Essays on Statistical Discrimination in a Dynamic Framework

Rebecca M. Glawtschew

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 Economics.

Chapel Hill
2012

Approved by:
Peter Norman
Gary Biglaiser
R. Vijay Krishna
Sérgio O. Parreiras
Helen Tauchen
Abstract

(Under the direction of Peter Norman.)

This dissertation consists of two papers in the field of statistical discrimination. In the first paper, I determine the potential causes of persistent labor market inequality by developing a dynamic model of statistical discrimination in a competitive environment. In this dynamic model, the forward-looking behavior of economic agents determines the dynamic paths to the steady states. By characterizing these dynamic paths, I am able to establish the initial conditions that can lead to each steady state and to determine if it is possible to move from one steady state to another. I find that the model can be broken down into two classes of parameterizations. In the first class, history alone determines the final outcome. In the second class, the expectations of forward-looking may agents determine the final outcome. In both of these cases, moving from one steady state to another is not possible. Using a simplified version of the dynamic model, I also examine how the parameter values impact both the existence of multiple steady states and the importance of expectations in determining the final outcome.

In the second paper, I consider the effectiveness of three government policies designed to eliminate persistent statistical discrimination in the framework of the dynamic model developed in the first paper. I determine the paths that workers will take after a policy is instated as well as how long a policy needs to be in place to guarantee the successful elimination of discrimination. The policies I consider are (1) a hiring subsidy that promotes the hiring of disadvantaged workers to the better job, (2) an investment voucher that defrays the monetary cost of human capital investment, and (3) an equal treatment policy under which firms are required to treat workers equally across groups. I find that all three policies have the potential to eliminate persistent discrimination if certain conditions are met.
Acknowledgments

I would very much like to thank the members of my dissertation committee for their guidance, patience and advice. I also owe a special debt of gratitude to my husband, Nick, whose loving support made this journey possible.
Table of Contents

List of Tables ......................................................... vi

List of Figures ........................................................ vii

1 Introduction ......................................................... 1

2 Persistent statistical discrimination in a competitive economy ........ 3
   2.1 Introduction .................................................. 3
   2.2 The static model .............................................. 7
      2.2.1 The economic environment ............................ 7
      2.2.2 The game .............................................. 9
      2.2.3 Equilibria ........................................... 9
   2.3 The general dynamic model ................................ 15
      2.3.1 The economic environment ............................ 16
      2.3.2 The game .............................................. 17
      2.3.3 Equilibria ........................................... 17
   2.4 The simple dynamic model ................................ 32
      2.4.1 Firms and workers .................................... 32
      2.4.2 The updated dynamic system .......................... 32
      2.4.3 Example 1: Spiralling paths ......................... 33
      2.4.4 Example 2: Deterministic paths ..................... 37
   2.5 Rationalizing the existence of multiple saddle points ............. 39
   2.6 Concluding remarks ........................................ 41
3 Eliminating persistent statistical discrimination: an analysis of several policy options

3.1 Introduction

3.2 The model

3.3 Equilibria

3.4 Simple parameterization

3.5 Potential Policies

3.5.1 Hiring subsidy program

3.5.2 Investment voucher program

3.5.3 Government enforced equal treatment

3.6 Welfare analysis of the three policies in a parametric example

3.7 Conclusion

A Proof of Proposition 1

B Mathematical Methods

B.1 Calculating paths to the steady states

B.2 Welfare calculations

B.2.1 Welfare in the base model

B.2.2 Welfare under a hiring subsidy

B.2.3 Welfare under an investment voucher

B.2.4 Welfare under the equal treatment policy

Bibliography
List of Tables

3.1  Details of the discriminatory equilibrium ........................................ 86
3.2  Total change in welfare while the policies are in effect ....................... 89
List of Figures

2.1 Three equilibria in the static model ........................................ 14
2.2 Loci of the general dynamic model ........................................... 21
2.3 Phase diagram of the dynamic system ........................................ 22
2.4 Deterministic paths to the steady states .................................... 28
2.5 Spiralling paths to the steady states .......................................... 29
2.6 Phase diagram for example 1 .................................................... 34
2.7 Spiralling paths to the steady states .......................................... 34
2.8 Impact of the parameter values on the existence of the steady states . 36
2.9 Impact of the parameter values on the shape of the dynamic paths ... 36
2.10 Deterministic paths to the steady states .................................... 38
2.11 Impact of the parameter values on the existence the steady states . 39
2.12 Impact of the parameter values on the shape of the dynamic paths ... 39
2.13 The dynamic system and the paths to $\Sigma_h$ and $\Sigma_l$ when costs are discrete .... 40

3.1 Phase diagram of the dynamic system ........................................ 54
3.2 Spiralling paths to the steady states .......................................... 57
3.3 Deterministic paths to the steady states .................................... 58
3.4 Phase diagram for the parametric example .................................. 60
3.5 Spiralling paths to the steady states in the parametric example ......... 61
3.6 Phase diagram of the dynamic system under the hiring subsidy ...... 66
3.7 Path to $\Sigma^\xi$ ................................................................. 67
3.8 Phase diagram for the parametric example under the hiring subsidy .... 69
3.9 Path to the $\Sigma^\xi$ in the parametric example ............................... 70
3.10 Path to $\Sigma_h$ once the hiring subsidy policy is lifted .................... 70
3.11 Phase diagram of the dynamic system under the investment voucher .. 73
3.12 Path to $\Sigma^\nu$ ................................................................. 75
3.13 Phase diagram for the parametric example under the investment voucher . . . 76
3.14 Path to the $\Sigma^v$ for the parametric example . . . . . . . . . . . . . . 77
3.15 Path to $\Sigma_h$ once the investment voucher policy is lifted . . . . . . . . 77
3.16 Potential starting points under the equal treatment policy . . . . . . . . . 83
3.17 Path to $\Sigma_h$ given an optimistic population . . . . . . . . . . . . . . . 84
3.18 Path to $\Sigma_l$ given a pessimistic population . . . . . . . . . . . . . . . 84
3.19 Welfare effects of the hiring subsidy . . . . . . . . . . . . . . . . . . . . . 86
3.20 Welfare effects of the investment voucher . . . . . . . . . . . . . . . . . . . 87
3.21 Welfare effects of the equal treatment policy . . . . . . . . . . . . . . . . . 87
Chapter 1

Introduction

This document presents the two papers that form my dissertation in accordance with the Graduate School and Economics Department at UNC Chapel Hill.

The first paper is titled “Persistent statistical discrimination in a competitive economy”, wherein I develop a dynamic model of statistical discrimination that allows me to answer several questions that are not addressed in the static literature. While static models of statistical discrimination successfully explain why workers from different observable groups may earn different wages on average, these results are only examined in equilibrium. Static models cannot explain how the groups get to their respective equilibria nor are they able to address whether it is possible for a group to move from their current equilibrium to another. In order to answer these important questions, I develop a dynamic model based on the static models of Coate and Loury (1993) and Moro and Norman (2003). In particular, I introduce dynamics to the model by allowing workers to be long-lived. With this addition, I determine both the dynamic paths that can be taken to the relevant steady states and the conditions under which discrimination in this model will be persistent.

I find that there are two potential forms that the dynamic paths can take depending on the parameter values. In one class of parameterizations, history alone determines the final outcome. This situation is more likely to occur if workers heavily discount future wages. In the second class of parameterizations, the dynamic paths create what I refer to as the uncertain region. If a group’s initial conditions place them in this region, then the expectations of the forward-looking workers will determine the final outcome. This type of dynamic path will
occur if the workers place significant value on future wages. I also establish that, no matter what form the dynamic paths take, moving from one steady state to another is not possible.

The second paper is titled “Eliminating persistent discrimination: an analysis of several policy options”. In this paper I revisit the model of the first paper and make a few additional assumptions. Using this dynamic framework, I analyze the effectiveness of three government policies designed to eliminate existing discrimination. The policies that I investigate are (1) a hiring subsidy that promotes the hiring of disadvantaged workers to the better job, (2) an investment voucher that defrays the monetary cost of human capital investment, and (3) an equal treatment policy under which firms are required to treat workers equally across groups. By examining these policies in a dynamic framework, I am able to determine the dynamic path that the workers will take once a policy is put in place and the point at which it is safe for the government to remove the policy and still guarantee the eventual elimination of discrimination. With the inclusion of competitive wages, I am also able to complete meaningful analysis of the welfare effects of the three policies.

The analysis indicates that the hiring subsidy policy will always effectively eliminate discrimination if a large enough subsidy is offered. However, while the policy is in place the advantaged group experiences a significant loss in welfare. On the other hand, the investment voucher policy results in very little loss in welfare to the advantaged group, but it can only be successfully applied if a large enough proportion of the cost of investment is monetary in nature. The equal treatment policy is attractive as it requires no transfers from the advantaged to the disadvantaged workers and welfare losses are relatively small when it is successful. However, there are many situations in which the equal treatment policy will not be successful, for example, if there are relatively more disadvantaged workers than advantaged workers in the population.
Chapter 2

Persistent statistical discrimination in a competitive economy

2.1 Introduction

This paper develops a dynamic model of statistical discrimination that can rationalize persistent labor market inequality across observable groups as a result of either differing initial firm beliefs or differing group expectations. Just as in related static settings, discrimination in this model is a result of different groups coordinating on different steady states. However, with the addition of dynamics I am able to determine how the different groups get to their respective steady states as well as the conditions under which they become stuck in that steady state. Understanding the dynamic system also allows for richer analysis of government policies designed to mitigate labor market inequality.

With the introduction of long-lived workers, I am able to add dynamics to a special case of the Moro and Norman (2003) model of statistical discrimination. Statistical discrimination refers to the situation in which a decision-maker uses observable characteristics of individuals as a proxy for unobservable, but outcome relevant, characteristics. In this model, the human capital investment decisions of workers are unobservable, as a result firms will condition wages and job assignments on group identity. Consequently, workers from different groups, that are otherwise identical, may be assigned to different jobs and be paid different wages in equilibrium.

Multiple steady states are necessary for the statistical discrimination in this model to be persistent, otherwise, the groups are treated identically as soon as they both reach the single steady state. I focus on the case in which there are three steady states. Under this assumption, the middle steady state is unstable and the dynamic paths to the high and low steady states
emanate from that point. These paths can take several forms depending on the parameter values.

For one class of parameterizations, initial firm beliefs about the average skill level of a group are decisive and a group’s expectations about future payoffs are irrelevant. In this case, a group will converge to the low steady state with certainty if initial firm beliefs are below those of the middle steady state. Conversely, if firm beliefs are above those of the middle steady state the group will converge to the high steady state. In this situation it is not possible for a group to move from one steady state to another.

In another class of parameterizations, the dynamic paths create what I refer to as the uncertain region of firm beliefs. If initial firm beliefs about a group fall into this region, then it is possible for that group to converge to either the high or the low steady state. The final outcome depends on the group’s expectations about future payoffs. If the group is optimistic about the future, they will converge to the high steady state. If they are pessimistic, they will find themselves on the path to the low steady state. In this class of parameterizations groups are not able to move from one steady state to another.

By characterizing the dynamic paths to the steady states, I am able to establish the initial conditions that can lead to discrimination. I find that discrimination will arise if initial firm beliefs about one group are relatively high and for the other group they are relatively low, or if one group is optimistic about future payoffs while the other group is pessimistic. Furthermore, the fact that groups may become stuck in a steady state implies that re-coordination on a different steady state is not possible without a structural change in the model. Consequently, if present, discrimination will be persistent. These are important results that related static models have been unable to address.

I also consider a simple version of the dynamic model and examine two parametric examples: one in which historical firms beliefs are decisive and one in which the expectations of workers potentially determine the final outcome. In the framework of this simple model, I find that multiple steady states are more likely when signals about the quality of workers are of intermediate precision. I also find that expectations are less likely to determine a group’s final outcome if the rate of population turnover is high or if the population discounts future payoffs.
at a high rate.

The paper is structured as follows. The related literature is discussed in the remainder of Section 2.1. The static model is introduced in Section 2.2. Section 2.3 develops the dynamic model and addresses the conditions under which discrimination will be persistent. In Section 2.4 a simple version of the dynamic model is described and two parametric examples are examined. I provide a brief rationalization for the existence of two paths in Section 2.5. Finally, Section 2.6 concludes.

Related Literature

There is a large literature on statistical discrimination that builds upon the seminal contributions by Arrow (1973) and Phelps (1972). In models of statistical discrimination, uncertainty in the labor market makes it rational for firms to condition on group identity. In particular, firms in these models are unable to perfectly observe worker productivity so they base hiring and wage decisions on a noisy signal and on prior beliefs about the productivity of workers from different groups. In such environments inequality can occur as the result of exogenous group differences in the precision of information\(^1\) or as a result of multiple equilibria\(^2\). The model presented here is the latter type.

The model developed in this paper is highly related to both Coate and Loury (1993) and Moro and Norman (2003). Coate and Loury (1993) present a static model of statistical discrimination where two ex ante groups may end up in different, Pareto ranked, equilibria. Their model formalizes many of ideas that were originally presented loosely in Arrow (1973), but it makes the additional assumption that wages are set exogenously. They find that inequality will occur if the two groups of workers play different equilibria.

Moro and Norman (2003) relax two of the assumptions in Coate and Loury’s static model. In particular, they remove the linearity of the production technology and the exogeneity of wages. Allowing for curvature in the production function eliminates the separability of groups,\(^1\)As in Phelps (1972), Aigner and Glen (1977), Cornell and Welch (1996), Lundberg and Startz (1983) and Oetttinger (1996).

thus, creating the possibility for true interaction effects between groups. Introducing endogenous wages allows for more meaningful welfare analysis. Under these relaxed assumptions, they determine that the dominant group will actually gain from discrimination and, as a result, they will be motivated to ensure its continuation. In the model presented in this paper, I incorporate endogenous wages, however, for simplicity I assume that the production technology is linear.

There is relatively little literature that addresses discrimination in a dynamic environment. Fryer (2007), Blume (2006) and Kim and Loury (2009) all develop dynamic models that are adaptations of the original Coate and Loury (1993) paper. Fryer adds dynamics by introducing a promotion stage after workers are initially hired. In this environment he finds that a firm may discriminate against a group in the initial hiring process and then prefer to promote members of that group once they are hired. This idea of “belief flipping” originates from the fact that those workers in the disadvantaged group who are available to promote were initially held to a higher standard. Consequently, during the promotion stage firms believe they are more likely to be skilled than workers in the advantaged group.

Blume (2006) takes a different approach and investigates the effects of two exogenously specified learning dynamics that are applied to the static Coate and Loury (1993) model. He characterizes long-run stable patterns of discrimination under both learning dynamics and is able to connect them to the equilibria of the static model.

Kim and Loury (2009) add long-lived workers to the Coate and Loury (1993) static model in order to introduce dynamics. They determine that the equilibria of this dynamic model are characterized by a two variable dynamic system. Using the dynamic system, they describe the paths to the relevant steady states. Their analysis has similarities to the model presented here; however, they do not incorporate a competitive environment and their main focus is on the analysis of a highly simplified model.

Antonovics (2006) considers a dynamic model of statistical discrimination that accounts for intergenerational income mobility. Like the model I present, the focus of Antonovics’ analysis is to show how a group can become trapped in a bad equilibrium. However, she achieves this result by allowing for income to be transmitted across generations through parental investments
in the human capital of children.

While not related in subject matter, Krugman (1991) is related in its approach to analyzing dynamic models with multiple steady states. In this paper, he determines that the existence of two equilibrium paths leading to the two steady states has the potential to create an “overlap”. Within this overlap, expectations determine the final steady state, and outside of it, the final steady state is determined by history alone. This interpretation is central to my analysis of persistent discrimination in Section 2.3.

2.2 The static model

In this section, I present the static model that is the foundation of the dynamic analysis in Section 2.3. This static model is adapted from Moro and Norman (2003); I generate competitive wages just as they do. Given that my primary goal is to focus on the dynamic paths to the steady states, I use a simple linear production function so that groups remain separable as in Coate and Loury (1993). A thorough analysis of this model allows me to highlight the questions that are left unanswered in the static environment and to draw comparisons across the equilibria of the static model and the steady states of the dynamic model.

2.2.1 The economic environment

The firms

Consider a market where two firms compete for workers. Output is generated by the completion of two types of tasks, a simple task and a complex task. All workers are able to perform the simple task, while only those workers that invest in becoming qualified are able to successfully perform the complex task. Consequently, the effective input of labor into the simple task is the total number of workers a firm employs in that task. I represent this with $S$. The effective input of labor into the complex task is the number of qualified workers a firm employs in that task. I represent this with $C$. Both firms’ output is linear in the two tasks; it is given by $Y(S, C) = \alpha_s S + \alpha_c C$, where $\alpha_c > \alpha_s > 0$. 

The workers

There is a continuum of workers with mass normalized to unity. Each worker belongs to one of two identifiable groups, B(lack) or W(hite). The proportion of workers that are members of the B group is \( \lambda^B \) and the proportion of workers that are members of the W group is \( \lambda^W \). All workers have the option to invest in human capital prior to entering the labor market. If a worker chooses to invest in human capital, he becomes qualified and is therefore able to complete the complex task. Workers who do not make this investment are unqualified and can only successfully complete the simple task. In order to invest in human capital a worker must pay a cost, \( c \). This cost varies across workers and is distributed over \([c, \bar{c}] \subseteq \mathbb{R}\) according to a continuous and strictly increasing distribution \( G(c) \).

Workers are risk neutral and care only about the net wages that they receive. This means that a worker has no preference over the two tasks outside of the different wages they provide. Pay-offs are additively separable in income and cost of investment. That is, a worker with a cost \( c \) who chooses to become qualified and earns a wage \( \omega_q \) will receive utility of \( \omega_q - c \), and a worker who is unqualified and earns a wage \( \omega_u \) will receive utility of \( \omega_u \).

Information technology

Firms are unable observe whether an individual worker invested in human capital, instead for each worker they observe a signal \( \theta \in [0, 1] \). This signal is distributed according to density \( f_q \) if the worker is qualified and \( f_u \) otherwise. Both densities are bounded away from zero and, without further loss of generality, \( f_q(\theta)/f_u(\theta) \) is increasing the \( \theta \). Using Bayes’ rule, the firms form beliefs about the likelihood that a worker is qualified given his signal. If \( \pi^J \) represents a firm’s prior belief about whether a worker from group \( J \) is qualified, then the posterior probability that a worker from group \( J \) with signal \( \theta \) is qualified is

\[
p(\theta, \pi^J) \equiv \frac{\pi^J f_q(\theta)}{\pi^J f_q(\theta) + (1 - \pi^J) f_u(\theta)}.
\] (2.1)

The monotone likelihood ratio property implies that \( p(\theta, \pi^J) \) is increasing in \( \theta \) so that a high signal reflects positively on a worker and a low signal negatively. I denote the associated
cumulative distributions by $F_q$ and $F_u$. I assume that a law of large numbers holds so that these are also the realized frequency distributions of signals for qualified and unqualified workers, respectively.

2.2.2 The game

The timing of the game is as follows: In the first stage of the game firms post wage and task assignment rules and workers simultaneously decide whether or not to invest in becoming qualified. An investment strategy for a worker in group $J$ is a map $v^J : [c, \bar{c}] \rightarrow [0, 1]$. Put simply, a strategy maps cost of investment into the probability of investment.

Firms may condition wage and task assignments on $\theta$. A strategy for a firm $i$ is to select some wage schedule $\omega^J_i : [0, 1] \rightarrow \mathbb{R}_+$ and a task assignment rule $\tau^J_i : [0, 1] \rightarrow [0, 1]$ for each group $J$. This task assignment rule maps worker signals into the probability of assigning a worker to the complex task.

In the second stage nature assigns a signal $\theta$ to each worker. The signal for a qualified worker is distributed according to density $f_q$ and for an unqualified worker it is distributed according to density $f_u$. In the third and final stage workers observe the posted wage and task assignment rules and decide where to work.

Workers care only about potential wages when comparing job offers and their investment costs are sunk. So they will choose to work for the firm whose wage and task assignment rule provides them with the highest wages.

2.2.3 Equilibria

I consider Nash equilibria satisfying the additional requirement that workers choose firms in a sequentially rational way after any history of play. These equilibria are perfect Bayesian. There are two main components to the equilibria of this game: First, each firm must employ both an optimal task assignment rule and an equilibrium wage assignment rule. Second, the workers must make their investment decisions optimally; they should invest only if their expected benefit exceeds their cost. The equilibria are determined by taking into account the optimal wages and task assignments and ensuring that the actual investment rate is equal to the firms’
beliefs. Discrimination in this model is possible as long as there are multiple equilibria.

**Equilibrium task assignments and wages**

The group specific fraction of qualified workers is denoted by \( \pi = (\pi^B, \pi^W) \). In equilibrium the firms’ beliefs about the group specific fraction of qualified workers will equal the true fraction so I use \( \pi \) to represent both of these values. I also refer to this value as the investment rate of a group.

Since (2.1) is increasing in signal, it is without loss of generality to focus on task assignment rules with a threshold property. Under this type of rule, workers with a signal above the threshold are assigned to the complex task and workers with lower signals are assigned to the simple task. Given a threshold \( \bar{\theta}^J \) for group \( J \), the effective input of labor into the complex task is \( C = \pi^J (1 - F_q(\bar{\theta}^J)) \) and the effective input of labor into the simple task is \( S = \pi^J F_q(\bar{\theta}^J) + (1 - \pi^J) F_u(\bar{\theta}^J) \). Each firm will maximize its output with a threshold solving:

\[
\sum_{J=\text{B},W} \lambda^J \max_{\theta^J \in [0,1]} \left( \alpha_c \left[ \pi^J (1 - F_q(\theta^J)) \right] + \alpha_s \left[ \pi^J F_q(\theta^J) + (1 - \pi^J) F_u(\theta^J) \right] \right).
\]  

This task assignment problem can be solved separately for each group because the firms face a linear production function.

The fact that production is linear in \( C \) and \( S \) creates the potential for \( \hat{\theta}(\pi^J) \) to be a correspondence. In order to eliminate this issue and ensure that there is a unique solution to (2.2), I make an additional assumption on \( f_q(\theta)/f_u(\theta) \).

**Lemma 1.** Suppose that \( f_q(\theta)/f_u(\theta) \) is strictly increasing in \( \theta \). Then for each group \( J \) there is a unique \( \hat{\theta}(\pi^J) \in [0,1] \) that solves (2.2) for any \( \pi^J >> 0 \).

