the Citizen are for all intents and purposes two

---

122 Sagoff, “At the Shrine of Our Lady of Fatima, or Why Political Questions Are Not All Economic,” *Arizona Law Review* 23 (1981), p. 1286. These same basic views are articulated by Sagoff in other pieces. See, for example, “Free Market Versus Libertarian Environmentalism,” *Critical Review* 6, pp. 211-30. Though the contrast here is worded as one between consumer *preferences* and citizen *judgments*, this is not a wording to which Sagoff adheres, as other passage cited will show. The distinction between preferences and judgments is an interesting and important one. Indeed, I explore it elsewhere (*citation deleted for blind review*), and so will not repeat that discussion here.


124 Sagoff, ibid.
different individuals. It is not a question, therefore, of rejecting individual preference maps; it is, rather, that market and political preference maps are inconsistent.\textsuperscript{125}

Sagoff says that for this reason, “social choices optimal under one set of preferences will not be optimal under another,”\textsuperscript{126} and the economists’ notion of “optimality” is effectively undermined.

In their recent work of substantial influence, Frank Ackerman and Lisa Heinzerling also embrace the consumer/citizen distinction. They believe that CBA fails “to take into account the difference between citizens and consumers” and “omits the possibility that people will have different preferences when they take on a different role.”\textsuperscript{127} They write that CBA is an approach on which “valuation of environmental benefits is based on individuals’ private preferences as consumers or workers, not on their public values as citizens.”\textsuperscript{128} They cite as the “classic example of this distinction” none other than Sagoff, when he “found that his students, in their role as citizens, opposed commercial ski development in a nearby wilderness area, but, in their role as consumers, would plan to go skiing there if the development were built.”\textsuperscript{129} Ackerman and Heinzerling say that “there is no contradiction between these two views: as individual consumers, the students would have no way to express their collective preference for wilderness preservation. Their individual willingness

\begin{footnotesize}
\begin{enumerate}
\item Sagoff, ibid, p. 1287.
\item Ibid, p. 1566.
\item Ibid, pp. 1566-67.
\end{enumerate}
\end{footnotesize}
to pay for skiing would send a misleading signal about their views as citizens.\textsuperscript{130} They also say:

Americans are notoriously bad at saving money on their own, apparently expressing a disinterest in the future. Still, Social Security is arguably the most popular entitlement program in the United States. The tension between Americans’ personal saving habits and their enthusiasm for Social Security implies a sharp divergence between temporal preferences of people as consumers and as citizens.\textsuperscript{131}

Elizabeth Anderson makes essentially the same argument. Focusing specifically on decisions that have environmental implications, she says:

Because market prices and willingness-to-pay statistics generally reflect individuals’ valuations of things only as satisfying their private wants and interests, they do not capture all the ways people value environmental goods. The preferences people express in their roles as consumers therefore do not capture all the concerns they have. So people in their roles as citizens debating public policy do not and should not take the preferences they express in their market choices as normative for public purposes.\textsuperscript{132}

This same language has shown up even in a \textit{New York Times} Op-Ed piece by Robert Reich, who writes:

Today’s economy offers us a Faustian bargain: it can give consumers deals largely because it hammers workers and communities.

We can blame big corporations, but we’re mostly making this bargain with ourselves. The easier it is for us to get great deals, the stronger the downward pressure on wages and benefits. Last year, the real wages of hourly workers, who make up about 80 percent of the work force, actually dropped for the first time in more than a decade; hourly workers’ health and pension benefits are in free fall. The easier it is for us to find better professional services, the harder professionals have to hustle to attract and keep clients. The more efficiently we can summon products from anywhere on the globe, the more stress we put on our own communities.

\textsuperscript{130} Ibid, p. 1567.

\textsuperscript{131} Ibid, p. 1573.

But you and I aren’t just consumers. We’re also workers and citizens. How do we strike the right balance? To claim that people shouldn’t have access to Wal-Mart or to cut-rate airfares or services from India or to Internet shopping, because these somehow reduce their quality of life, is paternalistic tripe. No one is a better judge of what people want than they themselves.

The problem is, the choices we make in the market don’t fully reflect our values as workers or as citizens. I didn’t want our community bookstore in Cambridge, Mass., to close (as it did last fall) yet I still bought lots of books from Amazon.com. In addition, we may not see the larger bargain when our own job or community isn’t directly at stake. I don’t like what’s happening to airline workers, but I still try for the cheapest fare I can get.

The only way for the workers or citizens in us to trump the consumers in us is through laws and regulations that make our purchases a social choice as well as a personal one. A requirement that companies with more than 50 employees offer their workers affordable health insurance, for example, might increase slightly the price of their goods and services. My inner consumer won’t like that very much, but the worker in me thinks it a fair price to pay. Same with an increase in the minimum wage or a change in labor laws making it easier for employees to organize and negotiate better terms.133

I cite so extensively in order to emphasize that the distinction between consumer preferences and citizen values is an influential one. This same basic view, articulated by each of these different authors, is certainly provocative. It is, I believe, also confused. The aforementioned arguments and distinctions do not help us clearly identify CBA’s shortcomings.

(a) The Explanatory Power of the Prisoners’ Dilemma

The cases mentioned by Sagoff and the others need explaining; but I believe that distinctions between consumers and citizens or between preferences and values will not be

133 Reich, “Don’t Blame Wal-Mart,” The New York Times, Feb. 28, 2005. (Reich was the 22nd United States Secretary of Labor and served under President Clinton. He is now Professor of Public Policy at the University of California at Berkeley.) I cannot resist pausing to draw attention to Reich’s claim that “No one is a better judge of what people want than they themselves.” The preceding chapter should make clear what a strange claim this is. (It is also grammatically strained, but nevermind that.) The more traditional claim – and one I spent considerable time analyzing – is “Each person is the best judge of her well-being.” This truth of this latter claim is disputable, but its relevance is not. On the other hand, the truth of Reich’s claim is not in dispute – it seems clearly true – but its relevance is. It is hardly clear why it matters that each person is the best judge of what she herself wants! Surely not even opponents of CBA would deny that claim.
part of the best explanation. The first thing to note is that many of these cases are plausibly understood as instances of a Prisoners’ Dilemma—or at least of some closely related collective-action problem. Seeing them in this light gives us a compelling alternative diagnosis of the case of Sagoff’s driving the oil-dripping car, of the students’ skiing, of Reich’s online book shopping, and many others. To see this more clearly, let us consider a new but related case. Imagine a person (call her P) who drives a gas-guzzling sport-utility vehicle (SUV) but who wants a clean environment and wishes that gasoline and SUVs would be heavily taxed. One diagnosis is simply that P contradicts herself. Another is the one the aforementioned authors would seemingly give, namely that P simply prefers one thing as an individual consumer but values something different in her role as a citizen; her consumer preference pulls apart from her citizen values or citizen preferences. But the Prisoners’ Dilemma, or something like it, provides us with an alternative explanation that is clearer and more intuitive.

Consider first an explanation put just in terms of P’s preferences (and that does not turn in any way on a distinction between consumer and citizen preferences). Now there is no denying that many people simply like the look and feel of driving an SUV; they are not particularly bothered that these produce more polluting emissions than compact cars. But not

---

134 For those who are not familiar, or who need a refresher, I will briefly summarize the Prisoners’ Dilemma. In its standard formulation, two individuals are arrested and accused of committing a serious crime together. As it turns out, the police have sufficient evidence to convict both prisoners on a minor charge, but insufficient evidence to convict either on the more serious charge. Lacking more substantial evidence, the police try to get each prisoner to testify against the other. If either one testifies against the other and that other remains silent, the silent accomplice receives a long sentence (10 years, say) and the who has testified goes free. If both prisoners stay silent, both will receive only short sentences (six months, say) for the minor charge. If both testify against each other, they each receive the same medium-length sentence (two years, say). The outcome for each depends on the choice she makes as well as the choice of the accomplice. However, neither prisoner knows the choice of her accomplice; and even if they were able to talk to each other, neither could be sure she could trust the other. Each prisoner reasons as follows: if my accomplice testifies against me, I do best for myself by testifying against her; and if my accomplice stays silent, I do best for myself by testifying against her. Each is thus compelled (by what is called “dominance reasoning”) to testify against her accomplice, and thus each receives the two-year sentence rather than each receiving the mere six-month sentence.
everyone is like this; \( P \), for one, is not. Still, it seems plain that she and others like her are ensnared in a collective-action problem that can push them into SUVs. \( P \) knows, let us imagine, that SUVs pollute more than compact cars do, and pollution is not something she prefers. But she also has a strong preference for keeping herself and her family safe. She knows that the presence of so many big vehicles on the road makes driving dangerous and she believes that being in a smaller, more fuel efficient car when so many others are in big vehicles puts her and her passengers in jeopardy. Since she cannot compel others to drive smaller vehicles, she reasons that being in an SUV keeps her safest, and safer by a wide enough margin to outweigh the higher gasoline costs she incurs in an SUV. The most preferred outcome for \( P \) may be one where she drives an SUV but everyone else is in a small car; here she is spending more on gas but is maximally safe on the road. (As will be explained shortly, relaxing this assumption does not modify the upshot.) A scenario in which she and all others are in small, fuel-efficient cars is clearly preferable for her to one in which she and all others drive big, gas-guzzling SUVs. But – importantly – the worst case is where she is in a small car and everyone else is in an SUV. Other drivers reason in exactly the same way as \( P \) does. Many of them therefore drive SUVs, and \( P \) does too. As a result, \( P \) and the others like her miss out on an available outcome (namely, everyone in compact cars) that they would have preferred to the one they get (namely, many SUVs).

The only difference between this case and the standard formulation of the Prisoners’ Dilemma is that the standard version involves only two people while \( P \)’s situation involves many, many people. We can even introduce an important change to the account and still end up with an outcome in which \( P \) and many others are in SUVs. Imagine now that because of \( P \)’s strong preference for a clean environment, the most preferred scenario for her is one in
which everyone ($P$ included) drives compact cars. In this case, there is not a Prisoners’ Dilemma, strictly speaking, but there still exists a collective-action problem that may well compel $P$ to drive an SUV. $P$ and others like her reason – not implausibly or irrationally – as follows: “There are many, many SUVs on the road. The fact is, my driving a more fuel-efficient car is not going to make any real difference to the state of the environment (the problem of global warming, etc.) unless I am joined by many others. The worst-case scenario for me is to be in a small car surrounded by all these SUVs. Alas, I cannot compel them to drive smaller cars. Thus, in spite of my preference for a clean and healthy environment, my overriding preference for keeping myself and my loved ones safe compels me, on the whole, to prefer driving an SUV rather than a compact car.”

The above telling of the story captures $P$’s predicament, and does so in a way that not only avoids invoking a distinction between consumer and citizen preferences, but also avoids explicit use of ‘value.’ If, instead of talking just in terms of preferences, one insisted on talking about $P$’s values – trying, as Ackerman and Heinzerling do, for example, to establish a distinction between (consumer) preferences and (citizen) values – I would say simply that $P$ values a healthy environment but also values keeping herself and her family safe. Given the collective-action problem plaguing $P$, it is precisely these values that lead her to support the taxes but actually drive the SUV. Any claim that she sees no value in driving the SUV or that what pushes her to drive such a vehicle has nothing whatsoever to do with her values (and instead only her preferences, consumer wants, or whatever) is disingenuous or confused.

This reasonable telling of the story does not depend on a distinction between preferences and values and does not turn on different senses of ‘value.’ We do not need to talk both about $P$’s preferences and her values in order for the story to be intelligible and plausible.
Moreover, nothing is gained here by talking only in terms of $P$’s values \textit{rather than} only in terms of her preferences.\footnote{This is doubly true when the preferences we are talking about are an individual’s well-informed, well-considered, and rational preferences.}

Now, it is true that the preferences the subject in each of the aforementioned cases reveals in the market are different from those she or he would – to borrow Anderson’s phrase – endorse “as normative for public purposes.” If one tries, for example, to gauge $P$’s preference regarding the taxation of gasoline and SUVs by observing her behavior in the market – e.g., what vehicle she actually drives, how much gasoline she actually buys, how hard she looks for the lowest gasoline price in town, etc. – one will likely infer that $P$ opposes the taxes; but this would be the wrong inference. Does this show that people are “schizophrenically” split into consumers and citizens? Does it show that there is something other than a preference on which public policies should be based? Does it suffice to show that there is a problem with using WTP as a proxy for her preferences? The answer to each question is no. Looking at $P$’s behavior in the market indicates what she – not just some \textit{part} of her, or just what she in some particular \textit{role} – prefers, i.e., what she values, but only within a certain context of constraints. Were these constraints different, $P$ would express different preferences. The lesson, then, is that analysts whose task it is to gauge individual preferences need to be careful to avoid doing so in contexts where the preferences are inappropriately constrained by collective-action problems.

This is something it \textit{is} possible to avoid, even when the analyst is capturing preferences in terms of WTP. I think most proponents of CBA are sensitive to this issue; in any case, the important point for my purposes is that such sensitivity does not require a fundamental abandonment of CBA. Recall that analysts are not limited to using the
revealed-preference method, i.e., to making inferences from people’s behavior in the market. They also have the stated-preference method available. This provides opportunities for analysts to formulate survey questions, including willingness-to-pay (WTP) questions, in ways that avoid the aforementioned kinds of collective-action problems. This point seems to have been missed by Ackerman and Heinzerling, who ask:

What could it mean [for a contingent valuation survey] to ask how much you are personally willing to pay to clean up a major oil spill? If no one else contributes, the clean-up will not happen regardless of your decision. As the Nobel Prize-winning economist Amartya Sen has pointed out, if your willingness to pay for a large-scale public initiative is independent of what others are paying, then you probably have not understood the nature of the problem. Instead, a collective decision about collective resources is required.\(^\text{136}\)

In fact, CBA need not stumble here. Rather than asking simply what an individual would be willing to pay (e.g., for the oil cleanup) without reference to what others are paying, each respondent could be asked what amount she would be willing to pay if she had an assurance that everyone else would pay that amount as well. This way of putting it is still too rough, as it is not attentive enough to the reasonable contention that those who have more should pay more, and other such distributional matters. Still, the basic point is clear: the WTP question could be about what one would be willing to pay in the form of a tax, where everyone else would have to pay the tax as well. This gives us precisely what Ackerman and Heinzerling say we need, namely “a collective decision about collective resources,” and we get it precisely by soliciting individual preferences.\(^\text{137}\)

\(^{(b)}\) Consumer Choices are Moral Choices

\(^{136}\) Ackerman and Heinzerling, ibid.

\(^{137}\) This should not be taken as an outright endorsement of WTP. It is meant just to suggest that WTP can get around the specific problem to which Ackerman and Heinzerling point.
I suspect that even when collective-action problems are addressed as recommended above, Sagoff and the others will persist in their belief that WTP reflects only one’s consumer preferences and not one’s preferences (views, judgments, values, etc.) as citizens. I am ready to admit both that there are cases in which an individual is not faced with a collective-action problem but still spends her money in ways for which she feels contempt, and that Sagoff is among those who have genuine contempt for some of their preferences. (If he says he does, he should be believed.) I will argue, however, that this admission does not force me to countenance the consumer/citizen distinction that Sagoff, Heinzerling and Ackerman, Anderson, and Reich are urging and does not suffice to show a problem with CBA’s use of WTP. That is to say, these authors’ confusion is not limited to a failure to recognize the substantial explanatory power of the Prisoners’ Dilemma.

What is particularly startling is the great comfort Sagoff seems to have with his schizophrenia as well as the readiness of the other authors cited to follow suit. Schizophrenia is, after all, typically taken to be pathological! Sagoff’s claim that his behavior – at least to the extent that it is not explained by the Prisoners’ Dilemma – can be explained and excused as resulting from a split that exists within him between a consumer and a citizen is unconvincing. In situations where he is not involved in collective-action problems the likes of which were discussed in the previous section – and, as I will explain, up to a point even in situations where he is – he should work to eliminate his supposed schizophrenia by overcoming those preferences for which he has contempt. He should work to spend money in ways that are consistent with the things he supports in his so-called role as a “citizen.” The same environmentalism that compels him to display that “Ecology Now” bumper sticker gives him reasons for fixing the oil-leaking car on which it is displayed. He should urge his
senator to outlaw one-way containers, but he should admit that this does not excuse him from also buying and returning returnable bottles.

Sagoff and the others make it sound as though our purchasing preferences, even when we express them in response to appropriately worded WTP survey questions, portray nothing about our policy-relevant values. But why accept this view? (To merely appeal here to a citizen/consumer distinction is to beg the question. And again, the fact that the latter involve collective-action problems has already been addressed and thus is, I am supposing, not germane here.) Recall that Sagoff himself has admitted to paying dues to the Sierra Club in support of the Alaskan wilderness despite the fact that he will not visit there. This seems, frankly, to be a tacit admission that how he spends his money has at least something to do with his citizen preferences. Would he really maintain that the fact that he (and any others) are willing to give money towards this cause is altogether irrelevant to public policy decisions regarding the cause?

The reason Sagoff ultimately offers for why policy questions should not be based upon individual purchasing preferences is unconvincing. He cites policy questions about segregation, abortion, the war in Vietnam, and capital punishment as ones that it would be inappropriate to settle, or even consider, through “a market or quasi-market approach.” He says:

What separates these questions from those for which markets are appropriate is this. They involve matters of knowledge, wisdom, morality, and taste that admit of better or worse, right or wrong, true or false—and these concepts differ from that of economic optimality. Surely environmental questions—the protection of wilderness, habitats, water, land, and air as well as policy toward environmental safety and health—involves moral and aesthetic principles and not just economic ones.138

138 Sagoff, ibid, p. 1296.
Sagoff is absolutely right that questions about all of these involve moral and aesthetic principles. But why does he refrain from saying – why does he actually seem committed to denying – that questions regarding how one spends one’s money will also regularly involve matters of knowledge, wisdom, morality, and taste that admit of better or worse, right or wrong, true or false? He should acknowledge that we regularly face what he terms “environmental questions” as consumers, and that our purchasing decisions – our decisions regarding how much to spend, where, on what, etc. – often involve moral principles, not just economic ones.

The question of whether to buy 10 pens for $2.00 or 20 otherwise identical pens for $3.00 is an example of one involving only economic principles. But the question of whether to repair one’s oil-leaking car, to return returnable bottles, to “fix” parking tickets, etc, are among those that also involve moral principles. So too, for example, are the questions of whether to buy clothes made by overseas sweatshop laborers, to shop at a massive retail chain that refuses to give many of its employees health insurance, to purchase meat coming from factory farms raising animals in cruel conditions, and even to drive that gas-guzzling SUV. In each of these cases, and many others like them, nothing absolutely prevents Sagoff (or anyone) from bringing moral considerations to bear in deciding how to spend money. This is not to deny that in each case there may be moral considerations pulling in different directions. Furthermore, the collective-action problems that arise in connection with these decisions may suffice to explain and perhaps even excuse certain decisions, such as the one to buy the SUV. It seems reasonable to say that there is no moral obligation to “go it alone” in cases where doing so is not going to make even a dent in some large problem and where it becomes undeniable that few others are going to join in. At the same time, however, it is
plausible to posit non-consequentialist reasons for action, such that one may, for example, have compelling reasons to avoid buying the SUV, the clothes made in a sweatshop, etc., even where most others will not avoid doing so. Refraining from spending one’s money on certain things and in certain places may have an “expressive function.” Relatedly, one’s spending decisions can make a morally important “symbolic protest.” The significance of non-consequentialist reasons is an issue on which there is considerable substantive debate, which I cannot address more fully here. Again, though, the issue is actually largely skirted when CBA uses the stated-preference method to gauge how much an individual would be willing to pay in the form of a tax paid by all.

The preceding discussion shows problems with identifying consumer preferences with either those preferences regarding how one spends one’s money or those preferences that guide one where no moral principles are in play. Perhaps the idea is simply that consumer preferences are self-interested preferences and citizen preferences are selfless or other-regarding preferences. The distinction between self-interested and selfless preferences is obviously reasonable enough; but proponents of CBA will rightly ask how this distinction tells against CBA. At least offhand, it does not seem to be the case that CBA can only take account of self-interested preferences, nor is it even necessarily the case that it tends only to do so. Nothing would seem to require that self-interested preferences rather than other-regarding preferences dictate one’s actual spending behavior or one’s stated WTP. In situations involving actual market behavior and no collective-action problems, or collective-

---

139 For a classic discussion of the idea of an “expressive function,” see Joel Feinberg, “The Expressive Function of Punishment,” *The Monist* 49 (1965), pp. 397-423. It is an appeal to an expressive function that does a great deal to explain why a person should vote even in light of the collective-action problem plaguing voting that will be noted in the next section of this chapter.

action problems but appropriately worded WTP survey questions, someone whose WTP regularly fails to reflect her selfless preferences will sound awfully disingenuous when she claims that her selfless preferences must be taken into account and that CBA is flawed for failing to do so. Whatever specific language he uses to express them, Sagoff’s complaints against CBA sound just as disingenuous.

(c) Double-counting Objections:

At this point, the critics may note that the distinction between self-interested and other-regarding preferences looks a lot like Ronald Dworkin’s important distinction between personal preferences and external preferences, and that the correspondence between these distinctions ultimately means, contrary to what I have argued, that CBA can in fact only take account of self-regarding preferences. Dworkin’s distinction arises in a discussion of preference utilitarianism, but is relevant to CBA insofar as the latter, like the former, “asks officials to attempt to satisfy people’s preferences so far as this is possible.” According to Dworkin:

The preferences of an individual for the consequences of a particular policy may be seen to reflect, on further analysis, either a personal preference for his own enjoyment of some goods or opportunities, or an external preference for the assignment of goods and opportunities to others, or both. A white law school candidate might have a personal preference for the consequences of segregation, for example, because the policy improves his own chances of success, or an external preference for those consequences because he has contempt for blacks and disapproves of social situations in which the races mix.142

It should be noted that external preferences could just as well be based in affection for others (as opposed to contempt, as is the case in Dworkin’s example). And it should be reasonably clear that personal preferences seem to be essentially what I have called self-interested

142 Ibid.
preferences, while the external preferences seem quite similar to what I have called other-regarding preferences.

Dworkin then introduces (at least one version of) the “double-counting” objection. He has us imagine that a community could build a swimming pool or a theater, but not both. Furthermore, “Many citizens, who themselves do not swim, prefer the pool to the theater because they approve of sports and admire athletes, or because they think that the theater is immoral and ought to be repressed.”¹⁴³ The effect will be that “each swimmer will have the benefit not only of his own preference, but also of the preference of someone else who takes pleasure in his success,” and “actors and audiences will suffer because their preferences are held in lower respect by citizens whose personal preferences are not themselves engaged.”¹⁴⁴ Counting other-regarding preferences, the argument (in effect) goes, is inconsistent with treating all people as equals, and this is particularly problematic, according to Dworkin, because utilitarianism (and CBA too, he would presumably say) “owes much of its popularity to the assumption that it embodies the right of citizens to be treated as equals.”¹⁴⁵ This kind of worry is surely partly what underpins a view that creeps up in the economics literature that, in order for CBA to work properly, the individual preferences it takes account of should “reflect only [each individual’s] own personal economic motives and not altruistic motives, or sense of duty, or moral obligation.”¹⁴⁶

¹⁴³ Ibid.
¹⁴⁴ Ibid.
¹⁴⁵ Ibid.
There are some complicated issues involved here, but in the end this objection seems to fail for exactly the reasons that H. L. A. Hart gives. Hart does not deny that there is a distinction between personal and external preferences; he does not deny that double counting can occur, or that it is bad where it does. A political system that permits each woman two votes and each man only one vote would be guilty of double counting, and this would be wrong. Hart’s response, though, is that the cases Dworkin describes and calls problematic are not ones where anyone’s preferences are counted twice. Regarding the swimming pool example, Hart rightly notes:

It is only the case that the proposal for the allocation of some good to the swimmer is supported by the preferences both of the swimmer and (say) his disinterested non-swimmer neighbour. Each of the two preferences is counted only as one; and surely not to count the neighbour’s disinterested preference on this issue would be to fail to treat the two as equals. It would be “undercounting” and presumably as bad as double counting.147

A full treatment of this issue would require more space than I can allocate here, but numerous economists effectively join Hart in rejecting the exclusion of preferences that are not purely self-regarding. As Michael Hanemann says, “When estimating demand functions for fish prior to Vatican II, no economist ever proposed removing Catholics because they were eating fish out of a sense of duty. Nor, when estimating collective choice models, do we exclude childless couples who vote for school bonds because they lack a personal economic motive.”148


148 Michael Hanemann, “Valuing the Environment Through Contingent Valuation,” The Journal of Economic Perspectives 8 (1994), p. 33. For support of his (and my) view, Hanemann quotes Kenneth Arrow as saying, “It need not be assumed… that an individual’s attitude toward different social states is determined exclusively by the commodity bundles which accrue to his lot under each. The individual may order all social states by whatever standards he deems relevant.” Similarly, Gary Becker is quoted as saying that “individuals maximize welfare as they conceive it, whether they be selfish, altruistic, loyal, spiteful, or masochistic.”
Section 3: A Comparison of Voting and Cost-Benefit Analysis

I have sought thus far to challenge the employment of the citizen/consumer distinction in criticisms of CBA. However “consumer preference” is in the end understood, it is not the case that CBA improperly registers these alone. Behavior that appears to evince a schizophrenic split within individuals can frequently be understood and explained away simply as reasonable behavior by individuals trapped inside a collective-action problem. In cases where this is not true – and again, up to a point even in cases where it is – we must insist that moral principles (and not just economic ones) will often make claims on how one spends one’s money. Individuals are not excused from expressing a WTP that demonstrates their so-called citizen preferences or citizen values.

I want now to undertake a careful analysis of the similarities and differences between CBA and voting as mechanisms for public decisions. Many people are not nearly as critical of voting as of CBA. Idea voting is universally esteemed, while CBA is frequently vilified. The question is: Why might one who does not believe that it ought to be the aim of public policy to promote the satisfaction of individual preferences believe that it is important that citizens be able to vote? What principle could support the importance of voting and of taking people’s votes seriously, but at the same time oppose the identification of policy-level benefits with the satisfaction of individual preferences? Undertaking the comparison will help check (as it were) my assessment from the previous section of the merits (or more aptly, the demerits) of the arguments by critics of CBA that were discussed there. Sagoff, we know, favors voting over CBA on the grounds that he believes voting but not CBA to be a decision mechanism which registers citizen preferences (i.e., citizen values). My view is that
the fact that the consumer/citizen distinction is so muddled is closely related to the fact that – as will be argued – his argument for the superiority of voting over CBA as a public policy decision mechanism fails. More generally, exploring the similarities and differences between voting and CBA will help us better identify the true strengths and weaknesses of CBA. (The discussion in this section of the chapter stays focused on the general issue of whether CBA is to be criticized for the role it gives to individual preferences. It is in virtue of the fact that at least certain of the criticisms of CBA discussed in the previous chapter contrasted CBA with voting that I take the opportunity here to explore that issue in detail, even though the discussion in this section will eventually go beyond the specific criticisms of CBA discussed in the previous section.) I believe that CBA and voting are more alike than is commonly recognized, and that at least some of the differences are ones that favor CBA. There is, I will explain, one difference that strongly favors voting over CBA; but I will argue that this is not indicative of a problem with CBA’s commitment to identifying benefits with preference satisfaction (at least the satisfaction of informed and rational preferences), but rather with CBA’s insistence that all preferences, and in turn all costs and benefits, be expressed monetarily.

(a) Voting Can Be Self-Regarding Too

Let us begin by looking at some purported differences between CBA and voting that are, I believe, not in fact real differences. The first of these comes from Sagoff and has already been discussed: voting is, while CBA is not, suited to dealing with public decisions that have a moral dimension. This is mistaken partly because – as I have argued – while CBA is interested in each individual’s WTP, how one spends one’s money is often a moral issue. Sagoff’s charge is connected to his belief – as I understand him – that we spend
money only with our own narrow self-interests in mind while we vote with the well-being of the community in mind. I have claimed that one need not, and in fact should not, spend one’s money in all selfish ways. Relatedly, though, it must be noted that people who vote need not, and perhaps should not, always do so in entirely _selfless_ ways. Often times a voter is not motivated by the well-being of her community, but rather by what is best for her individually (perhaps including her family and other loved ones). It is not necessarily inappropriate to vote in one way and not another based partly on private financial considerations. Suppose a referendum or candidate promises to institute a program the funding of which would require $P$ to pay more in taxes but will not benefit her or her family, even indirectly. This gives $P$ a reason – though not obviously an overriding one – to vote against that referendum or candidate. If one’s vote may express selfless preferences, then one’s WTP can too; and if CBA is guilty of double-counting when it permits individuals to register selfless preferences, then voting is guilty of the same when it does so.

(b) _A Collective-Action Problem Plaguing Voting_

It has been noted that individuals in the market are subject to a collective-action problem that presents problems for CBA’s use of the revealed-preference method. One reason why those who give special praise to voting as a decision mechanism should tread cautiously in criticizing CBA is that voting is subject to a serious collective-action problem of its own. Its collective-action problem looks like this: $P$ does not know how many people will vote, but reasons that if many people vote, _her_ vote will not make a difference to the outcome of the election. After all, how many referendums or elections are decided by one vote? This diminishes $P$’s incentive to vote. But many other people reason like $P$, and the number of people who actually do vote is much smaller than is ideal. I have indicated how
CBA can get around the collective-action problem it faces; frankly, the one that plagues voting is considerably more difficult to resolve.

(c) The Issue of Standing

One comparison between voting and CBA was briefly made in the last chapter’s discussion of transformative value (Section 3). The basic issue was that sometimes we do not develop preferences until we have had particular experiences, and one example used was of a teenager who prefers pop music to classical music until she actually has the experience of listening to classical music (at a classical concert, say). This experience transforms her preferences. Among my claims was that CBA need not have any greater problem than voting does in taking account of transformative value and the transformation of preferences.

In that same section, the issue of existence value was also considered. This too is an issue that ultimately connects up with the question of the proper role for individual preferences. It was observed that our preferences can extend to things that we do not use or even interact with, and this included things that are geographically very far from us. These are things that do not benefit us in any direct or material sense but with which our well-being is clearly wrapped up. ANWR was the main example employed in this connection. At least some people who do not plan to visit ANWR, and who do not derive (and do not think they derive) any “material” benefits from ANWR, including even very indirect ones, have preferences regarding what goes on there. Some of these people would, for example, count themselves as worse off if ANWR were opened for commercial development, including drilling for oil. Voting arose as a topic here too, as it was observed that the problem of existence value, or of so-called non-use preferences, that confronts CBA also confronts voting. I claimed that the apparent problem for CBA in dealing with existence value is
resolved through addressing the problem of standing, and that CBA is no different from voting in needing to address this problem. For both decision mechanisms, this is an important issue and not always an easy one to resolve. There are regularly problems that transcend political boundaries, and environmental problems often fit this mold. Farmers irrigating from a river in one state reduces the water available for farmers in another state, or changes the ecology of a delta region in another state. Pollution from factories in one state carries and creates acid rain in other states. Sewage dumped into a creek by residents in one rural county poisons the water that residents of the adjacent downstream county drink. The heavy burning of fossil fuels by one nation changes the climate of the entire planet, which affects every other country, and some potentially very significantly.