**Proof.** Let \( \hat{\theta}(\pi^J) \) be any solution to (2.2). The effective factor inputs into the complex and simple task are

\[
C^J(\pi^J) = \pi^J (1 - F_q(\hat{\theta}(\pi^J))] \text{ and } S^J(\pi^J) = \pi^J F_q(\hat{\theta}(\pi^J)) + (1 - \pi^J) F_u(\hat{\theta}(\pi^J)) \text{ for } J = \text{B}, \text{W},
\]  

\[
\text{(2.3)}
\]
respectively. Combining the equations of (2.3) yields

\[ S^J = \pi - C^J + (1 - \pi)F_u \left( F_q^{-1} \left( \frac{\pi - C^J}{\pi} \right) \right), \tag{2.4} \]

which is the maximal level of \( S^J \) given \( C^J \). Using (2.4) I can restate the task assignment problem in (2.2) as

\[
\max_{(C^J,S^J)} \left[ \alpha_c(\lambda_B C^B + \lambda_W C^W) + \alpha_s(\lambda_B S^B + \lambda_W S^W) \right] \\
\text{s.t. } S^J \leq \pi^J - C^J + (1 - \pi^J)F_u \left( F_q^{-1} \left( \frac{\pi^J - C^J}{\pi^J} \right) \right) \text{ for } J = B, W. \tag{2.5}
\]

The monotone likelihood ratio implies that the right hand side of the constraint is a concave function of \( C^J \) and strictly concave under the strict monotone likelihood assumption. Problem (2.5) is a matter of maximizing a linear function, that is increasing in both arguments, subject to a strictly concave constraint. This yields a unique solution.

From here on I assume that \( f_q(\theta)/f_u(\theta) \) is strictly increasing in \( \theta \) so that Lemma 1 applies and \( \hat{\theta}(\pi^J) \) is guaranteed to be unique. The optimal threshold for group \( J \) is

\[
\hat{\theta}(\pi^J) = \begin{cases} 
1 & \text{if } \alpha_c p(1, \pi^J) \leq \alpha_s \\
0 & \text{if } \alpha_c p(0, \pi^J) \geq \alpha_s, \\
& \text{the unique solution to} \\
& \alpha_c p(\theta, \pi^J) = \alpha_s & \text{if } \alpha_c p(0, \pi^J) < \alpha_s < \alpha_c p(1, \pi^J). 
\end{cases} \tag{2.6}
\]

This optimal threshold can be interpreted in the following way: All workers will be assigned to the simple task (a threshold of 1) if prior beliefs are such that even a worker with a signal of 1 has higher expected productivity in the simple task. All workers will be assigned to the complex task (a threshold of 0) if prior beliefs are such that even a worker with a signal of 0 has higher expected productivity in the complex task. In all other cases, the optimal threshold is the \( \theta \) at which a worker’s expected productivity in the complex task is equivalent to that of the simple task. Note that \( \hat{\theta}(\pi^J) \) is a non-increasing function of \( \pi^J \) and that \( \hat{\theta}(0) \leq 1 \) and
Following Moro and Norman (2003) I call a strategy profile a *continuation equilibrium* if all equilibrium conditions except the requirement that investments are best responses to wages are satisfied. The following result states that wages are given by expected marginal products and task assignments are constrained efficient in any continuation equilibrium.

**Proposition 1.** Suppose that a fraction \( \pi^J \) of workers in group \( J \) invest so that \( \hat{\theta}(\pi^J) \) is the unique solution to (2). Then there exists a continuation equilibrium where both firms offer the following wages:

\[
\omega^J(\theta, \pi^J) = \begin{cases} 
\alpha_s & \text{if } \theta < \hat{\theta}(\pi^J) \\
\alpha_c p(\theta, \pi^J) & \text{if } \theta \geq \hat{\theta}(\pi^J),
\end{cases}
\]  

(2.7)

and assign a worker with characteristics \((J, \theta)\) to the complex task if and only if \( \theta \geq \hat{\theta}(\pi^J) \), where \( \hat{\theta}(\pi^J) \) is as in (2.6). Moreover, in any continuation equilibrium where a fraction \( \pi^J \) of the workers invest the wage schedule posted by firm \( i \) for group \( J \), \( w^J_i(\theta) \), agrees with (2.7) for almost all \( \theta \in [0,1] \) for each firm \( i \).

**Proof.** See Appendix A

---

**Equilibrium investment decision**

In equilibrium the optimal behavior for a worker is to invest in human capital if and only if the gain in expected earnings is higher than the cost. I refer to this gain in expected earnings as the *gain from investment*. Given that a proportion \( \pi^J \) of the \( J \) population invests and that wages are consistent with the continuation equilibrium in (2.7), the gain from investment for a worker in group \( J \) is

\[
\gamma(\pi^J, \hat{\theta}(\pi^J)) = \int_0^1 \omega(\theta, \pi^J) f_q(\theta) d\theta - \int_0^1 \omega(\theta, \pi^J) f_u(\theta) d\theta
\]

\[
= \alpha_s [F_q(\hat{\theta}(\pi^J)) - F_u(\hat{\theta}(\pi^J))] + \alpha_c \int_{\hat{\theta}(\pi^J)}^1 p(\theta, \pi^J) [f_q(\theta) - f_u(\theta)] d\theta.
\]  

(2.8)

The gain from investment depends on \( \pi \) in two ways, through \( p(\theta, \pi) \), as a result of the optimal wage structure, and through the optimal task assignment rule \( \hat{\theta}(\pi) \). For expositional simplicity
I refer to the gain from investment as \( \gamma(\pi^J) \) from now on.

It is useful to briefly discuss the potential values of the gain from investment function. From (2.6) I can show that there exists a \( \pi > 0 \) such that all workers are assigned to the simple task when \( \pi^J \leq \pi \). There is also some \( \bar{\pi} < 1 \) such that all workers are assigned to the complex task if \( \pi^J \geq \bar{\pi} \). I can also show that when \( \pi^J \leq \bar{\pi} \), then \( \gamma(\pi^J) = 0 \). However, the gain from investment is strictly positive in the range \((\bar{\pi}, 1)\) because \( p(\theta_t, \pi^J) \) is increasing in signal. Thus, there is wage inequality even when all workers are assigned to the complex task.

Given the cost structure, the proportion of workers in group \( J \) that benefit from investing is \( G(\gamma(\pi^J)) \). In a Nash equilibrium, firms have rational beliefs about the proportion of qualified workers and workers’ investment decisions are rational given the optimal wage structure. These equilibrium conditions imply that any equilibrium is fully characterized by all pairs of beliefs \((\pi^B, \pi^W)\) satisfying

\[
\pi^J = G(\gamma(\pi^J)) \text{ for } J = B, W. \tag{2.9}
\]

In any equilibrium the proportion of each group that is qualified is a solution to (2.9) and from any solution to (2.9) I can construct wage schedules, task assignments and investment decisions consistent with the equilibrium. The existence of at least one equilibrium is immediate since \( \gamma(\pi^J) \) is composed of continuous functions.

If there is only one solution to (2.9), then there is a unique equilibrium where groups must be treated identically. More interesting is the case where there are multiple solutions to (2.9), and thus the possibility of discrimination.

**Existence of multiple equilibria**

Figure 2.1 demonstrates a case where there are three equilibria. The hump-shaped curve is the graph \( \{(\gamma, \pi) | \gamma = \gamma(\pi)\} \); it represents the expected gain from investment given an investment rate of \( \pi \). The shape of this graph comes from the previous discussion of how the value of the gain from investment changes with \( \pi \). The s-shaped curve is the graph \( \{(\gamma, \pi) | \pi = G(\gamma)\} \); it indicates the proportion of workers who will invest given a gain from investment of \( \gamma \). The
The existence of multiple solutions is not always guaranteed and depends on the shape of $\gamma$ and $G(\cdot)$. The possibility of multiple solutions can be proven by construction: If I fix the parameters $f_u$, $f_q$, $\alpha_c$ and $\alpha_s$, then I can find an appropriate $G(\cdot)$ function such that (2.9) has multiple solutions. Graphically, if I fix the hump shaped curve in Figure 2.1, I can ensure multiple equilibria simply by selecting a $G(\cdot)$ function that yields multiple intersections of the two graphs. More generally, for any $f_u$ and $f_q$ that satisfy the monotone likelihood ratio and any technology parameters, I can find a non-trivial set of $G(\cdot)$ functions that ensure multiple equilibria.

Multiple equilibria are more likely to occur when signals about investment decisions are of intermediate precision. On the other hand, if these signals are very noisy, then they provide the firms with little additional information. In this case the benefit to becoming qualified is negligible and one equilibrium at a relatively low investment rate is likely to occur. If signals are not noisy enough, then firms can easily distinguish between workers that are qualified and workers that are not. This makes investment very attractive, and one equilibrium at a
relatively high investment rate is the likely outcome. I investigate this idea further in Section 2.4 where I develop a simple version of the model and directly examine the impact of signal precision on the existence of multiple equilibria.

When there are multiple solutions to (2.9) it is easy to show that they are Pareto rankable.

**Proposition 2.** Let \( \pi^* \) be the largest solution to (2.9). Then \((\pi^B, \pi^W) = (\pi^*, \pi^*)\) Pareto dominates all other equilibria of the model.

*Proof.* See proof of Proposition 4.

In this static model discrimination is sustained only as a result of coordination failure and the existence of multiple steady states is explained by self-confirming prior beliefs. When workers from one group are believed less likely to be qualified, this belief will be self-confirmed at the lower steady state. Similarly, when workers from one group are believed more likely to be qualified, this belief will be self-confirmed at the higher steady state. However, this static model cannot explain the case where the initial beliefs of the firms are not at one of the steady states nor can it explain why firms start to have different beliefs about the different identifiable groups of workers. This static framework is also unable to address why inequality is persistent. These types of questions can only be answered in a dynamic environment.

### 2.3 The general dynamic model

Now I adapt the static model of Section 2.2 into a dynamic environment. I start by constructing a discrete dynamic system and from there I approximate the continuous dynamic system by taking limits with respect to the length of the period. Starting with a discrete system is preferred as there is a distinct timing to the way that the game unfolds in each period.

With this dynamic system I am able to address the many questions left unanswered in the static framework. Specifically, I find that the paths to the steady states create the possibility for persistent inequality and that differing group expectations can explain the situation where groups end up at different steady states despite initial firm beliefs being equal across groups. I also address the relationship between the equilibria of the static model and the steady states of the dynamic model, finding that they are equivalent.
2.3.1 The economic environment

Firms and workers

Many of the details of the static model carry over to the dynamic model. Specifically, there are two firms competing for workers from two observable groups. These workers have the option to make a one-time investment in human capital that is unobservable to the employers. The cost associated with becoming qualified varies across workers and is distributed according to $G(c)$. In order to fully develop a dynamic model, some additional details are necessary.

I index the time by $t$, which is a continuous variable. Changes in this model occur at the start of each period. Firms are short-lived, and workers are long-lived and subject to a Poisson death process with parameter $\beta$. Given that periods have length $\Delta$, a fraction $\Delta \beta$ of the workers are replaced by newborn individuals in each period. Individuals also discount future pay-offs at a rate $r$, which is strictly positive. Workers have the one time option to invest in human capital and become qualified in the period they are born; once a worker is qualified he can never become unqualified. Just as in the static model there are two identifiable groups of workers and the total population of each group is constant at $\lambda^J$, where $\lambda^W + \lambda^B = 1$.

The length of the period also affects the productivity of the firms. Each firm’s per period output is given by $Y_\Delta(S_t, C_t) = \Delta[\alpha_s S_t + \alpha_c C_t]$, where $C_t$ is the number of qualified workers employed in the complex task in the period starting at time $t$ and $S_t$ is the total number of workers employed in the simple task in the period starting at time $t$.

Information technology

Firms cannot observe whether a worker is qualified, nor are they able to distinguish between new and old workers. The only information a firm has about a worker at any given time, aside from his group identity, is his current signal $\theta_t \in [0, 1]$, which is distributed according to $f_u$ if a worker is unqualified or $f_q$ if a worker is qualified. $\Pi^J_t$ represents the firms’ belief about what fraction of the $J$ population is qualified in a given period. The posterior probability that
a worker from group $J$ with signal $\theta_t$ is qualified given beliefs $\Pi^J_t$ is

$$p(\theta_t, \Pi^J_t) \equiv \frac{\Pi^J_t f_q(\theta_t)}{\Pi^J_t f_q(\theta_t) + (1 - \Pi^J_t) f_u(\theta_t)}.$$  

(2.10)

2.3.2 The game

The timing of each period of the dynamic game is as follows: In the first stage, firms post wage and task assignment rules just as in the static model. At the same time, $\Delta \beta$ new workers enter the workforce and decide whether or not to invest in human capital and become qualified. Nature then assigns a signal to each worker in accordance with the appropriate density. Finally, all workers observe the task and wage assignments and decide where to work. Tasks are performed and then $\Delta \beta$ of the workers exit the workforce.

2.3.3 Equilibria

There are two main components to an equilibrium of this game. First, firms and workers must behave optimally in each period. That is, the firms employ both an optimal task assignment rule and an equilibrium wage assignment rule in each period. While, the members of each new cohort of workers invest only if their expected lifetime benefit at the time of birth exceeds their cost, and all workers choose firms in a sequentially rational way after any history of play. Second, firms have rational beliefs about the proportion of qualified workers in each period. Formally, an equilibrium of this model is described by a sequence of rational firm beliefs regarding the proportion of qualified workers in each group, $\{\Pi^J_t\}_{t=0}^\infty$ for $J = B, W$.

The dynamic system

In order to develop the dynamic system that describes the equilibria, I must first address the optimal behavior of both the firms and the workers. In any equilibria the firms’ beliefs about the proportion of a group that is qualified will equal the true investment rate; I use $\Pi_t$ to represent both of these values.

From the perspective of the firms, each period of the dynamic game looks like the static
game of Section 2.2 with the exception of the aforementioned production technology adjust-
ment. Consequently, the equilibrium task assignment rule for each period will be identical to
that of the static model and the wage assignment rule is only slightly modified to account for
the period length. It is without loss of generality to focus on task assignment rules with the
threshold property. In the period starting at time $t$ the optimal threshold is

$$
\hat{\theta}(\Pi^j_t) = \begin{cases} 
1 & \text{if } \alpha_c p(1, \Pi^j_t) \leq \alpha_s \\
0 & \text{if } \alpha_c p(0, \Pi^j_t) \geq \alpha_s,
\end{cases}
$$

(2.11)

the unique solution to

$$
\alpha_c p(\theta_t, \Pi^j_t) = \alpha_s \quad \text{if } \alpha_c p(0, \Pi^j_t) < \alpha_s < \alpha_c p(1, \Pi^j_t),
$$

and the continuation wage is

$$
\omega_{\Delta}(\theta_t, \Pi^j_t) = \begin{cases} 
\Delta \alpha_s & \text{if } \theta < \hat{\theta}(\Pi^j_t) \\
\Delta \alpha_c p(\theta_t, \Pi^j_t) & \text{if } \theta \geq \hat{\theta}(\Pi^j_t).
\end{cases}
$$

(2.12)

Note that $\omega_{\Delta}(\theta_t, \Pi^j_t) = \Delta \omega(\theta_t, \Pi^j_t)$, where $\omega(\theta_t, \Pi^j_t)$ is as in (2.7).

Some slight adjustments to the other relevant functions are also necessary. I define the
unit period gain from investment as a worker’s one period gain from investment given a period
with length one. This is equivalent to the gain from investment in the static model and can
be written as

$$
\gamma(\Pi^j_t) = \alpha_s [F_q(\hat{\theta}(\Pi^j_t)) - F_u(\hat{\theta}(\Pi^j_t))] + \alpha_c \int_{\hat{\theta}(\Pi^j_t)}^1 p(\theta, \Pi^j_t) [f_q(\theta) - f_u(\theta)] d\theta.
$$

(2.13)

Given that the length of a period is $\Delta$, I can write the per period gain from investment as

$$
\Delta \gamma(\Pi^j_t).
$$

The lifetime gain from investment from the perspective of a worker at time $t$ is simply the
discounted sum of all future per period gains from investment. It can be written as

\[
\Gamma_t^J = \sum_{g=0}^{\infty} (1 - \Delta \beta - \Delta r)^g \Delta \gamma (\Pi_{t+g\Delta})
\]

\[
= (1 - \Delta \beta - \Delta r) \Gamma_{t+\Delta}^J + \Delta \gamma (\Pi_t^J).
\]  

(2.14)

Rearranging (2.14) yields

\[
\frac{\Gamma_{t+\Delta}^J - \Gamma_t^J}{\Delta} = (\beta + r) \Gamma_{t+\Delta}^J - \gamma (\Pi_t^J).
\]  

(2.15)

If I let \( \Delta \to 0 \), then I can express how the lifetime gain from investment evolves over time with the following differential equation:

\[
\dot{\Gamma}_t^J = (\beta + r) \Gamma_t^J - \gamma (\Pi_t^J).
\]  

(2.16)

A worker’s optimal strategy is to invest in human capital if his lifetime gain from investment at the time of birth exceeds his cost of investment. Workers can invest only in the period that they are born, so the investment rate of any cohort is fixed after the period of birth. Using this fact, I can address how the investment rate of each group’s population will develop over time. The proportion of new J population births in the period starting at time \( t \) that will invest is \( G(\Gamma_t^J) \).

Given the Poisson death process and the assumption that the total population is constant, the investment rate of the entire J population at time \( t \) is

\[
\Pi_t^J = (1 - \Delta \beta) \Pi_{t-\Delta}^J + \Delta \beta G(\Gamma_t^J).
\]  

(2.17)

Rearranging (2.17) yields

\[
\frac{\Pi_t^J - \Pi_{t-\Delta}^J}{\Delta} = \beta (G(\Gamma_t^J) - \Pi_{t-\Delta}^J).
\]  

(2.18)

If I let \( \Delta \to 0 \), then I can express how the overall investment rate of group J evolves over time
with the following differential equation:

\[ \dot{\Pi}_t^J = \beta \left( G(\Gamma_t^J) - \Pi_t^J \right). \tag{2.19} \]

I conclude that the equilibria of the dynamic model are characterized by a dynamic system consisting of (2.16) and (2.19).

**Proposition 3.** An equilibrium of this dynamic model is fully characterized by the following two-variable differential equations:

\[
\begin{align*}
\dot{\Gamma}_t^J &= (\beta + r) \Gamma_t^J - \gamma(\Pi_t^J), \\
\dot{\Pi}_t^J &= \beta \left( G(\Gamma_t^J) - \Pi_t^J \right). \tag{2.20}
\end{align*}
\]

This proposition indicates that the difference between the investment rate of the newborn cohort and the overall investment rate of the group \( J \) determines the speed of a firm’s beliefs about the fraction of workers in group \( J \) that are qualified. The change in lifetime gain from investment is determined by the difference between the discounted lifetime gain from investment and the unit period gain from investment.

The steady states of the dynamic system can be described as all beliefs satisfying

\[ \Pi_t^J = G \left( \frac{\gamma(\Pi_t^J)}{\beta + r} \right). \tag{2.21} \]

Describing the steady states like this makes it easy to compare the equilibria of the dynamic model to those of the static model. It is also useful for addressing the conditions under which there are multiple steady states.

**Existence of multiple steady states**

Discrimination in this dynamic model is a result of coordination failure just as in the static model. Consequently, in order to address the persistence of discrimination it is necessary to establish that it is possible for the dynamic system in Proposition 3 to have multiple steady states.
Figure 2.2: Loci of the general dynamic model

Figure 2.2 shows the two loci of interest in the general dynamic model and the associated directional arrows. The first panel depicts a potential $\dot{\gamma}_J = 0$ locus. See Section 2.2.3 for a discussion of the potential values of $\gamma(\Pi)$. The arrows indicate that to the right of the locus, when $\gamma_J > \frac{\gamma(\Pi_J)}{\beta + \gamma}$, $\gamma_J$ is increasing. To the left of the locus $\gamma_J$ is decreasing. The second panel of depicts the $\dot{\Pi}_J = 0$ locus. The shape of this locus is a result of the assumption that $G(\cdot)$ is continuous and strictly increasing. The arrows indicate that above this locus, when $G(\gamma_J) < \Pi_J$, $\Pi_J$ is increasing and below it the locus it is decreasing.

In Figure 2.3 I combine the two graphs of Figure 2.2. Examining the two loci of the dynamic system demonstrates that, as in the static case, for any $f_u$ and $f_q$ that satisfy the monotone likelihood ratio and any technology parameters I can find a non-trivial set of $G(\cdot)$ functions that ensure multiple steady states. I have drawn the figure to demonstrate a case in which there are three steady states; from now on I denote these three steady states as $\Sigma_l(\Gamma_l, \Pi_l)$, $\Sigma_m(\Gamma_m, \Pi_m)$, and $\Sigma_h(\Gamma_h, \Pi_h)$ referring to the low, middle, and high steady states, respectively.

Now that I have established the possibility of multiple solutions to (2.21), I can show that the steady states are Pareto rankable.

**Proposition 4.** Let $\Pi^*$ be the largest solution to (2.21). Then $\Sigma(\Gamma, \Pi) = \Sigma \left( \frac{\gamma(\Pi^*)}{\beta + \gamma}, \Pi^* \right)$ Pareto dominates all other steady states of the model.
Proof. Suppose that $\Pi^1$ and $\Pi^2$ solve (2.21) and let $\Pi^1 < \Pi^2$. If the system is initially at the low investment level $\Pi^1$, then I can show that if beliefs increase to $\Pi^2$ all workers are strictly better off even before changing their investment behavior. The change in the lifetime expected wage of a qualified worker when beliefs increase from $\Pi^1$ to $\Pi^2$ is

$$
\frac{1}{\beta + r} \int \omega^f(\theta; \Pi^2) f_q(\theta) - \frac{1}{\beta + r} \int \omega^f(\theta; \Pi^1) f_q(\theta)
= \frac{\alpha_s}{\beta + r} \int_{\hat{\theta}^{\Pi^2}}^{\hat{\theta}^{\Pi^1}} f_q(\theta) d\theta - \frac{\alpha_s}{\beta + r} \int_{\hat{\theta}^{\Pi^1}}^{\hat{\theta}^{\Pi^2}} f_q(\theta) d\theta
+ \frac{\alpha_c}{\beta + r} \int_0^1 p(\theta, \Pi^2) f_q(\theta) d\theta - \frac{\alpha_c}{\beta + r} \int_0^1 p(\theta, \Pi^1) f_q(\theta) d\theta
= \frac{1}{\beta + r} \left( \frac{\alpha_c}{\hat{\theta}^{\Pi^1}} \int_{\hat{\theta}^{\Pi^2}}^{\hat{\theta}^{\Pi^1}} [p(\theta, \Pi^2) - p(\theta, \Pi^1)] f_q(\theta) d\theta
+ \alpha_c \int_{\hat{\theta}^{\Pi^1}}^{\hat{\theta}^{\Pi^2}} (p(\theta, \Pi^2) - p(\theta, \Pi^1)) f_q(\theta) d\theta \right) > 0.
$$

Thus, the lifetime wage of those who invested in the original steady state is strictly improved by a move to the higher steady state. A similar argument can be made to show
that \( \frac{1}{\beta + r} \int \omega^J(\theta; \Pi^2) f_u(\theta) > \frac{1}{\beta + r} \int \omega^I(\theta; \Pi^1) f_u(\theta) \), which implies that the lifetime wage of those who did not invest in the original steady state is also strictly improved by a move to the higher steady state.