When situations like these arise, it is not always obvious that the members of all the affected counties, or states, or countries, or whatever, should be excluded from having a say in what goes on in the offending county/state/country. This is a complicated issue, and not one I can hope to resolve here. But the important point for the purposes of this project is simply that the institution of voting is, as a matter of fact, generally not brutally impugned when it is maintained (as it regularly is) that at least sometimes there will be people who want “a say” in particular policy issues but who are denied a vote because they lack standing. Sometimes a person who is affected by a policy decision is not an eligible voter in the jurisdiction that makes the decision. CBA faces the same need to limit who has a say, that is, whose preferences count, and should not for this reason alone be impugned any more severely than voting is. It is possible, of course, that cost-benefit analyses regularly draw the boundary around those who have standing too tightly, and more tightly than it is drawn in voting. If so, then it is rightly criticized. But I do not think this is a charge of which CBA is
always (or even regularly) guilty, and indeed it is not the charge that is typically leveled. (It should be remarked that the courts, a third mechanism for public decisions, also face the issue of standing, and like both voting and CBA, will often exclude people who have clear preferences regarding the particular issue/decision at hand.)

There is another important aspect of the discussion of standing. Some claim that CBA fails adequately to represent the preferences that people in the future will have for the goods regarding which we are currently making policy decisions. The worry is that making decisions through CBA is inconsistent with living up to our obligations to future generations. I am not convinced that CBA is guilty of this charge, but if it is, it is certainly no more – and is perhaps less – guilty than voting. Critics of CBA complain, particularly when CBA is applied in “big ticket” environmental policy connections, that our policy decisions have long-lasting consequences – perhaps permanent would not be entirely too strong. These consequences will affect those who live in the future. CBA goes wrong, these critics say, by not taking into account, by not including in the calculus, the preferences that people who will live in the future will have regarding these effects and ipso facto for the projects we are currently considering.

In fact, it is a mistake to say that CBA does not take into account the preferences of those who will live in the future. Firstly, it tries to register the future-oriented preferences of those who live now (and are within the relevant political jurisdiction). That is to say, nothing prevents those who live now from articulating particular preferences regarding what will transpire well into the future. Nothing prevents a respondent in a current CV from articulating a WTP to ensure that those in the future do not suffer unduly. But this does not quite amount to giving standing to those who live in the future; and indeed it is not – or at
least need not be – all that CBA does to give such people standing. In addition, it often also tries to anticipate and register the preferences that those who will be alive in the future will have. If a project is going to cause illnesses to people who will live in the vicinity of the project 100 years from now, a CBA conducted today can of course measure the WTP of people who are alive now to prevent those who will be alive 100 later from getting sick. But it can also try to gauge what people in the future would themselves be willing to pay to avoid getting sick. This counts as genuinely giving standing to future generations. Again, I am sure that not every CBA does this, but some do and any could, if it were deemed appropriate to do so.

Of course, proponents of CBA roundly recommend discounting the goodness of future preference satisfaction (and the badness of future preference frustration). It is a fact that if people are asked whether they would rather be given $100 today or $100 one year from now, they will say they would rather have it today. This is for good reason. If one takes the $100 today and invests it in an account that gets even just very modest interest, in one year she will have more than $100. And in virtue of inflation, $100 one year from now will actually be, in real terms, less than $100. The inference that most people want to draw from this and many similar examples is that the goodness of a future benefit, i.e., of a benefit received in the future, has to be discounted. From our perspective, a future benefit is not as good as one received now. And similarly the badness of a future cost has to be discounted; from our perspective, a future cost is not as bad as one incurred now. Nearly every CBA will involve discounting of future costs and benefits. The rate at which they are discounted is determined by people’s behavior. The analyst looks at, e.g., how much more of a benefit people require in order to take that benefit a certain amount of time into the future. The
present value, i.e., the value to us now, of a future cost or benefit is expressed by the following formula (where \( d \) stands for the discount rate, and \( t \) for the number of years into the future that the benefit/cost obtains):

\[
\text{Present value} = \frac{\text{Value in year } t}{(1 + d)^t}
\]

Some charge that the mere act of discounting is incompatible with respecting (or fulfilling obligations to) future people, or future generations. Imagine that a project delivers great benefits today, but the costs that it generates are not incurred for several generations. Nuclear power is a good example: we benefit from the power right away, but the dangerous waste lingers for a long, long time, and in fact becomes more dangerous the longer it is stored (as the chance for problems with the storage containers increases with time). Consider CBA’s appraisal of this scenario. Imagine that 1,000 years from now people are exposed to the stored nuclear waste (the stuff that we are generating today). How bad is this from our perspective? How seriously should we take this when we, today, make decisions about producing/storing nuclear waste? Well, the numerator of the equation above will be a very large negative number, since the waste 1,000 years from now is very harmful. But at whatever discount rate we employ, the denominator will be much larger; it is, after all, raised to the 1,000\(^{th}\) power (because we are looking at a scenario occurring 1,000 years from now). So the present value will be extremely small, which is in effect to say that we should count this future disaster as trivial. The present value of the damages will be outweighed by the present value of the benefits (which in virtue of being realized now or in the very near future are not discounted much). Thus, says Freeman, “the implication of discounting is that we
care virtually nothing about the damages that we inflict on future generations provided that
they are postponed sufficiently far into the future.”

This concern is only partly well-founded. After all, the formula for discounting is, if
turned around, the formula for figuring interest. If we take a the “trivial” present value of the
future costs of nuclear waste and multiply it by \((1 + d)^t\), i.e., invest it at rate \(d\) for \(t\) number of
years, then the money will equal that value of the cost in year \(t\). So, perhaps the real concern
is that it is extremely unlikely that this “trivial” amount will actually be set aside to generate
interest and used to compensate those who incur the costs in year \(t\). This is a complicated
issue, and not one I am capable of or concerned with resolving here.

My principal interest is to point out that voting encounters some of the same
problems. Voting does not involve discounting in this technical sense. But voting certainly
fails to directly register the preferences of people who will live in the future. Voting, like
CBA, allows people who live now to express preferences regarding those who live in the
future (including that future people not incur costs in order that those of us alive now might
be benefited). They can do this through a vote: current people can vote for what they think
future people will prefer. But in voting this is the extent of the consideration of preferences
of future people. It is not as if extra ballots are dropped into the box on behalf of those
people who will live in the future and be affected by the current decision. Thus, voting
plainly does not take future people/generations to have standing on current policy questions.
This is in contrast with CBA. As I have said, CBAs will at least sometimes try to anticipate
the preferences of those who will live in the future – not just the preferences that those who
live now have for those who will live in the future – and to build these into the calculus.

Freeman, ibid, p. 299. Freeman is not endorsing, but rather merely giving voice to, a charge that is leveled
against CBA.
(d) Public Debate and Informed, Considered Preferences

Thus far, I have endeavored to reject certain arguments given by proponents of CBA that purport to show that voting is superior to CBA. I have argued that in these respects in which CBA is criticized, CBA and voting are not in fact so different. However, I have not intended to insist that there are no differences between CBA and voting. There are, and they are no doubt sometimes differences that favor voting. In at least one such case, though, the difference is in virtue of the way that CBA is traditionally conceived of and practiced. Perhaps, however, CBA could be modified such that it still hangs on to its principles (i.e., the commitments in virtue of which it counts as CBA rather than as something else) but becomes more like voting in the respect that now favors voting.

What I have in mind specifically is the idea of public debate (mention of which was again made in the previous chapter). There are differences in connection with this issue between voting and CBA. As the previous three chapters were given to arguing, anyone who wants public policy decisions to be based on the preferences of the individuals affected should want those preferences to be as well informed and well considered as possible. Voting may do a better job than CBA typically does of registering informed and considered preferences. Voting seasons are constructed in order to create an electorate that is at least decently educated about the issues. Prior to voting, citizens typically have ample opportunity to gather information about issues and candidates. Newspaper articles and editorials, televised debates between candidates or at least between partisan pundits, commercials, door-to-door campaigns, public rallies, and so on, all provide this opportunity. It is not just that a set of facts is put out there, but rather that the perceived merits and demerits of each position and candidate are publicly discussed and debated at length, from a wide variety of
perspectives, and in a range of forums. Moreover, each individual is encouraged to reflect on the preferences she is preparing to express via her particular vote. One cannot just rush out and vote the instant she has a strong sense of how she wants to vote (unless, of course, election day itself is the first time she has this strong sense); instead, one has to continue reflecting until the day of voting rolls around.\textsuperscript{150}

This is all worth noting, but the point should not be overstated. After all, no one is categorically \textit{required} to be well-informed before she votes. It is not as though officials check in with each voter before she votes to ensure she is filled in on the issues and the candidates, ready to supply information in the event she is not. Ill-informed people have the same right to vote that well-informed people do, and it is undoubtedly the case that plenty of people go to the poll without being well-informed. Moreover, in practice of late, it sometimes seems as though the purpose of campaigns has been to spread \textit{mis}information, and that is disheartening. Of course when this happens those on the other side of this issue have the chance to counter with the truth, but alas they typically promulgate some misinformation of their own.

Now, how does this compare to CBA? If the analyst is using the revealed-preference method, there is no assurance that the market behavior from which the analyst attempts to infer individual preferences actually reflects anything other than people’s rash spending decisions. Not everyone is a rash spender, but that the concept of rash spending exists while the concept of rash voting does not is telling. Then again, if a person is going to spend a considerable sum of money, she will often take time to research her decision. On many occasions, though certainly not all, people have time leading up to a purchase in which they

\begin{footnote}{150 This is changing somewhat as more and more places make early voting available. Still, this is typically at most a few weeks earlier than it would otherwise have been that one can go and vote.}
can educate themselves; and thus when they make a purchase, they register a well-informed preference that squares with their other preferences. This is akin to the time and opportunity one has to learn about the issues and candidates leading up to a vote. Even where a spender is being deliberate, however, one difference is that typically when one makes a purchase – even a big one – the pros and cons of the issue, viz., the good- and bad-making features of the thing one is thinking of purchasing, whether one really needs it, what other uses one could put the money to, etc., are not discussed, or at least not in the same way, to the same extent, or with the same vigor, as are candidates or issues leading up to a vote. This is a significant difference. To approach a remedy, the state can work to ensure that people have easy access to relevant information before they make their purchasing decisions. How successful such an effort could be is unclear. This all serves to underscore the point made in the earlier discussion of collective-action problems, namely that cost-benefit analysts should be very careful about using the revealed-preference method.

In fact, these limitations exist not just when the revealed-preference method is used, but also when the stated-preference method is employed. Surveys – even well-constructed ones – have the potential to register snap decisions. A question that is put to a respondent may be one she has never considered before, and she will typically not have long to consider it before an answer is expected (or at least perceived to be expected). Preferences articulated in such a context are apt to be uninformed and/or inconsistent with one’s other preferences. Surveys can be much more liable to this than voting is. Moreover, as surveys tend to be employed, there will be no opportunity for public debate before each respondent’s preference is solicited. CBA is used precisely as a way to get policy recommendations without having to have “messy,” drawn-out public debates. The tendency here is understandable, but the
effect is to make contingent valuation less similar to voting, and indeed more liable to reasonable criticism. Public debate is important not just before people register their preferences in a traditional vote, but also before they register their preferences in response to contingent valuation surveys.

Fortunately for proponents of CBA, contingent valuation is not altogether incompatible with public debate. Issues regarding which people’s preferences are sought could be subject to public debate in any or all of its manifestations, and then if a vote is too time consuming or energy intensive, the stated-preference method could be used to capture people’s preferences. Moreover, the provision of information by the analyst herself (in order to help ensure that informed preferences are registered) is not incompatible with the stated-preference method. If contingent valuation surveys are constructed, and those administering them are trained, so that policy-relevant information is shared with a respondent who is given time for and assistance in organizing and processing this information, then CBA does not lag so far behind voting with respect to registering appropriately informed preferences. If proponents of CBA do not advocate this, then critics of CBA are right to disparage it while praising voting as a public decision mechanism.

(e) Registering the Direction and Strength of a Preference

The final differences between CBA and voting to which I want to draw attention are connected to CBA’s use of WTP as the measure of how strong individual preferences are. Voting typically utilizes a dichotomous nominal variable; one votes either “yea” or “nay.” WTP, on the other hand, is a numerical measure that is more or less continuous. While voting aims to capture only the direction of a preference (i.e., for or against), CBA aims to capture both the direction and the strength of a preference. This difference more than
occasionally makes a difference. The archetypal democracy is one in which each citizen gets one and only one vote. Each citizen’s vote counts equal to that of each other vote, regardless of the strength of the preference underlying the vote. For example, regardless of whether a person really, really wants the school levy on the ballot to pass, or only mildly wants it to pass, this person gets only one vote. So far as I know, no one recommends letting those who care more about an issue vote multiple times, counting their vote as more important, or anything of this nature. Why then should CBA work to measure not only the direction but also the strength of preferences?

In fact, it is not entirely unreasonable to see it as a shortcoming of voting (at least as it is commonly practiced) that it registers only the direction of a preference. Voting is in this respect a rather blunt tool. If a group of friends were going to vote on where to go for lunch, it would certainly seem odd to ignore how strongly each feels about the place she favors. I would not want to push too hard here, for I recognize that allowing policy decisions to be made on the basis not just of preferences’ direction but also their strength has the effect of putting the outcome of policy decisions too far into the hands of zealots, and thus even encouraging zealotry. But the basic point stands: it is not necessarily a knock on CBA that it aims to measure the strength of preferences. It is surely not always inappropriate to do so.

(f) The Monetization of Preferences

Even if it is appropriate to measure both the direction and the strength of a preference, there is obvious concern about a policy tool that gives more say in public decisions to those people with more money, which is precisely what CBA’s use of WTP does. The fact that CBA does, while voting does not, require that preferences be monetized means that the preferences one is able to register via a vote are less encumbered by one’s financial situation
than are the preferences one registers via CBA. Can anything be said in defense of CBA on this issue? First, it can be noted that while the preferences one is able to register via a vote are less encumbered by one’s financial situation than those one registers via one’s WTP, it would be naïve to say that the former are altogether independent of one’s financial situation. I do not have in mind here anything quite as insidious as the buying of votes, but clearly those with more money are (up to a point) legally able to make influence-garnering campaign donations. Additionally, there is again the fact that WTP is used so that policy decisions can be made on the model of the market in cases where market failure has led to an outcome that differs from the one a perfect market would have delivered. Those who favor voting over CBA on the grounds that the latter gives undue influence to the wealthy will thus need to explain why decisions like those specified at the very outset of this paper should not be made on the market model. The aforementioned arguments distinguishing consumers from citizens were seemingly attempting to give precisely this kind of explanation; but I hope to have shown that they fail.

Ultimately, this defense can be pushed only so far, since reasonable criticisms of CBA’s use of WTP in the making of public decisions exist. I will not rehearse these here, since they are given full treatment in Chapters 8 and 9. (Note, in fact, that even if the more novel objection to WTP that I press in Chapter 9 is unconvincing, those that are rehearsed in Chapter 8 are, I believe, quite solid, and also sufficient for establish that voting is superior to CBA in virtue of not requiring that individual preferences (and in turn costs and benefits) be expressed monetarily.) Importantly, none of these criticisms of WTP – and indeed none of the general differences I have acknowledged between CBA and voting – shows that the distinction between consumer preferences and citizen values is a useful one. None of the
criticisms of WTP shows that a person’s WTP has nothing to do with her policy-relevant preferences. Finally, they do not show that there is something wrong with connecting policy-level benefits up with the satisfaction of individual preferences. Liberal democracies ensure the right to vote in part because they take so seriously the importance of registering individuals’ preferences, which is the very same thing that CBA aims to do. What the criticisms do show is that WTP does not tell us everything there is to know about a person’s preferences. CBA is deeply flawed for its use of WTP as a measure of or proxy for people’s preferences.

**Section 4: Individual Preferences Versus Social Values**

One final distinction will be considered in this chapter. In certain places, Sagoff gets away from the notion of citizen preferences and even of citizen values and instead favors the concept of “social values,” which he contrasts with private preferences. “Social value” sounds like merely another way of saying “citizen preference” or “citizen value,” but I believe that Sagoff actually has something different in mind for that first notion. I argue in this section that Sagoff’s argument that CBA missteps by aiming to satisfy individuals’ preferences rather than to promote (respect, honor, etc.) what he calls “social values” is ultimately unsuccessful.

According to Sagoff, “policy analysis that is rooted in welfare economics makes little room for social values.” He says that “policy analysis, at least according to standard accounts, seeks to satisfy private preferences, whatever they are, whether or not they are...

---

consistent with the principled social values or commitments of society.”

What are social values? They are, he says, “those values ascribed directly to society as a whole, whether or not a majority of Americans at any moment agrees with them.” But there is ample cause for wondering what it means for a value to be ascribed to a “society as a whole” when the majority of that society’s members do not embrace the value. Sagoff tries to clarify by saying that “social values, which are the moral goals to which we commit ourselves as a community, are by no means derived mathematically or mechanically from private desires, wants, prejudices, fears, and so on. Rather, they are determined through a complex and deliberative cultural, legal, and political process.” He maintains that when the Supreme Court found racial segregation of public facilities to be unconstitutional, the goal of racial integration “officially” became a social value, and did so despite the fact that such integration may not have appealed to the majority of citizens. Another “clear example of a social value,” he writes, “would be the separation of church and state, that is, the principle that Congress shall make no law to establish a religion. Cost-benefit analysis—or the comparison of relative willingness to pay for different outcomes—is not even remotely relevant to the principle of church-state separation.”

Finally, he says that:

Occasionally, public opinion surveys find that by large majorities, Americans would disapprove of one or another of the Bill of Rights—for example, the Fifth Amendment guarantee against self-incrimination. A determination of the amounts that individuals are willing to pay, for example, to enforce religious uniformity or limit freedom of the press hardly seems relevant to the legitimacy of our legal rights, democratic values, and social commitments—freedom of conscience, equality before the law, habeas corpus, trial by jury, universal suffrage, and so on.”

---

152 Ibid.
153 Ibid.
154 Ibid.
155 Ibid, p. 93.
Sagoff is right that WTP hardly seems relevant in these cases. The reason is precisely that these are issues that are to be settled, ultimately, on constitutional grounds. Martha Nussbaum writes that “economists and others who defend preference-based views rarely make a clear distinction between their uses in social choice generally and their use in selecting basic principles that can be embodied in constitutional guarantees. Instead, they tend to make general pronouncements about social choice.” She is right that there is a distinction, and that it is too infrequently noted. It turns out that it is not just defenders of preference-based views, but also opponents of such views, who fail to make it clear. Surely most of those who endorse CBA deny neither that the Constitution lays out certain basic restrictions on what is and is not legal, nor that legislative acts and even particular public policy decisions cannot be upheld if they are not consistent with these restrictions. Proponents of CBA would, for example, certainly not claim that some particular accused criminal should be denied the ability to refrain from giving self-incriminating testimony in any case where the overall WTP for her to be required to incriminate herself exceeds the overall WTP for her not to be so required. The mere fact that, in some particular case, preferences for ignoring or overriding some constitutional guarantee exceed those for respecting and observing this guarantee does not make ignoring or overriding the guarantee justifiable.

While the role that individual preferences play in public policy is limited in the aforementioned respect, preferences can nonetheless play an important role even in the cases

---

156 Ibid.

that Sagoff cites. If, among a sufficiently large majority of the voting population, there develops a strong preference for overturning (say) the constitutionally-granted right to abstention from self-incriminating testimony via an additional amendment to the Constitution, then there is a powerful reason for so amending it. The justification is simply that this is preferred, and so strongly. The justification would not be that such an amendment has become a social value, unless something becomes a social value just when it is strongly and widely preferred (which is clearly not Sagoff’s view of social values). The same thing can be said about the separation of church and state: if there are enough people with an individual preference for a connection between church and state, and these people express these preferences by voting for candidates who favor a connection between church and state, and a sufficient number of such candidates, when elected, vote in favor of amending the Constitution to reflect such preferences, and the amendment is ratified by a vote in the requisite number of states, then the Constitution will be amended. So, while a determination of the amounts that individuals are willing to pay for a connection between church and state hardly seems relevant, individual preferences regarding the separation or connection clearly are relevant.

Individual preferences may even influence how the Constitution is interpreted; and I do not mean by this just the preferences of individual Justices. The Supreme Court case *Trop v. Dulles*, for example, established that the Eighth Amendment standard of what counts as “cruel and unusual punishment” is not fixed but must reflect current judgment among the populous regarding what counts as cruel, and what as fitting, punishment. This hermeneutic principle was upheld in the more recent case of *Roper v. Simmons*, which invalidated the execution of juveniles who were 16 or 17 at the time they committed their crimes. In this

228
more recent case, one significant point of contention was how to measure the current judgment of the populous. I believe it is not is misnomer to say that what is being measured is people’s preferences regarding what is and what is not to count as cruel and unusual punishment.

Sagoff says that “policy analysis—insofar as it pursues preference satisfaction, utility, and economic efficiency—appears to have little to contribute to the abortion question or to the discussion of any other principled social value or decision.” Abortion is indeed not an issue properly addressed using CBA; and there are plenty of other issues of which this is true. Again, these are issues that are to be decided by interpretation of the law and ultimately the Constitutional (though individual preferences can contribute to these issues in the ways I have explained in this section). Importantly, however, not every issue is such an issue. Contrary to what Sagoff asserts, there are plenty of other decisions – “principled” ones, even – regarding which policy analysis (by which I mean analysis that aims to record individual preferences by some means other than a vote) can make contributions. The five questions with which this paper opened all seem to point to decisions of this kind. Now, among the difficulties involved in interpreting the Constitution can be simply ascertaining what particular issues the Constitution even addresses. If it became clear, for example, that how much arsenic it is permissible to leave in drinking water is an issue addressed (even implicitly) in the Constitution, and it could further be determined that X parts-per-billion is what is constitutionally required, then it would not be appropriate to initiate a CBA in order to measure people’s preferences and see whether the state is in fact justified in providing water with (say) X+25 parts-per-billion of arsenic. But if these are not decisions to be addressed on constitutional grounds, then proponents of CBA are right to urge that they

158 Sagoff, ibid, p. 95.
should be decided by people’s preferences, at least their well-informed and sufficiently-deliberated-over ones. In these cases, WTP is not wholly irrelevant, for (as was argued in Sections III and IV) it surely can, if solicited properly, tell us something about these preferences; though because (as I suggested in Section V) WTP is not a good measure of people’s preferences, it is certainly not all that is relevant. This constitutes one of what are surely many legitimate causes for concern that the frequent use of CBA in public decisions regarding investments and regulations presents.

Sagoff’s next move in the paper from which I have been quoting is to introduce the work of James G. March, who “distinguishes between economic (preference based) and ethical (principle based) decision making. When decision makers adopt the first or economic approach to collective choice—the model associated with formal policy analysis—they choose among alternatives by evaluating their consequences in relation to the prior preferences of individuals,” says Sagoff paraphrasing March. According to Sagoff:

This approach seeks to maximize the well-being of society as a whole by satisfying the preferences of its individual members, based on the assumption that individuals always prefer what they believe will benefit them. When decision makers take a principled approach, March writes, they “pursue a logic of appropriateness, fulfilling identities or roles by recognizing situations they encounter.” Individuals typically do not ask, What outcome benefits me most as an individual? but, What do I believe is appropriate for society as a whole, given our shared principles, beliefs, and commitments?159

At this point it starts to look like the distinction between social values and preferences bears a great deal in common with a distinction I have already addressed and with which I have exposed problems, namely the distinction between consumer and citizen preferences. To the extent that Sagoff’s argument here appeals to this distinction, or aims to make essentially the same distinction using other words, it will not be successful. Just when it

159 Ibid, p. 94.
seems as though this is the direction Sagoff is going, he throws a curveball. Rather than straightforwardly connecting up the distinction between social values and preferences to the one between citizen and consumer preferences, Sagoff moves to the distinction “between utilitarian and rule-based or deontological conceptions of collective choice.” This is a strange move, in no small part because one does not need to have introduced March in order to discuss the distinction between utilitarian and deontological approaches to collective choice. It is also strange, of course, because it seems to represent a retreat from the distinction between preferences and social values. In any case, Sagoff writes that:

Pollution control provides a clear example of the opposition [between these two conceptions of collective choice]. A utilitarian approach regards pollution primarily as an external or social cost or production and therefore may seek to internalize this cost in prices paid for goods and services that pollute. A principle-based perspective, in contrast, regards pollution as a form of coercion, invasion, or trespass to be regulated as a violation of the rights or person and property. From this perspective, pollution constitutes a tort or nuisance—like a punch in the nose. On this view, pollution control protects the rights of individuals against trespass, which is not the same thing as—and it may conflict with—satisfying consumer desires or preferences.¹⁶⁰

Suppose that pollution does constitute a tort. Does this mean that there may not be any pollution, ever? Surely it would spell trouble for Sagoff’s account if it did, because life without pollution is simply impossible. What it would actually mean, it seems to me, is that you cannot expose me to pollution without compensating me for being exposed. This is an interesting claim, and one worth considering; but it looks an awful like the claim that I called Claim 1 early in this project (i.e., the first claim that I mentioned among the three that comprise the basic theory of CBA). Whether compensation is required when one person benefits and another loses is unquestionably an interesting question, but I already said that I do not want to assess this claim in this project. Rather, my aim – as I hope I have made clear

¹⁶⁰ Ibid, p. 96.
from the beginning – is to focus on CBA’s theory of value. Since Sagoff’s argument appears to veer off into a discussion of CBA’s theory of right action, this seems to be a fitting place to end my discussion.
CHAPTER 6: A PLAUSIBLE META-ETHICAL UNDERPINNING FOR NON-MARKET VALUATION

Section 1: Introduction

Throughout this project, the primary question on which I have focused has been: With respect to the analysis of public policy proposals, especially those concerning regulations and public investments, what should count as a benefit, what should count as a cost, and why? The preceding discussion has suggested that it is overly simplistic to say that in every case the satisfaction of (an actual) preference should count as a benefit. I have argued that the most clear-cut benefit is the satisfaction of fully-informed and rational preferences, and have explained that the state’s role is to provide as much empirical information as possible, to help (as far as possible) individuals to organize and process this information, and then to aim to satisfy these preferences (though again I have not endorsed the Kaldor-Hicks Criterion, or any other theory of right action for that matter). Pursing an answer to that primary question quickly led me to a discussion of value, and more specifically the question of whether the satisfaction of a preference has value, and if so, why and what kind. The notion of value dropped out of play for a while, but then made a reappearance in the last chapter, where a distinction between individual preferences and social values was considered, but not endorsed. Still, someone could read, understand, and agree with everything I have said thus far, but still not see how exactly it all bears on the economic approach to valuing goods, especially non-market goods, including many environmental goods.
I have perhaps been somewhat suckered by economists into choosing the primary question I did and avoiding straightforwardly presenting all of this as an account of value. Most economists are frustratingly hard to pin down when it comes to talk of what things have value, and what it is in virtue of which valuable things are valuable. CBA may stay neutral regarding which particular things have value, and how much, but it is surely not meta-ethically neutral, and those who endorse CBA would do well to lay their meta-ethical ones on the table. The primary question I considered was an important one, and it was not a mistake to focus on it, but I too need to give an account of value – meaning, in this case, not an account of what particular things have value (how much, of what kind, etc.), but an account of the nature of value. It makes sense for me to give that account now. As I hope will be clear, it is possible to translate my claim about what ought to count as a policy-level benefit into, or at least to pair it up with, a credible meta-ethical account of value. The account of value that I would endorse in conjunction with all that has preceded it is akin to an ideal observer account, but with preferences in place of observations. It is subjectivist, and, it will be argued, just the right extent. It avoids all of the problems of a simple (read: naïve) subjectivism, but it also avoids the problems and implausibilities of objectivism and even intersubjectivism. In the account of well-being I have defended, something contributes to an individual’s well-being just in case the individual prefers it (or would prefer it) when she is possessed of full information and is not plagued by lapses in instrumental rationality. My account of value parallels this very closely. On my view, a thing, outcome, or service – and henceforth, for the sake of brevity, I will just say thing – counts as valuable to or for an individual roughly when it is the case that it would be preferred by a fully-informed and rational individual.
This is just a brief sketch. In this chapter, I develop this account and defend it principally by explaining how it stands up to some of the most common criticisms of the account of value on which CBA is typically understood to be predicated (henceforth condensed as: “CBA’s theory of value,” “the economic account of value,” or some other subtle variant). My account is not entirely contrary to the economic account. Where my account and that of CBA are in agreement, criticisms of CBA’s theory of value will obviously be criticisms of my account too. Thus, there are some elements of CBA’s theory of value that I will defend against the critics’ charges. I argue that many of those who criticize CBA on metaethical grounds fail to map properly the conceptual terrain; and even where the meta-ethical commitments of CBA are accurately characterized, the criticisms of these commitments are often unconvincing. However, CBA’s theory of value and mine do differ in several important ways. When the account I give parts ways with CBA’s account, considering important criticisms that have been leveled against the economic approach to valuing goods gives me a way of showing the merits of my view. On the whole, what I am doing in this chapter can be understood in effect as arguing for a conception of the nature of value of goods that ought to underpin economic valuation, and arguing that though this differs from the theory of value typically taken to underpin CBA, the view I advocate is one that proponents of CBA can and should adopt as a theory of value for CBA. In other words, where the economic theory of value has meta-ethical shortcomings, modifying rather than outrightly rejecting it is what is in order.
Section 2: Unsuccessful Criticisms of the Meta-Ethics of Economic Valuation

Some critics claim that there is an important distinction – and one, these critics say, that CBA ignores – between expressing preferences and making judgments. Value claims, this argument goes, are judgments, not merely expressions of preferences. Sagoff, for example, says that “it is the characteristic of cost-benefit analysis that it treats all value judgments other than those made on its own behalf as nothing but statements of preference, attitude, or emotion, insofar as they are value judgments.”161 If the suggestion here is (as it seems to be) that value claims must be understood as something more than expressions of preferences, then I agree. Instead, they must be expressions of rational and well-informed preferences. But if the view is that value claims must articulate something entirely different from preferences, namely judgments, then this a mistake. It is a mistake in part because to distinguish all preferences from all judgments is to present a false dichotomy. It is absolutely true that value claims must be something more than mere expressions of psychic urges. Any view to the contrary too deeply impoverishes our moral discourse. But preferences are more than psychic urges. Preferences have a cognitive component. It is a mistake to conflate preferences with emotions, but the fact is that both preferences and emotions involve judgments. To illustrate the point, I turn – once more! – to music. Consider that my preferring The Who to The Spice Girls is at least in part a judgment that The Who are more skilled as musicians, songwriters, vocalists, performers, and so on, than the Spice Girls. My preference is not a mindless pull towards The Who. The reader should refer back to what was said at the beginning of Chapter 2 about how the notion of ‘preference’ avoids falling on one side or the other of the old belief/desire divide. Consider even that mental states

161 Sagoff, “At the Shrine of Our Lady of Fatima,” p. 1292. There is an aspect of this quote with which I will not engage now but to which I will return at the end of this chapter.
standedly considered to be even more “visceral” than preferences are not wholly distinct from judgments. My being afraid, for example often involves my judgment that I am in danger. More precisely, my being afraid is, a number of scholars these days want to say, partly constituted by this judgment.\(^{162}\)

The above criticism gets connected up with other worries about CBA. According to Sagoff, CBA makes “the assumption that valuation is subjective, that judgments of good and evil are nothing but expressions of desire and aversion.” He suggests that “expressions of desire and aversion,” that “statements of preference, attitude, or emotion,” must be contrasted with those that can be “regard[ed] as true.”\(^{163}\) This criticism too is confused. Sagoff accuses proponents of CBA of denying the “objectivity or cognitivity of valuation”\(^{164}\) as if objectivism and cognitivism are the same thing. They are not. Objectivism, intersubjectivism, and subjectivism are all distinct, but they are three ways of being a cognitivist. Cognitivism regarding value claims is the view that value claims, “when literally construed, are literally true or false.”\(^{165}\) Subjectivism is not hostile to cognitivism; rather, it is a view about what needs to be the case in order for value claims, when literally construed, to be literally true (and also, then, what needs to be the case in order for them to be false). Geoffrey Sayre-McCord gives a clear, concise, and (I believe) accurate explanation of what moral subjectivism is. He writes that:

---


\(^{163}\) Sagoff, ibid.