If workers are given the option to alter their investment decisions after the change in firm beliefs, then all workers who choose to do so will be strictly better off. This is implied as workers must behave optimally when making investment decisions. I conclude that if workers are given a choice of \( \Pi^1 \) or \( \Pi^2 \), they all prefer the higher \( \Pi^2 \). By revealed preference, the steady state with the largest \( \Pi \) that solves (2.21) is Pareto dominant.

\[ \square \]

**Relationship between the dynamic and static equilibria**

Proposition 3 indicates that all steady states of the dynamic system are characterized by beliefs satisfying \( \Pi^J = G \left( \frac{\gamma(\Pi^J)}{\beta + r} \right) \). Section 2.2 concludes that all equilibria of the static model are characterized by beliefs satisfying \( \pi^J = G(\gamma(\pi^J)) \). The only difference between these two conditions is that in the dynamic model a worker’s optimal decision to invest is based on the lifetime gain from investment rather than the unit period gain from investment. This similarity in conditions can also be seen by comparing Figure 2.1 and Figure 2.3. In these two figures the s-shaped curves are identical. The hump-shaped curve of Figure 2.3 is just a scaled, by \( \frac{1}{\beta + r} \), version of the curve in Figure 2.1.

A meaningful comparison of the static equilibria and the steady states of the dynamic model requires that the expected benefit from investment be equivalent across models. A slight adjustment to the production function of the static model achieves this goal. If I scale the production of the static model to \( \bar{Y}(S,C) = \frac{1}{\beta + r} \left[ \alpha_s S + \alpha_c C \right] \), then the optimal threshold is unchanged and the continuation wage is updated to

\[
\omega^J(\theta, \pi^J) = \begin{cases} 
\frac{\alpha_c}{\beta + r} & \text{if } \theta < \hat{\theta}(\pi^J) \\
\frac{\alpha_s}{\beta + r} p(\theta, \pi^J) & \text{if } \theta \geq \hat{\theta}(\pi^J)
\end{cases}
\]

(2.23)

This updated static wage is comparable to the lifetime wage of the dynamic model. Consequently, the new gain from investment in this static model is \( \frac{\gamma(\pi^J)}{\beta + r} \), where \( \gamma(\pi^J) \) is as in
Under this slightly adjusted static model the equilibria are characterized by all pairs of beliefs, \((\pi^B, \pi^W)\), satisfying

\[
\pi^J = G \left( \frac{\gamma(\pi^J)}{\beta + r} \right) \quad \text{for} \; J = B, W.
\]  

(2.24)

This new condition is identical to the steady state condition for the dynamic model. So, when the expected benefit from investment is equivalent across the two models, the static equilibria will equal the steady states of the dynamic model.

**Stability of the steady states**

Before I can talk about the situations in which discrimination is persistent, it is necessary to determine which of the steady states are stable and what the paths to those stable steady states look like. From now on I assume that \(G(\cdot)\) is such that there are three steady states \(\Sigma_l(\Gamma_l, \Pi_l)\), \(\Sigma_m(\Gamma_m, \Pi_m)\), and \(\Sigma_h(\Gamma_h, \Pi_h)\). This is the case of most interest. However, all of the results that follow can be generalized to cases with more than three steady states.

**Lemma 2.** Among the three steady states \(\Sigma_l(\Gamma_l, \Pi_l)\) and \(\Sigma_h(\Gamma_h, \Pi_h)\) are saddle points and \(\Sigma_m(\Gamma_m, \Pi_m)\) is a source.

*Proof.* The linearization of the dynamic system in Proposition 3 around a steady state \((\bar{\Pi}, \bar{\Gamma})\) is

\[
\begin{align*}
\dot{\Gamma} & = (\beta + r)\Gamma - \gamma(\Pi) + (\beta + r)(\Gamma_t - \Gamma) - \gamma'(\Pi)(\Pi_t - \Pi) \\
\dot{\Pi} & = \beta(G(\Gamma) - \bar{\Pi}) - \beta(\Pi_t - \bar{\Pi}) + \beta G'(\bar{\Gamma})(\Gamma_t - \bar{\Gamma}).
\end{align*}
\]  

(2.25)

Substituting in for \(\bar{\Gamma} = \frac{G(\bar{\Pi})}{\beta + r}\) and \(\bar{\Pi} = G(\bar{\Gamma})\) gives

\[
\begin{align*}
\dot{\Gamma} & = \gamma'(\Pi)(\Pi - \Pi_t) + (\beta + r)(\Gamma_t - \bar{\Gamma}) \\
\dot{\Pi} & = \beta G'(\bar{\Gamma})(\Gamma_t - \bar{\Gamma}) + \beta(\Pi - \Pi_t).
\end{align*}
\]  

(2.26)
The Jacobian matrix evaluated at a steady state is

\[ J_E \equiv \begin{bmatrix} \beta + r & -\gamma' \\ \beta G' & -\beta \end{bmatrix}. \]  \hspace{1cm} (2.27)

The trace of this matrix is \( trJ_E = r \), and the determinant is \( |J_E| = -\beta(\beta + r) + \beta'G' \). Since \( trJ_E \) is positive, every steady state is unstable. In order for a steady state to be a saddle point the eigenvalues must be of opposite signs. The eigenvalues of this system are:

\[
\phi_1, \phi_2 = r \pm \sqrt{r^2 - 4\beta[(\gamma'G') - (\beta + r)]} \hspace{1cm} (2.28)
\]

If \( 4\beta[\gamma'G' - (\beta + r)] < 0 \) the eigenvalues will be of opposite signs. This condition can be rewritten as: \( \frac{\gamma'}{\beta + r} < \frac{1}{G'} \).

For a steady state to be a saddle point, the slope of \( G \) must be less than that of the inverse of \( \Gamma_t \) at the steady state. Referencing Figure 2.3, it is apparent that both \( \Sigma_l \) and \( \Sigma_h \) are saddle points. This inequality does not hold for the middle steady state. Both eigenvalues are positive at \( \Sigma_m \), so it is either an unstable node or an unstable focus.

As \( \Sigma_m(\Gamma_m, \Pi_m) \) is a source, it is possible to determine the conditions under which the dynamic paths around it spiral out. Spiralling paths are of interest because they create a region of initial firm beliefs that can lead to either of the saddle stable steady states. If paths do not spiral, this region does not exist and historical beliefs completely determine the outcome for each group. I refer to such paths as deterministic paths.

**Lemma 3.** If \( \frac{(r+2\beta)^2}{\beta} < 4\gamma'(G')|_{(\Gamma_m, \Pi_m)} \) holds, then equilibrium paths spiral out in the neighborhood of \( \Sigma_m \). Otherwise, paths are deterministic.

**Proof.** For a steady state to be a unstable node the eigenvalues must be complex conjugate. This occurs when

\[
r^2 - 4\beta[(\gamma'G') - (\beta + r)] < 0. \hspace{1cm} (2.29)
\]

from Lemma 2.
Rearranging (2.29) yields

$$\frac{(r + 2\beta)^2}{\beta} < 4\gamma'G'.$$

(2.30)

**Corollary 1.** As $r$ decreases it becomes increasing more likely that the dynamic paths spiral out in the neighborhood of $\Sigma_m(\Gamma_m, \Pi_m)$.

**Proof.** The left hand side of (2.30) is increasing in $r$. So the smaller $r$ is, the more likely it is that (2.30) will hold. Note also that as $r$ decreases $4\beta[\gamma'G' - (\beta + r)] > 0$ will still hold.

This Corollary suggests that the less workers discount future pay-offs the less likely it is that initial firm beliefs will control the final outcome for each group.

The analysis of the paths to the saddle points is done under the assumption that workers discount the future at a positive rate $r > 0$. It is interesting to briefly consider what happens without this assumption. If $r = 0$ the paths leading from $\Sigma_m$ are no longer spiralling, instead they form closed circles around the steady state. In this case $\Sigma_m$ is a center and it is stable and $\Sigma_h$ and $\Sigma_l$ are saddle points. The situation in which the steady state changes stability and a unique limit cycle bifurcates from it is called a subcritical Andronov-Hopf bifurcation. See Kuznetsov (2006) for more information on dynamic systems with this property.

These stability results can easily be extended to a case where there are more than three steady states. In this particular model, as a result of the shape of the $\gamma$ function, there will always be an odd number of steady states. Suppose for example that there are five steady states. In this case the first, third and fifth steady states, counting up from the one with the lowest investment rate, will be saddle paths while the second and fourth will be sources. Depending on the parameter values, the paths from these two sources may be deterministic or spiralling.
Persistence of inequality

Now that I have identified which of the steady states are saddle points and the forms that the dynamic paths to the high and low steady states can take, I am able to determine the initial conditions that have the potential to yield persistent discrimination. In this dynamic setting, discrimination occurs when one group is at the high steady state while the other is at the low steady state. Persistent discrimination occurs when discrimination is present and neither group is able to move from their current steady state to another.

Since paths to the high and low steady states can take two forms depending on the parameters, there are two cases to address. When paths are deterministic, historical firms beliefs are the only factor that determine a group’s final outcome. On the other hand, when paths form interlocking spirals, historical firm beliefs may be indecisive and the expectations of the workers will determine the final outcome.

I refer to a group as disadvantaged if initial firm beliefs are such that the group can only converge to the low steady state. Conversely, a group is advantaged if firm beliefs lead them only to the high steady state. When paths are deterministic, all groups are either advantaged or disadvantaged. When paths are spiralling, it is possible for a group to be neither advantaged nor disadvantaged. In this case, the final outcome is determined by whether the group is optimistic or pessimistic about the future.

When the dynamic paths are deterministic, the steady state that a particular group converges to is entirely dependent on initial investment behavior. Figure 2.4 depicts potential deterministic paths to the high and low steady states. From this figure it is clear that as long as initial firm beliefs about one group are above those of the middle steady state while initial firm beliefs about the other group are below, there will be discrimination as defined above. Put more simply, discrimination occurs if one group is advantaged while the other is disadvantaged. Discrimination will always be persistent when paths are deterministic since there is no way for a group to move from one steady state to another.

**Proposition 5.** _If the dynamic paths to the high and low steady states are deterministic, then discrimination is persistent._
Figure 2.4: Deterministic paths to the steady states

Proof. This result is proven directly by the paths in Figure 2.4.

When the dynamic paths are spiralling, initial firm beliefs are potentially indecisive. If this is the case, then the steady state a group converges to will be determined by the group’s expectations about future payoffs rather than by historical firm beliefs.

Figure 2.5 depicts potential spiralling paths to the saddle points. I denote the upper boundary of the dynamic path to the low steady state as $\hat{\Pi}$. Note that this upper bound occurs as the path crosses the $\dot{\Pi} = 0$ locus. The figure indicates that if initial firm beliefs about a group are in the range $[0, \hat{\Pi}]$, then that group can converge to the low steady state. I denote the lower boundary of the dynamic path to the high steady state as $\check{\Pi}$. This bottom bound also occurs as the path crosses the $\dot{\Pi} = 0$ locus. If initial firm beliefs about a group are in the range $[\check{\Pi}, 1]$, then it is possible for that group to converge to the high steady state.

When $\hat{\Pi}$ is greater than $\check{\Pi}$, the spiralling paths create a range of initial firm beliefs that can lead to either the high or the low steady state depending on what path the group is on. This range is indicated by the purple area in Figure 2.5. The directional arrows of Figure 2.3
Figure 2.5: Spiralling paths to the steady states

and the assumption that paths to the steady states are spiralling imply that this range will always exist.

**Definition 1.** The range of initial firm beliefs $[\hat{\Pi}, \tilde{\Pi}]$ is called the “uncertain region”. If initial firm beliefs about the investment rate of a group are in this range, then that group can converge to either $\Sigma_h(\Pi_h, \Gamma_h)$ or $\Sigma_l(\Pi_l, \Gamma_l)$.

If initial firm beliefs are in the uncertain region, then the steady state that a group converges to is determined by the expectations of the workers in that group. If a group develops and shares an optimistic view of the future, then they will converge to the high steady state; if they develop a pessimistic view, they will converge to the low steady state. So, these expectations determine the path that a group will take. Consequently, I refer to the path that leads to the high steady state as the *optimistic path* and the path that leads to the low steady state as the *pessimistic path*. The existence of these two paths explains the situation where firms initially view the two groups as identical, in terms of average investment rate, but the groups end up at different steady states.

Given initial firm beliefs that are outside of the uncertain region, the steady state that a
group converges to will depend entirely on historical firm beliefs. In particular, when initial beliefs about a group’s average investment behavior are in the range \((\hat{\Pi}, 1]\), indicated by the pink region, then the group will converge to the high steady state with certainty. This is because the optimistic path is the only reasonable path to take from this starting point; put more simply, \(\Sigma_h\) is the only achievable saddle point.

**Definition 2.** The range of initial firm beliefs \([\hat{\Pi}, 1]\) is called the “certain region for \(\Sigma_h\)”.

If initial firm beliefs about the investment rate of a group are in this range, then that group can only converge to \(\Sigma_h(\Pi_h, \Gamma_h)\). Groups with initial firm beliefs in the certain region for \(\Sigma_h\) are advantaged since history, rather than optimism, causes them to converge to the high steady state.

If initial firm beliefs are in the range \([0, \tilde{\Pi})\), indicated by the blue region, then group \(J\) will converge to the low steady state with certainty.

**Definition 3.** The range of initial firm beliefs \([0, \tilde{\Pi}]\) is called the “certain region for \(\Sigma_l\)”.

If initial firm beliefs about the investment rate of a group are in this range, then that group can only converge to \(\Sigma_l(\Pi_l, \Gamma_l)\). Groups with initial firm beliefs in the certain region for \(\Sigma_l\) are disadvantaged since history, rather than pessimism, causes them to converge to the low steady state.

If the two saddle stable steady states lie inside of their respective certain regions, then, once reached, a group is permanently stuck in their steady state. The following lemma indicates that it is always the case that the high and low steady states lie outside of the uncertain region, implying that discrimination is persistent when paths are spiralling.

**Lemma 4.** If dynamic paths to the high and low steady states are spiralling, then \(\Pi_h > \hat{\Pi}\) and \(\Pi_l < \tilde{\Pi}\).

**Proof.** Fix the optimistic path. Note that \(\Pi\) is increasing to the right of the \(\hat{\Pi} = 0\) locus and decreasing to the left of it. Consequently, \(\hat{\Pi}\) must occur when the pessimistic path crosses the \(\hat{\Pi} = 0\) locus. Suppose that \(\Pi_h \leq \hat{\Pi}\), then the pessimistic path must cross the optimistic path (see Figure 2.5 for clarification). Dynamic paths cannot cross (see de la Fuente (2000), page
402), so it must be the case that the uppermost part of the pessimistic path lies below $\Pi_h$. I conclude that $\Pi_h > \hat{\Pi}$.

Fix the pessimistic path; $\hat{\Pi}$ will occur when the optimistic path crosses the $\hat{\Pi} = 0$ locus. Suppose that $\Pi_l \geq \hat{\Pi}$, then the optimistic path must cross the pessimistic path. Dynamic paths cannot cross, so it must be the case that the lowermost part of the optimistic path lies above the $\Pi_l$. I conclude that $\Pi_l < \hat{\Pi}$.

When paths to the high and low steady states are spiralling there are four classes of initial conditions that lead to discrimination. If initial firm beliefs are such that one group is disadvantaged while the other is advantaged, then discrimination is certain (just as in the case of deterministic paths). Alternatively, if initial firm beliefs are in the uncertain region for both groups, then discrimination can occur as long as one group is optimistic while the other is pessimistic. There will also be discrimination if beliefs and expectations are such that one group is advantaged and the other is pessimistic or if one group is disadvantaged while the other is optimistic. These last two cases are interesting because both history and expectations are impacting the final outcome. No matter how the discrimination occurs it will be persistent.

**Proposition 6.** If the dynamic paths to the high and low steady states are spiralling, then discrimination is persistent.

**Proof.** This result is directly implied by Lemma 4.

Propositions 5 and 6 indicate that discrimination, once present, will always be persistent in this model. The analysis also shows that both historical firm beliefs and group expectations can lead to inequality, depending on the shape of the dynamic paths. Furthermore, the existence of spiralling paths to the high and low steady states provides an explanation for the situation where firms have similar or identical beliefs across groups but the two groups converge to different steady states. Thus, the questions that were unanswerable in the static model have now been addressed in this dynamic framework.
2.4 The simple dynamic model

In this section, I develop a simple form of the dynamic model and I examine two specific parameterizations that generate three steady states. A group’s expectations about the future may play a role in the final outcome in the first parameterization. While, initial firm beliefs are decisive in the second parameterization. Using these two parameterizations, I am able to examine the impact that the precision of signals, the productivity of workers, the discount rate and the rate of population turnover have on the existence of three steady states and on the importance of group expectations.

2.4.1 Firms and workers

Firms face the same linear per period production function as in Section 2.3.1. For simplicity, workers have only two possible signals: $\theta_H$ and $\theta_L$, where $\theta_H > \theta_L$. If a worker is qualified, he receives a signal of $\theta_H$ with probability $P_q > \frac{1}{2}$ and a signal of $\theta_L$ with probability $1 - P_q$. If a worker is unqualified, he receives a signal of $\theta_L$ with probability $P_u > \frac{1}{2}$ and a signal of $\theta_H$ with probability $1 - P_u$. I assume the cost of investment is distributed uniformly over $[c, \bar{c}]$.

Time in this simple model is discrete, so $\Delta = 1$.

2.4.2 The updated dynamic system

The equilibria of this example are characterized by the same dynamic system as in Proposition 3. Given the parameterization, I can find more specific forms of the relevant functions. The updated optimal threshold rule is

$$\hat{\theta}(\Pi_l^f) = \begin{cases} 1 & \text{if } \Pi_l^f \leq \Pi \\ \theta_H & \text{if } \Pi < \Pi_l^f < \Pi \\ 0 & \text{if } \Pi_l^f \geq \Pi, \end{cases}$$

(2.31)

where $\Pi$ and $\bar{\Pi}$ are calculated by updating $\hat{\theta}(\Pi_l^f)$; they are

$$\Pi = \frac{\alpha_s(1 - P_u)}{\alpha_s(1 - P_u) + (\alpha_c - \alpha_s)P_q} \quad \text{and} \quad \bar{\Pi} = \frac{\alpha_sP_u}{\alpha_sP_u + (\alpha_c - \alpha_s)(1 - P_q)}.$$
Given the optimal task assignment rule in (2.31), the unit period gain from investment is updated to

\[
\gamma(\Pi_J^t) = \begin{cases} 
0 & \text{if } \Pi_J^t \leq \Pi \\
(P_q + P_u - 1)(\alpha_c P(\theta_H, \Pi_J^t) - \alpha_s) & \text{if } \Pi < \Pi_J^t < \Pi \\
\alpha_c(P_q + P_u - 1)(P(\theta_H, \Pi_J^t) - P(\theta_L, \Pi_J^t)) & \text{if } \Pi_J^t \geq \Pi.
\end{cases}
\] (2.33)

The posterior probability that a worker with signal \(\theta_H\) is qualified is

\[
P(\theta_H, \Pi_J^t) = \frac{\Pi_J^t P_q}{\Pi_J^t P_q + (1 - \Pi_J^t)(1 - P_u)},
\] (2.34)

and the posterior probability that a worker with signal \(\theta_L\) is qualified is

\[
P(\theta_L, \Pi_J^t) = \frac{\Pi_J^t(1 - P_q)}{\Pi_J^t(1 - P_q) + (1 - \Pi_J^t)P_u}.
\] (2.35)

The two loci of the dynamic system are now \(\Gamma_J^t = \gamma(\Pi_J^t)\beta + r\), where \(\gamma(\Pi_J^t)\) is as in (2.33), and \(\Pi_J^t = G(\Gamma_J^t)\), where \(G\) is the uniform distribution over \([c, \bar{c}]\).

### 2.4.3 Example 1: Spiralling paths

I start by examining a parametric example of the simple dynamic model that yields three steady states and dynamic paths that spiral out from the middle state to the high and low steady states. The values of the relevant variables are: \(P_q = 2/3\), \(P_u = 2/3\), \(\alpha_s = 1\), \(\alpha_c = 2\), \(\beta = 0.2\), \(r = 0.05\), \(\zeta = -0.1\) and \(\bar{c} = 0.9\). Figure 2.6 demonstrates the phase diagram for the simple dynamic model under these parameters. There are three steady states and, following the convention from the previous section, I denote these as \(\Sigma_l(\Gamma_l, \Pi_l)\), \(\Sigma_m(\Gamma_m, \Pi_m)\), and \(\Sigma_h(\Gamma_h, \Pi_h)\). The directional arrows of the phase diagram indicate that \(\Sigma_h\) and \(\Sigma_l\) are saddle points and \(\Sigma_m\) is an unstable source.

Figure 2.7 demonstrates the spiralling paths to the two saddle points. The steady state beliefs in this example are \(\Pi_l = 0.1\), \(\Pi_m = 0.47\), and \(\Pi_h = 0.76\). The paths are generated by inducting backwards from the high and low steady states using the equations in (2.14) and...
Figure 2.6: Phase diagram for example 1

Figure 2.7: Spiralling paths to the steady states
(2.17). The figure indicates that there is an uncertain region (the purple area) that results from the spiralling paths. There are also certain regions for both $\Sigma_h$ (the pink area) and $\Sigma_l$ (the blue area). Most importantly, because the high and low steady states are outside of the uncertain region, discrimination in this simple example will be persistent.