\(^{164}\) Ibid, p. 1293.

According to one version of moral subjectivism, judgments of value make sense only relative to the desires, preferences, and goals of the judge, so that the claim that ‘x is good’ should be treated as elliptical for ‘x is good-for-me’; and this claim in turn is true or false depending on the desires, preferences, and goals of the person doing the judging… Alternatively, a subjectivist about value might hold that judgments about value are not relative to judgers, even though their truth is dependent on the subjective states of individuals. Such a subjectivist view might say value claims are about what would satisfy someone’s (but not necessarily the judge’s) desires, preferences, or goals… Or again, a subjectivist about value might hold that judgments of value are not relative to judgers but are still true, when they are true, only because of the desires, preferences, or goals of some particular person (rather than anyone).\textsuperscript{166}

The first thing to note, then, is that subjectivism can treat “statements of preference, attitude, or emotion” as true. The subjectivist sets out truth conditions on moral claims: moral claims, the subjectivist maintains, are true when these conditions are satisfied, and false when they are not. Statements of preference, attitude, or emotion need not be \textit{contrasted} with statements that can be regarded as true. Now the subjectivist’s view of what makes normative statements true or false may itself be false, but that is a conclusion that would need to be argued for. In order that other important concepts are out on the table and in plain view, let me add (borrowing a bit more from Sayre-McCord) that moral \textit{objectivism}, like subjectivism, “is a view about what the truth conditions are for moral claims as literally construed. The differences… lie in what each takes to be the proper literal construal. Objectivists hold that the appropriate truth conditions make no reference to anyone’s subjective states or to the capacities, conventions, or practices of any group of people.”\textsuperscript{167}

And intersubjectivism is the view that it is the capacities, conventions, or practices of some relevant group of individuals to which a normative statement’s truth-conditions make reference. These concepts will be important.

\textsuperscript{166} Ibid, p. 17.

\textsuperscript{167} Ibid, p. 19.
Once we see that subjectivism is a form of cognitivism, we are in a position to entertain concerns in a more clearheaded way. Some who criticize CBA may do so on the grounds that it is non-cognitivist, while others may complain that it is subjectivist. Let us reflect first on the charge that CBA’s value theory is non-cognitivist. (It is not implausible to treat Sagoff as an example of someone who wants to register this complaint, but there is no reason to focus too myopically on him or his formulations of the concern.) One wanting fully to understand CBA – including to understand whether or where it is reasonable – will want to know whether it has the capacity for treating certain value claims as true and certain others as false. Just as it was conceded that any plausible view will be one on which value claims involve judgments, so too should it be conceded that value claims must not be altogether without truth values. And frankly, I see no reason why a proponent of CBA would want or need to argue for a non-cognitivist value theory. She might simply say, for example, that a value claim regarding a thing is false if the thing is not actually preferred in the way and to the extent that the person making claims to prefer it. On this view, the person making a value claim is doing something other than saying “Yea!” or “Boo” or “Join me in being made happy by such and such,” since these things cannot strictly be false under any conditions. Proponents of CBA should thus make clear that the “expression of a preference” is thus not simply the expression of a preference. One is stating what her preference is. She is making a claim about what she prefers, and ipso facto about the relative value of things to her. Hers is a statement that can be true or false, though so far the truth condition specified – which is thus far technically only a falsity condition, i.e., a necessary condition for truth – is fairly simplistic. It requires that one be sincere in her statement of her preference. Let us be a bit more concrete. Assume that any preferences can be expressed in monetary terms.168

168 Again, the claim that any preference can be so expressed is one with which I disagree, and I lay bare my
Working on this assumption, one’s “expression” of her monetized preference, of how much she is willing to pay for a thing or outcome, will be false if she is willing to pay $N for $X but she insincerely says that she is only willing to pay $N-5 for $X. If one is insincere in specifying how much she is willing to pay for $X, then her expression of WTP will not capture what to her is the value of $X. Only the very, very naïve proponent of CBA would deny this. For this reason, the charge that economic valuation is implausible in virtue of being entirely non-cognitivist can be dismissed.

So, it is a necessary condition for the truth of an expression of preference as a value claims that in order to count as true, the expression must be sincere. Is this also a sufficient condition? The fact that a value claim can be false if the preference it expresses is not sincere does not automatically mean that any value claim that sincerely states one’s preference will be true. Nor, of course, does it preclude this possibility. Let us suppose then that we have an expression of a preference that we know to be sincere. Suppose that some person $P$ sincerely says she is willing to pay $N for $X. Can this claim be false? Well, insofar as it is just a descriptive claim about how much the individual making the claim is willing to pay for something, then no. If $P$ claims to be willing to pay $N for $X, and is sincere about being willing to pay $N for $X, then it is true that $P$ is willing to pay $N for $X. That is clear. But the claim is not supposed to just be a descriptive claim about how much $P$ will pay for $X. It is meant to be a claim about the value that $X$ has for $P$. Can it be false as a value claim? Some proponents of CBA would probably say no. That is, they will say that even *qua* value claim, i.e., *qua* claim about how much $X$ is worth to $P$, it will be a true claim if $P$ is sincerely willing to pay $N for $X. What we have now is the view that a value claim regarding a thing reasons for disagreeing in Chapters 8 and 9; but let us assume it is true for the time being.
is true if and only if the thing is actually preferred in the way and to the extent that the person making claims to prefer it. This is simple subjectivism about morality.

If the critics’ worry is that CBA is necessarily non-cognitivist, then it is reasonable insofar as there are genuine concerns about the tenability of non-cognitivism, but unreasonable insofar as it looks like proponents of CBA can easily avoid endorsing non-cognitivism. The more appropriate criticism, however, is that economic valuation embraces simple subjectivism, which is in fact an unacceptably naïve subjectivism. Though I have not taken a poll, I believe that most proponents of CBA actually recognize that simple subjectivism is flawed, and so want not just for specifications of preferences (in the form of WTP, or whatever) to be sincere, but also for them to be at least moderately well-considered. Certainly, sincerity and well-consideredness are not necessarily the same thing; I could, e.g., sincerely prefer something, but without having much considered the matter. The quest for well-consideredness is why, for example, Hanemann maintains that contingent valuation studies should avoid self-administered surveys (like mail-in surveys), and should instead use in-person interviews. “Furthermore,” he says, “interviews should occur in a setting that permits respondents to reflect and give a considered opinion, such as their home.”

But even this, I suspect, is not going to satisfy most critics of the meta-ethical underpinnings of CBA’s approach to valuing goods. Perhaps critics of CBA do not merely want value claims to be able to be considered true or false, but also want them to be backed (or able to be backed) by reasons. Note, however, that even naïve subjectivism can offer this. Consider the claim “Ice cream is better than frozen yogurt.” This claim, on the simple subjectivist view, means roughly “I prefer ice cream to frozen yogurt,” and would then be true if the person uttering it actually has this preference. What is more, this preference can

be backed with reasons. If you ask me why I prefer ice cream to frozen yogurt, I can give you reasons; the primary one that I personally will give is that ice cream tastes better to me than frozen yogurt. And if you want reasons why ice cream tastes better to me than frozen yogurt, I could probably offer some of those too. My ability to give reasons will bottom out at some point, but that is no surprise. When they bottom out, they may bottom out in certain brute facts about me – be they psychological, physiological, or whatever. I will argue in the coming pages that this is not a problem. In any case, the simple example illustrates that value claims, even if they are simple expressions (which in my sense is akin to statements) of preferences, can be true, and can also be backed by reasons.

Many critics of CBA will still not be mollified. Even while conceding that a subjectivist can be a cognitivist (and in turn that the proponent of CBA can be a cognitivist, and even one who can back moral claims with reasons), many critics of CBA are deeply disturbed by CBA’s apparent subjectivism. They may argue at this point that a compelling account of value will be one on which value claims are able to be backed by reasons, and are either true or false, and whose truth conditions are outside of, i.e., make reference to more than or something other than, the individual making the claim. The critics want value claims to have truth-conditions that make reference to something beyond – something independent of – a person’s (or people’s) subjective states. Sincerity and well-consideredness of statements of preferences still clearly make reference to the subjective states or capacities of the person making the claim. For this reason, the critic says, even sincerity and well-consideredness together do not suffice as truth conditions. If CBA cannot offer truth conditions that are more robust, then CBA’s value is untenable.
How convincing this criticism is will depend on just how far removed from the subjective states of the individual making the value claim the critic wants the truth conditions for the value claim to be. Imagine, for example, that you are asked to specify your WTP for a project that will endanger a certain kind of plant. Imagine that you really consider the matter: you think about how much money you have, and how much you spend on other similar goods, etc. Furthermore, you are perfectly sincere when you give your specification. But now suppose that you do not know that the plant is very rare and also uniquely useful in treating an illness that plagues your child. In these cases, your sincere specification would likely be one with which you yourself would be uncomfortable were you to learn about the plant’s usefulness or arsenic’s effects on the body. If you learned about such things, you would likely want a new opportunity to specify your preferences. The outcome that you claimed to prefer to a certain degree is in fact not one that you prefer to that degree. You would maintain that your specification of this WTP was problematic. Your specification is true insofar as it is a claim about what you were willing to pay in the condition you were in when you made the claim; but it is not true insofar as it is a claim about what you would have been willing to pay had you known more about the plant, its pharmacological properties, and their relation to your sick child. Insofar as it is a claim about the plant’s value to you, your claim is false.

This vindicates the critics’ charge that sincerity and well-consideredness do not suffice as truth conditions. There is no point in my speculating on how many economists would endorse truth conditions on value claims that are so robust as to require more than sincerity and well-consideredness. Whatever the case, the account of value that I defend requires more than this. It is significantly more sophisticated. I do tie value and preferences
together, but the truth-conditions that I defend require more than sincerity and well-consideredness in one’s specification of one’s preferences. An individual can express a sincere and well-considered preference regarding a thing, but if this preference is uninformed or the product or irrational (in my sense), then there is no guarantee that the preference lines up with the thing’s value. The value of a thing for an individual is identified with the preferences that the individual would have regarding the thing were she rational and fully informed. If preferences could be fully expressed in terms of WTP – and, again, I will argue later that they cannot – then an expression of my WTP for X will be true as a claim about the value (to me) of X when and only when it specifies an amount that I would be willing to pay if I were rational and fully informed. More generally, a value claim is true if and only if it expresses preferences that the individual making the claim would have were she rational and fully informed about the relevant empirical facts. Those things that are valuable to an individual are the things that the individual would prefer were she rational and fully informed. (One could add ‘sincerely’ to each of the three previous sentences, but I believe it functions implicitly in each case. One could also add something about well-consideredness to those three sentences, but I think of it as being built into full-informedness.)

It is indeed reasonable for critics to want truth conditions on value claims that are not entirely within the individual making the claim. My account satisfies this want. The requirement that a preference not be irrational does not move us away from the subjective states of the individual making the value claim. We do, however, move away from this by requiring that the preference be well-informed. Statements of preferences (whether or not they are put in terms of WTP) are, when they are construed as value claims, not without truth conditions, and moreover, these truth conditions make reference to something beyond a
person’s (or people’s) subjective states or capacities: subject-independent facts about the
good in question. The reader will recall that in the course of articulating the informed-
preference account of well-being (Chapter 2, Section 5), it was required, in order to keep the
account from being circular, the being fully informed must not include knowing one’s level
of well-being in the various states of affairs over which her preferences range. A similar
constraint must exist here in the context of a compelling value theory for CBA; the difference
is simply that the information that is ruled out is not information about the individual’s well-
being, but rather information directly about the value of the various goods regarding which
the individual has preferences. A good’s value to an individual is identified with the
preferences she would have regarding that good were she fully informed and rational, but
being fully informed must not include being informed directly about the good’s value or else
the account is circular. One way of narrowing the information condition is to say that the
individual should be accurately informed about the physical properties of the different states
over which her preferences range; but this contentiously presupposes both that there is a
distinction between physical and non-physical properties and that the property of being
valuable is not a physical one. Still, some normative/non-normative distinction is going to be
needed. It is by no means a deep flaw of an informed-preference account that it requires
some such distinction. The most plausible way of cashing out this distinction is to say – and
this will unavoidably still be somewhat rough – is that full information is limited to the
empirical facts, where this is construed as those facts at least theoretically within the purview
of the natural, life, and social sciences.

Again, this is meant to echo in important ways what through the history of modern
moral philosophy been called an ideal observer account. On my view, however, what the
“ideal” individual is doing is not just observing, but preferring. I have said that the value of a thing for an individual is “identified with” the preferences that the individual would have regarding the thing were she rational and fully informed, and we are now in a position to make this more specific. The appropriately idealized individual’s preferences regarding things determines what things have value, and how much. But the word “determines” is ambiguous: the view could be that rational, full-informed preferences create value, or it could be that such preferences are they are just incredibly – perfectly, even – adept at detecting/responding to value. Economists who tie value to preferences are typically not clear on this point. Philosophers too can occasionally be guilty of not addressing this issue clearly. Consider, for example, David Sobel, who talks about the need to specify a privileged epistemic standpoint, “such that the preferences of the knowledgeable agent (or some subset of them) definitively determines the value of those goods or lives to the nonidealized agent.” How frustrating it is that Sobel misses the ambiguity in the word “determines.” Consider that what a parent does when she determines that her misbehaving son should be sent to bed without dinner rather than have his mouth washed out with soap is different from what a banker does when she determines that a customer has overdrawn an account. The parent’s determination is a matter of making something be the case, while the banker’s determination is a matter of discovering something to be the case. It is not as though the child’s punishment already exists and the parent has merely to go discover what it is. The parent may of course be able to get advice (from “professionals”) regarding punishment, but even so she determines the punishment in the sense that she creates the punishment; she makes it the case that the child gets one punishment and not another. The banker is not making it the case that the client has overdrawn an account. The banker is not, in other words, creating the
overdraw. Rather, the banker is discovering and verifying this. The ideal preferer is like the parent, not the banker. By preferring a thing, she makes it the case that the thing is valuable. She does not discover a value that already exists. This is essentially my view as well.

In explicating his ideal-observer account – which I take to be the standard bearer for such accounts – Roderick Firth says that “the ethically-significant reactions of an ideal observer must be psychological in nature, and… some of the evidence for the occurrence of these reactions could be directly accessible only to [the] ideal observer himself [or herself].”\(^{170}\) That such reactions must be psychological in nature helps illustrate why it makes sense to link my account to an ideal-observer account, even if it is not quite right to call mine an ideal-observer account. I would add – echoing claims that I made in Chapter 3 – that my ideal preferer is not so idealized as to preclude the possibility that two different individuals might count as ideally situated and yet have different preferences, which will mean that a might have different value across these individual’s. This is not to deny that two individuals who have exactly the same mental/psychological sets and who are in the exact same external circumstances will have all the same preferences, and as a result the value of \(X\) to one individual will, in substantive terms, be identical to the value of \(X\) to the other. Persons \(P\) and \(Q\) might both count as fully informed and rational and yet have different preferences such that the value of good \(X\) to \(P\) might nonetheless be different from \(X\)’s value to \(Q\). Thus, the manner in or extent to which what is morally relevant is “directly accessible” only to the idealized individual herself on my view is perhaps more extreme than Firth has in mind for his account. In any case, it is because there is an ineliminable psychological element in my account that I consider it to be subjectivist. But it is because the

truth of the information the idealized individual must have is independent of this individual’s impressions of its truth that the account is not subjectivist through and through, but is rather what I will call quasi-subjectivist, or a limited subjectivism. If the critic wants truth conditions on value claims that make no reference whatsoever to the subjective states of the individual making the claim, then my account will be a disappointment. But I will argue in the next section that this critic is pining for something she cannot have; the truth conditions she wants are too demanding, and she is going to run into insuperable ontological and epistemic problems.

One oft-anthologized criticism by Sagoff of CBA and its approach to valuing goods involves him presenting this approach as parallel, in at least a certain sense, to psychoanalysis. He says:

The assumption that valuation is subjective, that judgments of good and evil are nothing but expressions of desire and aversion, is not unique to economic theory. There are psychotherapists—Carl Rogers is an example—who likewise deny the objectivity or cognitivity of valuation. For Rogers, there is only one criterion of worth: it lies in “the subjective world of the individual. Only he knows it fully.” The therapist shows his or her client that a “value system is not necessarily something imposed from without, but is something experienced.” Therapy succeeds when the client “perceives himself in such a way that no self-experience can be discriminated as more or less worthy of positive self-regard than any other…” The client then “tends to place the basis of standards within himself; recognizing that the ‘goodness’ or ‘badness’ of any experience or perceptual object is not something inherent in that object, but is a value placed in it by himself.”

Rogers points out that “some clients make strenuous efforts to have the therapist exercise the valuing function, so as to provide them with guides for action.” The therapist, however, “consistently keeps the locus of evaluation with the client.”171

Frankly, I do not know whether Rogers’ picture of psychoanalysis is an accurate one. Some recent literature suggests at least that Rogers’ view is not universally held.172 In any

171 Sagoff, ibid, pp. 1293-1294.
case, there are certainly economists who would claim of economics what Rogers has claimed of psychoanalysis. For example, economist Alan Randall says that the “mainstream economic approach is doggedly non-judgmental about people’s preferences.”173 But note the contrast between my view of value and Rogers’ view of psychoanalysis. If CBA embraces the meta-ethical theory of value I have offered then the purported parallel between psychoanalysis and CBA does not hold. If CBA endorses an ideal-preferer account of the kind I have suggested, the criterion of worth need not lie – at least not exclusively – in “the subjective world of the individual.” If they go along with my account, proponents of CBA can insist that there is a “basis of standards,” a “locus of evaluation,” that is external – at least partly external – to the individual specifying a WTP. Rogers says that the psychoanalyst refuses to “exercise the valuing function” and practices an “unconditional positive regard” for all the affective states of the client. The psychoanalyst may do so, but the cost-benefit analyst who endorses the account I have recommended will not do so. The cost-benefit analyst should not hold specifications of WTP made under conditions of highly imperfect information in unconditional positive regard. There are still the considerations of respect for autonomy and the value of experiences of satisfaction that I discussed in Chapter 3, but purely from the point of view of the value of a good, a preference specified when one is ignorant of a plant’s rarity and curative properties, or ignorant of arsenic’s harmfulness when ingested, or whatever, can be appropriately “discriminated” as “less worthy” of positive regard than it would have been had you been, with respect to the empirical facts, better


informed when you expressed your preference. Such specifications are dubious as value claims.

I have argued that my ideal preferer makes value judgments, and have pointed to at least certain truth conditions that are outside of the individual making the value claim, but I have admitted to taking the ethically significant reactions of this observer to be psychological in nature. Is CBA properly criticized if it embraces this limited subjectivism? From anyone who answers “yes,” I will want to know what the alternatives to this view are, and why any of these should be counted as superior. Sagoff, for one, talks approvingly of “a Kantian conception of value,” which he believes contrasts with that of CBA. He claims the following:

The individual, for Kant, is a judge of values, not a mere haver of wants, and individual judges not for himself or herself merely, but as a member of a community or group. The central idea in a Kantian approach to ethics is that some values are more reasonable than others and therefore have a better claim upon the assent of members of the community as such. The world of obligation, like the world of mathematics or the world of empirical fact, is intersubjective, it is public not private, so that objective standards of argument and criticism apply. Kant recognizes that values, like beliefs, are subjective states of mind, but he points out that like beliefs they have an objective content as well; they are either correct or mistaken. Thus Kant discusses valuation in the context not of psychology but of cognition. He believes that a person who makes a value judgment—or a policy recommendation—claims to know what is right and not just what is preferred. A value judgment is like an empirical or theoretical judgment in that it claims to be true, not merely to be felt.\(^{174}\)

What exactly, we should ask, is known by someone who “know[s] what is right and not just what is preferred”? What is it for a value claim “to be true, not merely to be felt”? Additionally, Sagoff slides between talk of objectivity and of intersubjectivity. These are different. Which of these is the right account, and why? These questions will be taken up in the next section.

\(^{174}\) Sagoff, ibid, p. 306. Emphasis is in original text.
Section 3: Why Objectivism and Intersubjectivism Should Be Rejected

One possibility is that critics of CBA are drawn to that class of ethical positions that holds that how valuable a thing is, is altogether independent of how people appraise the thing, of how people value the thing, of the extent to which people prefer the thing or take the thing to matter. On the view I have in mind, that people – even fully-informed and rational ones – do not value some good or outcome X in no way proves that X is not valuable. But the view is not simply that there are subjects (i.e., appraisers) other than humans whose appraisals matter; rather, it is that the value of X is not essentially tied to appraisals of X made by subjects of any kind. This is an extreme form of moral objectivism. (The view that I mention by way of distinction, namely that it is not just how X is appraised by humans, but also how it is appraised by other animals, that matters, is ultimately the view I wish to defend. Again, later in this project I will work to defend the claim that a thing’s value is connected not just to humans’ appraisals but to animals’ as well.)

Back in Chapter 2, I briefly described part of G. E. Moore’s view. Moore counts as a moral objectivist. In the early part of the 20th century, Moore was working to analyze the concept “goodness” and to understand the relationship between value and valuers. He took issue with Sidgwick’s claim that “no one would consider it rational to aim at the production of beauty in external nature, apart from any possible contemplation of it by human beings.” Moore disagreed, and carried out a thought experiment that he believed shows Sidgwick to be wrong. Moore asks us to imagine two worlds, one exceedingly beautiful, and the other “the ugliest world you can possibly conceive.” We are to imagine further that

---

“no human has, or ever, by any possibility, can, live in either, can ever see and enjoy the beauty of the one or hate the foulness of the other.” “Well,” Moore says, “even so, supposing them quite apart from any possible contemplation by human beings: still, is it irrational to hold that it is better that the beautiful world should exist, than the one which is ugly?” Moore believes that this is not irrational to hold; it is, in fact, exactly what he holds!

As I said in Chapter 2, Moore contended that certain objects or states of affairs possess intrinsic value, which he took to be an objective, non-natural property. Humans can be skilled or not skilled at detecting the property, but its being there (or not being there, as the case may be) does not depend on its being detected. Consider, by analogy, the shape of Earth. Whether Earth is actually flat or round does not depend on anyone’s appraisal of its shape. Even when it was universally held that Earth was flat, Earth was nonetheless round. Likewise, Earth would be no less round tomorrow if tomorrow all people – or all conscious creatures with the concepts of flatness and roundness – began believing, however sincerely or ardently, that it is flat. It would, in fact, be no less round even if all conscious appraisers were somehow to disappear from Earth (and indeed from the universe altogether). The truth conditions on a claim about the roundness or flatness of Earth make no reference whatsoever to the subjective state of any appraiser (in particular or in general); there do not even need to be appraisers – human or otherwise – in order for the truth conditions to be satisfied.

Moore’s view about value is similar to this view about Earth’s roundness. The difference is that while Moore (like most people) would count Earth’s roundness as a natural property (by which I roughly mean a property within the purview of the natural sciences), he would count intrinsic value as a non-natural property. It is not within the domain of the natural sciences to uncover the presence or absence of intrinsic value, in general or in any

particular case, but its presence or absence is no less objective. While the instruments of
science are not well-equipped for this purpose, human beings are capable of detecting this
objective, non-natural property, though doing so will require practice just as the detection of
natural properties does for the scientist. In the same way that the average person will admit
that Earth’s roundness does not depend on there being people to perceive this roundness, so
Moore would seem to say that an intrinsically valuable thing’s intrinsic value does not
depend on any actual appraisal of the thing’s value, or even on there being people (or any
conscious being, for that matter) to perceive this value.

This kind of view received support (of a sort) from early academic environmental
ethicists. Those familiar with the literature in environmental ethics will be familiar with the
so-called “last man” argument. The argument – really a science fiction thought experiment –
was proposed in 1973 by Richard Sylvan (then known as Richard Routley) and “helped to
launch environmental ethics as a branch of academic philosophy.” It is a thought
experiment that owes much to Moore’s “two worlds” thought experiment. Sylvan’s
experiment has you imagine that you are the last human being left on Earth and that you will
soon die. I want to modify this slightly: imagine that you are actually the last conscious
creature. When you are gone, the only life remaining will be plants, microbes, invertebrates,
etc. In one version of the experiment, Sylvan asks you to imagine that it is within your
power, just before you die, to destroy the last remaining redwood tree. As no conscious
creature remains, destroying it will not “hurt” anyone. It will certainly not contravene
anyone’s preferences. So, would it be wrong to destroy the redwood? The question can be
put slightly differently: Is there any reason for you not to destroy the redwood?

Some people will maintain that my question can be restated in another way still, in the following terms: Does the redwood have intrinsic value? (Indeed, I think that Routley himself took the thought experiment to demonstrate that the redwood – and many other things too, for that matter – have intrinsic value.) I am reluctant to word the question in this way. Moore was clear enough in explaining what he meant by the term “intrinsic value” (even if his arguments regarding its existence are not convincing). Though Moore is gone and not many philosophers today are Mooreans, “intrinsic value” is still one of the most used technical terms in environmental philosophy. It is also one of the most misused! \footnote{178} Since Moore, an insidious ambiguity has come to plague the term (and some terms that are akin to it). \footnote{179} Some people, in claiming that something is intrinsically valuable, mean, as Moore apparently did, that the thing would be valuable even if there were no conscious creatures, no valuers, around. But others who say that something is intrinsically valuable would object to being interpreted as saying that the thing would be valuable even if there were no evaluator around.

Christine Korsgaard is responsible for an excellent diagnosis and treatment of the ambiguity in the notion. In one especially important paper, Korsgaard aims to “describe two distinctions in goodness which are often conflated, and [to] try to show the importance of keeping them separate. The two distinctions are: the distinction between intrinsic and extrinsic goodness, and the distinction between ends or final goods, and means or instrumental goods.” \footnote{180} She notes that philosophers standardly make a distinction between

\footnotesize
\begin{itemize}
\item \footnote{178} The reader should refer back to Chapter 2, Section 2, for more on this.
\item \footnote{179} The point that I am aiming to make here could, for example, likely be made just as legitimately about the term “inherent value.”
\item \footnote{180} Korsgaard, “Two Distinctions in Goodness,” \textit{The Philosophical Review}, XCII, No. 2 (April 1983), 169.
\end{itemize}
two types of value, namely “intrinsic” and “instrumental” value. Typically, a thing (object, activity, or whatever) is said to have instrumental value if it is valued for the sake of something else. On the other hand, a thing is said to be intrinsically valuable if it is valued for its own sake. But as Korsgaard rightly notes, this is not the meaning that the words “intrinsic value” actually suggest. Rather, she says, “to say that something is intrinsically good is not by definition to say that it is valued for its own sake: it is to say that it has goodness in itself. It refers, one might say, to the location or source of the goodness rather than the way we value the thing.”181 She points out that the correct contrast to intrinsic value (i.e., to the value a thing has “in itself”) is extrinsic value, i.e., the value a thing gets from some other source. And “the natural contrast to a thing that is valued instrumentally or as a means is a thing that is valued for its own sake or as an end.”182 There is nothing to say that intrinsic and instrumental cannot be contrasts, that ends and intrinsic goods cannot be the same, or that extrinsic goods cannot be equated with instrumental goods; but if any of these are to be the case, it is by reference to a theory. For example, Korsgaard notes, “one might equate ends and intrinsically good things by claiming that those things which have intrinsic value are or ought to be treated as ends. In this case we have a significant, and rather metaphysical, claim about ethics and moral psychology: namely, that choice is or ought to be a response to an attribute that we perceive in things—the attribute of intrinsic value.”183 But then that claim will need to be defended.

Korsgaard is right, and makes it all seem relatively simple. On this understanding, “intrinsic value” is mind-independent value (unless the thing that is supposed to be

181 Ibid, 170.
182 Ibid.
183 Ibid, p. 171.
intrinsically valuable is itself a mind. But leaving aside that complication…) The intrinsic value of a thing that is intrinsically valuable does not have its source in a mind. The possibility of mind-independent value is at least among the issues in which Moore should be understood to have been interested. It is, moreover, central among the issues on which the “last man” thought experiment focuses us. The fact remains, however, that too few people have heeded her wisdom; the term “intrinsic value” is still used in different ways by different people, and in most cases the user does not make clear what he/she means by it. Because of this ambiguity, I simply try to avoid, as far as possible, using the term “intrinsic value.”184 I thus eschew wording the relevant question above as: “Does the redwood have intrinsic value?”, preferring either of the two wordings I suggested: “Would it be wrong for the last conscious creature to destroy the last redwood?” or alternatively: “Is there a reason for this individual not to destroy the last redwood?” Here, the reader may ask, “What reason would I have – were I this last individual – for doing so?” But that is to change the subject in a subtle but important way. The fact that one does not have a reason to destroy the tree does not in itself mean that one therefore has a reason not to destroy it! Do you have a reason not to do so? Adopting a Moorean non-naturalistic view of value is one possible route towards a “yes” answer; though it would be a mistake to too quickly conclude that a naturalized ethical theory cannot provide the same answer. The subjectivist – even the simple subjectivist all the way on the other end of the meta-ethical spectrum – might say the following: “What if I actually do not want to destroy the last redwood? Suppose that I do not like violence, even when it is directed towards things that have no capacity for consciousness and that cannot feel pain. To destroy the last tree would be to act violently towards it, and acting violently towards the last redwood would make me sad. Does wanting to avoid being sad not count as a reason to not

184 An obvious exception will be where I am discussing the work of someone who him- or herself used the term.
destroy it?” Sure enough, this does count as a reason. This surely scores a point for the subjectivist.

Let us restate the question once more. Is there any reason for the last individual not to destroy the redwood that is altogether independent of how she would feel about doing so? Here, as far as I can tell, the subjectivist must answer no. This is surely the answer that the typical cost-benefit analyst will give. The cost-benefit analyst will necessarily maintain that where there are no conscious creatures to appraise a thing, there is nothing of value. Even on my more nuanced view, if the individual is well informed and rational and truly the last conscious creature, then it will be the case that if the individual has no reason not to destroy the Redwood that is altogether independent of how she would feel about doing so. To put the point in other words: it is an implication of my view that if there were no conscious creatures, there would be no value. Those who disagree with this seem to have a reason to endorse a mind-independent conception of value, perhaps Moore’s view specifically, since for Moore, the fact that no one is willing to pay for a thing does not necessarily show that the thing is valueless; it could have intrinsic value that people – even sincere people who have given the matter significant consideration, are reasoning well, and are possessed of complete empirical information – are simply failing to recognize.