Using the simple dynamic model, I can also examine how changes in the various parameters of the model impact the presence of the three steady states holding the other parameter values constant. Figure 2.8 has three panels indicating the regions of $P_u$ and $P_q$; $\alpha_s$ and $\alpha_c$; and $\beta$ and $r$ that lead to the existence of a high, middle, and low steady state. In each of these panels the light red region indicates the parameter values that generate a low steady state. The grey region indicates the parameter values that generate a high steady state. Where these regions overlap, as indicated by the dark red region, the simple dynamic model has three steady states.

The first panel of Figure 2.8 demonstrates the values of $P_u$ and $P_q$ for which three steady states will be present. The figure indicates that there are three steady states when signals are of intermediate precision. When signals are very precise, only the high steady state is present and when signals are very noisy, only the low steady state is present.

The second panel of Figure 2.8 demonstrates the impact that the productivity of workers has on the existence of multiple steady states. I assume to start that $\alpha_s < \alpha_c$ or else the model is trivial. This figure indicates that when the productivity of workers in the two tasks is very similar, then there is only one steady state with a relatively low investment rate. This makes sense as the productivity determines the wage in this model and similar wages across tasks make investment unattractive. If the productivity of the complex task is much larger than that of the simple task, then there is only one steady state where a relatively large proportion of the population is qualified. This stems from the significant wage differential that will arise across tasks when productivity differences are large.

The third panel of Figure 2.8 demonstrates how the speed of population turn over, $\beta$, and the workers’ discount rate, $r$, impact the existence of multiple steady states. I assume to start that $\beta + r \leq 1$. The figure indicates that there is no $\beta$ or $r$ that can eliminate the low steady state, so there is no exclusively red region in this panel. If $\beta$ or $r$ is very large, then there is only one steady state at a low level of investment. If, on the other hand, $\beta$ or $r$ is low, there
Figure 2.8: Impact of the parameter values on the existence of the steady states

Figure 2.9: Impact of the parameter values on the shape of the dynamic paths

are three steady states.

Figure 2.9 details how the various parameters of the model impact the shape of the paths to the high and low steady states. There are three panels indicating the values of $P_u$ and $P_q$; $\alpha_s$ and $\alpha_c$; and $\beta$ and $r$ that generate spiralling paths and deterministic paths to the high and low steady states. Each of these three panels shows the region of the parameter values that allow for three steady states, as determined in Figure 2.8, and breaks it down into the values that meet the conditions of Lemma 3 and the values that do not. The dark purple indicates values that create the potential for group expectations to determine final outcome, while the light purple indicates the values for which historical firm beliefs are decisive.
The first panel of Figure 2.9 demonstrates the values of \( P_u \) and \( P_q \) there will be spiralling paths to the high and low steady states. The second panel does the same for \( \alpha_s \) and \( \alpha_c \). These panels indicate that, in this particular example, all values of \( P_u \) and \( P_q \) and of \( \alpha_s \) and \( \alpha_c \) that allow for three steady states also meet the requirements for spiralling paths.

The third panel of Figure 2.9 demonstrates the impact that \( \beta \) and \( r \) have on the shape of the paths to the high and low steady states. This relationship was already mentioned in Corollary 1. This panel is slightly different from the first two because there are values of \( \beta \) and \( r \) that yield three steady states but do not meet the conditions necessary for spiralling paths. This panel makes it clear that for high \( r \) values or low \( \beta \) values there can be three steady states, but the initial firm beliefs will determine the steady state a group converges to. An example with deterministic paths is discussed in the next subsection.

### 2.4.4 Example 2: Deterministic paths

It is also interesting to examine a parametric example of the simple dynamic model that yields three steady states with deterministic paths to the high and low steady states. The values of the relevant variables are: \( P_q = 19/20, P_u = 1/2, \alpha_s = 1, \alpha_c = 2, \beta = 0.5, r = 0.5, \bar{c} = -0.1 \) and \( \bar{c} = 0.4 \). Figure 2.10 demonstrates the relevant loci and the dynamic paths to the two saddle stable steady states. Following the convention from the previous section, I denote the three steady states as \( \Sigma_l(\Gamma_l, \Pi_l) \), \( \Sigma_m(\Gamma_m, \Pi_m) \), and \( \Sigma_h(\Gamma_h, \Pi_h) \). The steady state beliefs in this example are \( \Pi_l = 0.2, \Pi_m = 0.54, \) and \( \Pi_h = 0.92 \). The paths are generated by inducting backwards from the high and low steady states using the equations in (2.14) and (2.17). In this parameterization the paths are deterministic; there is no uncertain region and historical beliefs are decisive.

Again, I can examine the different parameter values that generate three steady states and that generate spiralling paths to the high and low steady states. Figure 2.11 demonstrates the regions of \( P_u \) and \( P_q \); \( \alpha_s \) and \( \alpha_c \); and \( \beta \) and \( r \) that yield three steady states. These results look very similar to the previous example with the grey regions indicating parameter values that lead to a high steady state and the light red regions indicating parameter values that lead to a low steady state.
This second example confirms the previous conclusion that moderately precise signals are more likely to generate three steady states. It also confirms the fact that three steady states are more likely the larger is the difference between the productivity of the complex task and the simple task. The third panel of Figure 2.11 indicates that for all values of $\beta + r \leq 1$ there will be three steady states. In this example, the rate of population turn over and the discount rate of the population have no effect on the existence of three steady states. This stems from the fact that signals for qualified workers are very precise while signals for unqualified workers are imprecise.

Figure 2.12 demonstrates the regions of the different parameter values that generate the two types of dynamic paths. The results are more interesting than those for the previous example. For all three panels there are regions in which there can be three steady states and deterministic paths, indicated by the light purple region of the panel. There are also regions where there can be three steady states and spiralling paths, indicated by the dark purple region. This was not the case in the Example 1. This result stems from the different set of parameter values in this example. Particularly the fact that the signals are more precise for
qualified workers than for unqualified and the fact that the rate of population turnover is large make it more likely that historical firm beliefs will determine the steady state that a group converges to.

### 2.5 Rationalizing the existence of multiple saddle points

It may not seem clear how to interpret the existence of two saddle points or how a group can converge to one or the other depending on expectations. In order to rationalize the two paths to the saddle points, I make a slight adjustment to the simple dynamic model. Rather than
costs being distributed uniformly, I assume that some proportion of the population, $l$, is born qualified, or have zero cost of investment so that they always invest. While some proportion of the population, $h$, has an extremely high cost of investment and will never invest. The rest of the population, $1 - l - h$, has an intermediate cost of investment, $c$.

By selecting the values: $l = 0.1$, $h = 0.76$, $c = 0.37$, $P_q = 2/3$, $P_u = 2/3$, $\alpha_q = 1$, $\alpha_c = 2$, $\beta = 0.2$ and $r = 0.05$, I can generate the same three steady states as in Example 1. The only change to the dynamic system is that the $\dot{\Pi} = 0$ locus is now a step function, as depicted in Figure 2.13.

Suppose that initial firm beliefs are those of the unstable middle steady state. That is $\Pi_0 = 0.47$. The only workers that are of interest in this demonstration are those that have a cost $c$. The behavior of the other types of workers is predetermined. The optimal behavior of a worker with an intermediate cost is to invest if lifetime gain from investment, $\Gamma(\Pi^J)$, exceeds $c$.

To start, consider a worker with an intermediate cost of investment, and suppose that this worker is extremely pessimistic and believes that all workers with intermediate costs will
choose not to invest. I can calculate the series of firm beliefs that the worker anticipates given this pessimism as \( \{\Pi^P_t\}_{t=0}^\infty \), where initial beliefs are \( \Pi^P_0 = 0.47 \) and all following beliefs are calculated as \( \Pi^P_t = (1 - \beta)\Pi^P_{t-1} + \beta l \). Under these most pessimistic beliefs, the worker has a lifetime gain from investment of \( \Gamma^P_0 = 0.13 \). This is less than the cost of investment, \( c = 0.37 \), so the worker will not invest. Thus, given that a worker believes that all workers with intermediate costs will not invest, his optimal behavior is to not invest. Figure 2.13 demonstrates the path that will be taken to the low steady state if all workers with cost \( c \) are pessimistic in this way.

Consider instead that the worker is optimistic and believes that all workers with intermediate costs will invest. The series of expected firm beliefs under these assumptions are \( \{\Pi^O_t\}_{t=0}^\infty \), where again initial beliefs are \( \Pi^O_0 = 0.47 \) and all following beliefs are calculated as \( \Pi^O_t = (1 - \beta)\Pi^O_{t-1} + \beta h \). In this case, the worker has a lifetime gain from investment of \( \Gamma^O_0 = 0.59 \), which exceeds his investment cost, so he will choose to invest. I can conclude that if a worker believes that all workers with intermediate costs will invest, his optimal behavior is to invest as well. Figure 2.13 demonstrates the path that will be taken to the high steady state if all workers with cost \( c \) are optimistic in this way. This simple demonstration shows clearly how group expectations can determine the path that a group takes to a steady state.

### 2.6 Concluding remarks

The two main contributions of this paper are to establish the initial conditions that lead to discrimination and to determine when this discrimination will be persistent. There are two potential forms that the paths to the relevant steady states can take. If the paths are deterministic, then there will be inequality as long as initial firms beliefs are above \( \Pi_m \) for one group and below for the other; this discrimination will always be persistent. If paths are spiralling and initial firm beliefs are in the uncertain region, it will be the collective group expectations that determine the final outcome. If beliefs are outside of this region, then group expectations have no bearing on the final outcome and as in the deterministic case, historical firm beliefs are decisive. Discrimination is always persistent when paths are spiralling.
In the framework of a simplified version of the dynamic model, I determine how the parameter values affect the existence of three steady states and the shapes of the paths to those steady states. I find that the rate of population turnover and the discount rate of the workers are the main driving forces behind the form that the dynamic paths take. Similarly, the precision of signals regarding worker quality and the productivity of workers in the complex and simple tasks are the parameters that have the most impact on the existence of three steady states.

Another contribution of this paper is to address the dynamics of discrimination in a competitive environment. Allowing for competitive wages means workers in different groups that are assigned to the same task and have the same signal may have different wages. So, inequality in this model can occur via both the task assignment and the wage. A competitive environment also allows for a more realistic analysis of potential policies designed to eliminate discrimination and makes it possible to perform meaningful welfare analysis in future research.
Chapter 3

Eliminating persistent statistical discrimination: an analysis of several policy options

3.1 Introduction

This paper revisits the dynamic model of statistical discrimination developed in Chapter 2 and analyzes the effectiveness of three government policies designed to eliminate existing persistent discrimination. By examining these policies in a dynamic framework, I am able to determine the path that workers in a group will take once a policy is put in place and the point at which it is safe for the government to remove the policy and still guarantee the eventual elimination of discrimination. Furthermore, the inclusion of competitive wages in the model means that I am able to complete meaningful analysis of the welfare effects of the three policies.

I use the dynamic model developed in Chapter 2 as the framework to analyze the three policies. The model may have three steady states, in which case the middle steady state is unstable and paths to the high and low steady states emanate from that point. These paths can take several forms depending on the parameter values. For certain parameterizations, the dynamic paths create a region of initial firm beliefs, about the average skill level of a group, in which a group can converge to either the high or the low steady state depending on whether they are pessimistic or optimistic. For other parameterizations, this region does not exist. When this is the case, initial firm beliefs about the average skill level of a group are decisive and group expectations are irrelevant. No matter what form the dynamic paths take, a group becomes “stuck” in a steady state soon as it is reached.

For the purpose of this analysis, I assume that the economy is experiencing discrimination.
This means that one group of workers, which I refer to as *disadvantaged*, is stuck in the low steady state and the other group of workers, which I refer to as *advantaged*, is stuck in the high steady state. I examine the impact of three government policies from this starting point. The goal of these policies is to move the disadvantaged group from the low steady state to the high steady state.

The first policy I examine is one in which the government offers a *hiring subsidy* to firms for correctly assigning skilled workers in the disadvantaged group to the more productive job. This policy is always effective at eliminating discrimination as long as a large enough subsidy is offered.

The second policy of interest is an *investment voucher* offered to workers in the disadvantaged group. This voucher defrays the monetary cost associated with skill acquisition. As long as the monetary costs are high, this policy can eliminate discrimination. If monetary costs are low, then it may be the case that even the largest effective voucher is unable to eliminate discrimination.

The last policy is one in which the government enforces *equal treatment* of workers across groups. In particular, firms are not allowed to condition wages or job assignments on group identity. Under this policy, the dynamic system has three steady states. If the proportion of the population that is in the disadvantaged group is very small, then this policy is guaranteed to eliminate discrimination. If the groups are close to equal in size, then the policy may eliminate discrimination, depending on the shape of the paths to the steady states and on whether the population as a whole is optimistic. If the disadvantaged group is relatively large, then this policy will have disastrous results.

Using a simple parameterization, I examine the welfare effects of the three policies. I complete this analysis in a general equilibrium framework. Workers in the advantaged group are taxed in order to fund the policies. I find that all three policies create a net increase in welfare while they are in place. However, there are significant differences across policies with respect to how much the advantaged group is impacted over the course of the policy. The hiring subsidy results in a very large loss in welfare to the advantaged group, but it moves the disadvantaged group to the high steady state the fastest. The investment voucher reduces
the welfare of the advantaged group the least of all the policies and for this reason it may be viewed as optimal by some policy makers. The equal treatment policy has a moderate impact of the advantaged group’s welfare and may be viewed as optimal by a policy maker that does not want to redistribute.

The paper is structured as follows. The related literature is discussed in the remainder of Section 3.1. The details of the general dynamic model are described in Section 3.2. Section 3.3 develops the equilibria of the dynamic model and addresses the conditions under which discrimination will be persistent. In Section 3.4, a simple parameterization of the dynamic model is developed. Section 3.5 describes the three policies and examines their impact on both the general dynamic model and the simple parameterization. In Section 3.6, I consider the welfare effects of the three policies for the simple parameterization. Finally, Section 3.7 concludes.

Related Literature

The model presented in this paper is closely related to Coate and Loury (1993) and Moro and Norman (2003). Coate and Loury (1993) present a model of statistical discrimination where two ex ante groups may end up in different, Pareto ranked, equilibria. This model formalizes many of ideas that were originally presented loosely in Arrow (1973), but it makes the additional assumption that wages are set exogenously. They find that inequality will occur if the two groups of workers play different equilibria.

Moro and Norman (2003) relax two of the assumptions in Coate and Loury’s model. In particular, they remove the linearity of the production technology and the exogeneity of wages. Allowing for curvature in the production function eliminates the separability of groups, thus, allowing for true interaction effects between groups. Introducing endogenous wages allows for more meaningful welfare analysis. They find that under these relaxed assumptions the dominant group will actually gain from discrimination and, as a result, they will be motivated to ensure its continuation.

There is relatively little literature that addresses discrimination in a dynamic environment. Fryer (2007), Blume (2006) and Kim and Loury (2009) all develop dynamic models that are
adaptations of the original Coate and Loury (1993) paper. These models all differ from the one presented here in two key areas: they do not incorporate endogenous wages nor do they consider potential policies to eliminate or mitigate discrimination.

3.2 The model

Consider a market where short-lived firms engage in Bertrand style competition for workers in each period. Time is continuous in this dynamic model and all changes occur at the beginning of each period of length $\Delta$.

There are two identifiable groups of workers, B(lack) and W(hite), and the total population of each group is constant at $\lambda^J$ for $J = B, W$. Workers are long-lived and subject to a Poisson death process with parameter $\beta$. Consequently, at the beginning of each period $\Delta \beta$ of the existing workers are replaced by newborn individuals. Workers also discount future pay-offs at a strictly positive rate $r$. Prior to entering the labor market, each worker has the option to invest in human capital. Workers who make this investment become qualified and those that do not remain unqualified. These pre-market investment decisions are final; a worker cannot choose to invest in human capital later in life.

A worker who invests in human capital incurs a total cost, $c$. This cost of investment consists of two important parts: a monetary cost, $c_m$, which is strictly positive and fixed across all workers and non-monetary cost, $c_e$, which varies across workers. The non-monetary cost represents both the effort required to become qualified and any value that a worker places on being qualified (outside of the potential increase in expected wage); it may be positive or negative.

Total cost of investment will vary across workers because of the differences in $c_e$. I assume that total cost of investment, $c$, is distributed over $[c, \bar{c}] \subseteq \mathbb{R}$ according to a continuous and strictly increasing distribution $G(c)$. Workers are risk neutral and care only about the wage that they receive and the cost they incur. Pay-offs are additively separable in income and cost of investment.

Output in each period is generated by the completion of two types of tasks, a simple task
and a complex task. All workers can perform the simple task, but only those workers who are qualified can successfully perform the complex task. For a given period starting at \( t \), the effective input of labor into the simple task, \( S_t \), is the total number of workers employed in that task. The effective input of labor into the complex task, \( C_t \), is the number of qualified workers employed in that task. Both firms’ per period output is given by the linear production function \( Y_\Delta(S_t, C_t) = \Delta[\alpha_s S_t + \alpha_c C_t] \), where \( \alpha_c > \alpha_s > 0 \).

Firms cannot observe whether a worker is qualified, nor are they able to distinguish between new and old workers. The only information a firm has about a worker in a given period, aside from his group identity, is his current signal \( \theta_t \in [0,1] \), which is distributed according to \( f_u \) if a worker is unqualified or \( f_q \) if a worker is qualified. Both densities are bounded away from zero and, without further loss of generality, \( f_q(\theta)/f_u(\theta) \) is strictly increasing in \( \theta \).

A firm’s prior belief about what proportion of the total \( J \) population is qualified at time \( t \) is represented by \( \Pi_{Jt}^J \). The posterior probability that a worker from group \( J \) with signal \( \theta_t \) is qualified given beliefs \( \Pi_{Jt}^J \) is

\[
p(\theta_t, \Pi_{Jt}^J) \equiv \frac{\Pi_{Jt}^J f_q(\theta)}{\Pi_{Jt}^J f_q(\theta) + (1 - \Pi_{Jt}^J) f_u(\theta)}.
\]

(3.1)

The strict monotone likelihood ratio property implies that \( p(\theta_t, \Pi_{Jt}^J) \) is strictly increasing in \( \theta_t \) so that a high signal reflects positively on a worker and a low signal negatively. I denote the associated cumulative distributions by \( F_q \) and \( F_u \). I assume a law of large numbers holds so that these are also the realized frequency distributions of signals for qualified and unqualified workers, respectively.

The timing of each period of the dynamic game is as follows: In the first stage, \( \Delta \beta \) new workers enter the workforce and decide whether or not to invest in becoming qualified. At the same time, firms post wage and task assignment rules. Firms may condition wage and task assignments on \( \theta_t \). Formally, a strategy for each firm is to select some wage schedule \( \omega_t^J : [0,1] \to \mathbb{R}_+ \) and a task assignment rule \( \tau_t^J : [0,1] \to [0,1] \) for each group \( J \). In the second stage, nature assigns a signal \( \theta_t \) to each worker according to the appropriate density. Workers observe the posted wage and task assignment rules and decide where to work. Then tasks are
performed, wages are paid and output is realized. In the last stage, $\Delta \beta$ of the workers exit the workforce.

Workers only care about potential wages when comparing job offers and their investment costs are sunk. So they will choose to work for the firm whose wage and task assignment rule provides them with the highest wage.

### 3.3 Equilibria

There are two main components to an equilibria of this game. First, firms and workers must behave optimally in each period. This means that firms must employ both an optimal task assignment rule and an equilibrium wage rule in each period. While members of each new cohort of workers invest only if their expected lifetime benefit at the time of birth exceeds their cost, and all workers choose firms in a sequentially rational way after any history of play. Second, firms have rational beliefs about proportion of qualified workers in each period. Formally, an equilibrium of this model is described by a sequence of rational firm beliefs regarding the proportion of qualified workers in each group, $\{\Pi^J_t\}_{t=0}^\infty$ for $J = B, W$.

Let $(\Pi^B_t, \Pi^W_t)$ denote the group specific investment rates in the period starting at time $t$. Given this arbitrary investment behavior by workers, I can address the optimal behavior of firms in each period. The probability that a worker is qualified, as defined in (3.1), is strictly increasing in signal. This implies that threshold rules are optimal when assigning workers to tasks. Under this type of rule, workers with a signal above the threshold are assigned to the complex task and workers with lower signals are assigned to the simple task. Given a threshold $\bar{\theta}^J_t$ for group $J$ at time $t$ the effective input of labor into the complex task is $C^J_t = \Pi^J_t (1 - F_q(\bar{\theta}^J_t))$ and the effective input of labor into the simple task is $S^J_t = \Pi^J_t F_q(\bar{\theta}^J_t) + (1 - \Pi^J_t) F_u(\bar{\theta}^J_t)$.

The optimal threshold is the one which maximizes a firm’s expected output conditional on investment decisions. Formally, the task assignment problem is

$$\sum_{J = B, W} \lambda^J \max_{\theta^J_t \in [0, 1]} \left( \Delta \alpha_c \left[ \Pi^J_t (1 - F_q(\theta^J_t)) \right] + \Delta \alpha_s \left[ \Pi^J_t F_q(\theta^J_t) + (1 - \Pi^J_t) F_u(\theta^J_t) \right] \right), \quad (3.2)$$

which will always have a unique solution.
Lemma 5. For each group there is a unique $\hat{\theta}(\Pi^I_t) \in [0, 1]$ that solves (3.2) for any $\Pi^I_t > 0$.

Proof. Let $\hat{\theta}(\Pi^I_t)$ be any solution to (3.2). The effective factor inputs into the complex and simple task are

\[ C^I_t(\Pi^I_t) \equiv \Pi^I_t [1 - F_q(\hat{\theta}(\Pi^I_t))] \quad \text{and} \quad S^I_t(\Pi^I_t) \equiv \Pi^I_t F_q(\hat{\theta}(\Pi^I_t)) + (1 - \Pi^I_t) F_u(\hat{\theta}(\Pi^I_t)) \quad \text{for} \quad J = B, W, \]

(3.3)

respectively. Combining the equations of (3.3) yields:

\[ S^I_t = \Pi^I_t - C^I_t + (1 - \Pi^I_t) F_u \left( F_q^{-1} \left( \frac{\Pi^I_t - C^I_t}{\Pi^I_t} \right) \right), \]

(3.4)

which is the maximal level of $S^I_t$ given $C^I_t$. I can restate the task assignment problem in (3.2) as

\[
\max_{(C^I_t, S^I_t)} \left[ \Delta \alpha_c(\lambda^B C^B + \lambda^W C^W) + \Delta \alpha_s(\lambda^B S^B + \lambda^W S^W) \right] \\
\text{s.t.} \quad S^I_t \leq \Pi^I_t - C^I_t + (1 - \Pi^I_t) F_u \left( F_q^{-1} \left( \frac{\Pi^I_t - C^I_t}{\Pi^I_t} \right) \right) \quad \text{for} \quad J = B, W. \quad (3.5)
\]

The strict monotone likelihood ratio implies that the right hand side of the constraint is a strictly concave function of $C^I_t$. Problem (3.5) is a matter of maximizing a linear function, that is increasing in both arguments, subject to a strictly concave constraint. This yields a unique solution.

The linear production function allows for the task assignment problem to be solved separately for each group. Consequently, there is a unique optimal threshold for each group. It is

\[
\hat{\theta}(\Pi^I_t) = \begin{cases} 
1 & \text{if} \quad \alpha_c p(1, \Pi^I_t) \leq \alpha_s \\
0 & \text{if} \quad \alpha_c p(0, \Pi^I_t) \geq \alpha_s, \\
\text{the unique solution to} & \quad \alpha_c p(\theta, \Pi^I_t) = \alpha_s & \text{if} \quad \alpha_c p(0, \Pi^I_t) < \alpha_s < \alpha_c p(1, \Pi^I_t). 
\end{cases} \quad (3.6)
\]

Under this task assignment rule, all workers are assigned to the simple task (a threshold of 1)
if prior beliefs are such that a worker with a signal of 1 has higher expected productivity in the simple task. All workers are assigned to the complex task (a threshold of 0) if prior beliefs are such that a worker with a signal of 0 has higher expected productivity in the complex task. Otherwise, the optimal threshold is the \( \theta \) at which a worker’s expected productivity in the complex task is equivalent to that of the simple task.

Following Moro and Norman (2003), I call a strategy profile a \textit{continuation equilibrium} if firms are implementing their optimal wage and task assignment rules given some arbitrary investment behavior by workers. This first result suggests that wages are given by expected marginal products and job assignments are constrained efficient in any continuation equilibrium.

**Proposition 7.** Suppose for a given period beginning at \( t \) that fractions \( \Pi_t = (\Pi^B_t, \Pi^W_t) \) of the workers invest and that \( (\hat{\theta}_t(\Pi^B_t), \hat{\theta}_t(\Pi^W_t)) \) is as in (3.6). Then there exists a continuation equilibrium where both firms post wage schedules

\[
\omega_{\Delta}(\theta_t, \Pi^J_t) = \begin{cases} 
\Delta \alpha_s & \text{if } \theta_t < \hat{\theta}(\Pi^J_t), \\
\Delta \alpha_c p(\theta_t, \Pi^J_t) & \text{if } \theta_t \geq \hat{\theta}(\Pi^J_t),
\end{cases}
\]

(3.7)

and assign a worker with characteristics \((J, \theta_t)\) to the complex task if and only if \( \theta_t \geq \hat{\theta}_t(\Pi_t) \). Moreover, in any continuation equilibrium where fractions \( \Pi_t = (\Pi^B_t, \Pi^W_t) \) of the workers invest the posted wage schedule for group \( J, \omega_{\Delta}(\theta_t, \Pi^J_t) \), agrees with (3.7) for almost all \( \theta_t \in [0, 1] \) for each firm.

**Proof.** See Appendix A

In a full equilibrium, investments must be best responses to wages. The optimal investment decision for a worker is to invest in human capital if and only if the lifetime gain in expected earnings is higher than the cost \( c \). It is helpful to define what I refer to as the \textit{unit period gain from investment.} This is a worker’s one period expected gain from investment if periods have a length of one. For wages consistent with the continuation equilibrium, where a fraction \( \Pi^J_t \)
of the $J$ population is qualified, the unit period gain from investment is written as

$$
\gamma(\Pi^J_t) = \frac{1}{\Delta} \left( \int_0^1 \omega_{\Delta}(\theta, \Pi^J) f_q(\theta) d\theta - \int_0^1 \omega_{\Delta}(\theta, \Pi^J) f_u(\theta) d\theta \right)
= \alpha_s [F_q(\hat{\theta}(\Pi^J_t)) - F_u(\hat{\theta}(\Pi^J_t))] + \alpha_c \int_{\hat{\theta}(\Pi^J_t)}^1 p(\theta, \Pi^J) [f_q(\theta) - f_u(\theta)] d\theta.
$$

(3.8)

I can also define the *per period gain from investment* as $\Delta \gamma(\Pi^J_t)$.

It is useful to briefly discuss the potential values of the gain from investment function. From (3.6), I can show that there exists a $\Pi > 0$ such that all workers are assigned to the simple task if $\Pi^J_t \leq \Pi$. There is also some $\Pi < 1$ such that all workers are assigned to the complex task if $\Pi^J \geq \Pi$. I can also show that if $\Pi^J_t \leq \Pi$, then $\gamma(\Pi^J_t) = 0$. However, the gain from investment is strictly positive in the range $(\Pi, 1)$ because $p(\theta, \Pi^J)$ is increasing in signal. Thus, there is wage inequality even when all workers are assigned to the complex task. Also, $\gamma(\Pi^J_t) = 0$ when $\Pi^J_t = 1$. So, in general, $\gamma(\Pi^J_t) = 0$ for some low range of $\Pi^J_t$ values, and then it is strictly positive until $\Pi^J_t = 1$.

Using the per period gain from investment, I can define the *lifetime gain from investment* from the perspective of a worker at time $t$ in group $J$. It is simply the discounted sum of all future per period gains from investment. It can be written as

$$
\Gamma^J_t = \sum_{g=0}^{\infty} (1 - \Delta \beta - \Delta \tau)^g \Delta \gamma(\Pi^J_{t+g\Delta})
= (1 - \Delta \beta - \Delta \tau) \Gamma^J_{t+\Delta} + \Delta \gamma(\Pi^J_t).
$$

(3.9)

Rearranging (3.9) yields

$$
\frac{\Gamma^J_{t+\Delta} - \Gamma^J_t}{\Delta} = (\beta + \tau) \Gamma^J_{t+\Delta} - \gamma(\Pi^J_t).
$$

(3.10)

If I let $\Delta \to 0$, then I can express how the lifetime gain from investment evolves over time with the following differential equation:

$$
\dot{\Gamma}^J_t = (\beta + \tau) \Gamma^J_t - \gamma(\Pi^J_t).
$$

(3.11)
I can also address how the investment rate of each group’s population will develop over time. In each period, workers in the new cohort invest if their lifetime gain from investment exceeds their cost. The proportion of new J population births that invest in the period starting at time $t$ is $G(\Gamma^J_t)$, where $G(\cdot)$ is the distribution of costs.

Given the Poisson death process and the assumption that the total population is constant, the investment rate of the whole J population at time $t$ is

$$\Pi^J_t = (1 - \Delta \beta)\Pi^J_{t-\Delta} + \Delta \beta G(\Gamma^J_t).$$

(3.12)

Rearranging (3.12) yields

$$\frac{\Pi^J_t - \Pi^J_{t-\Delta}}{\Delta} = \beta \left( G(\Gamma^J_t) - \Pi^J_{t-\Delta} \right).$$

(3.13)

If I let $\Delta \to 0$, then I can express how the overall investment rate of group J evolves over time with the following differential equation:

$$\dot{\Pi}^J_t = \beta \left( G(\Gamma^J_t) - \Pi^J_t \right).$$

(3.14)

I conclude that any equilibria of this model is characterized by a dynamic system consisting of (3.11) and (3.14).

**Proposition 8.** An equilibrium of this dynamic model is fully characterized by the following two-variable differential equations:

$$\dot{\Gamma}^J_t = (\beta + r)\Gamma^J_t - \gamma(\Pi^J_t)$$

$$\dot{\Pi}^J_t = \beta \left( G(\Gamma^J_t) - \Pi^J_t \right).$$

(3.15)

This proposition indicates that the difference between the investment rate of the newborn cohort and the overall investment rate of the group J determines the speed of the firms’ beliefs about the fraction of workers in group J that are qualified. The change in lifetime gain from investment is determined by the difference between the discounted lifetime gain from
investment and the unit period gain from investment.

The steady states of the dynamic system can be described as all beliefs satisfying

$$\Pi^J = G\left(\frac{\gamma(\Pi^J)}{\beta + r}\right).$$

(3.16)

The existence of a solution to (3.16) is immediate since $\gamma(\Pi^J)$ is composed of continuous functions. There may be a unique solution, in which case groups are treated identically when they are both in the steady state. For the purposes of this analysis, multiple steady states are desirable since persistent discrimination is only possible if this is the case.

Figure 3.1 shows an example of the dynamic system with three steady states. From now on I denote these three steady states as $\Sigma_l(\Gamma_l, \Pi_l)$, $\Sigma_m(\Gamma_m, \Pi_m)$, and $\Sigma_h(\Gamma_h, \Pi_h)$, referring to the low, middle, and high steady states, respectively. The hump shaped curve is the $\dot{\Gamma}_t^J = 0$ locus. The arrows indicate that to the right of the locus, when $\Gamma_t^J > \gamma(\Pi_t^J)/(\beta + r)$, $\Gamma^J$ is increasing. To the left of the locus, $\Gamma_t^J$ is decreasing. The s-shaped curve is the $\dot{\Pi}_t^J = 0$ locus. The arrows indicate that above this locus, when $G(\Gamma_t^J) < \Pi_t^J$, $\Pi_t^J$ is increasing and below the locus it is decreasing. The shape is from the assumption that $G(\cdot)$ is continuous and strictly increasing.

The existence of multiple solutions is not guaranteed and depends on the shape of $\Gamma$ and $G(\cdot)$. The possibility of multiple solutions can be proven by construction: if I fix the parameters $f_u, f_q, \alpha_c$ and $\alpha_s$, then I can find an appropriate $G(\cdot)$ function such that (3.16) has multiple solutions. Graphically, if I fix the $\dot{\Gamma}_t^J = 0$ locus in Figure 3.1, I can ensure multiple equilibria simply by selecting a $G(\cdot)$ function that yields multiple intersections of the two loci. I can conclude that for any $f_u$ and $f_q$ that satisfy the monotone likelihood ratio and any technology parameters, I can find a non-trivial set of $G(\cdot)$ functions that ensure multiple steady states.

It is necessary to establish which of these three steady states are stable and to determine the potential shapes of the paths leading to steady states. Understanding the behavior of the system outside of the steady states is important both for determining if discrimination is persistent and for analyzing the effectiveness of the policies in Section 3.5.

The following Lemma describes the conditions that are necessary for a steady state in this
Lemma 6. A steady state \((\bar{\Gamma}, \bar{\Pi})\) is a saddle point if \(\frac{\gamma'}{\beta+r} < \frac{1}{G'(\bar{\Gamma}, \bar{\Pi})}\) holds. Otherwise, it is a source.

Proof. The linearization of the dynamic system in Proposition 8 around a steady state \((\bar{\Pi}, \bar{\Gamma})\) is

\[
\begin{align*}
\dot{\Gamma} &= (\beta + r)\bar{\Gamma} - \gamma(\bar{\Pi}) + (\beta + r)(\Gamma_t - \bar{\Gamma}) - \gamma'(\bar{\Pi})(\Pi_t - \bar{\Pi}) \\
\dot{\Pi} &= \beta(G(\bar{\Gamma}) - \bar{\Pi}) - \beta(\Pi_t - \bar{\Pi}) + \beta G'(\bar{\Gamma})(\Gamma_t - \bar{\Gamma}).
\end{align*}
\] (3.17)

Substituting in for \(\bar{\Gamma} = \frac{\gamma'(\bar{\Pi})}{\beta+r}\) and \(\bar{\Pi} = G(\bar{\Gamma})\) gives

\[
\begin{align*}
\dot{\Gamma} &= \gamma'(\bar{\Pi})(\bar{\Pi} - \Pi_t) + (\beta + r)(\Gamma_t - \bar{\Gamma}) \\
\dot{\Pi} &= \beta G'(\bar{\Gamma})(\Gamma_t - \bar{\Gamma}) + \beta(\bar{\Pi} - \Pi_t).
\end{align*}
\] (3.18)
The Jacobian matrix evaluated at a steady state is

\[ J_E \equiv \begin{bmatrix} \beta + r & -\gamma' \\ \beta G' & -\beta \end{bmatrix} \] (3.19)

The trace of this matrix is \( trJ_E = r \) and the determinant is \( |J_E| = -\beta(\beta + r) + \beta \gamma' G' \). Since \( trJ_E \) is positive, every steady state is unstable. In order for a steady state to be a saddle point the eigenvalues must be of opposite signs. The eigenvalues of this system are:

\[ \phi_1, \phi_2 = \frac{r \pm \sqrt{r^2 - 4\beta[(\gamma' G') - (\beta + r)]}}{2} \] (3.20)

If \( 4\beta[\gamma' G' - (\beta + r)] < 0 \), the eigenvalues will be of opposite signs. This condition can be rewritten as: \( \frac{\gamma'}{\beta + r} < \frac{1}{\beta G'} \).

This Lemma suggests that a steady state is a saddle point when the slope of \( \Gamma_t \) is less than that of the inverse of \( G(\cdot) \) at the steady state. Referencing Figure 3.1, it is apparent that both the high and low steady states are saddle points. This inequality does not hold for the middle steady state so it is either an unstable node or an unstable focus.

The existence of two saddle points means that members of a group may rationally conjecture that firms will eventually have positive beliefs about their group’s investment rate, \( \Pi_h \), or have negative beliefs about their group’s investment rate, \( \Pi_l \). The only reasonable steady states for a group to approach are \( \Sigma_h \) and \( \Sigma_l \). It is possible to characterize the shape that the paths to these saddle stable steady states will take. Depending on the parameter values the dynamic paths will be either deterministic or spiralling. When paths are deterministic, historical firm beliefs completely determine the final steady state of a group. On the other hand, when paths are spiralling, there is a region of initial firm beliefs that can lead to either of the saddle stable steady states.

The shape of these paths is relevant in both determining the initial conditions that can lead to persistent discrimination and in determining at what point a policy can be removed and still guarantee the elimination of discrimination.
Lemma 7. If \( \frac{(r+2\beta)^2}{\beta} < 4\gamma'G'|_{(\Gamma_m,\Pi_m)} \) holds, then equilibrium paths spiral out in the neighborhood of \( \Sigma_m \). Otherwise, paths are deterministic.

Proof. For a steady state to be a unstable node the eigenvalues must be complex conjugate. This occurs when

\[
r^2 - 4\beta[(\gamma'G') - (\beta + r)] < 0.
\]

Rearranging (3.21) yields

\[
\frac{(r + 2\beta)^2}{\beta} < 4\gamma'G'.
\]

Figure 3.2 depicts a case where there are three steady states and paths are spiralling. Spiralling paths to the steady states create a range of initial firm beliefs that can lead to either the high or the low steady state, this is indicated by the purple region. I call this range of initial beliefs the uncertain region. It is formed by the bottom bound of the path to the high steady state, \( \hat{\Pi} \), and by the upper bound of the path to the low steady state, \( \check{\Pi} \). If initial firm beliefs about the investment rate of a group are in this range, \([\check{\Pi}, \hat{\Pi}]\), then that group can converge to either \( \Sigma_h(\Pi_h, \Gamma_h) \) or \( \Sigma_l(\Pi_l, \Gamma_l) \).

When initial firm beliefs are in the uncertain region, the steady state that a group converges to is determined by the expectations of the workers in that group. If a group develops and shares an optimistic view of the future, then they will converge to the high steady state. If they develop a pessimistic view, they will converge to the low steady state. Consequently, I refer to the path that leads to the high steady state as the optimistic path and the path that leads to the low steady state as the pessimistic path.

If initial firm beliefs for a group fall outside the uncertain region, then the steady state that the group converges to will depend entirely on history. In the case of spiralling paths, when initial beliefs about investment rate are such that \( \Pi_0^J \in (\hat{\Pi}, 1] \), then group \( J \) will converge to
the high qualification steady state with certainty. This is because the optimistic path is the only reasonable path to take. When initial firms beliefs are such that $\Pi^i_0 \in [0, \hat{\Pi})$, then group $J$ will converge to the low qualification steady state with certainty. As discussed in Chapter 2, the high and low steady states will always lie outside of the uncertain region, which means that there is no way for a group to move from one steady state to another.

Figure 3.3 depicts a case where there are three steady states and paths are deterministic. In this situation there is no uncertain region, as a result initial firm beliefs are the only factor that determine the final steady state of a group. If initial beliefs about investment rate are such that $\Pi^i_0 \in (\Pi_m, 1]$, then group $J$ will converge to the high qualification steady state with certainty. If $\Pi^i_0 \in [0, \Pi_m)$, then the group will converge to the low qualification steady state with certainty. Again, when paths are deterministic, groups are unable to move from one steady state to another.

From now on, regardless of the shape of the paths, I refer to the range of initial firm beliefs that lead only to the high steady state as the *certain region for $\Sigma_h$*, and I refer to the range of initial firm beliefs that lead only to the low steady state as the *certain region for*
Figure 3.3: Deterministic paths to the steady states

$\Sigma_l$. These certain regions make it possible for a group to get stuck permanently at either the high or low steady state given appropriate initial firm beliefs. This suggests that once present, discrimination will persist indefinitely without a structural change to the model. Furthermore, if a policy is able to move firm beliefs into the certain region for $\Sigma_h$, then it can be removed and the group will continue to converge to the high steady state on their own.

3.4 Simple parameterization

I now develop a simple form of the dynamic model that I use for demonstrative purposes in later sections. In particular, I use this parameterization to complete more specific analysis of the effectiveness of the policies in Section 3.5. Then, in Section 3.6, I use the parameterization to examine the welfare effects of the three policies.

Firms face the same linear per period production function as in the original model. For simplicity workers have only two possible signals: $\theta_H$ and $\theta_L$ where $\theta_H > \theta_L$. If a worker is qualified, they receive a signal of $\theta_H$ with probability $P_q > \frac{1}{2}$ and a signal of $\theta_L$ with probability $1 - P_q$. If a worker is unqualified, they receive a signal of $\theta_L$ with probability $P_u > \frac{1}{2}$ and
a signal of $\theta_H$ with probability $1 - P_u$. The cost of investment is distributed uniformly over $[c, \overline{c}]$. Time is discrete, so that $\Delta = 1$.

The equilibria of this example are characterized by the same dynamic system as in Proposition 8. Given the parameterization, I can find more specific forms of the relevant functions. The updated optimal threshold rule is

$$\hat{\theta}(\Pi_{jt}) = \begin{cases} 
1 & \text{if } \Pi_{jt} \leq \Pi \\
\theta_H & \text{if } \Pi < \Pi_{jt} < \overline{\Pi} \\
0 & \text{if } \Pi_{jt} \geq \overline{\Pi},
\end{cases}$$

(3.23)

where $\Pi$ and $\overline{\Pi}$ are calculated by updating $\hat{\theta}(\Pi_{jt})$ in (3.6); they are

$$
\Pi = \frac{\alpha_s(1 - P_u)}{\alpha_s(1 - P_u) + (\alpha_c - \alpha_s)P_q} 
\quad \text{and} \quad 
\overline{\Pi} = \frac{\alpha_sP_u}{\alpha_sP_u + (\alpha_c - \alpha_s)(1 - P_q)}. 
$$

(3.24)

The unit period gain from investment is updated to

$$
\gamma(\Pi_{jt}) = \begin{cases} 
0 & \text{if } \Pi_{jt} \leq \Pi \\
P_q + P_u - 1)(\alpha_cP(\theta_H, \Pi_{jt}) - \alpha_s) & \text{if } \Pi < \Pi_{jt} < \overline{\Pi} \\
\alpha_c(P_q + P_u - 1)(P(\theta_H, \Pi_{jt}) - P(\theta_L, \Pi_{jt})) & \text{if } \Pi_{jt} \geq \overline{\Pi}.
\end{cases}
$$

(3.25)

The posterior probability that a worker with signal $\theta_H$ is qualified is

$$
P(\theta_H, \Pi_{jt}) = \frac{\Pi_{jt}P_q}{\Pi_{jt}P_q + (1 - \Pi_{jt})(1 - P_u)},
$$

(3.26)

and the posterior probability that a worker with signal $\theta_L$ is qualified is

$$
P(\theta_L, \Pi_{jt}) = \frac{\Pi_{jt}(1 - P_q)}{\Pi_{jt}(1 - P_q) + (1 - \Pi_{jt})P_u}.
$$

(3.27)

The two loci of the dynamic system are now $\Gamma_{jt} = \frac{\gamma(\Pi_{jt})}{\beta + \gamma}$, where $\gamma(\Pi_{jt})$ is as in (3.25), and $\Pi_{jt} = G(\Gamma_{jt})$, where $G$ is the uniform distribution over $[c, \overline{c}]$.

Figure 3.4 shows the phase diagram for this example under the following parameter values:
Figure 3.4: Phase diagram for the parametric example

\[ P_q = 2/3, \quad P_u = 2/3, \quad \alpha_s = 1, \quad \alpha_c = 2, \quad \beta = 0.2, \quad r = 0.05, \quad c = -0.1, \quad \bar{c} = 0.9 \]. In this simplified version of the dynamic model it is possible for there to be one or three steady states depending on the parameter values. From now on I focus on parameter values that generate three steady states. Following the convention from the previous section, I denote the three steady states as \( \Sigma_l(\Gamma_l, \Pi_l) \), \( \Sigma_m(\Gamma_m, \Pi_m) \), and \( \Sigma_h(\Gamma_h, \Pi_h) \), referring to the low, middle and high steady states, respectively. The directional arrows of the phase diagram indicate that \( \Sigma_h \) and \( \Sigma_l \) are saddle points and \( \Sigma_m \) is an unstable source.