Sagoff’s objection to the meta-ethical underpinnings of economic valuation is not obviously motivation by an intuition that it would be wrong for the last conscious creature to destroy the last redwood. In fact, I have no evidence at all that Sagoff even has this intuition. But a view like Moore’s could be used to argue that things have a determinate value that careful thinking can discover. It could be used to argue that value claims are in fact based on judgments and that they have cognitive content. These are all clearly things that Sagoff and
other critics of CBA want to be able to say. It could also be used to argue that specifications of WTP are either true or false as value claims. Perhaps, e.g., they are true when they accurately specify a monetary amount that is proportional (however that is defined) to the presence of some certain non-natural property; and they are false when they fail to specify a monetary amount that is in proportion to the presence of that non-natural property. And yet Sagoff’s view does not seem to be a Moorean one. He does not want to completely separate the truth or falsity of value claims from those who are making them, i.e., he does not want to altogether separate moral appraisals from moral appraisers. Perhaps this is for the best for Sagoff, because the Moorean objectivist account has not gone in the way since it was offered a century ago, this is more because it is fascinating than because it is convincing. It is actually not terribly convincing account. While I have not seen an argument that suggests to me that it is altogether impossible for there to be Moorean intrinsic value, I badly want to avoid having to embrace an ontology that contains non-natural properties. All of the reasons cited at the end of Chapter 2, Section 4 for rejecting an “objective-list” account of well-being apply here as well. Additionally, there is no good solution to the problem of how we could ever know whether it is intrinsic value to which we are responding. How would we know whether someone who says that something has this or that amount of value in the Moorean sense is right or wrong? This is important because when the Moorean makes a value claim, she is seemingly not merely claiming that the valuable thing is valuable to her, but rather that it is valuable, period, and that this is something that others should recognize. When someone does this, we want her to be able to back up the claim with something more than reference to intrinsic value that – she assures us! – she detects even if we cannot.
It was worth considering whether a Moorean objectivism is the kind of that would satisfy Sagoff. Though it should not (and would not, I believe), it must be noted that there are non-Moorean versions of objectivism. That is to say, not everyone who holds that the appropriate truth conditions on value claims make no reference to anyone’s subjective states needs to maintain that what the appropriate truth conditions do make reference to is the presence of a non-natural property that would exist even if there were no conscious creatures to apprehend it. The truth conditions might be such as to require that there be appraisers, but not make reference to any subjective, psychological state of any appraiser. An objectivist might say, for example, that the claim “Pollution is bad!” is true independent of anyone’s preferences for or against pollution. Rather, this objectivist says, it is true by reference to reasons that people are obligated to recognize regardless of their psychological states, though it would not be true if there were no minds to apprehend these reasons.

This strikes me as much more reasonable, but it is not obvious that we must embrace even this more moderate form of objectivism. All of Mackie’s famous criticisms seem to tell against even it. What are these reasons, if they are not psychological in nature, and what claim do they make on people who do not recognize them? None of us believes that pollution is good, \textit{period}; that is, none of us believe that all else equal, pollution is better than no pollution, more is better than less, etc. The intuitiveness of the claim that “Pollution is bad” lures us into thinking it could be objectively true. But let us consider a more interesting and important question: How bad is it? It is not easy to say. Nor is it easy to say how bad arsenic in drinking water, or drilling in ANWR, or development of the old wooded tract in town of land, is. In fact, not everyone would seem to regard them as bad at all! This difficulty remains significant even when we have all of the scientific facts. One who makes a
very specific statement about the badness of pollution, or of drilling in ANWR, etc., and then insists that her claim is it is true by reference to reasons that people are obligated to recognize regardless of their psychological states, will be looked at with due skepticism—and only slightly less skepticism if she adds that it would not be true if there were no minds to apprehend these reasons. If we separate value from individuals’ subjective states, then what precisely settles the question of what is good, what is bad, and (in each case) to what extent (or in what way)?

My aim is not to show that objectivism must be false. Rather, it is (at least) to show that all of the apparent advantages of objectivism actually attach to my limited subjectivism as well, and that my view avoids much of the ontological and epistemological difficulties that plague even a moderate objectivism. Korsgaard says that “the advantage of objectivism is that it explains certain of our beliefs about that good that a subjectivist account cannot readily accommodate. We believe that people sometimes fail to care about what is good and sometimes have interests in or desires for things that are not good. Yet in subjectivist theories it seems as if anything one enjoys or desires is good at least ceteris paribus, and anything one does not enjoy or desire is not.”185 This is not an advantage that objectivism holds over my account. It is an advantage that objectivism holds over simple subjectivism, but I myself have argued that simple subjectivism is untenable. Sagoff sounds similar notes of complaint, but again these fail to demand objectivism. He says that “the analyst supposes in all such cases that ‘this is right’ and ‘this is what we ought to do’ are equivalent to ‘I want this’ and ‘this is what I prefer.’” This strikes me as a misrepresentation of cost-benefit analysts’ commitments; but in any case, the account of value that I defend – which, if it differs from the one on which CBA tends to be based, both can and should be adopted by

CBA – certainly denies these equivalencies. Sagoff maintains that “value judgments are beyond criticism if, indeed, they are nothing but expressions of personal preference; they are incorrigible since every person is in the best position to know what he or she wants.” Indeed, value claims must not be incorrigible. Not all wants are equally good or equally worthy of satisfaction. But these claims can be accommodated by an approach that recognizes that if an individual lacks important pieces of information, if the individual does not sufficiently consider this information, if the information is not vivid in the imagination, and if the individual suffers lapses in her instrumental reasoning, then the individual’s value claims are not incorrigible, and this in spite of the fact that “every person is in the best position to know what he or she wants.”

One remaining possibility is that the most plausible account is an intersubjectivist one. I do not want to dwell too long on Sagoff’s views, but I cannot resist pointing out again how messy his analysis is. He claims, in the final passage of his that I quoted in the previous section, that the world of mathematics is intersubjective. I would have thought that mathematics is the paradigmatic case of objectivity. (Perhaps this is a point in favor of objectivism. The ontological and epistemological worries associated with objectivism, it seems to show, are not such at always to demand a rejection of objectivism.) The appropriate truth conditions for mathematical claims make no reference to anyone’s subjective states, or to the capacities, conventions, or practices of any group of people. This is not to say that the claim that \(2+2=4\) would be true even if no one whatsoever existed; though it is not to deny it either. In any case, the world of mathematics is not intersubjective. Is the world of value intersubjective? Intersubjectivism about value is (roughly) the view that a thing’s value is determined (in the “make it the case” sense of the word, I believe) not by appeal to what any
single individual prefers (even if that individual is perfectly well informed and fully rational),
but rather by appeal to the preferences, conventions, or practices of some relevant group of
individuals. Economic valuation, the claim goes, should not look just at preferences had by
each of the (as it were “isolated”) individuals affected by the policy, but rather by looking to
a set of standards established by the relevant community for that community. The
intersubjectivist may, for example, maintain that something is valuable when there exists the
requisite kind of agreement that it is valuable; the agreement is essential to its being valuable.
The intersubjectivist seems committed to denying that any one person can determine – even
if just for herself, and no matter how well she reasons and/or how well informed empirically
she is – how valuable something is.

I do not believe that intersubjectivism provides the most plausible way of
understanding the truth conditions on value claims. It frankly seems to be a ramshackle
halfway house between objectivism and subjectivism. But again, what is really important to
me is to establish that my limited subjectivism can offer all that the intersubjectivist claims to
want. Sagoff, for example, criticizes economic valuation on the grounds that “all valuation,
according to this approach, happens \textit{inforo interno}; debate \textit{inforo publico} has no point.”\textsuperscript{186}
This is surely not an overt endorsement of intersubjectivism, but debate \textit{inforo publico} would
be a bizarre thing for an objectivist to cal for – at least without some story about how debate
between subjects gets us closer to what is objectively true – and we know Sagoff objects to
(naïve) subjectivism, so construing the claim as an endorsement of intersubjectivism seems
reasonable. Note that for reasons parallel to those I explained in previous chapters, the
account of value I have defended reserves an important place for debate \textit{inforo publico}. I
discussed how CBA should encourage the subjection of individual preferences to reflection,

\textsuperscript{186} Sagoff, ibid, p. 1293.
public discussion, etc., much as campaigns do before public votes occur. I said that it is important to provide forums in which individual preferences can be confronted with the competing preferences of other people, as this confrontation can lead to revision.

If affirming that debate is important makes me an intersubjectivist, then I am an intersubjectivist. In fact, though, I do not think it does. Let us suppose that each individual’s preferences are put out there, that each person is encouraged to learn about others’ preferences, that people challenge each other’s preferences, and that each person is given the chance, if he or she is so moved, to revise his/her preferences in light of this debate. To say that it is only after this happens that an individual’s preferences determine the value of a good is not to embrace intersubjectivism. The “locus of evaluation” is not entirely within the individual, because evaluation, in order to count as authoritative, must be made in light of full information and must not be plagued by cognitive errors, and what the empirical facts are is objective, not subjective or intersubjective. But there is no set of value standards that, though making essential reference to preferences, must necessarily cut across individuals. That is to say, once the aforementioned conditions are satisfied, there is no “locus of evaluation” outside the individual who makes the claim the truth of which is up for evaluation. Once those conditions are satisfied, the individual determines for herself what is valuable, and to what extent. This is surely not intersubjectivism. It is, however, quite reasonable. Indeed, it is all we can reasonably expect.

Section 4: Why a Limited Subjectivism Is Not Self-Defeating

There remains to be discussed one important challenge in the work by Sagoff with which we have spent time in this chapter. (It is, in fact, an issue I flagged early in the chapter
Many economists believe that policy analysis ought to be morally neutral and that where the market fails accurately to capture and express the direction and strength of individual preferences, the policy analyst succeeds in remaining neutral if she merely helps to reveal, and then to deliver social outcomes that reflect, these preferences. Sagoff writes that:

Economists likewise argue that their role as policy-makers is legitimate because they are neutral among competing values in the client society. The political economist, according to James Buchanan, “is or should be ethically neutral: the indicated results are influenced by his own value scale only insofar as this reflects his membership in a larger group.” The economist might be most confident of the impartiality of his or her policy recommendations if he or she could derive them formally or mathematically from individual preferences. If theoretical difficulties make such a social welfare function impossible, however, the next best thing, to preserve neutrality, is to let markets function to transform individual preference orderings into a collective ordering of social states. The analyst is able then to base policy on preferences that exist in society and are not necessarily his own.¹⁸⁷

Sagoff is right that the claim of political economists and policy analysts to moral neutrality is confused. They are not truly morally neutral. CBA is morally neutral only in a particular, somewhat limited, sense. The policy analyst may be aiming to take into account, accurately and fully, the preferences expressed by those individuals in the client society. She may be aiming to avoid imposing preferences on, and to avoid even influencing the preferences of, those in the client society. She may even succeed at doing so. But even if she does, this is not the same as being entirely “ethically neutral.” Political economists and policy analysts who endorse CBA are clearly not meta-ethically neutral. Whether these meta-ethical commitments are reasonable is another question, and one that I have tried to address. I have tried to draw economists away from any simple form of subjectivism, saying that not all of the relevant truth conditions on value claims make reference to the subjective

¹⁸⁷ Sagoff, ibid, p. 1294.
states of any ordinary individual. There must also be a truth condition that makes reference to, e.g., the fullness of the information that individuals have. This is a substantive ethical position. It is not straightforward a view about, e.g., how bad an extra 25 part-per-billion of arsenic is, or anything like this. Instead, it is a kind of view about what needs to be true in order for something to be valuable, and about what would need to be the case in order for a value claim to be true. Whether proponents of CBA will embrace this particular meta-ethical view, they do embrace some meta-ethical position (whether they realize it or not). And stepping down from the meta-ethical level, we can even see that CBA is committed to the claim that the right policy is the one that results (or that we reasonably expect will result) in the most optimal balance of gains to losses. In other words, it is committed to a consequentialist theory of right action; and as such loses even more claim to value neutrality.

All this is true; but Sagoff uses it as leverage to hoist a further conclusion with which I disagree. The worry expressed by Sagoff that we have not yet explored is that proponents of CBA – and he would implicate me here too, though I surely do not count as an ardent proponent of CBA – defend some form of subjectivism within the context of CBA, but are not subjectivists in saying that CBAs ought to be used to make policy decisions. This is purported to be an inconsistency. He says that “the cost-benefit analyst regards as true the judgment that we should maximize efficiency or wealth. The analyst believes that this view can be backed by reasons, but does not regard it as a preference or want for which he or she must be willing to pay. The cost-benefit analyst, however, tends to treat all other normative views and recommendations as if they were nothing but subjective reports of mental states.” One of the principal worries he is expressing here is – to put the point charitably – that the analyst is a subjectivist (of some kind or another) about all normative claims except

---

188 Ibid, pp. 1292-1293.
the ones that she utters in defense of CBA. According to Sagoff, the claims the analyst makes in defense of CBA are alleged by the analyst to be true, and objectively so, but the analyst does not think that CBAs ought to be used only if people prefer to use them, or only if the use of CBA would be recommended by a CBA. Indeed, this would lead to a problematic regress. This, Sagoff believes, is a flagrant inconsistency. The deep problem, he believes, is that this is an inconsistency that cannot be resolved; one cannot defend subjectivism in CBA without appeal to some non-subjectivist position, but then one is admitting that the non-subjectivist is not unacceptable after all.

This is an important criticism. Again, however, it is far from a devastating one. I do not think that proponents of CBA actually believe that their endorsements, i.e., the normative claims they make in support, of CBA are, strictly speaking, objectively true. I myself maintain my limited subjectivist position across the board. I believe I can do this consistently. Endorsing a limited subjectivism does not mean that it does not make sense to say, for example that each individual’s well-informed and rational preferences ought to be taken into account (and all the other normative claims I have stated and defended throughout this project). It is true that proponents of CBA do not think that a CBA ought to be done in order to determine whether policy decisions should be made according to CBA. Nor would they need to say such a thing in order for them to make a good case for policy decisions to be made according to CBA. Their obvious strategy in response to Sagoff’s charge is to resort to that old strategy of pointing out partners in guilt. Consider that no one says that we ought to have a vote to determine whether or not to determine public policy decisions by appeal to popular votes. This would lead to the same regress problem. But voting is not to be condemned on account of this. Neither should CBA be. There are indeed questions about
what kind of approval is required in order to start a political community. How would anything less than a unanimous initial vote suffice? And yet here we are, a world filled with political communities, in spite of the fact that there never has been such a unanimous vote. Voting is not sharply criticized, and similarly, CBA’s proponents should not be criticized merely on the grounds that they defend it as a legitimate mechanism for making public decisions while denying that its legitimacy depends on the outcome of a prior CBA.
CHAPTER 7: ANIMALS, PREFERENCES, AND VALUE

Section 1: Introduction

If there is just one point I have endeavored thus far to support, it is that individual preferences have a legitimate central role to play in the analysis of public policy proposals, especially those concerning regulations and public investments. But there is more to be said about whose preferences count. A *New York Times* op-ed piece by Jim Doherty opened in the following way:

When the Arctic National Wildlife Refuge was created a little more than four decades ago, the aim was to preserve, intact, what one of the original planners call a “completely undisturbed” cross-section of northeastern Alaska’s spectacular mountains, rivers, and tundra. Ever since, oil companies have been eager to tap the rich formulations they believe lie under its ecologically fragile coastal plain. This spring, the struggle between conservationists and developers over the nearly 20-million-acre refuge promises to heat up as Congress considers an energy bill that would open it for exploration and drilling.

But aren’t the creatures whose welfare the refuge is supposed to protect – all those grizzly bears, caribou, musk oxen and wolves – entitled to participate in the process too? Don’t they have a right to be heard?189

These are fine questions. Doherty answers in the affirmative, and I think he is right to do so. In this chapter, I explain how this puts us both in opposition to essentially all traditional forms of public policy analysis—though my focus has been, and continues to be, primarily on CBA, and especially its approach to valuation, i.e., what it counts as costs and as benefits. I also explain why this kind of opposition to CBA is justified.

---

There is one particularly odd feature of my general account of costs and benefits, and relatedly of my account of value, about which I have not yet said much, but which I wish to develop in this chapter. I have made an effort throughout this project, and especially in the last chapter, to talk about “individuals”—preferences of an individual, well-being of an individual, value to an individual, information possessed by an individual, and so on. I was careful to talk about individuals in a general sense rather than individual humans because my fully-articulated account denies that humans are the only relevant individuals. (Where I have talked about “people” or “humans” rather than “individuals,” it was frequently simply because I was not ready to introduce arguments about animal preferences, and to insist through the whole project on the term “individual” rather than “person” without explaining why would have been odd.) I want to make an important place in my view of costs and benefits and in my theory of value for the preferences of non-human animals. Costs and benefits should be linked not just to human preferences, but in a fundamental way to animal preferences as well. Like humans, animals – at least some of them – are sources of value.190

Throughout this project, I have pursued the question of whether CBA’s theory of value rests on compelling normative foundations. Thus far, I have found nothing that suggested that it is necessarily morally unsatisfactory. That is to say, nothing so far has suggested that CBA must be abandoned; though I have argued for several critical modifications to it at the level of both theory and practice. There is an additional question – and an important one, I believe – as to how public policy analysis should deal with the preferences of animals. How CBA itself fares in the final analysis will turn in part on how it deals – or on what resources it has for dealing – with animal preferences. Many proponents

---

190 Henceforth I will, for the sake of simplicity, generally just say “animals” rather than “non-human animals,” but I certainly do not mean to obscure the important fact that humans are animals too.
of CBA will deny that any animal has preferences; but it will be argued in this chapter that this denial is implausible. Even if proponents of CBA concede that animals have preferences, few such proponents would endorse trying to include these in a CBA. The reasons for excluding animal preferences from CBA – a particular kind of policy analysis in which all preferences are monetized and then added up and compared – are indeed considerable, though, as I will argue, not as strong as they may seem to be. In any case, if, in the end, CBA cannot adequately handle animal preferences, then this must be counted as a serious strike against CBA—at least if any other conceivable approach to policy decision making can do better. Our aim, in that case, should be to look for a form of policy analysis that continues to take individual preferences seriously, but that unlike CBA can accommodate animal preferences.

That CBA fails adequately to deal with animal preferences is not among the most common criticisms of CBA, but it is, I believe, no less serious than those more common criticisms. The criticism developed in this chapter is a fairly radical one: if it is right, the implications for all forms of policy analysis are probably quite serious. I do not expect that one chapter on this topic will suffice to fully address the topic, so what I do here should be understood only as a start at what would have to be a more sustained project—one I leave for another time. In any case, I ask my reader for patience and an open mind. And I offer the assurance, for whatever it may be worth, that whether or not what follows is correct, it is not mere sentimentalism.
Section 2: On Animals’ Preferences and Well-Being

In this section, I want to do two things. First, I want to establish that animals – at least some of them – have preferences. After doing so, I will aim to show that a preference-based account of well-being for animals is plausible, though it may differ slightly from the preference-based account of well-being for humans I defended. Need any ink really be spilled over the first issue? In most contexts, the claim that it is not humans alone who have preferences, but rather that at least some animals do so as well, would probably be too obvious to require mentioning, let alone defending through argumentation. The context of economic valuation is one of the exceptions: many proponents of CBA seem to work with a conception of preferences that actually entails that no animal has preferences. I continue wishing to avoid trying to offer a precise definition of ‘preference,’ since such definitions are generally artificial, trying to put a fine point on something that does not in fact have one. In any case, however the notion of a preference is precisely to be understood, I believe there is an understanding that is thin enough to make reasonable the attribution of preferences to (some) animals on some occasions, and that is at the same time thick enough to play an important role within policy analysis. Where a preference is understood simply as a pro-attitude, or as a relational kind of pro-attitude, animals are surely not altogether precluded from having preferences. Furthermore, it seems quite natural to attribute desires to (many) animals, so where preference is treated as synonymous with desire, many animals will count as having preferences. However, we know that economists will be displeased with these understandings. Economists have long tended toward a more behaviorist understanding of the concept of preference. Importantly, though, a behaviorist approach does not on its own necessarily preclude the attribution of preferences to (at least some) animals. Quite the
contrary: on this understanding, animals would seem certainly to count as *having* preferences. For example, whatever behavior it is in which the preference for getting food consists – a preference just about every human has multiple times a day – it seems plain that the average pet dog will at least sometimes evince such behavior. Economists typically understand preferences in terms of trading behavior, i.e., (roughly) a willingness to give up one thing for another. But even where a preference is defined in these terms, animals are clearly still not altogether precluded from having preferences, since animals can give up one thing for another. (The dog, for example, gives up a fetched stick for a treat. It prefers getting the treat to hanging onto the stick.)

As noted earlier in this project, the strict economic understanding of a preference is behaviorist in an even more bizarre way: the important kind of behavior is trading behavior, but a specific kind, namely trading (or being *willing* to trade) things for money. Where the notion of a preference is defined in this way – i.e., in terms of one’s WTP – then it is indeed far from obvious that animals count as having preferences. Animals have no conception of money and thus no obvious WTP. I will say more later about the possibilities for, and problems with, attributing a WTP to an animal. For now, the important point is simply that the strict economic understanding of preference is an extremely implausible one. It is clear why economists want to monetize value and all costs and benefits, and in turn why they want to talk about WTP: this allows them simply to add up costs and benefits and then make a straightforward comparison of the two. It is also clear why they want to connect up WTP with preferences: without doing so, they will struggle for an explanation as to why public policy decisions should be based on people’s WTP. But none of this gives them license to co-opt the intuitive or common-sense notion of a preference. I have argued for an important
place for preferences in the analysis of public policy decisions, but this argument only works when the notion of a preference is understood in a plausible way. It does not work if preferences must be understood in terms of WTP. In addition to arguments that will be offered in the next chapter, thinking about animals shows why this understanding is unacceptably problematic. Intuitively, and even on reflection, animals have preferences. However, they have no WTP. Therefore, preferences cannot be necessarily defined in terms of WTP. In other words, the fact that the economic understanding of preferences, identifying as it does preferences with WTP, is incompatible with animals having preferences shows not that animals are incapable of having preferences, but rather that it is a mistake to identify preferences with WTP (be it stated, expressed, or whatever).

To be fair, it is not just those who insist on understanding preferences in terms of WTP who deny that animals have preferences; but this is perhaps unsurprising. After all, even today some people deny that animals are conscious and able to feel pleasure and pain.\(^{191}\) This is not the appropriate place for an elaborate discussion of what pain is, but it seems plain that many animals feel pain and pleasure. This is not to say that every non-human animal feels pleasure and pain, or that all animals that feel pain feel it in the same way or in response to the same things. But that the higher mammals feel pain seems to me to be well supported both by common sense as well as by the best philosophical arguments and physiological evidence. There is conceptual space for the view that while some animals feel pleasure and pain, none has preferences. Sure enough, given the more-cognitivist-than-usual construal of the notion of a preference I urged in Chapter 2, I would not insist that it is impossible that there could be an animal that is capable of experiencing pain and pleasure but incapable of

having preferences. But the view that no animal can have preferences strikes me as terrifically implausible. Indeed, among the preferences that most animals have will be the preference for experiencing pleasure rather than pain.

It should go without saying that it is not the case that animals have preferences regarding all and only the things regarding which we humans have preferences. For example, a human may have preferences about the intellectual quality of her relationships, about what plays her local theater company performs next season, and so on. Animals will not have preferences about such things. But there will be plenty of preferences that both animals and humans have. Plainly, a dog who has been kicked repeatedly prefers not to be kicked again, and its preference here is quite like the preference a human who has been kicked would have for not being kicked again. A dog who is inside of a closed-up car on a hot day prefers not to be there, and again its preference here is quite like the preference a human has when inside of a closed-up car on a hot day for getting out of the car. (More will be said later in the chapter on the important issue of comparing animal and human preferences.) An animal does not have to say, or be capable of saying, “I have a preference,” in order for the animal to have a preference. In general, if one of the higher animals is aware of X and Y, and makes efforts to pursue X but no effort to pursue Y, then it seems to me correct (and perfectly intuitive) to say the animal prefers X to Y. This is not intended as a definition of an animal preference, or as a specification of necessary and sufficient conditions for the having of a preference. It is merely a statement, and one that seems true. Apes, bears, caribou, dolphins, elephants, tigers, dogs, deer, pigs, cows, cats of all kinds, and many birds are, I believe, among those animals that are clearly cognitively sophisticated enough to qualify as not just able to feel pleasure and pain, but also as having preferences. (I certainly
do not intend for that list to be exhaustive.) And that these kinds of animals qualify as having preferences is true in spite of the fact that there are certain other animals whose merely moderate amount of cognitive sophistication makes it unclear whether they have preferences, and in spite of the fact that there are certain “animals” – where the term is used in the loosest possible sense so as to include such things as insects, reptiles, crustaceans, etc. – whom we should probably regard as clearly not cognitively sophisticated enough to have preferences at all.

We move now to the second of the two tasks to which this section is devoted to accomplishing. Not only do (at least some) animals have preferences, but (at least some) animals also have interests, i.e., an individual well-being. Indeed, I believe these animals’ well-being should be analyzed in the same basic way in which human well-being was analyzed: an animal’s well-being connects up in important ways with the satisfaction of the animal’s preferences. Rocks have no well-being (or, to use a phrase some have favored as more precise, no well-being of their own). If I break a rock in half, its well-being is not thereby affected, because it has no well-being to affect. The rock may, upon breaking, be less or more suitable for the purposes to which I (or someone else) would put the rock. If I break it enough, it may even cease being a rock: it may turn eventually into, e.g., a pebble. But the thing – called a rock, or a pebble, or whatever – is not worse (or better) off for this. It has no interests of its own. It is a mere thing. Many entities are mere things. My desk, an automobile, the rug, a book, a dirt clod—these are all obvious examples. The Taj Mahal, while quite spectacular, is also a mere thing. As Joel Feinberg says:

Try as we might, we cannot think of mere things as possessing interests of their own. Some people may have a duty to preserve, maintain, or improve the Taj Mahal, but they can hardly have a duty to help or hurt it, benefit or aid it, succor or relieve it. Custodians may protect it for the sake of a nation’s pride and art lovers’ fancy; but
they don’t keep it in good repair for “its own sake,” or for “its own true welfare,” or “well-being.” A mere thing, however valuable to others, has no good of its own.192

(Whether plants, i.e., living but non-animal (non-conscious, non-sentient) entities, are “mere things” is a difficult question. I am inclined at least to say that they are not the kinds of things that have rights or to which we have obligations. More to the point, it seems to me that they are not the kinds of things that have interests—at least not interests that are themselves morally relevant. My car has, in a way of speaking, interests in having its oil changed every so often, but I do not fail morally when I fail to make such a change. Despite certain obvious differences between plants and cars – e.g., that plants are living kinds of things, while cars are not – I am inclined to say the same thing about plants: they have, in a certain way of speaking, interests in getting, e.g., sunlight and water, but they lack interests in the morally relevant sense. However, I am not confident about this. It is an issue regarding which there is a vexing and still-open debate, so I am flagging it as an interesting issue but then simply moving past it for the purposes of this chapter and project. It is worth saying more about at some other time, but cannot be dealt with more fully here.)

Animals – again, at least the “higher” ones – are clearly not mere things. Feinberg continues the preceding discussion by claiming that:

Many of the higher animals at least have appetites, conative urges, and rudimentary purposes, the integrated satisfaction of which constitutes their welfare or good. We can, of course, with consistency treat animals as mere pests and deny that they have any rights; for most animals, especially those of the lower orders, we have no choice but to do so. But it seems to me, nevertheless, that in general, animals are among the sorts of beings of whom rights can meaningfully be predicated and denied.193


193 Ibid.
Not all aspects of this passage are equally germane. My purpose is not to establish that animals have rights—for rights are, in general, not my focus in this project; I have, for example, tried as far as possible to avoid the question of whether or not the Kaldor-Hicks Criterion is consistent with respecting individuals’ rights. Rather, my purpose is to establish that animals have interests. Just as it seems clear to Feinberg that (at least some) animals are not mere things (i.e., have a well-being, have interests of their own), so too does this seem obvious to me. One cannot act for the sake of a rock, but one can act for the sake of, e.g. a dog. If I kick your dog, I harm you, its owner (assuming you do not have a fully-informed and rational preference that your dog, your property, be damaged), but I also harm the dog itself. And conversely, if I feed a hungry dog, I benefit it.

Feinberg’s view of what an animal’s well-being consists in is importantly similar to, but also importantly different from, my view. In the aforementioned passage, Feinberg speaks of animals’ “rudimentary purposes.” In giving an account of human well-being, I rejected attempts to tie this to humans’ purposes. Talk of purposes makes little more sense vis-à-vis animals than it did with respect to humans, and so for animals, as was the case for humans, I want an account of well-being that avoids appeals to teleology, i.e., to purposes. Feinberg also talks about animals’ “appetites” and “conative urges,” and this seems more promising. My account of well-being for animals will not be exactly like my account for humans, but it will have a great deal in common with it. In giving an account of human well-being, I said that an individual’s preferences can pull apart from her interests. There are, in other words, at least occasionally preferences the satisfaction of which does not increase an individual’s well-being. I said that an individual’s well-being consists in what she would prefer if she were rational and had full information (where this entailed that the information
was vividly apprehended and considered in a reasonably cool hour). Though I take the account of human well-being I defended in Chapter 2 to have been largely the right one, I do not insist that it was perfect in every respect. The fact is, the correct account of human well-being is difficult to specify precisely. It will be complex. Giving a successful account of well-being for animals will be no easier, and quite possibly more difficult. But the account of human well-being should be our guide. Given the many similarities between humans and the other higher mammals, it would be strange if animal well-being were to consist in something *entirely* different from that in which human well-being consists. We should understand an animal’s well-being as somehow connected to its preferences. That is to say, even with respect to animals, there is a conceptual connection between well-being and preferences. But for animals, as was the case with humans, a strict identification of preference satisfaction and welfare promotion seems overly simple. A more nuanced account of well-being was required for humans, and this is true for our account of animal well-being as well.

Just as inconsistent preferences – preferences, in effect, for *X* and for *not-X*, pose a problem for an account of human well-being, they also do so for an account of animal well-being. We can imagine, for example, a dog who prefers not to experience pain and terrible sickness but also prefers to eat chocolate bars when she can get ahold of them. Of course, the only inconsistent preferences that make trouble for credible informed-preference account of well-being are fully-informed ones. Perhaps the dog would not have both these preferences if she knew that the chocolate would cause the terrible stomach ache, but perhaps she would. In any case, if humans can have inconsistent preferences even when fully informed, it seems reasonable to say that at least certain other higher mammals can do so as well. The problem
of inconsistent preferences is in one respect less pronounced, and in another sense more pronounced, vis-à-vis animal preferences than it is vis-à-vis human preferences. It is less pronounced in that most animals will have a simpler set of preferences than the average human will. Animal preferences do not range over as many things as human preferences, are not influenced by advertising, probably do not exist at the “second order,” etc. All of this reduces the likelihood of irrationality (in the thin sense I intend, which again is simply inconsistency among one’s preferences). However, the problem is more pronounced insofar as it makes more sense with respect to humans than with respect to animals to talk about one who has inconsistent preferences coming to a realization that the inconsistency exists – perhaps by having this pointed out to one – and then resolving this inconsistency by jettisoning one of the preferences. Indeed, this was my basic suggestion regarding how we should think (in general terms) about the issue of what it is in which a human’s well-being consists when her preferences are inconsistent: it depends on which of the preferences she would (if informed and pushed to resolve the inconsistency) prefer to be rid of. Recall, though, that I said that if a person is not even willing or able to say which of the preferences she would prefer to be rid of, then her well-being is undefined. We are not able, from the outside, to say which preference should be jettisoned. When an animal has inconsistent preferences, the view that its well-being consists in the satisfaction of the preference that it would (when fully informed and pushed to resolve the inconsistency) prefer to retain is more controversial. Still, something like this has to be right. Though we will not be able to reason with the animal to figure this out, it still seems plausible, in terms of the metaphysics of animal well-being, to say that the animal’s well-being will be promoted by the satisfaction of the one preference (among the inconsistent set) that the animal would prefer to retain. The
alternatives to this – among which are principally moving entirely away from a preference-based account, or straightforwardly identifying preference-satisfaction with welfare-increase – are less enticing. Again, though, some comfort may be taken from the fact that occasions in which there is an inconsistency of this sort will be relatively few and far between for animals.

Somewhat more customary than inconsistent preferences will, I suspect, be preferences that arise in conditions where the information the animal has is problematically incomplete. The dog who eats chocolate that unbeknownst to the dog acts as poison in the dog’s system, or an elk who unknowingly walks into view of a hungry wolf, or a lion who puts her paw into a spring-loaded metal trap—each of these actions arises out of a preference the animal has, and yet through such actions the dog, the elk, and the lion are hurt. The interests of each are contravened. As was true in the human case, an animal’s preferences can pull apart from its interests when its preferences are based in incomplete information. Animals – some of them, at least – can surely possess information. (It may not even be infelicitous to say that these animals may possess knowledge; though I do not think this claim is necessary, since the account of well-being I developed earlier tended to talk simply about having or lacking information rather than knowledge.) Though there is no guarantee, it seems likely that if the dog had realized that the chocolate would be poisonous, the dog would not have had the same preference for eating it. When the dog lacks this information, the fact that it prefers the chocolate does not ensure that the satisfaction of this preference increases the dog’s well-being. Of course, it is hard to say with any preciseness what it is to which full information for an animal amounts. This was a difficult question – and one I surely did not fully answer – in the case of human well-being; and I do not want to gainsay
the difficulty of the question when it is posed regarding animals. We need to be sure that the amount and kind of information that “full information” requires is consistent with the animal being the kind of thing it is. That is to say, it cannot include knowledge of certain facts that, e.g., a bear (or cow, or dog) could possess only by being something other than a bear (or cow, or dog)! Still, it seems correct to say that the more information an animal is missing regarding what the consequences of its actions will be, the more reason there will be for doubting that the satisfaction of the animal’s actual preferences will increase its well-being. I leave the details of this account sketchy, but the general contours are clear enough: it makes sense to say that many higher order mammals have a well-being of their own, and that such well-being consists in what the animal would prefer were it possessed of important information and consistent in its preferences.

Section 3: Taking Animal Preferences to Count in Their Own Right

Even if convinced of the conclusions from the preceding section, many proponents of CBA will, I believe, still deny that animals have standing on policy issues. This means, in other words, that they will deny that animal preferences should count in their own right in the valuation of goods or outcomes and an assessment of a policy’s costs and benefits. Let us be clear: one can deny that animal preferences count in their own right is not to say that they do not count at all. Most proponents of CBA will allow that animal preferences can get into the cost-benefit calculus indirectly: if, but only if, humans have preferences regarding the satisfaction or frustration of animal preferences, then the satisfaction or frustration of animal
preferences matters to an assessment of costs and benefits. But if no human has a preference regarding animal preferences, then the satisfaction or frustration of the animal preferences does not get registered by CBA. If no human’s preference would be frustrated by the frustration of an animal’s preference, then the frustration of this animal’s preference is not a cost; and if no human’s preference would be satisfied by the satisfaction of an animal’s preference, then the frustration of this animal’s preference is not a benefit.

I disagree with this way of understanding the relationship between animal preferences and costs/benefits, and will endeavor to explain the basis of my disagreement. For now, let us ignore one complication: let us suppose that any animal preference I mention is one the satisfaction of which would promote the animal’s well-being. I will reintroduce the nuance in the next section; but for now it will be easiest to avoid stumbling if it is set off to the side. Recall from the previous chapter the discussion of Moore’s “two worlds” thought experiment. I am not so convinced that it would be better that the so-called beautiful world exist than that the so-called ugly one exist. It matters not, though, since beauty is not what interests me here. Rather, I am interested in the form of the thought experiment. A modified version of the experiment can help make the case for taking animal preferences to count in their own right in the analysis of public policy proposals. Let us imagine two worlds, $W_1$ and $W_2$. No humans reside, or will ever reside, in either of these worlds; in fact, neither is ever even going to be conceived of by humans (at least once the thought experiment has concluded). In both worlds, however, there are many, many non-human animals, including many of the higher mammals. $W_1$ is one where animals’ preferences are roundly frustrated and $W_2$ is one where animals’ preferences are frequently satisfied. Suppose, for example,

---

194 Others will maintain that only strictly self-interested preferences should be counted, so as to avoid double-counting problems. Recall, e.g., the discussion of Dworkin in Chapter 5, where I argued that other-regarding preferences should not be excluded.
that those animals capable of having preferences prefer experiencing pleasure to
experiencing pain, and that in $W_1$ the animals all experience nothing but torturous pain, while
in $W_2$ they regularly experience pleasure. In all other respects, $W_1$ and $W_2$ are the same. If I
were able to create (effortlessly, say) $W_1$ or $W_2$, I would surely create $W_2$. I cannot imagine
that anyone who takes the thought-experiment seriously – that is, who imagines that she
really is creating one or the other of these worlds – would respond differently from this.
(Anyone who did would be a frighteningly sadistic person.)

It is always difficult to know what exactly thought experiments such as this one really
demonstrate. I do not take it to demonstrate in any entirely categorical respect that everyone
must prefer $W_2$. It is not even obvious that a person who sincerely would create $W_1$ is
irrational in the strict sense I laid out back in Chapter 2; though I have a difficult time
imagining what would possess someone sincerely to do so, and I sure would not want to be
friends with such a person! The idea is not that the content of all people’s preferences must
be constrained in certain ways. It is not that when comparing two real-world states of affair
$A$ and $B$, where animals’ preferences are frustrated in $A$ but satisfied in $B$, each person must
prefer $B$ to $A$ (be willing to pay more for $B$ than $A$, etc.). It may well be true that anyone who
would create $W_2$ will be inconsistent if she does not prefer $B$ to $A$ (is not willing to pay more
for $B$ than $A$, etc.); but even that is not the important point. What is importantly
demonstrated is this: those (like me) who would create $W_2$ – and again, I take this to include
essentially everyone – are, in effect at least, already committed to the claim that it is not just
(well-informed and rational) human preferences that matter in their own right, but rather that
(at least some) animal preferences also count in their own right. To believe that one should
create $W_2$ rather than $W_1$ is to believe even where there would be no human beings around, it
would be better – in some important though as yet not precisely specified sense – if animals’ preferences were being satisfied than it would be if animals’ preferences were being frustrated. It is to believe that animal preferences matter – somehow – independently of humans’ preferences regarding these preferences.

Something more specific must be said. With regard to humans, I tried to show in the previous chapter how an account of value piggybacks on the arguments that I made regarding the nature of well-being. We can do the same with regard to animals. We can say that animals – at least some of them – are themselves sources of value. *This* is the first important sense is which animals’ preferences matter. Humans are not the only creatures that have preferences, humans are not the only creatures that have interests, and humans are not the only creatures that create value. Roughly put, if an animal is cognitively sophisticated enough to have what genuinely count as preferences, then what this animal would prefer if it were well informed and rational is valuable for this animal. Recall how Sylvan’s “last man argument” was worded when it was explained earlier. I intentionally set up the experiment so that you – the person entertaining the thought experiment – were not only the last human, but indeed the last remaining *conscious creature*. As no conscious creature remains, destroying it will not hurt anyone. Thus, while you may have no reason to destroy it, you also have no reason not to do so. When the world is described as one that would, when you are gone, altogether lack conscious creatures, it seems most reasonable to concede that then there is no value in this world. What is reasonably concluded is significantly different, however, when we tell the story in a slightly different way (and one a bit more faithful to the experiment’s name). If you are the last human, but not the last conscious creature, and more specifically, not the last creature that has preferences, then there *is* reason to say that the
world you leave behind, supposing that you do not set off the weapons and so do not kill the remaining animals, is one in which there is value. In that hypothetical world without humans, that which is valuable is (roughly) that which these animals would prefer.

My general view of the nature of value is, of course, not a new one. It is, essentially, the view Spinoza articulated: “We deem a thing to be good, because we strive for it, wish for it, long for it, or desire it.”\textsuperscript{195} R. B. Perry, who also defends a view to which mine is akin, correctly notes that what is fundamentally right about Spinoza’s view is the idea that “in the last analysis good springs from desire and not desire from good.”\textsuperscript{196} Perry talks about the “characteristic of living mind to be \textit{for} some things and \textit{against} others.” He calls “this all-pervasive characteristic of the moral-affective life, this state, attitude, or disposition of favor or disfavor,” an interest – in the sense, now, of \textit{taking an interest} – though it is essentially what I have called a preference; and he says “this, then, we take to be the original source and constant feature of all value. That which is an object of interest is \textit{eo ipso} invested with value.”\textsuperscript{197} Many others endorsing this general account of value could be cited. But especially germane are discussions that bring this general conception of value into contact with animals and (in effect, at least) their preferences. It is both a comfort and a disappointment to know that the view I advocate, according to which animals are a source of value, is not a novel one. For example, despite some slight terminological modifications that I would want, and despite an apparent tendency further in the direction of behaviorism than I would endorse, the view was essentially articulated by D. W. Prall in 1921 as follows:


\textsuperscript{197} Ibid, p. 56-57.
Anything is properly said to have value in case, and only in case, it is the object of the affective motor response which we call being interested in, positively or negatively… The being liked, or disliked of the object is its value. And since the being liked or disliked, is being the object of motor-affective attitude in a subject, some sort of a subject is always requisite to there being value at all—not necessarily a judging subject, but a subject capable of at least motor-affective response. For the cat the cream has value, or better and more simply, the cat values the cream, or the warmth, or having her back scratched, quite regardless of her probable inability to conceive cream or to make judgments concerning warmth.\(^{198}\)

So, animal preferences matter in the sense that they are sources of value. But there is another important respect in which animal preferences matter, and this is an even more straightforwardly policy-relevant one. To see this, recall some of the terminology introduced back in Chapter 1. One criterion for evaluating a proposed policy or project was the Pareto Criterion, which says to accept only those policies that benefit at least someone while harming no one. A distribution of goods is Pareto optimal if no one can be made better without someone else being made worse. A Pareto improvement is achieved if at least one individual is made better without anyone else being made worse. The recommendation made by the Pareto Criterion is to go ahead with a project if the state of affairs in which the project is realized is Pareto superior to the status quo. Because it would be very limiting – too limiting, we are told – to move forward only on those projects that made at least one individual better without making anyone worse off, the Kaldor-Hicks Criterion was introduced. It is enough, this criterion says, to move forward with a project if those who are benefited by a project could fully compensate those who are harmed by the project and even after this compensation is paid out still enjoy some benefit. It is enough, in other words, that the potential for a Pareto improvement exists.

The Pareto and Kaldor-Hicks Criteria are traditionally understood in entirely anthropocentric terms. In other words, for defenders of these criteria, “no one” in the previous paragraph really means “no one human,” “someone” really means “some one human,” “individual” really means “individual human,” “anyone” really means “any one human,” and “those” really means “those humans.” Such anthropocentrism is not acceptable even in criteria for evaluating public policy proposals. Consider the following (fantastical but dramatic) example. Suppose a proposed project will provide me with some relatively minor benefit and in no way harm any other human, but will (for some reason) involve the performance of extensive and hideously painful vivisection on a thousand chimpanzees! Suppose, moreover, that no human cares a thing for these chimpanzees, and so no human is willing to pay even a penny to prevent the project. No one is harmed by this project if the only individuals who “count” are humans. Going ahead with this project would bring about a Pareto improvement, and would thus be approved by both the Pareto and Kaldor-Hicks Criteria, at least as these are traditionally understood.

I believe this is enough to show that these criteria are morally indefensible. We know by now that many critics will say that the problem with these criteria that this thought experiment demonstrates is that they leave important policy decisions to people’s preferences. My concern is rather that the criteria leave such policy decisions to people’s preferences. I have argued at length in preceding chapters of this project that preferences are fundamentally the proper basis for public policy decisions—though I have said that whether it is actual preferences or rather somehow idealized preferences with respect to which costs and benefits should be assessed is a complicated question. Recall that in my Chapter 2 defense of the informed-preference account of well-being, I recommended a structural
constraint on preferences that does not amount to a constraint on the content/substance of the preferences. No preference may be disqualified (from being one the satisfaction of which would increase well-being) simply on the basis of its content, but preferences must, in order to count (as playing a constitutive role in the account of well-being), satisfy certain “formal” conditions. The same kind of move should be made here, but here the additional structural constraint is on whose preferences count. Whose preferences count is among those structural or formal issues that can and must, on my account, be constrained—or, more aptly in this particular case, unconstrained. So, to return to the chimpanzees: my claim is not that humans must prefer that the chimps not endure painful vivisection; that would be the kind of constraint that I want to avoid, namely one on the content of preferences. Instead, the claim is that it is not human preferences alone that count; this is a kind of structural or formal condition that must, if a preference-based account is to be plausible, be put in place.

If it is true, as I have argued, that costs and benefits are assessed relative to preferences regarding the thing, it must be human and animal preferences that are taken into account for the determination. Perhaps this amounts to a rejection of the aforementioned policy criteria on the grounds that they do not take animals’ preferences to count in their own right. Or perhaps it amounts merely to a rejection of certain interpretations of these criteria; that is, perhaps the concern I am expressing here is just with those who take “someone” to mean “some one human,” and so on. If these criteria were not problematic for other reasons as well, I would advocate simply using this broader, more inclusive interpretation. Because these criteria have other problems – many of which connect up with theories of right action rather than theories of value and are thus beyond the scope of this project – I refrain from advocating in this way. I simply want to say that animals’ preferences ought to be taken into
account by public policy. How exactly preferences for and against are weighed against one another in order to determine a policy is a topic about which I do not have a great deal to say in this project, though I will address this to some extent in the project’s concluding chapter.

There is, finally, an argument for the inclusion of animals’ preferences in the analysis of public policy decisions that works on the model made famous by Peter Singer and Tom Regan. If someone is committed to the view that it is not just her preferences that public policy ought to take into account, but rather that it must also take into account the preferences of other people who have preferences, then this person will be on extremely unstable ground if she denies that public policy does not need to take into account the preferences of animals who would be affected when we can clearly discern the animals’ preferences. If one says that it is her preferences and no one else’s that should be taken into account, she is a self-centered egomaniac. If she says that it is only her preferences and the preferences of members of her race that ought to be taken into account, she is a racist. If she says that it is only her preferences and the preferences of those of her gender that should be taken into account, she is a sexist. And if someone says that it is only her preferences and the preferences of those of her species that ought to be taken into account, she is a speciesist—at least if policy analysts can have some success in discerning what the preferences of the affected non-humans would be. Though Singer’s and Regan’s arguments differ from one another in important respects, they share an overarching similarity: they both argue that there is no reason for extending the wide consideration to humans – their preferences, interests, suffering, etc. – that would not also count as a reason for extending consideration to at least some non-humans. This is true even in policy contexts.
One issue with respect to which I want to stay agnostic is whether an animal itself has value – as opposed to, or in addition to, it being the case that the satisfaction of the animal’s preferences has/creates value – independently of humans’ preferences. Regan says:

What has value for the utilitarian is the satisfaction of an individual’s interests, not the individual whose interests they are. A universe in which you satisfy your desire for water, food, and warmth is, other things being equal, better than a universe in which these desires are frustrated. And the same is true in the case of an animal with similar desires. But neither you nor the animal have any value in your own right. Only your feelings do.

Here is an analogy to help make the point clearer: a cup contains different liquids, sometimes sweet, sometimes bitter, sometimes a mix of the two. What has value are the liquids: the sweeter the better, the bitterer the worse. The cup, the container, has no value. It is what goes into it, not what they go into, that has value. For the utilitarian you and I are like the cup; we have no value as individuals and thus no equal value. What has value is what goes into us, what we serve as receptacles for; our feelings of satisfaction have positive value, our feelings of frustration have negative value.\(^{199}\)

Regan believes that this is a serious deficiency in utilitarianism. Whether or not he is right will not be resolved here, and need not be. I want to say that it is \textit{at least} the satisfaction of preferences – animals’ as well as humans’ – that has value independently of whether we humans are willing to pay anything to bring such satisfaction about. Whether animals \textit{themselves}, and indeed whether humans themselves – as kinds of cup, to use Regan’s analogy – also have value I leave an open question.

**Section 4: The Policy Implications of Counting Animal Preferences**

Let us look at an argument that denies animal preferences a direct role in public policy decisions. Terry Anderson has been one of President George W. Bush’s primary advisors on environmental policy. Anderson is an advocate of so-called “free-market

\(^{199}\) Regan, “The Case for Animal Rights.”
environmentalism.” This is not the same as CBA, but as CBA is at least a quasi-market approach to environmental decision making, free-market environmentalism and cost-benefit analysis of environmental policy proposals share certain theoretical commitments. One claim made by Anderson (and his co-author, Donald Leal) is indicative of how unconvincing the standard economic rationale for ignoring animal preferences is:

We emphasize from the outset that [the free marketeer’s] way of thinking assumes that the environment’s only value derives from human perceptions. Under the anthropocentric conception, the environment itself has no intrinsic value. People cannot manage natural resources for the sake of animals, plants, or other organisms because there is no Dr. Doolittle to “talk to the animals” and find out what is best for them. As long as humans have the power to alter the environment, they will do so based on human values – the only values that are ascertainable.200

There are several important elements of this passage. Of these, some are worth simply noting. Others, however, must be not only noted but criticized in no uncertain terms.

The phrase “what is best for them” shows two things, both by way of implication: (1) that Anderson and Leal continue in that tradition of economists who give a central place to the concept of well-being; and (2) that Anderson and Leal believe that animals are not mere things, i.e., that things can go better and worse for them (that they have interests, i.e., a well-being of their own). Furthermore, the fact that our inability to “talk to the animals” gets mentioned as a kind of problem suggests that Anderson and Leal also believe – like the economists we encountered in Chapter 2 – that each individual, including, seemingly, each individual animal, is the best judge of her own well-being. Why else would this inability even be mentioned as a hindrance to our making judgments about “what is best for” the animals? So, though Anderson and Leal do not here explicitly mention animal wants or preferences, it does not seem unreasonable to see them as endorsing a preference-based view

of well-being, including of animal well-being. There is no evidence that Anderson and Leal deny that animals *have* preferences, or that animal well-being consists (somehow) in preference satisfaction. On their view, the reason why it is not the case that we ought to “manage natural resources for the sake of animals” is that we cannot do so, and the reason cannot, apparently, is that we cannot *know* what they prefer.

For several reasons, this is an odd and unsatisfying story. First, Anderson and Leal seem to be articulating an ontological point, *viz.*, that animals are not the kinds of things that are sources of value, but then it is not at all clear how the epistemic worry they go on to express – that we cannot talk to the animals, which (supposedly) means that we cannot *know* what they prefer and in turn cannot *know* what is best for them – is relevant. If animals have preferences, and preferences had by humans create value – and this is surely their view: in spite of the fact that they imprecisely say that value derives from human perceptions, they surely think that, among humans, preferring and not just perceiving is what is important – why would animal preferences not do so as well? If we do not, and cannot, know what animals’ prefer, then we cannot know what has value, or how much. We cannot know what, substantively, will promote the animals’ well-being. That will have important policy implications; but that “the environment’s only value derives from human perceptions” (or preferences) is not among these.

Second, while it seems true that we do not always know what is best for animals, it is ridiculous to say that our inability to talk to them *fully* undermines such knowledge. In fact, it is ridiculous whether or not we are working with a preference-based conception of animal well-being. However, since I have defended a preference-based account of well-being, let us suppose that what is fundamentally at issue is our knowledge of animal preferences. It may
well not always be easy to know precisely what an animal prefers. It is one thing to know that a certain animal is capable of having preferences (and this, as I have already noted, is itself sometimes difficult to know); it is another thing, and more difficult still, to know anything more specific about its preferences. In spite of these obstacles, we can often have a reasonably good sense of what animals prefer. Sometimes it will be obvious. At the least, the average animal surely counts as preferring to avoid pain. This alone is sufficient to offer significant guidance to those involved in public policy analysis. Moreover, it seems to me reasonable to say that the average animal also prefers not to be killed, prefers for its habitat not to be destroyed, and prefers for its food source not to disappear. The preferences mentioned above seem obvious, but the attribution of these preferences to most animals surely gets support from observing animal behavior. In cases where what an animal prefers is not so intuitively obvious, our need to observe animal behavior and make inferences from this will be more pronounced. In spite of what Anderson and Leal’s Dr. Doolittle comment suggests, the idea that preferences are revealed through non-verbal behavior should not be so foreign to practitioners of CBA, who are already well acquainted with the revealed-preference method. Our attributions of the aforementioned preferences surely also work somewhat by reasonable inference from what we humans prefer in situations that are sufficiently similar.

Suppose that a certain policy, the effects of which would wipe out an animal’s habitat, is proposed. Perhaps it is a proposal to build a dam that would flood an enormous area where the animal lives. Do we know that the animal prefer that this policy not be approved? No, we do not. In fact, I am confident that the animal does not, strictly speaking,

---

201 Note that saying they are “revealed” through behavior is not the same as saying they consist simply in behavior.
have that preference. In other words, though it seems reasonable to me to say (at least of a higher mammal) that the animal prefers having an accommodating habitat to lacking such a habitat, it is not clear that just because the dammed river would ruin the accommodating habitat that the free-flowing river provides, the animal therefore counts as preferring that the river stay free-flowing. Surely no preference with that precise content “I would rather have a free-flowing river than a dammed one” has ever run through the animal’s head. This matters not, though, for even in the human case, the reader will recall that it was not simply what an individual prefers, but what the individual would prefer if fully informed and rational, that matters. Insofar as one of the state’s ultimate aims is to promote individual well-being, what matters is what the animal would prefer if it understood the proposal to build the dam. It seems clear that if it did understand the proposal, it would prefer that the proposal be defeated. The building of the dam would frustrate the preference the animal would have it if were well informed. It is obvious that we do not have to “talk to the animal” to know that it would prefer some different policy and that this policy will not enhance the animal’s well-being. All this shows, I believe, that the putative reason Anderson and Leal offer for why we must endorse an anthropocentric conception of value (at least for public policy purposes) is no reason at all.

While Anderson and Leal’s reasoning is quite unconvincing, there are nonetheless important questions regarding the proper role of animal preferences in public policy that remain. Insofar as I see this chapter as just a first look at an issue that ultimately deserves substantially more thought than I can devote here, I want to keep things simple; but I would ultimately do a disservice to my own position by oversimplifying. When discussing humans, there was a real question as to whether the state should promote a person’s preferences or
well-being when the two pull apart. I said that there is not a simple answer: one legitimate ultimate aim of public policy decisions is to promote well-being, though others are to promote experiences or satisfaction and to respect individual autonomy, and the latter of these can recommend against promoting the person’s well-being over and against her actual preferences. I said, moreover, that the state best accomplishes these things – all three at once – by providing people with information, helping them to organize and process it, encouraging debate, etc., and then assessing costs and benefits relative to the preferences people then articulate. In the previous section of this chapter, I asked my reader just to suppose that all the preferences being discussed were ones the satisfaction of which promoted the animal’s well-being (and indeed in each case that supposition seemed to me entirely plausible). There is, however, no reason to think this will always be the case. We can at least say that it is not a conceptual necessity. So, though I urged in the previous section that animal preferences be taken to count in their own right, there is a question – again, at least a very theoretical one – as to whether this means actual (though uninformed or irrational) preferences or informed and rational (though hypothetical) preferences.\footnote{In the previous paragraph, the issue was not that the animal had a preference that competed with its well-being. Rather, it was that the animal had no preference at all regarding something that nonetheless surely affected its well-being.} We need to think briefly but carefully about what the three ultimate but irreducible aims of public policy decisions that I explored in Chapter 3 – promoting well-being, respecting autonomy, and promoting experiences of satisfaction – mean for the relationship between (on the one hand) the satisfaction or frustration of actual (as opposed to hypothetical, idealized) animal preferences and (on the other hand) an assessment of costs and benefits. At least one thing is clear up front: there is no version of the suggestion that the state provide information, help assist in its understanding and processing, etc., that will work for animals. There is simply no hope of
supplying an animal with all the information that we would want it to have in order for us to conclude that the satisfaction of its preferences would necessarily promote its well-being. I take it that the reasons why this would be impracticable are generally obvious.

If there were a case where there is very good reason to think that we know better than the animal what will promote its well-being, then the legitimate ultimate aim of promoting well-being obviously recommends counting the frustration of the animal’s actual preference as a benefit. In such a case, it would be nice to be able to make animals better informed and more rational and then solicit their preferences, since this would – if it were possible – give us a more accurate sense of what the precise effects of a policy would be on an animal’s well-being; but this is not possible, and so we will have to settle for making estimates. What do considerations of autonomy recommend? Importantly, I believe these do not compete – at least not to the same extent that they did in the human case – with considerations of well-being. That is, contrary to what we concluded in the human case, considerations of autonomy do not seem to provide reason for counting the satisfaction of actual though welfare-reducing preferences of an animal as a benefit. Animals surely deserve a measure of autonomy. They ought not be enslaved, locked up, overworked, experimented on invasively against their will, etc. Nonetheless, what is required in order to respect an animal’s autonomy is much different from what is required to respect a human’s autonomy. The state does not fail to respect an animal’s autonomy if the state promotes the animal’s well-being over and against the satisfaction of the animal’s (welfare-reducing) preferences. Moreover, the state does not fail to respect an animal’s autonomy if the state does not actually supply the animal with relevant information that the state has but the animals lacks. This suggests simply assessing costs and benefits relative to what the state reasonably concludes that
animals would prefer if well informed and rational. The problem remains that the state could, at least in theory, promote the well-being of an animal without the animal thereby experiencing any satisfaction. Just as promoting experiences of satisfaction in the human case was said to be one of the proper aims of public policy decisions, so too must this be said with respect to animals. Thus, even with respect to animals there is an ineliminable tension, at least at the conceptual level, between promoting well-being and promoting experiences of satisfaction. Considerable comfort can be taken in the fact that those preferences an animal actually has are typically – and much more frequently than is the case for humans – ones the satisfaction will promote the animal’s well-being. Indeed, I suspect that many times when we make judgment regarding what an animal’s actual preference is, we do so in large part from a prior sense of what it is in which the animal’s well-being consists! In cases (like the dam case above) where a policy stands to diminish an animal’s well-being but would not frustrate any preference of that animal’s, it will typically be because the animal lacks all preferences regarding the policy, not because it has some preference regarding the policy the satisfaction of which would diminish the animal’s well-being. However, were the animal’s actual preference and its well-being to pull apart, then there would be reasons for counting the satisfaction of the preference as a cost and reasons for counting it as a benefit. As was true in the human case, there is alas no algorithm for how to proceed.

There are numerous other reasonable questions that arise in connection with the role I recommend for animal preferences. One of these is that there are billions and billions of animals in the world – many more non-human than human animals – and many animals who are affected by, and thus should seemingly on my view be counted as having (actual or hypothetical) preferences regarding, public policy decisions. One worry is that governments
conducted censuses in order to find out precisely how many people live where, and this information can be used to aid in the extrapolation of results gleaned through the revealed-and-stated-preference methods across the entire (relevant) human population, but we have no such census numbers for animals, and this would seem to get in the way of our accurately assessing costs and benefits relative to each animal. This worry should, however, not be overstated, since population biologists work – fairly successfully, I might add – to provide essentially this kind of data on animals. A related but more serious worry is that there are so many animals that their preferences may (in any of several relevant respects) overwhelm the analysis. All I can say in response is that the transition to counting animal preferences in their own right would have to take place slowly. Policy methodology would need to develop in accordance with the idea that it is appropriate to take such preferences into account. This practical worry is a legitimate one, but it is not enough to undermine the idea that animal preferences should be taken into account, and that they should not be discounted in importance merely in virtue of the fact that they belong to animals. The important point is, conceptually speaking, a fairly basic one: in those cases where it is clear that animals have preferences, if we can figure out what an animal prefers, or would prefer if consistent and possessed of the relevant facts about the policy options confronting it, this information surely ought to have a place in a ledger of costs and benefits that is used for the evaluation of policies.

Let us be very clear, though: some forms of policy analysis will have more trouble than others in accommodating even that one basic point. CBA is among those that will have a very difficult time doing so. Proponents of CBA will want to assess the strength of animals’ preferences and to compare these preferences to humans’ preferences. Economists
will note that it is problematic enough to determine the strength of human preferences and to try to compare the strength of one human’s preference to the strength of another human’s preference. They will claim that it will be much more problematic to determine the strength of an animal’s preference and compare the strength of an animal’s preference to the strength of a human’s preference. WTP is used precisely for this purpose (though, as will be argued in the next chapter, it frequently fails even as a comparison of one person’s preference to another’s.) Obviously there is no straightforward way to measure the strength of an animal’s preference using the animal’s WTP. Animals apparently have no concept whatsoever of money. They do not buy and sell anything. No behavior of theirs will clearly tell us how much money something is worth to them.

This is a concern, no doubt. Must we, however, make the jump from this problem to the conclusion that animal preferences are altogether incommensurable with human preferences? In other words, is it really the case that there is no scale on which at least some animal preferences can be weighed or compared against some human preferences? Surely it is at least sometimes possible to make comparisons between animal and human preferences, and to make these comparisons with a great deal of confidence. I like to drink a hot cup of tea from time to tea. I prefer a hot (though not boiling!) cup of tea to a cup of tea that is at room temperature (at which point I simply do not like it much). If all human preferences can be compared to one another along one scale, i.e., if any human preference differs from another human preference just in terms of strength, not kind, then we must certainly also be able to say that my dog’s preference for not being kicked repeatedly with a steel-toed boot over being so kicked is commensurable with – and more specifically, stronger than – my preference for warm tea over tea at room temperature. This is intelligible. It means
something, even if it does not mean that the dog would be willing to pay more to avoid being
so kicked than I would be willing to pay to be able to drink a hotter cup of tea. I am, in fact,
much more comfortable comparing my preference for a hot instead of cool cup of tea to a
dog’s preference not to be kicked repeatedly than I am comparing my preference not to be
kicked repeatedly to my preference, e.g., that some species of majestic migratory bird not go
extinct. The dog’s preference not to be kicked repeatedly is generally like my preference for
a cup of tea that is a bit hotter, except that the dog’s preference in this case is much stronger.
It is much less obvious how my preference not to be kicked repeatedly compares to my
preference that the bird species not go extinct; both are quite strong, but as I will elaborate on
in the final chapter, they are very different—seemingly different in kind. The proponent of
CBA would have us believe, however, that the first comparison is fully illegitimate while the
second is, conceptually, problem-free. We are told that all human preferences can be
compared on a single scale, but that no animal preference can be accurately weighed on this
scale. This all strikes me as disingenuous, or unimaginative, or both.

I do not want to identify preference satisfaction with pleasure. In any case, it seems
reasonable to say that if not all pleasures are commensurable, then not all preferences are
commensurable. Relatedly, and more importantly, if no animal pleasure is commensurable
with any human pleasures, then no animal preference is commensurable with any human
preference. For this reason, it will be instructive to bring my discussion of animal and human
preferences into contact with John Stuart Mill’s discussion of animal versus human pleasure.
I see nothing wrong with the view that at least certain animal pleasures and at least certain
human pleasures can be measured on the same scale. Indeed, the preferences that I called
commensurable across species were on the whole preferences regarding the avoidance of
pain and the experience of pleasure (e.g., getting kicked repeatedly, being locked up in a car on a hot day, drinking a hot cup of tea, etc.).

What is Mill’s view? In *Utilitarianism*, he maintains that pleasures can differ from one another not just quantitatively but also qualitatively. Higher pleasures, which include pleasures of the intellect, of feelings and imagination (love, friendship, art, music, etc.), and of moral sentiments (e.g., contemplating one’s courage and nobility) are, Mill thinks, *better pleasures* than the lower, bodily pleasures. It is for this reason, and because animals are (according to Mill) incapable of the higher pleasures, that Mill famously says that it is better to be a dissatisfied human than a satisfied pig. He apparently believes this is true no matter how dissatisfied the human and no matter how satisfied the pig. He seems to think, in other words, that the higher pleasures are *ininitely* better than lower pleasures. This can be understood as, in essence, the view that the higher and lower pleasures vary not along some one single scale, but rather show up on two entirely different scales. They are, in other words, incommensurable on his view.