Figure 3.5 demonstrates the spiralling paths to the two saddle points. The steady state beliefs for this particular parametric example are \( \Pi_l = 0.1, \quad \Pi_m = 0.47, \quad \Pi_h = 0.76 \). The paths are generated by inducting backwards from the high and low steady states using equations (3.9) and (3.12). See Appendix B.1 for a discussion of how these paths were generated. The figure indicates that there is an uncertain region (indicated by the purple region) that results from the spiralling paths. There are also certain regions for both \( \Sigma_h \) (the pink area) and \( \Sigma_l \) (the blue area). These certain regions will help determine how long a potential policy needs to be in effect to guarantee the elimination of discrimination. It is also possible to select...
parameter values that generate three steady states and deterministic paths. The following parameter values generate deterministic paths: \( P_q = 19/20 \), \( P_u = 1/2 \), \( \alpha_s = 1 \), \( \alpha_c = 2 \), \( \beta = 0.5 \), \( r = 0.5 \), \( \xi = -0.1 \) and \( \tau = 0.4 \).

3.5 Potential Policies

The analysis of Section 3.3 suggests that, if present, statistical discrimination is persistent. Consequently, eliminating this discrimination requires some sort of government intervention. In this section, I consider three government policies that have the potential to eliminate existing statistical discrimination.

As a starting point for the analysis of these policies I assume that one group is in the high steady state and the other is in the low steady state, so that persistent discrimination is present. From now on I refer to the group in the low steady state as the *disadvantaged group* and the group in the high steady state as the *advantaged group*. I then determine under what circumstances the three policies are able to move the disadvantaged group from the low steady state to the high steady state. Specifically, I determine whether each policy allows...
the disadvantaged group to get on the path to the high steady state. I also address if the policy needs to be in effect permanently or if there is a point at which it can be removed and the discrimination will not return. I assume that the government commits to employing each policy at least until it can be safely removed and still successfully eliminate the discrimination.

I begin by considering the three policies and their effectiveness in the framework of the general dynamic model, and then I examine the impact of these policies in the simple parameterization. The policies that I investigate are (1) a subsidy paid by the government to the firms for every worker in the disadvantaged group that the firm correctly assigns to the complex task, (2) a government subsidization of investment costs for those workers in the disadvantaged group, and (3) a government policy that forces firms to assign workers in different groups with the same signal to the same task and pay them the same wage.

3.5.1 Hiring subsidy program

The first policy is a hiring subsidy that promotes the assignment of disadvantaged workers to the complex task. This subsidization is applied in the following way: For each worker in the disadvantaged group who is correctly assigned to the complex task, the government compensates firms with a subsidy of \( \Delta \xi \).

The government cannot actually observe whether workers are correctly assigned, instead this policy is implemented through a lump sum payment of \( \Delta \xi \Pi_{d}[1 - F_{q}(\hat{\theta})] \) to the firms. This is simply the value of the subsidy times the expected number of workers from the disadvantaged group that will be correctly assigned to the complex task. Under the subsidy, the firms’ profit from the disadvantaged group, given an arbitrary threshold \( \theta_{d} \), is

\[
\lambda^{d} \left( (\Delta \alpha_{c} + \Delta \xi) \left[ \Pi_{d} \left( 1 - F_{q}(\theta_{d}) \right) \right] + \Delta \alpha_{s} \left[ \Pi_{d} F_{q}(\theta_{d}) + (1 - \Pi_{d}) F_{u}(\theta_{d}) \right] \right). \tag{3.28}
\]

Since there are two firms engaged in Bertrand competition for workers and the firms have linear production functions, all of the benefits of the subsidy program are passed on to the workers. Workers in the advantaged group are not affected by this policy; they will remain at \( \Sigma_{h} \). For simplicity of notation, I drop the group subscripts in this section.
Effects of the hiring subsidy in the general dynamic model

The profit function in (3.28) indicates that I can represent the hiring subsidy with a change to the productivity of qualified workers in the disadvantaged group. Given a subsidy of $\Delta \xi$, the expected productivity of a disadvantaged worker with signal $\theta_t$ who is assigned to the complex task is $\Delta(\xi + \alpha_c)p(\theta_t, \Pi_t)$. The higher productivity of qualified workers in the disadvantaged group means that, given the same firm beliefs, the optimal threshold is reduced by the policy.

The optimal threshold for the disadvantage group given a subsidy of $\Delta \xi$ is

$$
\hat{\theta}_\xi(\Pi_t) = \begin{cases} 
1 & \text{if } (\alpha_c + \xi)p(1, \Pi_t) \leq \alpha_s \\
0 & \text{if } (\alpha_c + \xi)p(0, \Pi_t) \geq \alpha_s,
\end{cases}
$$

(3.29)

Since workers are paid their expected productivity, the wages of those workers assigned to the complex task will be higher as a result of the policy. The continuation wage under the hiring subsidy is

$$
\omega^\xi(\theta; \Pi_t) = \begin{cases} 
\Delta \alpha_s & \text{if } \theta < \hat{\theta}_\xi(\Pi_t) \\
\Delta(\alpha_c + \xi)p(\theta, \Pi_t) & \text{if } \theta \geq \hat{\theta}_\xi(\Pi_t).
\end{cases}
$$

(3.30)

The gain from investment will also be impacted. The new unit period gain from investment is

$$
\gamma^\xi(\Pi_t) = \alpha_s \left[ F_q(\hat{\theta}_\xi(\Pi_t)) - F_u(\hat{\theta}_\xi(\Pi_t)) \right] + (\alpha_c + \xi) \left[ \int_{\hat{\theta}_\xi(\Pi_t)}^{1} p(\theta, \Pi_t)[f_q(\theta) - f_u(\theta)]d\theta \right].
$$

(3.31)

The change from $\gamma(\Pi_t)$ to $\gamma^\xi(\Pi_t)$ also changes the lifetime gain from investment $\Gamma_t$.

Consequently, under the hiring subsidy, the equilibrium for the disadvantaged group is described by a new dynamic system.
Proposition 9. An equilibrium of the dynamic model under a hiring subsidy is fully characterized by the following two variable differential equations:

$$\dot{\Gamma}_t = (\beta + r)\Gamma_t - \gamma^\xi(\Pi_t)$$
$$\dot{\Pi}_t = \beta (G(\Gamma_t) - \Pi_t). \quad (3.32)$$

It is clear that the hiring subsidy policy only impacts the $\dot{\Gamma}_t = 0$ locus. The $\dot{\Pi}_t = 0$ locus is unchanged.

In order for this policy to eliminate discrimination it must move the disadvantaged group into the certain region of $\Sigma_h$ from the base model. This result is guaranteed if the dynamic system under the hiring subsidy has only one steady state that is within the relevant region.

Given that the baseline dynamic system has three steady states, the only way that the dynamic system under the hiring subsidy will have one steady state is if $\gamma^\xi(\Pi_t)$ is either much larger or much smaller than $\gamma(\Pi_t)$ for every value of $\Pi_t$. If $\gamma^\xi(\Pi_t)$ is smaller, then there is only one steady state at a low investment level. If $\gamma^\xi(\Pi_t)$ is larger, then the single steady state occurs at a high investment level.

The following Lemma indicates that $\gamma^\xi(\Pi_t)$ is non-decreasing in $\xi$ (the unit period subsidy). This means that under the hiring subsidy policy, $\gamma^\xi(\Pi_t)$ is larger than $\gamma(\Pi_t)$ when $\xi$ is positive.

Lemma 8. If $\Pi_t$ is such that $\dot{\theta}^\xi(\Pi_t) \in [0, 1)$, then the unit period gain from investment is strictly increasing in $\xi$. If $\Pi_t$ is such that $\dot{\theta}^\xi(\Pi_t) = 1$, then the unit period gain from investment is equal to zero and unaffected by an increase in $\xi$.

Proof. The partial derivative of the unit period gain from investment with respect to the hiring subsidy $\xi$ is

$$\frac{\partial \gamma^\xi(\Pi_t)}{\partial \xi} = \frac{\partial \dot{\theta}^\xi(\Pi_t)}{\partial \xi} \left[ \alpha_s - (\xi + \alpha_c)p(\dot{\theta}^\xi(\Pi_t), \Pi_t) \right] \left( f_q(\dot{\theta}^\xi(\Pi_t)) - f_u(\dot{\theta}^\xi(\Pi_t)) \right)$$
$$+ \int_{\theta^\xi(\Pi_t)}^{1} p(\theta, \Pi_t) [f_q(\theta) - f_u(\theta)] d\theta. \quad (3.33)$$
If $\Pi_t = 0$ or 1, then $\gamma^\xi(\Pi_t) = 0$ and consequently $\frac{\partial \gamma^\xi(\Pi_t)}{\partial \xi} = 0$. There are three additional ranges of $\Pi_t$ that need to be considered:

1. If $\Pi_t$ is such that $\hat{\theta}^\xi(\Pi_t) \in (0, 1)$, then $\alpha_s - (\xi + \alpha_c)p(\hat{\theta}^\xi(\Pi_t), \Pi_t) = 0$ so that $\frac{\partial \gamma^\xi(\Pi_t)}{\partial \xi} = \int_{\hat{\theta}^\xi(\Pi_t)}^1 p(\theta, \Pi_t)[f_q(\theta) - f_u(\theta)]d\theta > 0$.

2. If $\Pi_t$ is such that $\hat{\theta}^\xi(\Pi_t) = 0$, then $\gamma^\xi(\Pi_t) = (\alpha_c + \xi)\int_0^1 p(\theta, \Pi_t)[f_q(\theta) - f_u(\theta)]d\theta$ and $\frac{\partial \gamma^\xi(\Pi_t)}{\partial \xi} = \int_0^1 p(\theta, \Pi_t)[f_q(\theta) - f_u(\theta)]d\theta > 0$.

3. If $\Pi_t$ is such that $\hat{\theta}^\xi(\Pi_t) = 1$, then $\gamma^\xi(\Pi_t) = 0$ and consequently $\frac{\partial \hat{\theta}^\xi(\Pi_t)}{\partial \xi} = 0$.

**Corollary 2.** If $\Pi^{sup} = \sup\{\Pi|\hat{\theta}^\xi(\Pi) = 1\}$, then $\Pi^{sup}$ is decreasing in $\xi$.

**Proof.** In order for $\hat{\theta}^\xi(\Pi_t) = 1$, it must be that $(\alpha_c + \xi)p(1, \Pi_t) \leq \alpha_s$ (see (3.29)). As $\xi$ increases, the highest $\Pi_t$ for which this inequality holds decreases.

Lemma 8 and the associated Corollary indicate that the unit period gain from investment is strictly increasing in $\xi$ unless $\Pi_t$ is such that $\hat{\theta}^\xi(\Pi_t) = 1$, and that the range of $\Pi_t$ values that induces $\hat{\theta}^\xi(\Pi_t) = 1$ is decreasing in $\xi$. I conclude that, as $\xi$ increases, the $\dot{\Gamma} = 0$ locus is stretched to the right and if there is a region of $\Pi_t$ values for which $\gamma^\xi(\Pi_t) = 0$, its upper bound decreases as $\xi$ increases. At some value of $\xi$, which I refer to as $\xi^{min}$, only one steady state will remain and it will be at a relatively high investment rate. I refer to this steady state as $\Sigma^{\xi}(\Gamma^\xi, \Pi^\xi)$.

**Lemma 9.** If $G(0) > 0$, then there exists a $\xi^{min}$ such that for all $\xi > \xi^{min}$ the dynamic system under the hiring subsidy has only one steady state, which is a saddle point.

**Proof.** Lemma 8 indicates that $\gamma^\xi(\Pi_t)$ is strictly increasing in $\xi$. This result, along with the assumption that $G(0) > 0$, implies that there exists a $\xi^{min}$ such that for all $\xi > \xi^{min}$ there is only one steady state.

Figure 3.6 depicts a potential phase diagram for the dynamic system when the hiring subsidy is such that there is only one steady state. This steady state is a saddle point as $\frac{\partial \gamma^\xi}{\partial \Pi} \frac{1}{\beta + r} < \left(\frac{dG}{dt}\right)^{-1}|_{\Gamma, \Pi}$ holds (See Lemma 6). Consequently, when a hiring subsidy of sufficient size is put in place, the disadvantaged group will move from the low steady state to the
Figure 3.6: Phase diagram of the dynamic system under the hiring subsidy

path that leads to $\Sigma^\xi(\Gamma^\xi, \Pi^\xi)$, the only remaining steady state. In each period that the policy is in place, the group will move along the path that approaches the steady state.

If the hiring subsidy is removed when the investment rate of the disadvantaged group is in the certain region for $\Sigma_l$, then the group will return to their initial steady state. If, on the other hand, the policy is removed after the disadvantaged group reaches the certain region for $\Sigma_h$, then they will converge to $\Sigma_h$ and discrimination will be eliminated.

It is important that $\Sigma^\xi$ is not in the uncertain region, or else there is no way to guarantee the success of the hiring subsidy. The following lemma indicates that for a positive subsidy $\Sigma^\xi$ is always in the certain region of $\Sigma_h$.

**Lemma 10.** If $G(0) > 0$ and $\xi > \xi^{\min}$, then $\Sigma^\xi$ will be in the certain region for $\Sigma_h$.

**Proof.** From Lemma 8, $\gamma^\xi(\Pi^\xi)$ is strictly increasing in $\xi$. By assumption, $G(\cdot)$ is strictly increasing in its argument. Together these imply that $\Sigma^\xi$ will always involve a higher investment rate than $\Sigma_h$. From Lemma 4 in Chapter 2, the upper bound of the uncertain region is always below $\Pi_h$. As a result, $\Sigma^\xi$ must be in the certain region of $\Sigma_h$ because $\Pi^\xi$ is greater than $\Pi_h$, which is always in the certain region. 

\[\square\]
Taking Lemma 10 into account, it is clear that the hiring subsidy policy will be successful as long as a large enough subsidy is offered.

**Proposition 10.** If \( G(0) > 0 \) and \( \xi > \xi^{\text{min}} \), then the hiring subsidy will eliminate discrimination in the general dynamic model.

**Corollary 3.** This hiring subsidy can be removed as soon as \( \Pi_t \) is in the certain region for \( \Sigma_h \).

Figure 3.7 depicts the path to the remaining steady state from the disadvantaged group’s starting point at the low steady state, see Appendix B.2.2 for a discussion on how these paths were generated. For comparison the figure also includes the original \( \gamma(\Pi_J^f) \) function. I do not show the path to \( \Sigma_h \) once the policy is removed, but as long as the policy is removed after \( \Pi_t > \hat{\Pi} \) then the disadvantaged group will simply move to the path that leads to \( \Sigma_h \) as it is the only plausible path given the current investment rate of the group.

I conclude that the hiring subsidy will always effectively eliminate discrimination as long as a large enough subsidy is offered. This policy eliminates discrimination by changing the
optimal behavior of the firms. Specifically, the subsidy makes hiring disadvantaged workers more attractive, and this is passed on to workers with an increase in the gain from investment. The end result is that the workers in the disadvantaged group are able to move to the high steady state.

**Effects of the hiring subsidy in the parameterization**

I now apply the hiring subsidy to the parameterization described in Section 3.4. By doing so I am able to confirm that the hiring subsidy is effective at eliminating discrimination. In the framework of the parameterization, I am also able to determine the actual path that is taken once the policy is put in place and to calculate the minimum \( \xi \) required to eliminate the discrimination.

The equilibria of this simple model are described by the same dynamic system as in Proposition 9. Just as in Section 3.4 I can find more specific forms of the relevant functions. The policy only impacts the gain from investment function. It is updated to

\[
\gamma^\xi(\Pi_t) = \begin{cases} 
0 & \text{if } \Pi_t \leq \Pi^\xi \\
(P_q + P_u - 1)[(\alpha_c + \xi)P(\theta_H, \Pi_t) - \alpha_s] & \text{if } \Pi^\xi < \Pi_t < \Pi_\xi \\
(P_q + P_u - 1)(\alpha_c + \xi)[P(\theta_H, \Pi_t) - P(\theta_L, \Pi_t)] & \text{if } \Pi_t \geq \Pi_\xi,
\end{cases} \tag{3.34}
\]

where \( \Pi^\xi \) and \( \Pi_\xi \) are calculated by updating \( \hat{\theta}(\Pi_t^\xi) \) in (3.29). They are

\[
\Pi^\xi = \frac{\alpha_s(1 - P_u)}{\alpha_s(1 - P_u) + (\alpha_c + \xi - \alpha_s)P_q} \quad \text{and} \quad \Pi_\xi = \frac{\alpha_sP_u}{\alpha_sP_u + (\alpha_c + \xi - \alpha_s)(1 - P_q)} \tag{3.35}
\]

Compared to \( \gamma(\Pi_t) \), this new gain from investment is larger for all beliefs above \( \Pi^\xi \). In the region of firm beliefs below \( \Pi^\xi \), it is the same and equal to zero. As noted in the previous section, the range of \( \Pi_t \) values that leads to \( \gamma^\xi(\Pi_t) = 0 \) is decreasing in \( \xi \). This clearly holds in this parameterization as well since \( \Pi^\xi \) is decreasing in \( \xi \).

The phase diagram of the dynamic system under the hiring subsidy will therefore have a new \( \dot{\Gamma} = 0 \) locus of \( \Gamma_t = \frac{\gamma^\xi(\Pi_t)}{\beta + r} \). This locus is larger than that of the original model for all values of \( \Pi_t > \Pi^\xi \). The \( \bar{\Pi} = 0 \) locus is unchanged. Figure 3.8 depicts the phase diagram of the
Figure 3.8: Phase diagram for the parametric example under the hiring subsidy

new dynamic system given that $\xi = 6$. All other parameter values are as in Section 3.4. The result of this hiring subsidy is a dynamic system with only one steady state, which I refer to as $\Sigma^\xi(\Gamma^\xi, \Pi^\xi)$. The arrows indicate that $\Sigma^\xi$ is a saddle point.

In this parameterization, I am able to explicitly determine the minimum subsidy necessary for the dynamic system have only one steady state, $\xi^{\text{min}}$. The subsidy value must be such that $\Pi^\xi < G(0)$. In this simple model $\xi^{\text{min}} = \alpha_s \left[ \frac{1-P_u}{P_q} \frac{1-G(0)}{G(0)} + 1 \right] - \alpha_c$. Note that $\xi^{\text{min}}$ is decreasing in the precision of signals ($P_u$ and $P_q$) as well as in $\alpha_c$; it is increasing in $\alpha_s$.

**Lemma 11.** If $G(0) > 0$ and $\xi > \xi^{\text{min}}$, then the simple parameterization under a hiring subsidy has only one steady state, which is a saddle point.

**Proof.** This is directly proven by the directional arrows and the definition of $\xi^{\text{min}}$. \qed

The existence of a single steady state means that workers in the disadvantaged group have only one rational path to take. As soon as the policy takes effect, workers in the disadvantaged group move to the saddle path leading to $\Sigma^\xi$. Figure 3.9 demonstrates this path as well as the group's starting point.
Figure 3.9: Path to the $\Sigma^\xi$ in the parametric example

Figure 3.10: Path to $\Sigma_\eta$ once the hiring subsidy policy is lifted
From the analysis in Section 3.4, I know that $\Sigma^\xi$ is in the certain region for $\Sigma_h$ (see Figure 3.2 for clarification). Consequently, the hiring subsidy is guaranteed to eliminate discrimination as long as $\xi$ is large enough and it can be removed as soon as firm beliefs about the investment rate of the disadvantaged group exceed $\hat{\Pi}$. Once removed, the workers will converge to $\Sigma_h$ on their own. Figure 3.10 demonstrates the path the group will take to $\Sigma_h$ assuming that the policy is removed after the group reaches $\Sigma^\xi$, see Appendix B.2.2 for a discussion on how these paths were generated.

**Proposition 11.** If $G(0) > 0$ and $\xi > \xi^{\text{min}}$, then the hiring subsidy will eliminate discrimination in the parametric example.

**Corollary 4.** The hiring subsidy can be removed as soon as $\Pi_t$ is in the certain region for $\Sigma_h$.

For this simple version of the model the hiring subsidy is effective at eliminating the discrimination as long as the subsidy is large enough and the policy is left in place for long enough. These results confirm those of the general dynamic model.

### 3.5.2 Investment voucher program

The second policy of interest approaches the elimination of discrimination more directly. Rather than providing hiring incentives to the firms, an investment voucher program is designed to affect the decisions of the workers. The voucher is applied in the following way: The government provides all workers in the disadvantaged group with an investment voucher of $\nu$ dollars. This voucher is only good towards the purchase of human capital, so it defrays the monetary cost of investment, $c_m$. Consequently, a voucher larger than $c_m$ has no more impact than one equal to $c_m$. By implementing the voucher in this way, the government does not need to observe whether an individual worker invested or not. This policy is applied only to the disadvantaged group, so for simplicity of notation I drop the group subscripts in this section.

**Effects of the investment voucher program in the general model**

I can represent this investment voucher with a change in the effective lifetime gain from investment since the voucher is only used if a worker invests. The impact of the policy on
the dynamic system is a strict increase in the investment rate of new workers at time \( t \). In particular, the proportion of investors is \( G(\Gamma_t + \min\{v, c_m\}) \) under the voucher, rather than \( G(\Gamma_t) \), as in the baseline model.

Consequently, under the investment voucher the equilibrium for the disadvantaged group will be described by a new dynamic system.

**Proposition 12.** An equilibrium of the dynamic model under an investment voucher is fully characterized by the following two variable differential equations:

\[
\begin{align*}
\dot{\Gamma}_t &= (\beta + r) \Gamma_t - \gamma (\Pi_t) \\
\dot{\Pi}_t &= \beta (G(\Gamma_t + \min\{v, c_m\}) - \Pi_t).
\end{align*}
\] (3.36)

The only resulting change to the phase diagram will be in the \( \dot{\Pi}_t = 0 \) locus, which is updated to \( \Pi_t = G(\Gamma_t + \min\{v, c_m\}) \). Given that \( G(\cdot) \) is strictly increasing, a negative voucher shifts the \( \dot{\Pi}_t = 0 \) locus to the right, so that the investment rate is lower at every \( \Gamma_t \). A positive voucher shifts the locus to the left, so that investment rate is higher at every \( \Gamma_t \).

Just as in the analysis of the hiring subsidy, in order for this policy to eliminate discrimination it must move the disadvantaged group to the certain region of \( \Sigma_h \). The original dynamic model assumes a \( G(\cdot) \) function that generates three steady states. The investment voucher policy does not change \( \gamma \) or \( G(\cdot) \) but instead affects how investment decisions are made by increasing the lifetime gain from investment by \( \min\{v, c_m\} \). The dynamic system under the investment voucher will have only one steady state if \( \min\{v, c_m\} \) is sufficiently positive or negative.