If Mill is right about this, and if, moreover, animals are capable of only lower pleasures while humans are capable of only higher pleasures, then my claim that at least some animal pleasures can be weighed on the same scale as at least some human pleasures is mistaken, and the claim that at least some preferences are commensurable across species might need to go out the window too. But I do not take Mill to be claiming that humans are altogether incapable of lower pleasures. This is clearly *not* what he asserts. His claim seems to be that animals are capable only of the so-called lower pleasures, while humans are capable of both higher *and* lower pleasure. Again, Mill’s claim seems to be that the former are infinitely better than the latter. When he says it is better to be Socrates unsatisfied than a
fool satisfied, I take him to mean that a little of the higher pleasure (which Socrates experiences) is better than a great deal of the lower pleasures (which the fool experiences). The substantive point is debatable, but it should be noted that nothing here is incompatible with my view, since Mill is not denying that an animal’s pleasure – a lower pleasure – can be measured on the same scale as a lower pleasure experienced by a human. Mill’s view is that not all human pleasures can be measured on the same scale as animal pleasures; and nothing I have said requires denying this. In fact, Mill’s argument is more antithetical to the economists’ position than it is to mine. If Mill is right that even within the human community not all pleasures can be assessed on the same single scale, then economists seem to be in trouble, as economists generally want to argue that all human pleasures are commensurable. (They presumably want this at least as a precursor to the claim that all human preferences are commensurable.)

Mill’s claim that pleasures differ qualitatively, with some being better than others, is often criticized. The question for Mill is: What is it in which the goodness of pleasure consists? Pleasures differ in terms of intensity, duration, and costliness (i.e., the difficulty of obtaining the pleasure). These seem to make for quantitative differences in pleasures, but not in qualitative differences. So what is it in virtue of which pleasures differ qualitatively? What is more, even if there are qualitative differences, there seems to be no reason to think that higher pleasures are infinitely better than lower pleasures. More widely accepted is Bentham’s view, which maintains that all pleasures are qualitatively alike, all are measurable on the same scale, and differ only quantitatively, depending on their intensity, duration, and costliness. For this reason, Bentham’s view will in a certain respect be more to economists’ liking than will Mill’s. But if Bentham is right, then the possibilities for comparisons
between humans’ and animals’ pleasures, and in turn humans’ and animals’ preferences, proliferate, and the view I have defended in this chapter will then be even more germane. On Mill’s view, the comparisons were possible only with lower pleasures, since animals do not experience higher pleasures at all. Since Bentham rejects the higher/lower distinction, then all human pleasures and all animal pleasures can be assessed on the same scale. This is the dilemma for economists: hold that pleasures differ in kind, and then face the problem of incommensurable pleasures even among humans, or hold that pleasures do not differ in kind, and then face the problem of having to admit that animal pleasures will be commensurable with human pleasures.

That at least some animal pleasures are measurable on the same scale as at least some human pleasure provides some support for the view that at least some animal preferences and at least some human preferences can (in theory at least) be measured on the same scale. But what is this scale? For reasons I have already mentioned, it is difficult to see how it could be WTP. Before giving up entirely on this possibility, we might think about the idea of representation developed especially in two seminal papers, one by Christopher Stone203 and the other by Feinberg (which I quoted from earlier in this chapter). Both suggest that though a thing may not be capable of making claims, it may be a self-originating source of claims if it has interests that can be represented. An animal has interests that can be represented by a person who does her best to be sensitive, and give good representation, to the animal’s preferences. How much is an animal willing to pay (for this or that)? Well, the animal is taken as “speaking though” this representative. The animal is willing to pay however much its sensitive, sincere representative specifies. The specification made by the representative puts the animal’s preference on the scale used by economists for comparing one human’s

preferences with those of another human. This is still consistent with the idea that animal preferences count in their own right. The preferences that are getting into the calculus are the animal’s preferences. The representative is not imposing her own preferences, but is, in effect, merely translating (what she sincerely believes to be) the animal’s preferences into monetary units. The representative is not specifying how much of her own money she would be willing to pay in order to get (or prevent) a change. Rather, the representative is essentially imagining how much of the animal’s money the animal would prefer to give up (or accept) in order to get (or prevent) a change.

This is, of course, fairly rudimentary. One immediate and obvious difficulty is that the animal does not have money! Consequently, some decision would have to be made regarding how much money each animal should start with. How much this is, whether different animals get different amounts, etc., will be terrifically difficult questions. I am not sure that these are even questions it is worth trying to answer, since I have broader doubts about the monetizability of all costs and benefits. What is interesting and important to note, however, is that this same problem is one that advocates of free-market environmentalism face—where the question will be, e.g., who initially gets the rights to the air, the water, etc. It is, as I will elaborate on in the next chapter, also a question that plagues CBA. Since WTP is constrained by ability to pay, the question of how much money each human gets matters tremendously even when comparing one human’s WTP to another’s. The obvious answer here is that each person gets just as much money as she currently has. But why think that each person has exactly the amount she should have? Why think that there was ever a just initial distribution of money (goods, etc.), and even if there were such a distribution, why think it has not been upset via transactional injustices? The basic point is simply that the
problem regarding how much money animals would get is not obviously more vexing than one that proponents of free-market environmentalism, and even proponents of CBA, already face.

Some will say that though it is true that animals’ preferences do indeed count in their own right, we should not try to build animals’ preferences into economic analysis. We should, the objection goes, consider these preferences next to, rather than as a part of, a ledger of strict economic costs and benefits. I am certainly not unsympathetic to this view. I would take it to be a major step forward from where we are now if animals’ preferences, though not monetized and considered within an economic analysis, i.e., a CBA in the strict sense, were at least taken to be an important non-economic factor that demanded conscientious consideration from public decision makers. But I think that if we are going to perform CBAs at all – and it is a substantial “if” – then it would be better still for these preferences to be considered within the analysis. Too often when a consideration is not taken account of by economic analysis, it is simply neglected when it comes times for the decision to be made. And as I said earlier, it is wrong-headed to think that we could achieve a Pareto improvement through, e.g., that policy that no one would pay even a penny to avoid but that involved inflicting devastating pain on 1,000 chimpanzees. It is possible to count such a policy as delivering a Pareto improvement (or at least a potential Pareto improvement) but to acknowledge after finishing with the analysis that there are creatures whose preferences matter that we have left out of this analysis and that must be given consideration somehow. But how much more direct, and how much less speciesist, it seems to say simply that the example that I have described is not one that counts as yielding a Pareto improvement. If
animal preferences cannot be properly captured by CBA, then this should count as an important reason for moving away from CBA.

In short, the upshot of my argument is this: if CBA is to count as resting on a solid moral foundation, its proponents and practitioners should work to get to the point where animals’ preferences are taken to count in their own right in such analysis. If it is decided that this just cannot be done, let us at the very least get to the point where we – we who are interested in thinking about, teaching, learning, and/or doing policy analysis – acknowledge that economic analysis, in leaving animal preferences out of the calculus, leaves out something that matters, something that serious-minded public decision making should take into account in some form! But let us then also work on developing good alternative forms of policy analysis. It will undoubtedly be some time before politicians get on board with all of this, but that does not strike me a decisive consideration against the proposal. They will never get on board if no one is pressing this concern. Nearly anything is better than where things stand now, where animal preferences are taken to matter only if, and then only to the extent that, humans are willing to pay to satisfy them.
CHAPTER 8: WHY NOT ALL COSTS AND BENEFITS CAN BE MONETIZED (PART 1)

Section 1: Introduction

As has been said from the beginning, the aim of this project is to assess CBA’s theory of value. A great deal of time has been spent defending part of this theory of value, namely the claim that the satisfaction of individual preferences is what should count as a benefit and their frustration is what should count as a cost. Recall that this is what was, near the beginning of Chapter 1, labeled Claim 2. My contention has never been that this claim is patently true, but rather that there is a defensible formulation of it, even if this formulation is not the one standardly articulated by proponents of CBA. I argued at length that CBA is right to count the satisfaction of a preference as a benefit and the frustration of a preference as a cost, so long as the preferences are well informed and rational; and where they are not informed and rational, the state has an important role in working to make them so. This means that the economic approach to valuing non-market goods, environmental goods included, is right to tie a thing’s value to individuals’ informed and rational preferences regarding these goods. As I hope to have shown, it is possible to defend non-market valuation against a variety of objections, including: that it ignores people’s citizen preferences in favor of their consumer preferences; that it supposes that public policy should be dictated by individual preferences rather than social values; that it fails by basing policy
decisions on mere expressions of preferences, which are not even truth-assessable, rather than on actually-true judgments; and several others.

This is not to say, however, that CBA’s theory of value is acceptable. Claim 2 comprises part – but just part – of CBA’s theory of value. The other part of this theory of value is the claim that all costs and benefits can be expressed in monetary terms. (This was dubbed Claim 3. And Claim 1, recall, is CBA’s theory of right action, which I have sought to avoid addressing.) In this chapter, I explore the second of these two tenets comprising CBA’s theory of value. I argue that not all goods are such that their value can be expressed monetarily. Value is fundamentally pluralistic, which is to say that not all values are commensurable, or that there is no single scale – monetary or otherwise – on which all values can be measured. Differences in value are not all differences in amount, but are rather at least sometimes differences in kind. Recall my contention in Chapter 1 that the three claims characterizing CBA are all logically independent of one another. Showing that one is true does not show that either of the others is true; and showing that one is false does not show that either of the others is false. Having defended a version of Claim 2 does not mean Claim 3 is therefore true as well, just as refuting Claim 2 would not have sufficed to show that Claim 3 is false. (And defending Claim 2 also surely does not mean Claim 3 is false.) This is what I mean in calling the claims logically independent. Nonetheless, it is certainly the case that how one defends or criticizes Claim 3 is affected by whether one takes Claim 2 to be true or to be false.

I have defended a limited subjectivism (or quasi-subjectivism) regarding costs and benefits, which includes a limited subjectivism regarding the value of goods. This has implications for my argument in this chapter. How one argues for or against the claim that
any good’s value can be expressed monetarily will turn on the meta-ethical position regarding value that one endorses. This means, for example, that the arguments an objectivist would make for an incommensurability of values would probably be considerably different from those that I will make. To put the point in other words: If the critic of CBA could show there to be no connection between benefits and satisfied preferences, idealized or otherwise, then one who wants to know whether all costs and benefits are can be expressed monetarily need not wonder whether all preferences can be expressed monetarily. Having defended the view that it should be a goal of public policy to satisfy individual preferences, there is now ample cause for me to consider whether all preferences can be expressed monetarily.

I begin this chapter by discussing some of the unconvincing criticisms offered by those who (at least seem) to endorse a preference-based conception of value but who nonetheless deny that all costs and benefits can truly be monetized. But that some of the arguments against it are unsuccessful does not mean that there are not genuine problems with it. There are. I explored some of the problems in the previous chapter, where I said that at least some animals have preferences the satisfaction of which should be counted as benefits, though it will be essentially impossible to translate these into monetary terms. In the final section of this chapter, I discuss more of these problems. A number of these have been well documented, and so what I am doing is essentially just reviewing them. In the next chapter, then, I develop a line of argument in opposition to Claim 3 that has not been sufficiently explored.
Section 2: Flawed Arguments Against Monetizing Costs and Benefits

Everyone should admit that Claim 3 is not clearly dictated by common sense. It is not uncommon to hear people say “You cannot put a price on...” and then fill in the blank in any of a number of ways: human life, health, liberty, integrity, love, happiness, peace of mind, etc. In fact, this is likely to be the first objection offered by a student who is just being taught CBA. But what is common-sensical is not always true. The fact is, the case against Claim 3 is far from a slam dunk. While I believe that the claim that all benefits and costs can be measured on one scale is false, I do not take it to be patently false. Not all of the arguments against it are good ones, and many of those who give up on the claim do so too quickly.

An example of one popular but ultimately unsuccessful kind of argument against monetizing the value of certain goods is provided by Steven Kelman. Kelman writes that “one may oppose the effort to place prices on a non-market thing and hence in effect incorporate it into the market system out of a fear that the very act of doing so will reduce the thing’s perceived value. To place a price on the benefit may, in other words, reduce the value of that benefit.” Depending on how it is interpreted, this point is either false or too weak to pose any threat to CBA. Clearly it is true that one reduces the value of a benefit by assigning to it a price that is lower than its actual price. If a thing is actually worth $1,000 and I for whatever reason go around claiming it to be worth only $100, then I have reduced the thing’s perceived value and thus have done something wrong. This is obvious and would not be denied even by the most ardent defender or CBA. But if the thing is capable of having a monetary value, I do no wrong when I assign to it its true monetary value. The mere act of assigning a monetary value does not always reduce the value of the benefit; it does so when

---

the monetary value that is assigned is too low. The mere fact that a thing is not priced by the market does not mean that it is incapable of being priced by non-market means.

Take any mundane example of a benefit that is not priced through the market. My ability to take a walk on the undeveloped tract of public land down the street is such an example. Just because this benefit is a non-market benefit – that is, just because the market has not established a price for taking such a walk – does not mean that pricing the benefit will necessarily undervalue it. Suppose this benefit is monetized, and specifically priced as a $1,000,000,000 benefit to me. In this case, it is obviously terribly unconvincing to say that the benefit was cheapened by being priced. Contrary to Kelman’s apparent claim, placing that price on this benefit does not reduce the value of that benefit. If the value of a thing is incapable of being represented monetarily – and I believe this is true of some goods – then it is indeed wrong to assign a monetary value to the thing. But in this case, it is wrong because the thing has no monetary value. Though it does not always cheapen the thing, it does nonetheless misrepresent its value, and pricing it is wrong for this reason. It is wrong because it is fundamentally inaccurate, but not because it undervalues the benefit. Kelman needs another argument if he is to conclude that in every case, pricing a non-market good results in that good being undervalued as a result of being priced. There is simply no reason to think, with respect to every non-market good, that pricing the good cheapens it or will result in its being undervalued.

Kelman goes on to say that one “way in which placing a market price on a thing decreases its perceived value is by removing the possibility of proclaiming that the thing is ‘not for sale,’ since things on the market by definition are for sale. The very statement that something is not for sale affirms, enhances, and protects a thing’s value in a number of ways.
To begin with, the statement is a way of showing that a thing is valued for its own sake, whereas selling a thing for money demonstrates that it was valued only instrumentally.”

Again, Kelman’s argument is dubious. It seems true that if something has no monetary value, then it is not for sale. If it has no monetary value, then how much would or should it be sold for? In any case, it is surely not the case that anything that has a market price, i.e., that anything that has a monetary value, is thereby actually for sale. A thing gets its market price according to what most people (or the average person, or something like this) would be willing to buy and sell the thing for; but this does not preclude the possibility that a particular individual who has the good would not sell it for this amount, or for any amount for that matter.

Even supposing it is true that “placing a market price on a thing… [removes] the possibility of proclaiming that the thing is ‘not for sale,’” this observation challenges non-market valuation only if there are things that really ought not be for sale. There may be such things, but Kelman’s argument does not establish this. At best, it begs the question. Contrary to what Kelman argues, calling something “not for sale” or “priceless” (and as I have said, these are not the same, but the point I want to make holds for both) is not the same as saying that a thing is valued for its own sake. Consider again John Stuart Mill, who talks in terms of “desiring” a thing for its own sake, but seems to have in mind by this nothing different from what Kelman has in mind in talking of “valuing” a thing for its own sake. Mill observes that some people desire/value money for its own sake. He says, “The love of money is not only one of the strongest moving forces of human life, but money is, in

---

205 Kelman, ibid.

206 David Schmidtz works to make this point, but does not do so as convincing as he might have. See “A Place for Cost-Benefit Analysis,” Philosophical Issues 11 (2001): 148-171.
many cases, desired in and for itself; the desire to possess it is often stronger than the desire to use it, and goes on increasing when all the desires which point to ends beyond it, to be compassed by it, are falling off.” Note the obvious incoherence in saying that money has no monetary value. Thus, if Mill is right, then we have an example of something that is valued for its own sake but nonetheless has monetary value, and we in turn have an illustration that Kelman’s worry is misconceived.

While these arguments fall well short, there are nonetheless some more vexing problems for CBA’s commitment to monetizing all costs and benefits. Particularly important are various concerns with CBA’s use of WTP and WTA. As has been noted, an individual’s sincere expression of WTP or WTA is supposed to be an expression of that person’s preference. WTP and WTA are both supposed to be, as it were, measures of preferences. As mentioned back in Chapter 1 (Sec. 2), WTP and WTA sometimes fail – significantly, even – to line up with one another. Considerable empirical evidence demonstrates disparities between WTP and WTA. We should look carefully at differences between the two. Some of these differences – those considered in this section of the chapter – are not especially damaging to the project of monetizing costs and benefits; while others – those discussed in the next section – are indicative of a deeper problem with this project.

Consider the following story. I was given a large ceramic bowl as a gift recently. I was appreciative of the thoughtful gesture on the part of the giver, but ultimately the bowl was something I did not need. The giver included a gift receipt, so I could see where the bowl came from and would be permitted to return it there, though I could not see how much the giver had paid for it. I discovered when I went to return the bowl that it was priced at $20. I thought: “This is really a pretty nice bowl – if all I am going get for it is $20, I will

207 Mill, *Utilitarianism*, Ch. 4.
just keep the bowl.” This makes it seem as though my preference was to have the bowl rather than the $20. Interestingly, though, prior to receiving the bowl, I would not have been willing to pay $20 for it. If I had seen it in the store, I truly would not have pulled it off the shelf and marched to the register to fork over $20. When I did not already own the bowl, my preference on the whole was to have the $20 rather than the bowl. Now that I owned it, though, I was not willing to give it up for $20. And nothing in my circumstances had changed except for the fact that I now owned it whereas before I was, at most, merely looking at it on the store’s shelf. This story illustrates a phenomenon psychologist have called loss aversion: people have a reluctance to give up what they have that outweighs a desire they would have had to get that thing if they did not already have it. To adopt the language of utility, the idea is that the perceived disutility of giving up an object is typically greater than the perceived utility associated with acquiring it. Individuals have a strong tendency to remain at the status quo, because the disadvantages of leaving it loom larger than advantages. Researchers have seen that an alternative became significantly more popular when it was designated as the status quo, and the advantage of the status quo increases with the number of alternatives.\footnote{See Kahneman, Knetsch, and Thaler (1991), Kahneman and Tversky (1979), Samuelson and Zeckhauser (1988).}

This story – and the general phenomenon – has been thought by some to pose a problem for the monetization of all costs and benefits. The basic question the story above poses is: How much money is the bowl worth to me? More than $20, or less? It is not obvious. Does the fact that I never went out to spend $20 for this kind of bowl show that it is worth less than $20 to me? Or does the fact that I am unwilling to return it for $20 show that it is worth more than $20 to me? This is, in effect, a situation where my WTP is not the same
as my WTA, but where there is no fact of the matter about which measure should be used. The question of whether it is what I was (not) willing to pay or what I was (not) willing to accept that matters seems to be an open one. And again, this is not just a problem for the stated-preference method of monetizing costs and benefits. Certainly those formulating surveys must decide whether to put a question in terms of WTP or WTA; but even those making inferences from people’s behavior must confront the fact that what someone in practice pays (or would pay) to get something will often be different from what she in practice accepts (or would accept) to give that thing up. In any case, the general point of the objection is that for those seeking to monetize value, how valuable a thing is, i.e., the extent to which a proposal generates costs and benefits, turns on something that, intuitively, it should not turn on, namely on who happens to have what at the specific time the analysis is performed.

In fact, the problem is not just one for those who advocate monetizing all costs and benefits. (That is, it is not just a problem for those defending Claim 3.) Rather, it is a problem even for those who defend the view that benefits are to be identified with preference satisfaction, including those (like me) who add the requirement that the preferences be informed and rational. Even when the aim is not to translate all preferences into monetary terms, the phenomenon of loss aversion remains. (There is, in other words, also a challenge to those defending Claim 2—even the modified version of it that I defend.) But this is actually a problem that a preference-based conception of value can reasonably handle. Proponents of such a value theory can defend themselves in several ways, all of which acknowledge that an individual’s preference for some good thing would be stronger if she had the good and was being asked about giving it up than her preference would be if she did
not have the good and was being asked about procuring it. One possibility is to maintain that, with respect to any individual, both keeping the good when she has it and getting the good when she does not have it should count as benefits, but to insist that the question as to whether these do so to the same or to different extents is an open one. The idea, in other words, is to relate benefits to satisfied (informed and rational) preferences and costs to frustrated (informed and rational) preferences, but to deny that attending to such preferences enables us to straightforwardly draw conclusions about how big a cost or benefit is. But taking an even stronger line does not seem inappropriate. Proponents of a preference-based conception of value should argue that the phenomenon of loss aversion does not mean that there is no fact of the matter about how valuable a thing is. There is no a priori fact of the matter about how valuable a thing – the bowl, say – is; and there is no preference-independent fact of the matter about value. Relatedly, there is no fact of the matter that is independent of the actual distribution of goods. We are psychologically constituted so as to have a stronger preference, ceteris paribus, for keeping what we have than for getting what we do not have. This is how we are, and it is simply an implication of this that a thing is more valuable to us when we possess it than the same thing is when we do not possess it. There may be – are, I will argue – real problems with the use of WTP and WTA, but the fact that one’s WTP to get something one lacks differs from one’s WTA to give up something one has is not among such problems.

A related worry is in fact more severe. In the previous story, the difference seemed to turn on whether or not I already had the bowl: when I had it, I valued it more highly than I did when I did not have it. But consider the following pair of cases. In both, I already have the bowl, but in one I am asked how much I would be willing to pay to keep it and in the
other I am asked how much I would be willing to accept in order to give it up. If I am like most people, then even here there will be a difference between my WTP and my WTA. Research consistently shows that in such cases WTA questions generate values from three to nineteen times greater than those elicited by WTP questions.\textsuperscript{209} Importantly, loss aversion cannot explain this difference; in both cases, I have the good but am being asked about the prospect of losing it. Daniel Levy and David Friedman argue that the different sorts of inquiry – assessing my WTP to keep the thing or assessing my WTA for giving it up – imply different property rights. They say:

In a WTA format, the respondents are asked to specify the price for which they will sell their existing rights to an amenity, allowing some form of degradation of the good, or foregoing some improvement. WTA surveys thus assume that respondents already own implicit rights to the amenity at issue. WTP, in contrast, supposes that the respondents do not yet have the right to enjoy an amenity in a certain condition; instead, people are asked the most they would pay to obtain a right to have the amenity improved or to prevent its degradation. WTP questions presume that individuals other than the respondents have the right to the resource in question.\textsuperscript{210}

If I am the rightful owner of the bowl, then the question of how much I am willing to pay to keep it seems out of place. It is my bowl—I should not have to pay anything to keep it! It is, however, not unreasonable to ask me what I would accept to give it up. But if I am not the rightful owner, then the question of how much I am willing to pay to keep it seems much more appropriate.

In order for this argument to pose a problem for CBA, it must at least be paired with the claim that there is no obvious fact of the matter as to how the property rights should be assigned in the cases where CBA is used. Suppose analysts are conducting a contingent


valuation study in an attempt to assess the costs and benefits of (say) tighter regulations on arsenic in drinking water. They are, let us imagine, considering lowering the maximum permissible amount of arsenic from 50 ppb to 10 ppb. In this case, an important question will be whether or not people have a right to water with no more than 10 ppb of arsenic. If they do, then the question might be (roughly): How much money would you be willing to accept in order to continue enduring 50 ppb of arsenic in your water? If they do not have such a right, then the question (again roughly) might be: How much money would you be willing to pay in order that there be only 10 ppb of arsenic in your water? The charge is that if these questions generate different answers, but there is no fact of the matter regarding the assignment of the aforementioned right and thus no fact of the matter about which question should be asked, then there is a problem for economic valuation. A related (and somewhat more paranoid) worry is that which of the two measures is used will depend precisely on what those making the determination want the outcome of the CBA to be. If, for example, a project is one that would diminish people’s access to an amenity, then someone who wants the project to be approved could better ensure that the project’s costs are relatively small by using WTP (viz., “How much would you be willing to pay to keep the amenity?”) than by using WTA (viz., “How much would you be willing to accept to give up the amenity?”).

I believe that unlike the previous criticism, this one is not a criticism of both Claims 2 and 3. It is not a general objection to preference-based approaches to valuation, but rather a more specific objection to WTP and WTA as measures of preferences. I want to stay neutral on whether or not this objection succeeds. I refrain from arguing that it does or does not succeed as an objection to the monetization of all costs and benefits because I (1) do not know whether there are fair ways of establishing who has rights to which amenities (and in
turn to establishing whether, in any particular case, WTP or WTA ought to be used); and (2) believe this is a problem that gets over into that aspect of CBA that I am endeavoring as far as possible to avoid addressing, namely the issue of rights and right action. It is of course no problem to me if this problem is, in the end, a severe threat to the claim that all costs and benefits can be expressed monetarily, for – as will continue to be demonstrated – that claim is one that I myself do not believe. At the very least, it is a problem that provides reasons for expecting from CBA a theory of right action that is more complex and thoughtful than the Kaldor-Hicks Criterion is. If costs and benefits are so sensitive to the initial distribution of rights – and, more basically, the initial distribution of goods – then a theory of right action that aims simply to maximize the good without any appreciable attention to distributional issues is unacceptably crude.

Section 3: Well-Established Problems for Monetizing Costs and Benefits

Even if the preceding discussion does not suffice to show that neither WTP nor WTA succeeds as an accurate measure of preferences, I argue in this section that other considerations do so suffice; and these arguments, if they succeed, demonstrate serious problems with CBA’s commitment to expressing all costs and benefits monetarily. Consider first that on the supposition that WTP is the measure of the distance of a preference from indifference – a supposition that proponents of CBA would patently endorse – we can generate a contradiction. Assume that WTP is a good measure of this. WTP is said to be constrained by ability to pay, so someone who has no ability to pay has no WTP. This means that the person who has no money has no preferences. But the person who has no money clearly does have preferences, including in all likelihood the preference for having some
money! In fact, it seems reasonable to say that it is precisely this person’s preference for not being penniless that motivates her to go out and try to earn money. This gives us the contradiction: this person both has preferences and does not. In turn, we have to reject the assumption that generated the contradiction, namely that WTP is (at least in theory) a good measure of the distance of a preference from indifference. WTA fares better: it does not generate this same contradiction, since someone who has no ability to pay – that is, who has no money to her name – can still have a WTA that is greater than $0, which is to say that this person can still have preferences. This is an important conceptual difference between WTP and WTA, and one I will discuss further. But the general point is that there is a deep problem at least with WTP as a gauge of preferences.

The previous problem is one we can see without moving to interpersonal contexts. Other problems for the project of monetizing preferences arise more clearly when we try to look across people. WTP and WTA are employed precisely in order to handle the problem of making interpersonal comparisons of subjective states. This is always a difficult task, but the fact is that neither WTP nor WTA adequately deals with the problem. If Abe and Bob each sincerely specify a willingness to pay $100 for a good (service, outcome, etc.) but Abe has only $10,000 to his name while Bob is a multimillionaire, then it seems plain that Abe has expressed a stronger preference than Bob. This good is more important to Abe than to Bob. CBA tends not to be sensitive to this point. As CBA is typically taught, understood, and practiced, Abe and Bob have an equally strong preference for the object so long as they are both willing to pay the same raw amount (or so long as the two are willing to accept the same amount). CBA requires us to suppose, implausibly, that an extra dollar means the same to each individual. This ignores a critically important observation made by one of the
pioneers of the concept of economic value, Alfred Marshall. As Hanemann has noted, Marshall “emphasized the proviso of ceteris paribus, including constant tastes, constant income, and constant prices for all other goods, and he readily conceded that, when consumer’s surplus was calculated from an aggregate demand function, the resulting aggregation of values neglects the fact that ‘a pound’s worth of satisfaction to an ordinary poor man is a much greater thing than a pound’s worth of satisfaction to an ordinary rich man.”

We gain insight into problems with counting WTP and WTA as proxies for people’s preferences by noting another significant conceptual difference between the two. As noted previously, insofar as one’s preference for $X$ is identified with one’s WTP for $X$, one’s preference for $X$ will never exceed one’s ability to pay for $X$. If I have only $10 to my name, I cannot have a WTP that is higher than $10. This entails that the wealthier individual’s WTP is capped at a higher point than the less wealthy individual’s. What counts as a benefit and what counts as a cost should be a direct function of what informed, rational people prefer. But for reasons beyond those already mentioned, it is a sham to make each person’s preferences so directly and strictly encumbered by her financial situation. If WTP is used to gauge preferences, then if you have more money than me, I am by definition not capable of having preferences that are as strong as those that you are capable of having. This is terribly counterintuitive. The intuitive idea is not that the poorer person necessarily has weaker preferences, but rather that the poorer person has fewer financial resources available to her for backing up her preferences, or for translating her preferences into spending behavior – as the wealthy person has. We are able to say this because (as was discussed in Chapter 2) we

---

are not working with a strictly behaviorist conception of a preference; and even if we were, there would be no reason to take spending behavior to be the only relevant kind of behavior.

Again, it is worth noting that this counterintuitive result is not generated in quite the same way by WTA. Even if I only have $10 total, I could, e.g., refuse to accept any less than $100 in order to give up a certain good to which I presently have access. This is a fundamental difference between what are supposed to be two different ways of getting at the same underlying preference. However, let us not overstate the difference. Suppose I have only $10 to my name while you have $1,000,000, and WTP were used to gauge our preferences. There is a chance my WTP for a particular outcome could exceed yours; I could be willing to pay, say, $5 for that outcome while you are not willing to pay even a penny. The mere fact that you have a much higher ability to pay does not always mean that you will have the higher WTP. But we should also not understate the difference. Typically, the person with more money will have a higher WTP, and the fact that she has more money will frequently directly explain why this is so. If I have only $10 it is an extremely big deal for me to give up $5 of it. If you have $1,000,000, you will not miss $5. Suppose instead that WTA were used. Here my WTA is not necessarily capped below yours; but chances are good that my WTA for X will be much smaller than your WTA for X. After all, an extra dollar means much more to me than it does to you. WTA still very closely ties the preferences a person will express to her financial situation; and WTP and WTA both tie these together too tightly.

One of the unwelcome effects of ignoring Marshall’s emphasis on the ceteris paribus clause is that CBA, whether it uses WTP or WTA, will be biased against those of lower incomes. Suppose the public investment decision is about where to locate a waste
incinerator. People throughout some particular county, let us imagine, are asked how much they would be willing to pay to have the incinerator located away from their neighborhood. (Alternatively, people could be asked the question that is supposed to generate the same answers as the first question but typically does not, namely how much they would be willing to accept to have it located in their neighborhood). Let us assume that everyone knows the relevant information about incinerators (what would be burned in this one, what the health effects would be, etc.). (Of course if anyone is going to be ill-informed, it is probably going to be the poorer people, who are likely to be less educated, in general and with respect to this particular issue. This is not to insist, of course, that poor neighborhoods and communities cannot or will not ever get organized and very educated about an issue like the health effects of exposure to incinerator emissions. Cases in which poor communities get extremely educated and active regarding an issue like the one at hand here are well documented. This is not an issue worth getting hung up on here. Let us simply suppose here that everyone is well informed.) People in the wealthiest neighborhoods will typically be willing to pay more than those in the poorer neighborhood, often essentially just for the reason that they are able to pay more. And if the question is put in terms of WTA, the minimum amount those in the wealthiest neighborhoods are willing to accept will generally be much higher than the amount minimum amount those in the poorest neighborhoods are willing to accept. Now, if the incinerator is located in order to achieve the optimal balance of benefits over costs, it will end up in the low-income neighborhood. If the thing we were trying (again on the basis of costs and benefits) to locate were a desirable rather than an undesirable thing – a beautiful botanical garden, for example – we would expect it to be located in the wealthier neighborhood. If the decision is about whether to clean up a polluted beach, the benefits will
be much higher, and thus the chances much higher of the project’s benefits exceeding its costs and the project proceeding, when the beach is in a high-income area than when it is in a low-income area. Countless other real-life examples are available.

When these outcomes start to pile up in this manner, project after project and CBA after CBA, reasonable people become concerned. Much ink has been spilled in analysis of the idea of “one person, one vote.” There are certainly a great many reasonable questions regarding what this requires in practice. Is it violated, for example, by the constitutional arrangement that gives Rhode Island, a state with just over one million residents, the same number of U.S. Senators as is given to California, with its more than 35 million residents? It is a complicated question. It seems clear, in any case, that regularly giving more say in the outcome of policy decisions to individuals with more money than to those with less is not consistent with “one person, one vote.” As was discussed back in Chapter 5, it would be naïve to deny that the current political system – especially the part of it pertaining to campaign finance – gives more influence to individuals (firms, groups, etc.) with more money. But this is less pernicious than straightforwardly giving wealthier people more votes than poor people, which is essentially just what CBA does.

Defenders of CBA note that there is nothing that prevents occasional redistribution – via taxation, say – of goods (especially money) to people who are less well off. Some economists, who I can only describe as disingenuous, maintain that “if it were found that some groups were severely and systematically disadvantaged by the application of cost-benefit analyses that ignore distributional concerns, [the economists] would favor redressing
the balance,” but go on to add that “they do not believe this is generally the case.”

It is hard to imagine that anyone really believes this, let alone economists who claim to prize objective analysis. But the other worry with the proposal for occasional redistribution is that when compensation does not take place on a project-by-project basis, the fact that poorer people were burdened with costs through a number of projects tends to get lost. Politicians who understand the way CBA systematically imposes uncompensated costs on the poor have an extremely tough time explaining this effect to the general public. Those selling tax cuts are much more successful; and the result is that the redistribution never takes place.

One might object at this point that the real problem that I am addressing right now relates to another of those issues with which I said I was not going to engage, namely with CBA’s theory of right action, viz., Claim 1, the Kaldor-Hicks Criterion. The problem, the objection runs, is not with what counts as a benefit and what as a cost, but with the idea that policies should be approved when and only when benefits outweigh costs. If, for example, there is a policy that would create a situation in which benefits just slightly outweigh costs but the benefits are all received by the wealthy while the costs all fall on the poor, then a plausible theory of right action will not recommend this policy; and KHC is flawed in that it does recommend this. The problem does indeed relate to KHC, but it is not exclusively a problem with or for KHC. It is also a problem with the claim that all preferences can be expressed monetarily. As I have done a bit so far to suggest, and as I will do more throughout this chapter to demonstrate, not all preferences, and so not all costs and benefits, can be monetized. If this is true, it is going to bear on KHC, as it is going to affect (and significantly complicate) the issue of compensation (be it actual, or merely the possibility of,}

compensation). But the point that I want to make is that even if people are actually compensated by the amount they specify as being willing to pay or willing to accept, there will still be a problem for CBA regarding monetization, namely that their preferences were not really captured by the WTP or WTA specification.

One way to remedy the problem I am pointing out might be somehow to index each person’s WTP (or WTA) to her ability to pay. This could involve, e.g., taking WTP as a percentage of one’s annual income, or of the total value of all one’s assets. Short of limiting CBA to application only in cases where all the people involved really do have equal income, this would be the best way to satisfy Marshall’s demand. It would be a way of standardizing “a pound’s worth of satisfaction,” or the marginal utility of a dollar spent (or a dollar that an individual testifies as being willing to spend) across varying levels of financial wherewithal. There are many people far better at mathematics than I am, but it is not obvious to me how this suggestion would work in practical terms. Suppose the state is considering a project that will affect Persons A, B, C, D, and E. Imagine that A and B want the change, while C, D, and E do not. A has $100,000 total, B has $200,000, C $300,000, D $400,000, and E $500,000. Each, let us suppose, has a WTP – A for C, D, and E against – of $1000. If we just take these specifications as raw numbers – and, moreover, follow the Kaldor-Hicks Criterion – the project should not be approved. Going ahead would result in $2000 worth of benefits, but $3000 worth of costs. How does the math work on the proposal to index WTP to total monetary worth? A’s WTP is $1000/$100,000, or .01 (where the unit would be something like dollar willing to pay per dollar able to pay). B’s is $1000/$200,000, or .005. C’s is .0033, D’s is .0025, and E’s is .002. Thus, going ahead would generate .015 in benefits, and .0078 in costs. Suddenly, the recommendation is that the project should not be approved.
Perhaps this is the result that seems more intuitively correct. But now imagine that there is suddenly a sixth person affected, and (to be somewhat dramatic) this is someone who has only $100 total. Imagine that this person is willing to pay $1 for the project not to go ahead. This gets factored into the calculus as .01, and when added to the .0078, makes for costs that total .0178, which exceeds the benefits. In other words, this sixth person’s $1 is enough to shift the project back to a net gain (which means that if we straightforwardly follow KHC, it will again be approved). That is a worrisome fact.

Even if the mechanics could be tweaked so as to avoid this kind of worry, there will be problems. Consider the incinerator again. Even if those who have more money were not able to unduly influence CBA so that a weighing of costs and benefits dictated that the incinerator gets located in a poorer part of the county, the wealthy would still have greater ability to move homes so as to be far away from the incinerator if one were situated near their (original) home than would the poor. And if the wealthy were made sick by emissions from the incinerator, they would have a greater ability to purchase premium healthcare for themselves than would those who are poorer. Moreover, before making any move to index WTP to ability to pay, it must be recognized that CBA, in systematically favoring the well off, is not guilty of anything of which free market exchange is not also guilty. Indeed, CBA in its standard incarnation tries to mimic outcomes that would have been generated by the market had the market not been plagued by failures that led to inefficient outcomes. And of course the more money one has, the more influence one has through the market over how goods are distributed. So, if this is the aim of CBA, then considering WTP as a percentage of total ability to pay (or anything along these lines) is exactly the wrong move. Those who
wish to index WTP and WTA to ability to pay must then come up with an altogether different justification for CBA.

Those who do not wish so to index WTP and WTA have problems of their own. Rather than understanding the commonality between CBA (in its standard, non-indexed incarnation) and free market exchange as vindicating CBA, it seems reasonable to understand this commonality as (up to a point at least) an indictment of both CBA and unfettered free market exchange. The following argument is well-known but worth repeating: How much money one ends up with is significantly the product of factors that are ultimately out of one’s control. They turn, as a matter of fact, to no small extent on where one is born, the family one is born into, the education one gets at an early age, one’s race, one’s gender, one’s level of attractiveness, one’s “native” intelligence, and so on. But these things should not influence the extent to which one is able to demonstrate, express, or have preferences. They should not affect, e.g., the extent to which being subject to harmful emissions from an incinerator counts as a cost, or the extent to which avoiding subjection to these counts as a benefit. This must be part of the rebuttal of any attempt to defend CBA against that charge that it inappropriately ties preferences too tightly to ability to pay. In turn, it is part of the more general argument against the claim that all costs and benefits can be expressed monetarily. But it is not the entire response. The argument I will press in the next chapter is meant to further undermine that claim.
CHAPTER 9: WHY NOT ALL COSTS AND BENEFITS CAN BE MONETIZED (PART 2)

Section 1: Introduction

The problems discussed in the previous chapter are weighty, but the one that most interests me is not among them. There is an additional problem for the project of monetizing all cost and benefits that persists in both interpersonal and intrapersonal contexts, and it is one that I believe has not yet been adequately diagnosed. It is the diagnosing (and to a lesser extent, the treating) of this problem to which this chapter is devoted. This problem is a much more subtle one than those discussed previously, but its subtlety does not make it any less genuine or significant. Economists typically argue that substitutability implies commensurability. First supposing this is true, I explain what exactly the claim regarding the substitutability of goods would have to be in order for it the claim that all costs and benefits can be expressed monetarily to even have a chance at being true. I then suggest that this claim about the substitutability of goods is plausible; this involves making the concession that the non-market goods that economists are interested in pricing for the sake of public investment and regulatory decisions are indeed things that people will substitute for money, at least if right amount of the good and the right amount of money are involved.

Nonetheless, I then argue that even though this claim about the substitutability of goods is plausible, it does not suffice to establish the value of any good can be expressed monetarily. The pithy idea is: substitutability does not imply commensurability. More
specifically, I maintain that an individual’s willingness to pay or to accept a certain amount of money for a good does not suffice to show that this good’s value can be monetized. Such a willingness demonstrates a sense in which this individual prefers some particular amount of money (and all amounts above that, presumably) to the good, but this is a fairly superficial sense. The important sense of preferring has to do with an individual’s evaluative attitudes regarding various goods. But each individual is capable of evaluative attitudes that differ from one another along multiple irreducible dimensions. Different goods or outcomes will sometimes engage fundamentally distinct attitudes such as admiration, fascination, respect, agitation, disgust, and shame. When this happens, preferences will not generate a single ordering of value: there will be pairs of goods of which it can neither be said that one member is simply more valuable than the other nor that the members are equal in value. I suggest that this is no threat to CBA’s commitment to tying costs and benefits to people’s (informed and rational) preferences, and also no threat to the idea of rational or reasonable choice, though it is a threat to CBA’s commitment to expressing benefits and costs in monetary terms, and would in turn seem to be a threat to its ability simply to promote the largest possible benefit once costs are subtracted.

**Section 2: The Substitutability of Money for Other Goods**

Proponents of CBA typically say that the question of whether costs and benefits – i.e., whether preferences, and in turn the value of goods, services, outcomes, etc. – can be monetized is an empirical question. The economists’ argument goes roughly as follows. In order to see whether a good $Y$ has a monetary value (and if so, what that monetary value is), we need simply to look at what each (relevant) person’s preferences are regarding money and
the good. Questions about a person’s preferences regarding money and the good are really questions about whether money and the good are, for that individual, substitutable. So, if we want to know whether $Y$ has a monetary value, relative to some person, we need to see whether, for that person, money is substitutable for $Y$. Simply put, money is said to be substitutable, relative to some person $P$, for $Y$ when there is an amount of money $\$X$ that $P$ will substitute, i.e., will trade in or out, for $Y$. If there is no amount of money that $P$ will substitute for $Y$, then money is not substitutable for $Y$. Because what a person will (or will not) substitute for what is essentially an empirical question, then whether money and $Y$ are substitutable is an empirical question. If there is some $\$X$ that $P$ will substitute for $Y$, then (according to this argument by economists) $Y$ has a monetary value to $P$. The gauge for substitutability, and in turn for value, is $P$’s WTP or WTA. If $Y$ is a good thing, then its value is identical to $P$’s highest WTP in order to get or to avoid giving up $Y$, or $P$’s lowest WTA in order not to get or to give up $Y$. If $Y$ is a bad thing, then its value to $P$ is identical to $P$’s highest WTP in order to avoid getting or having to keep $Y$, or $P$’s lowest WTA in order to get or to keep $Y$, or where $P$. Thus, it is clear – the argument goes – that the question of whether $Y$’s value can be expressed monetarily is ultimately an empirical one. This is a specific application of a more general normative principle the economist employs, which we can call, following precedent, the substitution principle.\(^\text{213}\) According to this principle, a good’s value to an individual is identical to, or can be expressed in terms of, the goods that the individual will substitute for it.\(^\text{214}\) It could then be said that a good’s total value is equal to the aggregate of goods that people would substitute for it. A good’s total economic value

---

\(^\text{213}\) I do not claim that this is a term that every economist will recognize, but it is straightforwardly used by Freeman, so it is not a term that I have simply made up.

\(^\text{214}\) This is almost exactly the formulation that Freeman gives in “The Ethical Basis of the Economic View of the Environment.”
is the aggregate amount of money that people would substitute for it. I call this general argument the argument from substitutability to commensurability.

There are two initial problems with this argument that should be mentioned but that can be relatively easily resolved. The first problem is that the substitution principle, as worded above, makes no mention whatsoever of the extent to which the person faced with the decision regarding substitution is possessed of complete information and/or is rational. As I was at pains to explain in the earlier parts of this project, this exclusion is highly problematic. That these requirements are not included in the substitution principle helps vindicate the claim (on which I pressed for so long) that proponents of CBA are not sufficiently concerned with ensuring that preferences are well-informed and rational before these preferences are measured and factored into the calculus of benefits and costs. If the substitution principle is to have any shot at succeeding, it must at minimum be modified so as to say: a thing’s value is identical to the goods that well-informed and rational people will substitute for it. (Again, only relevant information is required, and rational in this context just means consistent.)

The second problem is that the substitution principle does not clearly demand that the substitution be a *reasonable* one. What does this mean? Suppose that I am put in a position of having to choose between some good $Y$ and some amount of money $X$. The economist will say that the smallest amount of money that I would substitute for $Y$ is the monetary value of $Y$. But suppose that I am utterly incapable of processing the tradeoff and so resort to simply pulling a number out of a hat, where I say ahead of time that the number I pull out is the $X$ that I will trade for $Y$. Or, suppose someone suggests a dollar value and I simply flip a coin in order to decide whether to make the substitution at that dollar value. This is not what
the economist has in mind when she talks about substitutability. Pulling numbers out of a hat or flipping a coin may make the decision of which $X$ to substitute for $Y$. However, that a random method (like pulling numbers out of a hat, or flipping a coin) can resolve the choice between $X$ and $Y$ is not going to show anything about the value to me of good $Y$. The substitution principle is more compelling – though, I will argue, still ultimately problematic – when the substitution is made in light of normative reasons. Henceforth, when I talk about the substitution principle, it is this kind of substitution I will have this in mind. When I want to make explicit mention of this issue, I will call it “reasonable substitution,” in the sense that it is substitution made for reasons, not just one made arbitrarily, i.e., by some random procedure like a coin flip of a hat pull. (I would call it “rational” substitution, but this would lead to this requirement being confused with the requirement that the individual making the decision about the substitution must be rational in the sense of having a consistent set of preferences.) Importantly, these reasons do not have to be ones that anyone (other than the person whose reasons they are, of course) counts or would count as “good reasons.”

For two reasons, it would be easy to misstate the extent of this second problem. First, suppose, for example, that $P$ is shopping for a new home and cannot decide whether to pay the $300,000 asking price or to insist on going no higher than $290,000; $P$ might say: “I am agonizing over this decision, but my deliberation has not managed to push me one way or the other here. To help me conquer this indecision, I am going to flip a coin; if the coin lands ‘heads up,’ I will pay the $300,000, and if it lands ‘tails up,’ I will stick to $290,000.” In this case, the mere fact that a coin was flipped may not suffice to make trouble for the substitution principle. $P$, after all, appealed to a coin flip for reasons (and reasons, moreover, that are quite easy to understand). The important point is simply that the substitution
principle is not plausible where the substitution is made on a purely arbitrary or random basis, not mediated at all by the agent’s reasons. Second, it may be that where \( P \) makes the substitution on a purely arbitrary basis, it cannot truly be said that the substitution is one that \( P \) prefers. \( P \)’s preferring one thing to another is obviously critical to the substitution principle, so if the arbitrary substitution does not involve a preference, then it is the fact that no preference is involved, not the fact that it the substitution is arbitrary, that is genuinely the hang up. Even if this is true, it does not show that the point that the substitution must not be arbitrary was somehow too strong and needs to be withdrawn, but rather simply shows that the point could have been made in a different (and perhaps simpler, more direct) way.

We move forward with a version of the substitution principle that takes account of these two problems. Suppose now that person \( P \) is, in accordance with the arguments defended earlier in this project, possessed of complete information (in the relevant respects) and is not plagued by lapses in her instrumental reasoning. The first question to consider in connection with the economists’ argument is whether there is any good \( Y \) and any \( P \) such that \( P \) will not substitute \( Y \) for money, no matter how great the amount of money. This is indeed, as the economist has maintained, an empirical question. When someone is confronted with a choice between \( X \) and \( Y \) – no matter what \( X \) and \( Y \) are, and whether or not one of them represents some amount of money – which of these the individual will choose is an empirical question. In the economists’ eyes, problems arise for CBA when there is a good that an individual will not substitute for money. This good, the economist will be forced to say, is

\[ \text{This second amendment to the substitution principle may, in a certain sense, be much ado about nothing. It may be that where } P \text{ makes the substitution on a purely arbitrary basis, it cannot truly be said that the substitution is one that } P \text{ prefers. } P \text{'s preferring one thing to another is obviously critical to the substitution principle, so if the arbitrary substitution does not involve a preference, then it is the fact that no preference is involved, not the fact that it the substitution is arbitrary, that is genuinely the hang up. Even if this is true, it does not show that the point that the substitution must not be arbitrary was somehow too strong and needs to be withdrawn, but rather simply shows that the point could have been made in a different (and perhaps simpler, more direct) way.} \]
one that does not have a monetary value. It is not capable of being assigned an accurate price. If such a good would be produced or destroyed by a project or regulation being subjected to CBA, then the economist will say we cannot compare this effect with those costs and benefits that could be monetized, and this destroys the applicability to the Kaldor-Hicks Criterion (and the Pareto Criterion, for that matter). For this reason, proponents of CBA hope that there is no good that is not substitutable for money.

It bears noting that some economists say, mistakenly, that the important empirical claim that needs to be true in order for CBA to function properly is that all goods are substitutable. To say that all goods are substitutable, relative to some person $P$, is to say that for any possible good $Y$ that exists, there is at least some other good or combination of goods that $P$ would trade for $Y$ (or at least for a marginal unit of $Y$, or for some chance of losing $Y$). It is to say, in other words, that there is no good that $P$ would not give up even a little bit of, no matter what other combination of goods was offered in exchange. In fact, CBA needs something different from this claim. CBA needs the claim that all goods are substitutable for money, but this can be false even if it is the case that all goods are substitutable, i.e., even if there is no good that would not be traded for some other package of goods. There is a tendency to assume that if, relative to some particular individual, all goods are substitutable, then for any good, there is an amount of money for which that person would give up that good. This in fact does not follow directly from the substitutability of all goods. The substitutability of all goods (as I have specified that notion) does not entail the substitutability of any good for money, though the substitutability of any good for money does entail that all goods are substitutable. Consider that the substitutability of all goods does not mean that for any good $Y$, there is some amount of ice cream for which an
individual will trade $Y$. Ice cream has diminishing marginal utility. The more of it one has, the less desirable the next unit of ice cream will be. As discussed earlier (Ch. 2, Sec. 5), I am an enormous Bob Dylan fan. There is no amount of ice cream for which I would entirely give up my ability to listen to Bob Dylan, though there is some package of goods that consists of more than just ice cream for which I will give up this ability. Similarly, saying that all goods are substitutable (for me) does not mean that it is the case that for any good $Y$, there is some amount of *money* for which I will trade $Y$. It is possible that there is a good that I would trade for some particular *package* of goods, but where this package of goods must contain something more than just money. After all, even money has diminishing marginal utility.

While this is possible in theory, I believe that with respect to most people and most goods, money alone, at least where enough of it is offered, *will* be enough to compel an individual to trade the good for the money. Because substitutability is an empirical matter, and I am from the armchair in a position to know just about my own tendencies and behaviors regarding substitution, I will continue to talk a bit about myself. It is an empirical fact about me, about my psychology, about my well-informed, rational preferences, that I would be willing to give up a great number of the things that I care about if I were offered enough money to do so. Hard as it may be for my friends to believe (and disconcerting as it in fact is for me to imagine), if I were offered enough money I would give up all of my Dylan recordings and agree to never listen to Dylan again. While money does indeed have decreasing marginal utility, the fact that it can be stored away and traded for so many things makes it such that its marginal utility diminishes much less quickly than does that of something like ice cream.
Goods of all kinds could be discussed here, but because my primary interest in this context is non-market goods, it is these on which I will focus. I certainly count myself as an environmentalist. I am very fond of the beautiful old trees that grow around my house. For enough money, however, I would cut them down to be turned into timber. It would need to be a significant amount of money, not a trivial amount, in order for me to do this. But for the right price, I would do it. I also really enjoy seeing sunsets. However, for enough money, I would be willing to give up my ability to ever see another one. I have never climbed, or ever even visited the Grand Teton; I badly want to do so at some point in the future, but for the right amount of money, I would consent never to do so. For enough money I would agree never to go to the Tetons, and for the right amount of money I personally would even consent to having the Tetons irrevocably altered, or even destroyed (supposing that it were somehow up to me). I like for my water to be free of harmful chemicals, but for the right amount of money, I would agree to drink water with slightly higher arsenic content than the water I drink now has. I love that the Alaskan National Wildlife Refuge exists in its relatively pristine condition, but for the right amount of money I would consent to its being developed. Money, for me, is (in the important marginal sense) substitutable with all of the goods I have just mentioned. This is true for a great, great many goods, including many of those for which I care a tremendous amount. And I do not suspect that I am terribly unique in this regard. My guess is that most others who do some genuine soul-searching would admit the same thing regarding many of the things about which they care a great deal.

Nonetheless, I do not insist that this is true of all people, or that it is true of all goods. It is clearly worth noting that there is actually empirical evidence that for certain goods and certain people \( P \), if \( P \) is offered the choice between the good and the money, \( P \) will take the
good, no matter how much money is being offered as the alternative. Some examples are trivial. People are unwilling to choose even great amounts of money over, e.g., having air to breathe. There is obviously no sense whatsoever in choosing the money instead of air, as doing so (and thus being without air) would quickly lead to death. I have other, more interesting, cases in mind. If you asked me what amount of money I would accept in order to hand my wife, my parents, or my brothers over to killers, I would tell you that there is no amount of money great enough. And you would be wrong to doubt my sincerity. It is, of course, an empirical matter; you could trust me or you could try me. For obvious reasons, I hope not to be tested here. Still, I firmly maintain that there is no amount of money great enough. There is also no amount of money that I would accept to subject them to torture, or to sell them into slavery.

Economists are surely aware of empirical findings of this sort. Such findings should not be considered to be altogether unimportant, but it would be a mistake to make much of them. After all, when we talk about the substitutability of one good for another, we are talking about marginal units of the goods. Thus, what proponents of CBA are really hoping is that there is no good that affected individuals would not give up a little of, or would not accept even some chance of losing, in exchange for money, no matter how much money is involved. In the above case, it is possible to make the “unit of loss” smaller. Imagine that instead of being asked what amount of money I would accept to subject my wife, parents, and/or brothers to certain death, I am asked whether there is an amount of money for which I would consent to a slight increase in their exposure to a risk of death. It would be disingenuous of me to say here that there is no amount of money great enough. There is some amount of money that is great enough. If I were being sincere, I would specify some
finite amount; and my behavior also occasionally demonstrates this. When my wife and I shopped for a new car for her to drive, we considered how safe various models are and how expensive each is. When I cast my vote (as it were) for a car that is both less safe and less expensive than another that is available, I was doing just this; that is, I was effectively accepting (or keeping) some amount of money in exchange for subjecting her to a slightly greater chance of death. I was specifying what amount of money – relative to the total amount of money we have, and with an eye towards all our other expenses – I am willing to pay to ensure her safety on the road. If I had been asked whether there’s an amount of money which I would accept to make her car a bit less safe, my sincere answer would be yes. And if you asked me whether there is some amount of money great enough for me to choose it, not over simply having air to breathe, but rather in exchange for having my air dirtied by a small amount, my sincere answer would be yes.

It must be remembered, though, that I do not currently have much money in general. There are of course some extremely wealthy people in the world. For these people, a little – or even a lot – more money is not going to make a difference in their life. The marginal utility of an additional dollar to them is extraordinarily small. If I had $1,000,000,000 dollars in my bank account, there would be likely be no amount of money that I would accept in exchange for the loss of sunsets, or the Tetons, or other such environmental goods. In fact, if I had that much money, there is likely no amount of additional money for which I would trade away my ability to listen to Dylan.216 Perhaps there is no amount sufficient to get me to agree to some marginal diminishment to these things—a bit of harm to them, or some risk of

---

216 I say “likely” because there is one important complication. Even if I had $1,000,000,000, were I convinced that I would get an enormous sum of money for giving up sunsets, or the Tetons, or Dylan’s music, I might do so on the grounds that I could give this money to other who need it. But then, why would I not have given away money for this same purpose such that I would not now have less than $1,000,000,000?
the loss of or damage to these things. This point is not trivial, but I will not try to wrestle much out of it. Instead, I simply want to concede to the economist that for the large majority of real-world policy proposals, each person who would be affected by the policy is such that there is an amount of money for which he or she would, if rational and well-informed, be willing (reasonably) to trade the goods that would be affected were the policy enacted (again, at least if it is a small enough unit of the good). Henceforth, these are the people and goods about which I will be talking.

Section 3: Substitution, Preferences, and Commensurability

Though the extent of substitutability is in fact a contentious empirical question, I have suggested that wide concessions be made regarding substitutability. The previous section conceded, in effect, that with respect to each person, we can order goods according to what the individual would substitute for what. If this person is rational, then there will be a single ordering, which is to say an ordering that is (essentially) complete and according to which there are no intransitivities in the ordering. For each individual, and for each pair of goods $X$ and $Y$, there is some amount of $X$ great enough and some amount of $Y$ small enough to compel the individual to substitute $X$ for $Y$ (to give up $Y$ in order to get $X$, etc.). Moreover, there will be no cases in which good $X$ is ordered ahead of $Y$ and $Y$ ahead of some other good $Z$ but where $Z$ is ahead of $X$. The claim that any good is substitutable for money – a claim that can also essentially be conceded – is in effect the claim that, for each individual, each good would be traded for at least some amount of money, and that if the individual calls good $X$ better than good $Y$ and good $Y$ better than good $Z$, then $Y$ will have a higher price than $Z$, and $X$ will have a higher price than both $Y$ and $Z$. 
The most general question now is: what are the implications of this concession for public policy? This is a difficult question. More specifically and immediately, we should ask: Does this concession entail that all gains and losses can be assessed in terms of a single measure of value? More specifically still, does it entail that all gains and losses can be fully captured in monetary terms? I maintain that the answer to latter two question is no, and for reasons that go beyond those articulated in the previous chapter. I have extensively defended a connection between (on the one hand) informed and rational individual preferences and (on the other hand) policies’ costs and benefits and goods’ values. There is a sense in which the aforementioned ordering of goods (which includes the corresponding ordering of monetary amounts) is an ordering according to the preferences of this individual. But this does not mean that all goods can be valued on one single scale. This is because there is another sense of preference according to which these same goods may, for this same person, not be capable of a complete and fully-transitive ordering, and it is this later sense that corresponds to the goods’ true value. This is a sense according to which preferring is not just about substituting, and so though value has to do with preferences, value commensurability is not just about substitutability. Anyone who wants to know about a good’s true value, and the relationship between the value of one good and another, should inquire into more than simply what the individual would trade for what. Suppose that, for non-arbitrary reasons, \( P \) chooses \( X \) over good \( Y \)—perhaps the situation is one in which \( P \) is asked whether she would be willing to pay \( X \) in order to get good \( Y \), and her sincere answer is no. Does this mean that \( P \) prefers \( X \) to good \( Y \)? Does it mean that \( X \) is more valuable to \( P \) than good \( Y \)? Does it mean that the value, to \( P \), of good \( Y \) can be expressed in monetary terms? There is an obvious sense in which \( P \)’s unwillingness shows that she prefers that amount of money to the good. Choosing
$X$ over good $Y$ would be an odd thing for $P$ to do if she in no sense preferred $X$ to $Y$. (It would be an especially odd thing to do if she categorically preferred $Y$ to $X$.) But the case is not thereby necessarily an open-and-closed one. There is a complication in connection with the question of what really is involved in preferring. The aforementioned sense of preferring is not an illegitimate one; but it is an emaciated one.

I have pointed out from the beginning of this project that there is a behaviorist sense of preference that economists have favored—though that it need not involve any actual behavior, but counts even where there is just a willingness to behave in a certain way is an odd kind of behaviorism. So, on essentially no account of preference is one’s actually choosing $X$ over $Y$ (where $X$ could be, but could also not be, a monetary sum) necessary in order for one to count as preferring $X$ to $Y$. In general, being disposed to choose $X$ over $Y$ is more plausibly necessary. There is, however, a sense of preference according to which neither choosing nor being disposed to choose $X$ over $Y$ is sufficient for preferring $X$ to $Y$. A deeper, less behaviorist, more nuanced conception of preference must be considered. My own (perhaps idiosyncratic) view is that a preference is a psychological state that has motivational force of the sort that desires are taken to have, but that also consists, in part, of certain beliefs. This is a more-cognitivist-than-usual conception of a preference; and I adhere to it because it seems to me most faithful to psychological phenomenology. Importantly, though, nothing I will go on to say about preferences depends on this cognitivism. What is critical for my story is that the preference for $X$ over $Y$ does not consist simply in the choosing, or in the disposition to choose, $X$ rather than $Y$. Preferring $X$ to $Y$ involves having particular attitudes regarding $X$ and $Y$. These attitudes or impressions often compel or lead one to choose $X$ rather than $Y$, i.e., to prefer $X$ to $Y$ in the behaviorist sense. But to say one
prefers $X$ to $Y$ is not just to comment on the choiceworthiness to this person of $X$ over $Y$. Notice that this is not to deny that preference is a relational concept. A preference is always for a thing or things over or instead of another or others. But again, preferring does not have to do simply with choice.

When all that is focused on is what one chooses (or would be willing to choose) then the opportunity to note and appreciate complexities in our psychology – complexities that matter vis-à-vis the issue of the value of goods, and indeed matter for the question of whether value is commensurable – is missed. Questions about a thing’s value (relative to some individual) are questions about this person’s well-informed and rational preferences, but then these are really questions about how this person appraises the good in question, i.e., about the evaluative attitudes the individual brings to bear on the good. To prefer $X$ to $Y$ is roughly to make a more favorable evaluation or have a more favorable impression of $X$ than $Y$. But things get very interesting here. It is a fact about human psychology that we are capable of a wide range of impressions or evaluative attitudes. They range widely in at least two respects. The first of these is very commonly recognized: many attitudes are such that they can be manifested to different degrees, which is to say they range in strength, going from indifferent to extremely negative or from indifferent to extremely positive, and covering all points in between. The second respect in which these attitudes range widely is less commonly appreciated, and the one that has implications on which I want to focus.

There are many different kinds of such attitudes, which is to say that not all differences between the attitudes are differences in degree/strength. We admire, adore, appreciate, desire, enjoy, love, respect, revere. We are amazed, astonished, amused, aroused, attracted, awed, charmed, entranced, grateful, inspired, interested, joyous, pleased. We find
things agreeable, amorous, appropriate, beautiful, breathtaking, calming, delectable, delightful, entertaining, enticing, exhilarating, fascinating, hilarious, liberating, wonderful. We abhor, admonish, dread, fear, lament, rue, worship. We are agitated, angered, annoyed, ashamed, disgusted, embarrassed, indignant, irritated, offended, saddened, shocked. We find things abysmal, agitating, asinine, atrocious, awful, boring, burdensome, cavalier, cloying, disconcerting, despicable, dreary, foolish, frightening, ghastly, heart-rending. I could go on and on. We have this large variety of evaluative terms for a reason, and the reason is not just that they make our language more colorful. The evaluative terms that are somehow positive are not all ways of stating the same fundamental favorable assessment; they are not all synonyms, and do not differ from one another just by being positive to a greater or lesser degree. The pejorative terms are not all ways of stating the same unfavorable assessment. It is not the case that the various terms differ just in the respect that some are stronger and some are weaker, or some more and some less intense, than others. They differ from one another along more than one single dimension. They do not fit onto one evaluative scale. The attitudes mentioned do not reduce to one single kind of attitude, and it is not as though only one of the kinds of evaluative attitudes matters morally.