Given the goal of moving the disadvantaged group to \( \Sigma_h \), an effective investment voucher must be positive. Moreover, assuming that the monetary cost of investment, \( c_m \), is high enough, at some value of \( v \), which I refer to as \( v^{\min} \), only one steady state will remain. I refer to this steady state as \( \Sigma^v(\Gamma^v, \Pi^v) \). If the monetary cost of investment is low, then it is possible that even the largest effective voucher, \( v = c_m \), is not able to eliminate the two lower steady states. In this case, the voucher policy is not effective at eliminating discrimination and no matter what \( v \) is employed there will be three steady states.
Lemma 12. If the dynamic system under the largest effective investment voucher, \( v = c_m \), has a single steady state, then there exists a \( v_{\text{min}} \) such that for all \( v > v_{\text{min}} \) the dynamic system under the investment voucher has only one steady state, which is a saddle point.

Proof. \( G(\Gamma_t + \min\{v, c_m\}) \) is strictly increasing in \( v \) when \( v < c_m \). As \( v \) increases, the \( \dot{\Pi}_t = 0 \) locus shifts to the left. Given a large enough \( c_m \), there must exist a \( v_{\text{min}} \) such that for all \( v > v_{\text{min}} \) there is one steady state. \( \square \)

Figure 3.11 depicts a potential phase diagram for the dynamic system when the investment voucher is large enough to induce a single steady state. According to Lemma 6, this steady state is a saddle point as \( \frac{\partial \gamma}{\partial \Pi} \left( \frac{1}{\beta + r} \right) < (\frac{dG}{d\Gamma})^{-1}|(\bar{\Gamma}, \bar{\Pi}) \) holds. Just as in the case of the hiring subsidy, the existence of a single steady state means that when the policy is put in place the disadvantaged group moves from the low steady state to the path that leads to \( \Sigma^v(\Gamma^v, \Pi^v) \). In each period the policy is in effect, the group moves along the path approaching that steady state.

In order to successfully eliminate the discrimination, the policy can only be removed when
the disadvantaged group is in the certain region for $\Sigma_h$. Again it is important that $\Sigma^v$ is not in the uncertain region, or else there is no way to guarantee the success of the investment voucher. The following lemma indicates that for any voucher greater than $v^{\text{min}}$, $\Sigma^v$ will be in the certain region of $\Sigma_h$.

**Lemma 13.** If $v > v^{\text{min}}$, then $\Sigma^v$ will be in the certain region for $\Sigma_h$.

**Proof.** $G(\Gamma_t + \min\{v, c_m\})$ is strictly increasing in $v$ when $v < c_m$. So, the $\dot{\Pi}_t = 0$ locus is always higher under the policy. This implies that $\Sigma^v$ will involve a higher investment rate than $\Sigma_h$. From Lemma 4 in Chapter 2, the upper bound of the uncertain region is always below $\Pi_h$. $\Sigma^v$ must be in the certain region of $\Sigma_h$ because $\Pi^v$ is greater than $\Pi_h$, which is always in the certain region.

Taking Lemma 13 into account, it is clear that the investment voucher policy will be successful as long as monetary costs make up a large enough proportion of total costs and a sufficient voucher is offered.

**Proposition 13.** If there exists a $v^{\text{min}}$, then for all $v > v^{\text{min}}$, the investment voucher will eliminate discrimination in the general dynamic model.

**Corollary 5.** This investment voucher policy can be removed as soon as $\Pi_t$ is in the certain region for $\Sigma_h$.

Figure 3.12 depicts the path to the remaining steady state from the disadvantaged group’s starting point at $\Sigma_l$, see Appendix B.2.3 for a discussion on how these paths were generated. For comparison, the figure also includes the original $\dot{\Pi} = 0$ locus. As long as the policy is removed after $\Pi_t > \hat{\Pi}$, the disadvantaged group will simply move to the path that leads to $\Sigma_h$ as it is the only plausible path given the firms’ current beliefs about the group’s overall investment rate.

The investment voucher approaches the discrimination issue in a different manner than the hiring subsidy. It works by changing the incentives of the workers and making it easier for a worker’s gain from investment to exceed his total cost. By adjusting the optimal behavior of workers, this policy allows the firms and workers to move to the high steady state.
Effects of the investment voucher program in the simple model

I can also apply the investment voucher to the parameterization described in Section 3.4. Examining the investment voucher policy in this simple framework allows me to explicitly calculate the minimum required voucher and to examine how the parameters affect this value. I am also able to see the actual path that the disadvantaged workers take to the steady state after the policy is put in place.

The equilibria of this model are described by the same dynamic system as in Proposition 6. The relevant functions are just as in Section 3.4. The only change resulting from the policy is that the $\dot{\Pi} = 0$ locus is updated to $\Pi_t = G(\Gamma_t^I + \min\{\psi, c_m\})$. In this simple parameterization, $G(\cdot)$ is uniform so the policy shifts the $\dot{\Pi} = 0$ locus to the left by the value of the voucher.

Figure 3.13 depicts the updated phase diagram given that $\psi = 0$ and $c_m = 0.5$. All other parameter values are as in Section 3.4. The result is a dynamic system with only one steady state. As with the hiring subsidy, this single steady state is not the $\Sigma_h$ from the parametric example.

In this simple framework, I can explicitly calculate the minimum $\psi$ necessary to produce
only one steady state, assuming that $c_m$ is large enough. In particular $G(\min\{v, c_m\})$, the equilibrium value of $\Pi_t$ when $\Gamma_t = 0$, must be larger than $\Pi$. Put more concisely, $v$ must be such that $G(v) > \frac{\alpha_s(1-P_u)}{\alpha_s(1-P_u)+(\alpha_c-\alpha_s)P_q}$. Some rearranging yields $v_{\text{min}} = (\bar{r} - \zeta)\Pi + \zeta$. This minimum voucher is decreasing in the precision of signals ($P_u$ and $P_q$) as well as in $\alpha_c$, and it is increasing in $\alpha_s$. Note that if $v_{\text{min}} > c_m$ then the policy will not be effective.

**Lemma 14.** If $v_{\text{min}} = (\bar{r} - \zeta)\Pi + \zeta < c_m$, then for all $v > v_{\text{min}}$ the dynamic system under the investment voucher has only one steady state, which is a saddle point.

**Proof.** This is directly proven by the directional arrows in Figure 3.13 and the definition of $v_{\text{min}}$. 

Figure 3.14 demonstrates the path that the disadvantaged group will take to the new steady state from their starting point at the low steady state. Following the introduction of the policy the disadvantaged group immediately moves to the path leading to the remaining steady state as this is the only rational action that can be taken by the workers facing this investment voucher policy.
Figure 3.14: Path to the $\Sigma^u$ for the parametric example

Figure 3.15: Path to $\Sigma_h$ once the investment voucher policy is lifted
Referring back to Figure 3.5, it is clear that $\Sigma^\nu$ is in the certain region for $\Sigma_h$ because $\Pi^\nu > \Pi_h$. In this parameterization, as long as $v^\text{min} < c_m$, any voucher that exceeds the minimum required value is guaranteed to eliminate discrimination. The investment voucher must be in effect until the investment rate of the disadvantaged group is in the certain region of $\Sigma_h$. As soon as this occurs the policy can be removed and the disadvantaged group simply moves to the path to $\Sigma_h$. Figure 3.15 depicts the path that will be taken if the policy is removed after the group reaches $\Sigma^\nu$.

**Proposition 14.** If $v^\text{min} < c_m$, then for all $v > v^\text{min}$ the investment voucher will eliminate discrimination in the parametric example.

**Corollary 6.** The investment voucher policy can be removed as soon as $\Pi_t$ is in the certain region for $\Sigma_h$.

I can conclude that this policy will eliminate discrimination in the parametric example as long as the monetary cost of investment is large enough and a sufficient investment voucher is provided. These results confirm those for the general dynamic model.

### 3.5.3 Government enforced equal treatment

The final policy of interest involves intervention by the governing body with respect to the optimal behavior of the firms. Specifically, under what I call the *equal treatment policy*, the government requires that both firms ignore a workers’ group identity when assigning tasks and wages. This policy affects both disadvantaged and advantaged workers.

The equal treatment policy can be formally stated as: Firms are forbidden to condition task and wage assignments on group identity. I assume that the government is able to observe the firms’ task and wage assignments and if they observe a deviation from the policy, they will charge the firm a substantial fine. As long as this fine is large enough, the firms will never have incentive to deviate.
Effects of the equal treatment policy in the general model

Under the equal treatment policy, the firms’ optimal behavior can no longer depend on group identity. So there will be a single \( \hat{\theta}(\Pi_t) \) and a single \( \omega_\Delta(\theta_t; \Pi_t) \) in each period, rather than one for each group.

The updated optimal threshold at time \( t \) is the one which maximizes a firm’s output conditional on investment decisions. Formally, the task assignment problem is

\[
\max_{\theta_t \in [0,1]} \Delta \alpha_c \left[ (\lambda^d \Pi_t^d + \lambda^a \Pi_t^a)(1 - F_q(\theta_t)) \right] \\
+ \Delta \alpha_s \left[ (\lambda^d \Pi_t^d + \lambda^a \Pi_t^a)F_q(\theta_t) + (1 - (\lambda^d \Pi_t^d + \lambda^a \Pi_t^a))F_u(\theta_t) \right].
\]

(3.37)

Just as in the original dynamic model there is a unique solution to the task assignment problem. But, in this case, it is not solved separately for each group. If I define a firm’s beliefs about the entire population of workers as \( \Pi_{pop} = \lambda^d \Pi_t^d + \lambda^a \Pi_t^a \), then the unique optimal threshold is

\[
\hat{\theta}(\Pi_t^{pop}) = \begin{cases} 
1 & \text{if } \alpha_c p(1, \Pi_t^{pop}) \leq \alpha_s \\
0 & \text{if } \alpha_c p(0, \Pi_t^{pop}) \geq \alpha_s, \\
\text{the unique solution to} & \\
\alpha_c p(\theta, \Pi_t^{pop}) = \alpha_s & \text{if } \alpha_c p(0, \Pi_t^{pop}) < \alpha_s < \alpha_c p(1, \Pi_t^{pop}).
\end{cases}
\]

(3.38)

The continuation wage under the equal treatment policy will also depend on \( \Pi_{pop} \) rather than the group investment rates. The new optimal wage schedule for both firms is

\[
\omega_\Delta(\theta_t; \Pi_t^{pop}) = \begin{cases} 
\Delta \alpha_s & \text{if } \theta_t < \hat{\theta}(\Pi_t^{pop}) \\
\Delta \alpha_c p(\theta_t, \Pi_t^{pop}) & \text{if } \theta_t \geq \hat{\theta}(\Pi_t^{pop}).
\end{cases}
\]

(3.39)

This change in the task assignment rule and the associated alteration of the continuation wage results in a new unit period gain from investment (and per period gain from investment). The updated unit period gain from investment is

\[
\gamma(\Pi_t^{pop}) = \alpha_s \left[ F_q(\hat{\theta}(\Pi_t^{pop})) - F_u(\hat{\theta}(\Pi_t^{pop})) \right] + \alpha_c \int_{\hat{\theta}(\Pi_t^{pop})}^{1} p(\theta, \Pi_t^{pop})[f_q(\theta) - f_u(\theta)]d\theta.
\]

(3.40)
Under the policy, the lifetime gain from investment for each group will be determined by the population investment rate rather than the group investment rate. This causes a subsequent change in the dynamic system.

**Proposition 15.** An equilibrium of the dynamic model under the equal treatment policy is fully characterized by the following differential equations:

\[
\dot{\Gamma}_t^{\text{pop}} = (\beta + r)\Gamma_t^{\text{pop}} - \gamma_t^{\text{pop}} (\Pi_t^{\text{pop}})
\]

\[
\dot{\Pi}_t^{\text{pop}} = \beta (G(\Gamma_t^{\text{pop}}) - \Pi_t^{\text{pop}}).
\]

The dynamic system in (3.41) is identical to that of the original dynamic model. The only difference is that it describes the entire population rather than each individual group. In order to talk about when the equal treatment policy can be removed, it is also necessary to determine how a firm’s beliefs about the investment rate of each group will change over time. Under the policy these beliefs are described by:

\[
\dot{\Pi}_t^J = \beta (G(\Gamma_t^{\text{pop}}) - \Pi_t^J).
\]

Since I have already assumed that the original dynamic model has three steady states, with the high and the low steady states being saddle points, I know that the dynamic system under the equal treatment policy will as well. I call these three steady states \(\Sigma_{\text{pop}}^l(\Gamma_{\text{pop}}^l, \Pi_{\text{pop}}^l)\), \(\Sigma_{\text{pop}}^m(\Gamma_{\text{pop}}^m, \Pi_{\text{pop}}^m)\), and \(\Sigma_{\text{pop}}^h(\Gamma_{\text{pop}}^h, \Pi_{\text{pop}}^h)\) referring to the low, middle and high steady states.

**Lemma 15.** Under the equal treatment policy the dynamic system of the general model has three steady states. \(\Sigma_{\text{pop}}^l\) and \(\Sigma_{\text{pop}}^h\) are saddle points and \(\Sigma_{\text{pop}}^m\) is a source.

The uncertain region, if present, is identical to that of the original dynamic model. With that in mind, it is straightforward to determine in which cases this policy will be effective. Given that the starting point of this analysis is one in which discrimination is present, when the policy is put in place the investment rate of the population is:

\[\Pi_0^{\text{pop}} = \lambda^c\Pi_h + \lambda^d\Pi_l.\]

If this \(\Pi_0^{\text{pop}}\) is in the certain region for \(\Sigma_h\), then the policy will be effective as the entire population will move to the only feasible path, which happens to lead to the high steady state. However if \(\Pi_0^{\text{pop}}\) is in the uncertain region of firm beliefs, then the policy will only be effective if the entire population of workers is optimistic. The policy is guaranteed to fail if \(\Pi_0^{\text{pop}}\) is in the
Proposition 16. If $\Pi_{\text{pop}}$ is in the certain region for $\Sigma_h$ or if $\Pi_{\text{pop}}$ is in the uncertain region and the population is optimistic, then the equal treatment policy will eliminate discrimination in the general dynamic model.

Corollary 7. This equal treatment policy can be removed as soon as $\Pi_t^a$ and $\Pi_t^d$ are both in the certain region for $\Sigma_h$.

Corollary 7 indicates that while the equilibria of the model are described by the system in (3.41) the movement of the individual group's investment rates is still an important factor in determining when the policy can be removed. This will be more clear when I examine the effects of the equal treatment policy on the simple parameterization.

I can restate the results of Proposition 16 in terms of the size of the disadvantaged population. Specifically, if $\lambda^d < \frac{\Pi_h - \Pi}{\Pi_h - \Pi_l}$, then the equal treatment policy is guaranteed to eliminate discrimination in the general dynamic model. I conclude that this policy is more likely to be effective if the disadvantaged population is small or if the difference between $\Pi_l$ and $\Pi_h$ is small.

This policy is different from the previous two because it impacts workers in both the advantaged and the disadvantaged groups. It is attractive to consider this policy because there is no direct cost associated with it. However, the policy is potentially disastrous if the initial population investment rate is in the certain region for $\Sigma_l$, or in the uncertain region and the population is pessimistic.

Effects of the equal treatment policy in the parameterization

I now apply this equal treatment policy to the parameterization of Section 3.4. As I determined in the previous subsection, this policy generates the same dynamic system as the original model; however, there is only one dynamic system that depends on the population investment rate, rather than two separate systems each depending on the group investment rate.

In the simply dynamic model under the equal treatment policy the optimal threshold rule
for both groups will be
\[
\hat{\theta}_t(\Pi_t^{\text{pop}}) = \begin{cases} 
1 & \text{if } \Pi_t^{\text{pop}} \leq \Pi \\
\theta_H & \text{if } \Pi < \Pi_t^{\text{pop}} < \Pi \\
0 & \text{if } \Pi_t^{\text{pop}} \geq \Pi,
\end{cases}
\] (3.42)
where \(\Pi\) and \(\bar{\Pi}\) are just as in the original parametric example. They are
\[
\Pi = \frac{\alpha_s(1 - P_u)}{\alpha_s(1 - P_u) + (\alpha_c - \alpha_s)P_q} \quad \text{and} \quad \bar{\Pi} = \frac{\alpha_s P_u}{\alpha_s P_u + (\alpha_c - \alpha_s)(1 - P_q)}.
\] (3.43)

Given the optimal task assignment rule in (3.42), the unit period gain from investment is
\[
\gamma(\Pi_t^{\text{pop}}) = \begin{cases} 
0 & \text{if } \Pi_t^{\text{pop}} \leq \Pi \\
(P_q + P_u - 1)(\alpha_c P(\theta_H, \Pi_t^{\text{pop}}) - \alpha_s) & \text{if } \Pi < \Pi_t^{\text{pop}} < \Pi \\
\alpha_c(P_q + P_u - 1)(P(\theta_H, \Pi_t^{\text{pop}}) - P(\theta_L, \Pi_t^{\text{pop}})) & \text{if } \Pi_t^{\text{pop}} \geq \Pi.
\end{cases}
\] (3.44)

The two loci of the dynamic system under the equal treatment policy are \(\Gamma_t^{\text{pop}} = \frac{\gamma(\Pi_t^{\text{pop}})}{\beta + r}\), where \(\gamma(\Pi_t^{\text{pop}})\) is as in (3.44), and \(\Pi_t^{\text{pop}} = G(\Gamma_t^{\text{pop}})\), where \(G(\cdot)\) is the uniform distribution over \([\underline{c}, \bar{c}]\).

I can draw the same conclusions for the simple parameterization under the equal treatment policy as I did for the general dynamic model. If the original parameterization has three steady states, so will the parameterization under the equal treatment policy. Similarly, if \(\Pi_0^{\text{pop}}\) is in the certain region for \(\Sigma_h\), then the policy will definitely eliminate discrimination. On the other hand, if \(\Pi_0^{\text{pop}}\) is in the uncertain region, then discrimination will be eliminated only if the population is optimistic. The following Proposition formalizes these results.

**Proposition 17.** If \(\Pi^{\text{pop}}\) is in the certain region for \(\Sigma_h\) or if \(\Pi^{\text{pop}}\) is in the uncertain region and the population is optimistic, then the equal treatment policy will eliminate discrimination in the parametric example.

**Corollary 8.** This equal treatment policy can be removed as soon as \(\Pi_t^e\) and \(\Pi_t^d\) are both in the certain region for \(\Sigma_h\).

Figure 3.16 shows the paths to the steady states for the same parameter values as in Section
3.4. I make the additional assumption that $\lambda^a = 0.75$ and $\lambda^d = 0.25$. From the analysis in Section 3.4 I know that the steady state beliefs in this example are $\Pi_l = 0.1$, $\Pi_m = 0.47$, and $\Pi_h = 0.76$. So the resulting initial beliefs about the proportion of the population that is qualified are $\Pi_0^{pop} = 0.60$; this starting point is indicated in the figure by the purple dotted line. For this parameterization, the uncertain region of initial firm beliefs is $\Pi_t \in [0.32, 0.65]$. It is clear that $\Pi_0^{pop}$ is in this region. As such, the policy may or may not be effective depending on whether the population is optimistic or pessimistic. The purple dots in Figure 3.16 indicate the three points that the population may move to after the policy is put in place. Note that there are two points on the pessimistic path and one on the optimistic path.

Figure 3.17 shows the path that the population will take if it is optimistic. I also include the paths that the two groups will take from their individual starting points, see Appendix B.2.4 for a discussion of how these paths were generated. The disadvantaged group starts with an investment rate of $\Pi_l$ because that is the steady state that they were previously operating in. The advantaged group starts with an investment rate of $\Pi_h$ for the same reason. It is clear, from the figure, that when the policy is put in place, each group moves to their individual
Figure 3.17: Path to $\Sigma_h$ given an optimistic population

Figure 3.18: Path to $\Sigma_l$ given a pessimistic population
path to the high steady state. Figure 3.18 shows one of the paths that the population may take if it is pessimistic.

When the policy is removed, the two groups are again treated separately. For this reason, the policy can only be removed, and still be successful, if both groups have an investment rate that is in the certain region for \( \Sigma_h \). Once the policy is removed, each group will simply move to the path leading to the high steady state, as this is the only feasible path. Note that there is no real reason to remove this policy because it has no cost, and the high steady state under the policy is the same as the high steady state without the policy.

Just as in the general dynamic model under the equal treatment policy, if the proportion of the population that is in the disadvantaged group is very small, then the policy will be successful at eliminating the discrimination. For this particular set of parameter values if \( \lambda^d < 0.17 \), then the policy is guaranteed to eliminate the discrimination.

### 3.6 Welfare analysis of the three policies in a parametric example

I am now able to look at the welfare effects of the three policies for the specific parameterization of the dynamic model described in Section 3.4. I complete this analysis in a general equilibrium framework. The hiring subsidy and investment voucher policies require that the government have funds to distribute to the firms and to the workers in the disadvantaged group, respectively. Given the general equilibrium framework, these funds are collected by the government via a tax levied on workers in the advantaged group.

For the purpose of more meaningful comparison between the three policies, I apply the minimum necessary voucher and minimum necessary subsidies. For the parametric example \( \nu^{min} = 7/30 \) and \( \xi^{min} = 7/2 \). I also assume that the population is optimistic so as to guarantee that the equal treatment policy is successful. The government also commits to employing each policy until the impacted groups reach their new steady state.

Table 3.1 describes the conditions of the advantaged and disadvantaged groups before a policy is applied. Notice that the advantaged group has a higher average per period expected welfare than the disadvantaged group does. See Appendix B.2.1 for a discussion of how these
Table 3.1: Details of the discriminatory equilibrium

<table>
<thead>
<tr>
<th>Discriminatory Equilibrium</th>
<th>Disadvantaged group (d)</th>
<th>Advantaged group (a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_u = P_q = 2/3, \lambda^a = 3/4, c \sim U[-0.1, 0.9]$</td>
<td>$\alpha_s = 1, \alpha_c = 2, \beta = 0.2, r = 0.05$</td>
<td>Steady state investment $\Pi^d = 0.1000$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lifetime gain from investment $\Gamma^d(\Pi^d) = 0$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wages $\omega(\theta_H, \Pi^d) = 1$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\omega(\theta_L, \Pi^d) = 1$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Average per period expected welfare 1.0010</td>
</tr>
</tbody>
</table>

Values were calculated.