What do all of these various attitudes have to do with our preferences? It is not quite right to say that they are themselves preferences; but they underpin, or are partly constitutive of, preferences. Again, in the sense I have in mind, to prefer X to Y is to evaluate X more favorably than Y. Because evaluative attitudes are inescapably and irreducibly diverse, then it is possible for X to be evaluated in fundamentally different terms than Y, i.e., for X to be evaluated against irreducibly different attitudes than those against which Y is evaluated. If this happens, then it is strictly mistaken to say that X is evaluated more favorably than Y,
mistaken to say that $Y$ is evaluated more favorably than $X$, and mistaken to say that $X$ and $Y$ are deemed equally favorable (or unfavorable). This means that when $X$ and $Y$ are evaluated (by $P$) in irreducibly different terms, there is thus a sense of the notion of preference according to which, relative to $P$, $X$ is not strictly preferred to $Y$, $Y$ is not strictly preferred to $X$, and $P$ is not indifferent between $X$ and $Y$. James Griffin maintains that WTP is a measure of “the distance of a desire from indifference.”\textsuperscript{217} I think many economists would disagree with this formulation. Eschewing talk of desire for talk of preference, they would say instead that WTP is a measure of the distance of a preference from indifference. But the fact that we are capable of so many evaluative attitudes and that these differ from one another, and from indifference, not along one single dimension, but rather along multiple irreducible dimensions means in turn that preferences differ from one another along multiple dimensions, such that differences between them can extend beyond their strength. At least some also differ from one another in \textit{kind}. The multitude of evaluative attitudes means, in effect, that we have a multitude of ways of preferring.

Consider, for example, the following three particular evaluative attitudes: loving, revering, and respecting. These are three kinds of (essentially) favorable impressions or evaluative attitudes; but they are not identical, and furthermore, they do not differ from one another only in terms of strength. I might both love and revere something, but if I love $X$ without revering it and revere $Y$ without loving it, then the difference between these things is, in some important sense related to their value to me, one of kind. Other examples are readily available. My feeling (say) of shame regarding $X$ and my feeling of fascination regarding $Y$ do not differ from indifference along some one dimension. In appraising $X$ and $Y$ in these

terms, I am not assessing $X$ and $Y$ on the same scale. These evaluative terms are not
antonyms. They are fundamentally different kinds of assessments, but both help constitute
my preferences. These are just examples. I could have focused on agitation and
embarrassment, on awe, amusement, and attraction, and so on. And we can, and often do,
combine kinds of attitudes even in our appraisal of a single good: I find seeing a bald eagle in
flight to be exhilarating and inspiring. Spending time with my friends is generally not
exhilarating, and often not inspiring either, but it is entertaining and relaxing, and I am often
amused by them and often find them to be interesting. Money, a piece of art, a high-end
stereo, a pristine landscape, a designer suit, an ivory-billed woodpecker (discovered after
having been thought for 60 years to be extinct), a family member, an artifact of an ancient
culture—this diverse set of goods makes an inescapably diverse set of claims on me.

Particularly noteworthy for the purposes of this project is the fact that the kinds of
evaluative attitudes many of us have regarding money tend to differ fairly sharply from the
kinds of attitudes that tend to be engaged by environmental goods. For many
environmentalists, the fact that certain environmental goods are not private goods is not just
another fact about these goods. It is a critical, defining fact. That they are not so easily
excludable (in the economist’s sense), that rights cannot easily be assigned to them, that they
are not rivalrous in the same way or to the same degree that many other things (like a bag of
potato chips in the vending machine) are, and that we must all work together to take care of
them—these are all parts of what makes environmental goods so special. Many of the
environmental goods that are most special to people are ones that were not made by humans,
were not made quickly, and could not be replaced if they were destroyed. Frankly, though,
there are hundreds and hundreds of naturalists, storytellers, poets, and artists who are more
capable than I am of articulating why environmental goods – individual animals and plants, trees and forests, species, meadows, rivers and streams, canyons and gorges, mountains and boulder fields, ecosystems, with the history behind each of these things and the various ways in and extents to which they are all connected to each other – are so special to so many people, why they are so powerful in our imaginations, why our experiences with these things move so many of us, and in a way different from how, for example, medium-sized consumer dry goods do. Also noteworthy for our purposes are human health and life. Our own health and life, as well as the health and life of others – others we care about, as well (at least occasionally) as people we do not know at all – elicit attitudes not generally otherwise elicited. Those who would equate money with evil are wrong to do so. Certainly, some people are led to do some awful things for money, and there are many uses of money that I would not hesitate to call frivolous and even shameful. But money can also achieve some wonderful things, for ourselves, our family and friends, for strangers, and even for the environment. Nonetheless, for whatever reason, it tends to fail to engage in us – many of us, at least – the same kinds of attitudes that are engaged by human health and life and by environmental goods of all sorts.

This all means that money has – again, for whatever reason, and for many though perhaps not all of us – a different kind of value from that had by many non-market goods. Though it can be conceded from the fact that some informed and rational person \( P \) is (reasonably) unwilling to pay \( $X \) for \( Y \) that there is a sense in which \( P \) prefers \( $X \) to \( Y \), this does not tell us all there is to know about the value of \( Y \) relative to \( $X \). It does not tell us whether \( Y \)’s value can genuinely be expressed monetarily. The fact that \( P \) reasonably chooses \( $X \) rather than good \( Y \) does not entail that, in the deeper sense I have worked to
articulate, this person prefers $X$ to $Y$. $X$ and $Y$ are appraised in very different terms. Their respective values are very different. Specifying $Y$’s true value to $P$ is different from simply specifying how much money $P$ would be willing to trade for $Y$. Again, there are some pairs of goods $X$ and $Y$ of which it can neither be said, with respect to $P$, that $P$ prefers – in the deeper sense of prefers – one member of the pair to the other, nor that $P$ is simply indifferent between the members. In this case, $X$ is not more valuable than $Y$, $X$ is not less valuable than $Y$, and $X$ and $Y$ are not equally valuable. Where $X$ is a monetary amount, then $Y$ simply has no genuine monetary equivalent. There is an amount of money that will be traded for $Y$, i.e., for which $Y$ is substitutable. But if money and good $Y$ engage irreducibly diverse attitudes, then $Y$’s value cannot be genuinely be monetized. This will at least sometimes be true of those preferences we have for money versus certain non-market goods. Of course for at least some goods $Y$ where $P$ reasonably chooses $X$ rather than $Y$, it will be true that $Y$’s value can be expressed monetarily. It would be ridiculous to claim that no one who reasonably chooses $X$ over $Y$ ever evaluates $X$ and $Y$ in the same basic terms; $P$ could bring the very same evaluative attitudes to bear on $X$ that she does on $Y$. My claims are simply that this is not conceptually necessary, and that it is not always true in actuality.

The point could be put in another way. The notion of indifference is typically taken to be an extremely important one in non-market valuation. The standard economic line is that, relative to $P$, a good’s value (or the extent to which a policy-induced change counts as a cost or a benefit) can genuinely be expressed in terms of the other goods that would make $P$ indifferent between the former good/change and the latter goods. A good’s monetary value, then, is the amount of money that would make $P$ indifferent between (on the one hand) the good or change and (on the other hand) the amount of money. I believe that when thinking
about whether the value of one good can be expressed in terms of another good or goods, indifference is indeed the critical issue. However, it involves something different, and is in fact harder to come by, than economists tend to believe. One is indifferent when her preferences fit a certain description: it might be said that indifference is the point at which one does not prefer X to Y or Y to X, but rather prefers X and Y equally. If preferring is interpreted in shallow, purely behaviorist terms, P is indifferent between X and Y just in case half the time she has a choice between X and Y she will choose X and half the time she will choose Y. (Or perhaps the behaviorist line is that P is indifferent when she is unable – by any means other than something like a coin flip, perhaps – to choose between X and Y.) In fact, genuine indifference requires more than this—and it does so because preferring involves more than the behaviorists acknowledge. I might be indifferent between two goods X and Y in either of those aforementioned purely behaviorist senses of indifference, but if I (say) love but do not revere X, and revere but do not love Y, then regardless of how much I love X and how much I revere Y, I will never be genuinely indifferent between them. Similarly, I might be indifferent between Y and some amount of money $X in either of those purely behaviorist senses of indifference, but if $X is engaging attitudes that are fundamentally, irreducibly distinct from those Y engages, then I am not genuinely indifferent. In this case, this point of behavioral indifference is not a point at which the two goods are genuinely equal in value. Y’s value cannot, in this case, be genuinely expressed monetarily. There are, in this case, incommensurable values.

Cass Sunstein seems to appreciate the point, though he puts more of an emphasis on types of relationships and ways of valuing and less on types of evaluative attitudes and ways of appraising or preferring than I do. He says:
Suppose that Jones has arranged to have lunch with a friend today, but that he has become very busy and would like to cancel. Suppose Jones thinks in this way: If I cancel, my friend will be disappointed, because he would like my company, and also a bit insulted, because it is cavalier and disrespectful to cancel lunch at the last minute. Maybe I should make it up to him, or provide compensation, by offering a nontrivial cash payment. This would clearly be a hopelessly inadequate response, even if the payment is very high. A cash payment would be inconsistent with the way that someone values a friend.\textsuperscript{218}

A second example regards sex, and more specifically sex for money. What explains the position that it is inappropriate for sex to be traded on the market? Sunstein says, and I agree, that the value of sex and the value of money differ in kind. At least for those who think that sex should not be bought and sold, the explanation of why this is so seems to be that money and sex engage different kinds of preferences. “The claim is not that markets value sexuality too much or too little; it is that markets value these activities in the wrong way,” he maintains. Another of Sunstein’s example is even more germane to this project. He says:

> Suppose that we feel awe toward something. If we do, we will not believe that it should be valued in the same way as its cash equivalent. The judgment that a mountain is “really worth” $10 million may well be right in an important sense, but it is inconsistent with the way that most people value the mountain.\textsuperscript{219}

In all three of these examples, I understand Sunstein’s notion of a \textit{way of valuing} to be identical to what I have in mind in talking of a \textit{way of preferring}. I am not even sure what the important sense is in which it is right to judge the mountain to be “really worth” $10 million (or whatever amount people are willing to trade for it). I would say if the sum total amount that people are willing to trade for the mountain is $10 million, then there is a sense


\textsuperscript{219} Ibid.
in which it is right to judge the mountain to be worth $10 million, but in no sense at all is this what it is “really worth.” There is no such amount of money. This difference between what Sunstein and I are willing to say is subtle but important. If someone is awed by the mountain but not by money, then it is infelicitous to say that the mountain has a “cash equivalent.”

Stepping back for a moment, the challenge of any argument for incommensurability is to avoid letting the claim that not all values of are the same kind spiral into the claim that no two values are of the same kind. That is, one wants an argument for incommensurability that is not an argument for radical incommensurability, whereby there are an overwhelming number of kinds of value. The argument I have worked to spell out in this chapter risks establishing – if it establishes anything – a fairly radical form of incommensurability. It seems to me that there are many irreducibly distinct evaluative attitudes. But if each such attitude corresponds, or gives rise, to an irreducibly distinct kind of value, then there do in fact seem to be many irreducibly distinct kinds of values. This is an issue that bears further thought, and alas more attention than I can devote here. One possibility is somehow to group evaluative attitudes so that it is only each different group, not each different attitude, that corresponds to different kinds of value. More fundamentally, one important question this generates is: What is the precise relationship between evaluative attitudes and kinds of value? Suppose that $P$ takes $X$ to be disgusting but $Y$ to be frightening. Make no mistake: disgust and fright are different kinds of reactions. But what exactly is the different kind of value (or in this case, perhaps disvalue) that corresponds to each of these two attitudes? This is a difficult question. In any case, irreducibly distinct attitudes are required for the qualitative differences in value that I have in mind. For example, suppose I am trying to decide which of two houses to buy: one is more conveniently located, but the other more spacious; one is
less expensive, but the other more attractive. These can legitimately be called different dimensions of value (in a house). And yet I do not seem to have a distinctive attitude toward the houses associated with the dimension of convenient location, and another associated with the dimension of spaciousness. It is in fact for precisely this reason that, though calling these different dimensions of value may not be illicit, they do not correspond to differences in kind in the value of the goods (the houses) at the deep ontological level that I have in mind.

In any case, my argument thus far has the effect of denying the truth of what is generally called the “trichotomy thesis.” The trichotomy thesis maintains that for any two goods $X$ and $Y$, the following three classes of relations between $X$ and $Y$ are exhaustive and mutually exclusive:

(i) $X$ is better than $Y$

(ii) $X$ is worse than $Y$

(iii) $X$ and $Y$ are equally good

In denying this thesis, I am endorsing an incommensurability of value, or – what I take to be the same thing – a value pluralism. The economists’ hope that all preferences, that the value of all goods, and thus that all costs and benefits, could be expressed monetarily is a hope in vain. This may be unfortunate. It may complicate public policy decisions (or it may not—the real policy implications of this incommensurability have not yet been discussed). But it is a hope in vain nonetheless.

Layard and Glaister note that “it is fair to ask [someone] what he means in saying that $[X]$ is better than $[Y]$, unless he has some means of comparing the various dimensions along which $[X]$ and $[Y]$ differ.”\(^{220}\) It is indeed fair to ask. The point Layard and Glaister are insinuating is a reasonable one. $X$ could be better than $Y$ in some respect but worse others. If

\(^{220}\) Layard and Glaister, ibid.
this is so, then there are means of comparing \( X \) and \( Y \) along all these dimensions, i.e., in all these respects—those in which \( X \) is better, as well as those in which \( X \) is worse. If \( X \) is called simply better than \( Y \)—better, simpliciter, or better “on the whole” than \( Y \), then (roughly) one of three things must be true. Either:

(1) \( X \) and \( Y \) differ from one another only along one dimension and this dimension is one with respect to which there is a means of comparing the two; or

(2) \( X \) and \( Y \) differ from one another in multiple respects but in all and only the same respects, and \( X \) fares better than \( Y \) in all of these respects; or

(3) \( X \) and \( Y \) differ from one another in multiple respects, and though \( X \) fares better in some respect(s) and \( Y \) better in another respect or respects, nonetheless there is some single overarching respect in which \( X \) and \( Y \) differ, and in this respect \( X \) fares better than \( Y \).

Thus, I think the following conditional must be counted as true: If a rational and well-informed person counts \( X \) as better than \( Y \) (in some respect, or on the whole), then \( X \) and \( Y \) are able to be valued (in that respect, or on the whole) in the same units (namely, whatever units \( X \) comes in and/or \( Y \) comes in). If someone counts a certain monetary amount as better (in some respect, or on the whole) than a certain good \( Y \) (or worse than \( Y \), or calls that monetary amount and \( Y \) equally good), then \( Y \)’s value (in that respect, or on the whole) can be expressed monetarily. We should not call one thing in a certain respect better than, worse, than or equal to—i.e., worth more than, worth less than, or worth the same as—any particular amount of money if we cannot compare them in that respect. We should not call one thing on the whole better than, worse than, or equal to—i.e., worth more than, worth less than, or worth the same as—any particular amount of money if neither (1), nor (2), nor (3) is true.
It is possible, of course, to call one thing on the whole better than another just because the former was preferred to the latter in the strict behaviorist sense. But there is the deeper sense of preference, and in turn the deeper sense of value (i.e., of better and worse), and when we are focusing in at this deeper level, we must admit that it is conceptually possible, and will in actuality at least sometimes be the case, that there are goods of which we cannot claim that one is better than the other and also cannot say that they are equal in value. It does make sense to say, e.g., that one thing is better than another with respect to how awe-inspiring it is, or how loveable it is, how fascinating it is, how worthy of respect it is, and so on. It also makes sense to call one thing worse than another in these and other respects; and it makes sense to call two things equal in particular respects. But we cannot say, in any more general way, that one thing is better or worse than another, or that two things are equally good. Is 10 yards better than, worse than, or equal in goodness to, 10 pounds? The answer is: it is none of these. 10 yards and 10 pounds differ from one another, but along different and irreducible dimensions. In fact, 10 yards and 10 pounds are themselves fundamentally different dimensions along which two things can differ from one another. Is the color orange better than, worse than, or equal to, pizza? Again, it is none of these. I am being provocative, but only to illustrate a point: the denial of the trichotomy thesis is not so outrageous. Is a one-in-10 risk of cancer over the next year better than, worse than, or equal in goodness to, $50,000? Is $200 better than, worse than, or equal to, the opportunity to experience the Grand Canyon without loud aircrafts shuttling around overhead disturbing the relative quietude? Is $1 million better than, worse than, or equal to a highly polluted beachfront? In each case, none of the three options is the right one.
Section 4: Incommensurability, Choice, and Practical Reason

It should be noted that the argument from the previous section of the chapter was descriptive in an important respect. The claim is not that for each good, there are certain kinds of appraisals that are appropriate. The assertion is not that certain goods *should* engage certain preferences, i.e., that mountains should inspire awe, or that bald eagles should engender reverence, or that pollution should create indignity. I am not suggesting that certain kinds of evaluative attitudes are in any way superior to other kinds, or even that certain kinds of attitudes are fitting (more fitting, or whatever) for certain goods or activities. The claim is just that our attitudes do as a matter of fact differ from one another qualitatively, and that different goods do as a matter of fact tend to engage different kinds of attitudes. The aim of the argument is, in the first instance, simply to represent a fact about us, about our psychology, about the structure of our preferences. Though the account is in that respect just descriptive, there are, I have argued, important normative implications. The previous section defended a rough account of incommensurability of values. An incommensurability was claimed to arise when different goods engage irreducibly distinct evaluative attitudes. We are immediately confronted with important questions regarding the implications of this incommensurability for practical reason and choice (or, perhaps, for the implications of choice for the account of incommensurability). This section of the chapter is given to exploring some of these questions.

One obvious objection to the view I have defended is (roughly) that though the behaviorist conception of a preference is not compelling, and though it is true that we have a variety of evaluative attitudes that differ from one another in kind, nonetheless, if a person can, *for reasons* (i.e., via some non-random procedure), choose between $X$ and $Y$, then there
must be some one broad evaluative attitude that arches over, subsumes, and essentially unifies all the more specific evaluative attitudes listed earlier. (Perhaps this attitude should itself simply be called preferring or favoring; but that is not essential. What is important, the objection goes, is simply that the claim that evaluative attitudes are irreducibly diverse is false.) My view is that there is in fact, strictly speaking, no such overarching attitude. The fact that \( P \) reasonably chooses \( X \) over \( Y \) does not require or demonstrate necessarily that she has brought some single attitude, or even the very same set of attitudes, to bear in appraising \( X \) and \( Y \). Again, this is not to say that \( P \)'s reasonably choosing \( X \) over \( Y \) does not entail anything about \( P \)'s preferences. It entails that \( P \) has preferences regarding \( X \) and \( Y \). It does not prevent us from saying that \( P \) prefers \( X \) to \( Y \), so long as this is interpreted in the appropriate shallow sense that she will, in virtue of her evaluation of \( X \) and \( Y \), choose, or at least be disposed to choose, \( X \) instead of \( Y \). This loose sense does not require that evaluative attitudes are all of a piece, i.e., that preferences differ from indifference along only one single dimension.

Suppose that someone asks me whether I would choose to watch a bald eagle in flight or spend time with my friends (if I had to pick), and that I say (on this occasion, at least) that I would choose seeing the bald eagle. For this reason, I count (on this occasion) as preferring in the strict behaviorist seeing the eagle to spending time with my friends: I choose to do the former. Does this mean that seeing the bald eagle is (on this occasion, for me) more valuable than spending time with my friends? No, it does not. Nothing prevents me from choosing as I do without the choice being based on, or redounding to, a comparison of these two activities along some single value scale. The fact that \( P \) (in this sense) prefers \( X \) to \( Y \) certainly does not prevent \( X \) from being better (in some overall, totally unified respect) than \( Y \), but it also does
not guarantee it. $P$ can in the behavioral sense prefer $X$ to $Y$ in the absence of any single measure of the deep value for the two. In no sense, however, does this erase the fact that we have an irrediclably diverse set of evaluative attitudes, or that these attitudes are connected to goods’ value. In order for one thing to be more valuable than another, the things must both show up on the same value scale; but the choice of one good rather than another, i.e., the substitution of one good for another, does not require, and does not prove, that they show up on the same scale. The mere fact that $P$ chooses $X$ over $Y$ does not tell us that $Y$’s value (to $P$) can be expressed in terms of dollars. One can reasonably choose one thing over another without having, and without the choice redounding to reflect or (as it were) create, a common measure of value for the two things.

The pressure to deny this comes only from a mistaken conception of what choice – more specifically, “rational” or “reasonable” choice – requires. The suggestion that such choice demands that there be, or proves that there is, some unifying, overarching attitude misrepresents – cheapens, it would not be unfair to say – the character of our lived experiences. Sunstein says that “probably we cannot make sense of our experience without reference to qualitatively diverse goods.”\(^{221}\) He is right. For support of the claim that there is no such overarching attitude, and \textit{ipso facto} of my general objection to what I have termed the argument from substitutability to commensurability, I turn to phenomenological considerations within moral psychology. Again, I have conceded that for all non-market goods, there is an amount money that we would accept in order to give up the good, at least where the unit of loss of the good is small enough. But for many of the examples I can imagine – accepting money in exchange for the damming of a river, the destruction of a species, the loss of a unique landscape, my health or that of someone I love, and so on – I am

\(^{221}\) Sunstein, ibid, p. 237.
not entirely comfortable with my monetary specification. Even if I were “compensated” by the amount I specified, I believe I would still be plagued by a residual discomfort. This discomfort, moreover, would not simply reflect my disappointment that I could not have the environmental good and the money. It would also not simply be the result of my specifying a monetary exchange amount for a good that I have been led just by convention to believe should not be traded for money.

And I know that I am not alone here. This same sense of discomfort plagues many people who have made or imagined making choices between money (on the one hand) and many non-market goods, including environmental goods of the kind I have mentioned, as well as human life, health, comfort, and so on (on the other hand). Just about everyone says that there are certain things on which we cannot put a price. They say this even though they might — indeed, sometimes even as they do — concede that they would substitute (at least marginal units of) these things for at least a large sum of money. It is certainly not the case that all these people have considered the argument from substitutability to commensurability, the substitution principle, and so on, and developed carefully reasoned rebuttals to these. They are simply uncomfortable with the idea of treating the value of certain kinds of goods as identical to the amount of money they would substitute for the goods, and they would feel uncomfortable even if they knew they would be compensated by the monetary amount they specify as needing in exchange.

While my appeal here is surely not tantamount to a deductive proof, the expectation is that the felt experience of discomfort should make more plausible the idea that when a project delivers a financial gain but an environmental loss, these cannot be expressed in common units. This phenomenological data, this widespread and genuine discomfort,
suggests that for many people, the money and the environmental good are not really being fit onto the same scale. The idea that there are incommensurable values goes a considerable distance in explaining the discomfort that many people do obviously have in certain cases of choice even where they are ultimately willing to substitute one good for another (with money for an environmental good being my prime interest and example, but not the only possible example). We are comparing them and choosing between them, but we do so without one measure of their value, and this can result in a sort of dissonance.

The fact of irreducible diversity among preferences does not mean that we cannot reasonably choose between these goods/outcomes that do not engage the same preferences, though it does mean that it will often not be easy to do so. It means that deliberation is more messy, more complicated, than it would be if there were just one kind of preference and in turn one kind of value, though there is no reason to conclude that it must necessarily leave people standing frozen, unable to decide and thus unable to move. That an individual brings to bear a set of irreducibly diverse attitudes is not a threat to the individual’s ability to make a rational decision, though ‘rational’ here will have to be construed in a difference sense from the one the average economist – and perhaps the average philosopher – will count as standard. I believe that individuals can make a rational choice between money and any of a number of goods, including environmental goods, without it being the case that there is some one scale on which the value of money and whatever good is in question are compared. But many people who have thought hard about this issue disagree with me. That is to say, many people have said that denying any of the three conditionals that I stated earlier is not only counterintuitive, but also poses significant problems for the possibility of rational choice between goods. As Elizabeth Anderson writes:
A venerable tradition in philosophy upholds three claims about the demands of practical reason: that reason can settle all questions about what to choose; that it requires the global maximization of value; and that the grounds for rational choice must be fully and decisively articulable, leaving no room for judgment and hence none for dispute. The first two claims imply that the values of all options are commensurable, that their values can comprehensively be measured on a single scale. More technically, two goods, A and B, are commensurable if and only if there is a scale of overall value by which they can be at least ordinally ranked. An ordinal ranking says that either A is better than B, B is better than A, or they are equal in value. Unlike a cardinal ranking, it doesn’t say how much more valuable one good is than the other... Welfare economists often appeal to the first two demands of rational choice to argue for a single measure of value. The necessity of choice between two goods, such as money and life, requires their commensurability, preferably in terms of money.222

This “venerable tradition” has an undeniable intuitive appeal. Interestingly, though, Anderson does not count herself as part of the tradition. She rejects the view that a rational decision between X and Y requires that they both be assessable on the same scale, i.e., that they be commensurable. Sunstein seems to reject it too, saying “certainly we can have comparability (and hence reason-guided choice) without commensurability.”223 He believes that “options may be incommensurable, in the sense that they cannot be aligned along a single metric, or assessed along a single, low-level dimension, without being incomparable, in the sense that we cannot choose among them for good reasons.”224 Bernard Williams is also among those outside of the so-called venerable tradition. Williams says that there is “an assumption about rationality, to the effect that two considerations cannot be rationally weighed against each other unless there is a common consideration in terms of which they can be compared. This assumption is at once very powerful and utterly baseless... The drive

222 Anderson, ibid, pp. 45-46. I would modify this last sentence to read: “The necessity of rational choice between two goods...” As I have said, it is not choice per se that is difficult to make sense of in the absence of commensurability, but rational choice.

223 Sunstein, ibid, p. 239.

toward a rationalistic conception of rationality comes… from social features of the modern world.”

Though much more would need to be said, it thus certainly starts to appear as though there is also a “venerable tradition in philosophy” that denies that commensurability is required for rational decision making.

This is an appropriate place to end the dissertation, but the next step in this broad research project will be the important but difficult one offering and defending a detailed account of how to understand the notion of rational choice in the face of incommensurable values. The task may be even more important, and more difficult too, when we focus not at the level of personal decision making but rather at the level of public decision making. I suspect that a very large part of the appeal of CBA is the difficulty involved in envisioning conceptions of public justification for a project if we cannot simply add up the gains that would be created by the project, add up the losses that would be generated by the project, and then straightforwardly compare the costs and the benefits. In addition to the need for a model for how public decisions could be made without all benefits and costs being reduced into monetary terms, there is also the issue of legal burdens of proof for, e.g., regulations. What will satisfy the courts when a corporation brings suit claiming that a regulation was unfairly burdensome, if not a demonstration that the regulations benefits (to the public, say) outweigh the costs, and what will “outweighing” mean if costs and benefits are not all monetized? Until a viable alternative is offered, the urge towards economic valuation of non-market goods will always be a strong one.

In the face of incommensurability, the terms “reasonable” or “reliable” choice may initially be easier to swallow than “rational” choice. Roughly, I think that we should say that

---

not every reasonable or reliable choice requires a mathematical balancing of quantified values, and that it is instead sufficient for the decision maker to have carefully scrutinized and forthrightly reflected on the options, taking account of and having shown herself to be responsive to all the relevant normative and empirical considerations involved. Alas, I am not prepared to offer an alternative to the standard conception that is any more nuanced, sophisticated, or polished than this. But even if I were so prepared, I suspect I would not offer that account here. The aim of this dissertation was, from the start, simply to assess whether CBA’s theory of value – the valuational aspect of CBA, i.e., the normative foundations of economic valuation of non-market goods – was satisfactory. I have done that, and so this project has basically accomplished what it set out to do. Even if I had argued that all values, all gains and loses, all costs and benefits, could be expressed monetarily, more work would be required for a full assessment of CBA, for it is by no means obvious that that other part of CBA – its theory of right action – is acceptable. Given the power of certain moral outlooks that hold that rightness is not simply about maximizing goodness, defending CBA’s theory of right action would be a formidable task.

If the substantial elements of my argument in this project are generally correct, then a challenge has certainly been posed to those who want to endorse the Kaldor-Hicks Criterion. Costs and benefits related to losses or gains in at least certain non-market goods cannot, in the deep sense I have articulated, be expressed monetarily. The maximizing project of CBA thus seems to have been threatened. This project gets back on track only if the economist’s argument can be successfully modified and redeployed. Economists are right to be interested in what people prefer – again, at least where these preferences are well-informed and rational – but given the viable, deep sense of preferring that I articulated in this chapter, economists
will, if they want to continue focusing only on what people prefer in the shallow, behaviorist sense, need to give an argument for why this is all that public policy analysts should take interest in. “Shallow,” after all, need not be a pejorative. Perhaps the shallow sense is the right one. If it is, why this is so is, I believe, not obvious.

Economists may object that my argument involves getting inside people’s heads, and may in turn claim that a conception of and approach to valuation should be favored if it avoids – as their approach does – the need to do this. They will claim to be recommending an approach to public policy analysis that involves just observing and make policies in response to people’s behavior in the market. This objection is clearly unconvincing. Despite possible protestations to the contrary, economists already try to get inside people’s heads. Cost-benefit analysts are not simply observing how people do behave. The standard methods of CBA involve trying to figure out how people would behave – how much they would be willing to pay, willing to accept, etc. – in particular situations. Sometimes this involves asking people, and when it does, analysts want to be sure the answers specified are sincere. I think that inquiries into the real value of various goods must involve asking about how exactly people assess the goods, not just what will be traded for what. The objection may be that this is getting “deeper” inside people’s heads than the economic approach gets, and that may be true; but then, why is this a problem? If the worry is that this too time consuming, and in turn too expensive, I am not sure I share it. In Chapter 4, I briefly discussed a role that might be played by randomly selected small groups of decision makers modeled after a jury. If it is just the heads of the individuals comprising these small groups that policy analysts need to get inside of, then there is little reason to think that doing so will be terribly burdensome in terms of time or money spent. If the answer is that analysts only go as deep
as they do because all public decision makers really care about is how much people want the state to spend, then I will answer that that is a shame. If the claim is that analysts only go as deep into the head as they do because how much people want to spend is the only relevant issue for public policy decisions, then I will first note that asking someone how much should be spent is different from asking how much she would be willing to spend, and will also point out that how much people want to spend is not the only relevant issue. Public decision makers should be interested to know how to treat those things, activities, places, opportunities, relationships, processes, etc. that people care about, and you do not learn everything there is to know about all this simply by asking about money.

So, the question is whether there is another reason for thinking only about preferences in the shallow, behaviorist sense. The economists’ original argument – that the purpose of public policy decisions is to promote social welfare, which is done by promoting some aggregate of individual well-being, which is done by promoting the satisfaction of (well-informed and rational) individual preferences – does not seem straightforwardly to offer such a reason. There is not reason to think that learning just about someone’s preferences in the shallow sense will tell us all that we want to know about how a policy would affect her well-being. Just as the evaluative attitudes that are partly constitutive of such preferences vary along multiple irreducible dimensions, so too will an individual’s well-being vary along multiple irreducible dimensions! This means that there is no one single scale on which one does better or worse. I have not ruled out the possibility that there is an argument for looking just at people’s choices (including perhaps what people sincerely claim they would be willing to substitute for what, at their WTP, or their WTA, etc.), but such an argument will need to be one for taking people’s *choices*, not simply their (well-informed, rational) *preferences*,
seriously, and it is far from clear to me how a successful argument of that sort would go. It is
worth thinking further about, but so too is it assuredly worth thinking about how to make
policy decisions regarding public investments and regulations without trying to monetize
each relevant preference.