Figures 3.19, 3.20 and 3.21 depict the average expected welfare of each group and of the population as a whole under the subsidy, voucher and equal treatment policies, respectively. Each figure shows the welfare before the policy is in place, how welfare changes when the policy is applied and then when the policy is removed. All three policies move those in the disadvantaged group to the high steady state, so the end result of each policy is that the welfare across groups is equal and the discrimination is eliminated. See Appendices B.2.2, B.2.3 and B.2.4 for a discussion of how these values were calculated.

Comparing across figures indicates that the hiring subsidy is the most costly of the three policies. The reason for this is two-fold. First, the transfer that is necessary to fund the policy is very large. Second, the hiring subsidy is provided for each worker correctly employed in the complex task in every period the policy is in place. The result is that the advantaged group loses a significant amount of welfare over the course of the policy. However, this policy will get the disadvantaged group to the steady state in the shortest amount of time.

The voucher policy is the least costly of the three policies. This is because the transfer necessary to support the policy is very low and in each period it is only paid to new workers that choose to invest. This policy needs to be in place for slightly longer than the hiring subsidy policy. The equal treatment policy also has a relatively small negative impact on the advantaged group, but it has the additional benefit of not requiring the government to redistribute income across groups. However, the equal treatment policy must be in place for the longest amount of time. Note that, across all policies, the welfare loss to the advantaged group would be greater if the disadvantaged group was larger.
Figure 3.19: Welfare effects of the hiring subsidy

Figure 3.20: Welfare effects of the investment voucher
Figure 3.21: Welfare effects of the equal treatment policy

Table 3.2 shows the total change in welfare for each group and the population while each policy is in place. These numbers are simply the sum of the welfare gains of the disadvantaged group and the sum of the welfare losses of the advantaged group for each period that the policy is in place. The net welfare effect is the weighted sum of these two values, taking into account the relative sizes of the two groups. This table indicates that, while the subsidy does increase the disadvantaged group’s welfare greatly, the net result over the course of the policy is the least favorable. The most favorable, in terms of net change in welfare is the equal treatment policy and the voucher policy falls somewhere in the middle.

Depending on a policy makers preferences, it may be that the investment voucher, equal treatment policy or hiring subsidy is preferred. If the policy maker wants to inconvenience the advantaged group the least, then the investment voucher is best. If his goal is to ensure that net welfare is improved the most, then the equal treatment policy is best. If the policy maker wants the policy in place for the shortest amount of time then the hiring subsidy is preferred.

Keep in mind that both the investment voucher and equal treatment policies can fail to be successful for some parameterizations. In particular, if the monetary cost of investment is low,
Table 3.2: Total change in welfare while the policies are in effect

<table>
<thead>
<tr>
<th>Group</th>
<th>Group d welfare gain</th>
<th>Group a welfare loss</th>
<th>Net change in pop. welfare</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hiring Subsidy</td>
<td>61.908</td>
<td>17.855</td>
<td>2.086</td>
</tr>
<tr>
<td>Investment Voucher</td>
<td>12.251</td>
<td>0.329</td>
<td>2.816</td>
</tr>
<tr>
<td>Equal Treatment</td>
<td>18.495</td>
<td>0.887</td>
<td>3.959</td>
</tr>
</tbody>
</table>

then it may be that there is no voucher that can eliminate the low steady state. Similarly, if the population is pessimistic, then the equal treatment policy will have the disastrous effect of moving both groups to the low steady state. The subsidy policy, on the other hand, will always have the intended effect given that a large enough subsidy is applied.

3.7 Conclusion

The main contribution of this paper is to analyze the effectiveness of several government policies designed to eliminate discrimination in a dynamic environment. Unlike policy analysis in a static environment, I am able to determine the paths that each group will take once a policy is put in place and to address how long a policy needs to be in effect. The base model allows for competitive wages, and as a result I am also able to perform meaningful welfare analysis of the three policies in a general equilibrium setting.

I find that the hiring subsidy policy will always effectively eliminate discrimination but it results in a significant loss in welfare to the advantaged group while in place. However, this policy does have the advantage of needing to be in effect for the shortest amount of time. I also find that the investment voucher is a successful policy as long as a large enough proportion of the investment costs are monetary in nature. This policy may be attractive to policy makers because it involves very little loss in welfare to the advantaged group.

The equal treatment policy may or may not be successful depending on several conditions. The policy is safe to implement in industries where a majority of the workers are in the advantaged group or if the population as a whole is optimistic. If this is the case, then it may be the preferred policy, as it requires no transfers from advantaged to disadvantaged workers and welfare losses are relatively small. However, the equal treatment policy should not be
applied if a large proportion of workers are disadvantaged or if the population as a whole is pessimistic.

I also determine that all three of these policies need only be in effect until the investment rates of both groups are in the certain region of the high steady state. Once this occurs the policies can be safely removed and the groups will converge to the high steady state on their own.
Appendix A

Proof of Proposition 1

The following proof is a slightly modified version of one found in Moro and Norman (2003, 2002). I make only slight adjustments to suit my purposes, as such all credit should be given to them. Due to the linear nature of the firms’ production function, I will address each group separately. For simplicity in notation I drop all group indicators $J$.

Proof. ( Sufficiency). Given $\pi \in (0,1]$, let $\theta(\pi)$ be the unique solution to the task assignment problem, $\tau : [0,1] \to [0,1]$ be the threshold rule with cut-off $\theta(\pi)$ and $(C(\pi), S(\pi)) = (\pi F_q(\theta), \pi F_q(\theta) + (1 - \pi) F_u(\theta))$ be the associated aggregate factor inputs. Suppose that each firm posts the wage schedule $\omega : [0,1] \to \mathbb{R}$ given by $\omega(\theta; \pi)$ in (2.7). Moreover, suppose that workers facing indifference between working for firm 1 and firm 2 based on wage and task assignments (in equilibrium this will be all workers) simply flip a fair coin. This implies that

$$ C_1(\pi) = C_2(\pi) = \frac{C(\pi)}{2} $$
$$ S_1(\pi) = S_2(\pi) = \frac{S(\pi)}{2}. \tag{A.1} $$

The profit for firm 1 is (firm 2 is equivalent)

$$ \Pi_1 = \alpha_s S_1(\pi) + \alpha_C C_1(\pi) - \frac{1}{2} \int_0^1 \omega(\theta; \pi)[\pi f_q(\theta) + (1 - \pi) f_u(\theta)]d\theta $$
$$ = \alpha_s S_1(\pi) + \alpha_C C_1(\pi) - \frac{1}{2} \int_{\theta(\pi)}^{\theta(\pi)} \alpha_s [\pi f_q(\theta) + (1 - \pi) f_u(\theta)]d\theta $$
$$ - \frac{1}{2} \int_{\theta(\pi)}^{1} \frac{\pi f_q(\theta)}{\pi f_q(\theta) + (1 - \pi) f_u(\theta)}[\pi f_q(\theta) + (1 - \pi) f_u(\theta)]d\theta $$
$$ \tag{A.2} $$
\[ \Pi_1 = \alpha_s S_1(\pi) + \alpha_c C_1(\pi) - \frac{1}{2} [\alpha_s \int_0^{\theta(\pi)} [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta + \alpha_c \int_{\theta(\pi)}^1 \pi f_q(\theta) d\theta] \\
= \alpha_s S_1(\pi) + \alpha_c C_1(\pi) - \frac{1}{2} [\alpha_s S(\pi) + \alpha_c C(\pi)] = 0. \quad (A.3) \]

Suppose one firm deviates to \((\omega', \tau') \neq (\omega, \tau)\). Let \(C'\) and \(S'\) denote the implied factor inputs and let \(a(\theta) \in [0, 1]\) denote the fraction of workers with signal \(\theta\) that accepts a job at the deviating firm. Since \(\omega'(\theta; \pi) \geq \omega(\theta; \pi)\) for all \(\theta\) such that \(a(\theta) > 0\) the profit for the deviating firm, \(\Pi_i'\), satisfies

\[ \Pi_i' \leq \alpha_s S' + \alpha_c C' - \int_0^1 \omega(\theta; \pi) a(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta, \quad (A.4) \]

where (here the assumption that ties are broken the same way by qualified and unqualified workers is used)

\[ C' = \int \tau'(\theta) \pi f_q(\theta) a(\theta) d\theta \text{ and } S' = \int (1 - \tau'(\theta)) a(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta. \quad (A.5) \]

Moreover \(\omega(\theta; \pi) = \max \{\alpha_c p(\theta, \pi), \alpha_s\}\) so,

\[ \int \omega(\theta; \pi) a(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \\
= \int \tau'(\theta) \omega(\theta; \pi) a(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \\
+ \int (1 - \tau'(\theta)) \omega(\theta; \pi) a(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \\
\geq \alpha_s \int (1 - \tau'(\theta)) a(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta + \alpha_c \int \tau'(\theta) \pi f_q(\theta) a(\theta) d\theta \\
= \alpha_s S' + \alpha_c C'. \quad (A.6) \]

This implies that \(\Pi_i' \leq \alpha_s S' + \alpha_c C' - [\alpha_s S' + \alpha_c C'] = 0. \)

To prove necessity of the conditions in Proposition 1 I proceed by proving a sequence of intermediate results:

**Lemma 16.** Each firm earns zero profit in equilibrium.
Proof. Let $\Pi_1$ and $\Pi_2$ denote the profits for firm 1 and firm 2 and assume for contradiction that $\Pi_1 > 0$. If $\Pi_2 < 0$ there would be a profitable deviation, so I assume without loss that $0 \leq \Pi_2 \leq \Pi_1$. Total industry profits are

$$\Pi_1 + \Pi_2 = \alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2 - \int_\theta \max\{\omega_1(\theta), \omega_2(\theta)\} [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta. \quad (A.7)$$

I observe that $\alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2 \leq \alpha_s S(\pi) + \alpha_c C(\pi)$. This simply means that aggregate output cannot exceed what a planner could achieve. Suppose firm 2 deviates by offering $\omega'_2$ given by $\omega'_2(\theta) = \max\{\omega_1(\theta), \omega_2(\theta)\} + \epsilon$ for some $\epsilon > 0$, implying that firm 2 attracts all workers. In addition suppose firm 2 assigns as in the solution to (2.2). The corresponding profit is

$$\Pi'_2 = \alpha_s S(\pi) + \alpha_c C(\pi) - \int_\theta \max\{\omega_1(\theta), \omega_2(\theta)\} [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta - \epsilon$$

$$= \alpha_s S(\pi) + \alpha_c C(\pi) - [\alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2] + \Pi_1 + \Pi_2 - \epsilon$$

$$\geq \Pi_1 + \Pi_2 - \epsilon. \quad (A.8)$$

Hence for $\epsilon$ sufficiently small $\Pi'_2 > \Pi_2$ is a profitable deviation. Thus proving the statement by contradiction.

$\square$

Lemma 17. $\omega_1(\theta) = \omega_2(\theta)$ for almost all $\theta \in [0,1]$ in any equilibrium.

Proof. For contradiction, suppose that $\omega_1(\theta) > \omega_2(\theta)$ for all $\theta \in \Theta' \subset [0,1]$ and let

$$\beta = \int_{\theta \in \Theta'} (\omega_1(\theta) - \omega_2(\theta))[\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta, \quad (A.9)$$

where $\Theta'$ has a positive measure, which implies that $\beta > 0$. Let $C_1, C_2, S_1, S_2$ be the effective factor inputs in the hypothetical equilibrium and suppose that firm 1 deviates and offers $\omega'_1$ given by $\omega'_1(\theta) = \omega_2(\theta) + \epsilon$ for all $\theta$ and assigns all workers (the deviation attracts all workers)
Then there is a pair $(\omega_1, \omega_2)$ such that $(1) \omega_i(\theta) = s_i$ for $i = 1, 2$ and for almost all $\theta < \theta(\pi)$, $(2) \omega_i(\theta) = p(\theta, \pi)k_c$ for $i = 1, 2$ and for almost all $\theta \geq \theta(\pi)$.

Proof. The two parts have almost identical proofs, I will prove only part (2), which may appear as less obvious. Let $\omega(\theta) = \max\{\omega_1(\theta), \omega_2(\theta)\}$ and $(C(\pi), S(\pi))$ be the factor inputs

\[
\Pi'_1(\epsilon) = \alpha_s S(\pi) + \alpha_c C(\pi) - \int_\theta \omega_2(\theta)[\pi f_q(\theta) + (1 - \pi)f_u(\theta)]d\theta - \epsilon \\
= \alpha_s S(\pi) + \alpha_c C(\pi) - \int_\theta \omega_1(\theta)[\pi f_q(\theta) + (1 - \pi)f_u(\theta)]d\theta + \beta \\
- \int_{\theta \in [0,1) \setminus \Theta'} \omega_2(\theta)[\pi f_q(\theta) + (1 - \pi)f_u(\theta)]d\theta - \epsilon \\
\geq \alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2 - \int_\theta \max \{\omega_1(\theta), \omega_2(\theta)\} [\pi f_q(\theta) + (1 - \pi)f_u(\theta)]d\theta + \beta - \epsilon \\
= \beta - \epsilon, \tag{A.10}
\]

where the last inequality follows from Lemma 16. The deviation is thus profitable if $\epsilon$ is small enough, thus the equilibrium where $\omega_1(\theta) > \omega_2(\theta)$ cannot be optimal. \hfill \Box

**Lemma 18.** $\alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2 = \alpha_s S(\pi) + \alpha_c C(\pi)$.

**Proof.** By feasibility $\alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2 \leq \alpha_s S(\pi) + \alpha_c C(\pi)$, so assume for contradiction that $\alpha_s S(\pi) + \alpha_c C(\pi) - \alpha_s S_1 - \alpha_c C_1 - \alpha_s S_2 - \alpha_c C_2 = \beta > 0$. Suppose firm 1 offers $\omega'_1(\theta) = \omega_2(\theta) + \epsilon$ for all $\theta$ and assigns all workers (the deviation attracts all workers) in accordance with (2.2), the solution to the task assignment problem. The implied profits are

\[
\Pi'_1(\epsilon) = \alpha_s S(\pi) + \alpha_c C(\pi) - \int_\theta \omega_2(\theta)[\pi f_q(\theta) + (1 - \pi)f_u(\theta)]d\theta - \epsilon \\
> \alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2 - \int_\theta \omega_2(\theta)[\pi f_q(\theta) + (1 - \pi)f_u(\theta)]d\theta - \epsilon. \tag{A.11}
\]

But (Lemma 17) $\omega_1(\theta) = \omega_2(\theta)$ almost everywhere so $\int_\theta \omega_2(\theta)[\pi f_q(\theta) + (1 - \pi)f_u(\theta)]d\theta$ is the sum of wages paid out by firms 1 and 2 before the deviation. By zero profits (Lemma 16) this implies that $\Pi'_1(\epsilon) = \beta - \epsilon$, so for $\epsilon$ small enough the deviation is profitable. \hfill \Box

**Lemma 19.** Suppose $(\omega_1, \omega_2)$ is a pair of equilibrium wage schedules and let $\theta(\pi)$ be as in (3). Then there is a pair $(k_s, k_c)$ such that (1) $\omega_i(\theta) = k_s$ for $i = 1, 2$ and for almost all $\theta < \theta(\pi)$, $(2) \omega_i(\theta) = p(\theta, \pi)k_c$ for $i = 1, 2$ and for almost all $\theta \geq \theta(\pi)$.

**Proof.** The two parts have almost identical proofs, I will prove only part (2), which may appear as less obvious. Let $\omega(\theta) = \max\{\omega_1(\theta), \omega_2(\theta)\}$ and $(C(\pi), S(\pi))$ be the factor inputs
corresponding to (2.2) the unique solution to the task assignment problem. For contradiction suppose there is a set $A \subset [\theta(\pi), 1]$, where $m = \int_A \pi f_q(\theta) d\theta > 0$, and for some $\beta > 0$ such that for all $\theta \in A$,

$$\frac{\omega(\theta)}{p(\theta, \pi)} \leq \frac{1}{1 - F_q(\theta)} \int_{\theta}^{1} \frac{\omega(\theta)}{p(\theta, \pi)} f_q(\theta) d\theta - \beta$$

$$= \frac{1}{\pi(1 - F_q(\theta))} \int_{\theta}^{1} \omega(\theta)[\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta - \beta$$

$$= \frac{1}{C(\pi)} \int_{\theta(\pi)}^{1} \omega(\theta)[\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta - \beta. \tag{A.12}$$

By continuity there exists a set $B \in [0, \theta(\pi))$ such that $\int_B [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta = \frac{S(\pi)}{C(\pi)} m$ and

$$\omega(\theta) \leq \frac{1}{\pi F_q(\theta) + (1 - \pi) F_u(\theta)} \int_{\theta(\pi)}^{1} \omega(\theta)[\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta$$

$$= \frac{1}{S(\pi)} \int_{0}^{\theta(\pi)} \omega(\theta)[\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \tag{A.13}$$

for every $\theta \in B$. Consider a deviation by firm $i$ where it offers $\omega_i'(\theta) = \omega(\theta) + \epsilon$ to workers with $\theta \in A \cup B$ and $\omega_i'(\theta) = 0$ for all other $\theta$, and assigns workers from $A$ to the complex task and workers from $B$ to the simple task. The profit from this deviation is

$$\Pi' = \alpha_c \int_A \pi f_q(\theta) d\theta + \alpha_s \int_B [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta$$

$$- \int_{\theta \in A \cup B} (\omega(\theta) + \epsilon)[\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta$$

$$\geq [\alpha_c C(\pi) + \alpha_s S(\pi)] \frac{m}{C(\pi)}$$

$$- \left( \frac{1}{C(\pi)} \int_{\theta(\pi)}^{1} \omega(\theta)[\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta - \beta \right) \int_{\theta \in A} p(\theta, \pi) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta$$

$$- \int_{\theta \in B} [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \left[ \frac{1}{S(\pi)} \int_{0}^{\theta(\pi)} \omega(\theta)[\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \right]$$

$$- \epsilon \int_{\theta \in A \cup B} [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \tag{A.14}$$
= [\alpha_c C(\pi) + \alpha_s S(\pi)] \frac{m}{C(\pi)} - \left( \int_{\theta(\pi)}^{1} \omega(\theta)[\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta - \beta \right) \frac{m}{C(\pi)} \\
- \frac{S(\pi)}{C(\pi)} \left[ \frac{1}{S(\pi)} \int_{0}^{\theta(\pi)} \omega(\theta)[\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta \right] - \epsilon \int_{\theta \in A \cup B} [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta \\
= \frac{m}{C(\pi)} \left( \alpha_c C(\pi) + \alpha_s S(\pi) - \int_{\theta \in [0,1]} \omega(\theta)[\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta \right) \\
= 0 \text{ by Lemmas 16 and 17} \\
+ \beta \frac{m}{C(\pi)} - \epsilon \int_{\theta \in A \cup B} [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta. \hspace{1cm} (A.15)

Hence, \( \Pi' \geq \beta \frac{m}{C(\pi)} - \epsilon \int_{\theta \in A \cup B} [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta > 0 \) for \( \epsilon \) small enough, which together with Lemma 17 establishes part (2) of the claim. The proof for the other half is symmetric. It’s simply a matter of removing \( \beta \) from (A.12) and inserting a \( \beta \) in the inequality in (A.13) and again constructing an \( A \) and \( B \) such that the factor ratio is as in the unique threshold in (2.2). The rest of the argument is unaltered. \( \square \)

**Proof. (Necessity).** It remains to be shown that \( k_s = \alpha_s \) and \( k_c = P(\theta, \pi)\alpha_c \). Firms would make positive profits if \( k_s < \alpha_s \) and \( k_c < P(\theta, \pi)\alpha_c \) and negative profits if the inequalities go the other way. Thus I must only consider cases where the inequalities are of opposing directions.

The arguments are symmetric so I only consider the case with \( k_s > \alpha_s \) and \( k_c < P(\theta, \pi)\alpha_c \). If \( \theta(\pi) = 0 \), (1) each firm makes positive profits (loss), so the only case to consider is when \( \theta(\pi) \) is interior. A necessary condition for optimality for problem (2.2) is that \( \alpha_c P(\theta(\pi), \pi) = \alpha_s \).

Hence there must be an interval \((\theta(\pi), \theta^*)\) where \( \omega_i(\theta) = p(\theta, \pi)k_c < k_s \) for all \( \theta \in (\theta(\pi), \theta^*) \). Consider the deviation

\[
\omega_i(\theta) = \begin{cases} 
  \omega(\theta) & \text{for } \theta \in (\theta(\pi), \theta^*) \\
  0 & \text{otherwise}
\end{cases} \quad \text{and} \quad \tau_i(\theta) = \begin{cases} 
  0 & \text{for } \theta \in (\theta(\pi), \theta') \\
  1 & \text{for } \theta \in (\theta', \theta^*)
\end{cases}, \hspace{1cm} (A.16)
\]

where \( \theta' \) is set so that the factor ratio is as in the solutions to (2.2),

\[
\frac{\int_{\theta(\pi)}^{\theta'} \pi f_q(\theta) d\theta}{\int_{\theta(\pi)}^{\theta'} [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta} = \frac{C(\pi)}{S(\pi)}.
\]
The profit is

$$
\Pi' = \alpha_c \int_{0'} \pi f_q(\theta) d\theta + \alpha_s \int_{\theta(\pi)}^{0'} [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta
$$

$$
- \int_{\theta(\pi)}^{0'} \omega(\theta)[\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta - \int_{0'}^{\theta(\pi)} \omega(\theta)[\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta
$$

$$
\geq \frac{(\alpha_c C(\pi) + \alpha_s S(\pi)) \int_{0'}^{0'} \pi f_q(\theta) d\theta}{C(\theta)} - \alpha_c \int_{\theta(\pi)}^{0'} [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta - \alpha_s \int_{0'}^{\theta(\pi)} \pi f_q(\theta) d\theta
$$

$$
= \frac{1}{C(\pi)} \int_{0'}^{0'} \pi f_q(\theta) d\theta [\alpha_c C(\pi) + \alpha_s S(\pi) - \alpha_s S(\pi) - \alpha_c C(\pi)] = 0.
$$

which completes the proof of Proposition 1. \qed
Mathematical Methods

B.1 Calculating paths to the steady states

Calculating the paths to the steady states is relatively straightforward. First, it is necessary
to determine all the steady state values by finding the solutions to the following system of
equations:

\[
\Pi^* = G(\Gamma^*) \quad (B.1)
\]
\[
\Gamma^* = \frac{\gamma(\Pi^*)}{\beta + r}. \quad (B.2)
\]

Once \(\Sigma_h(\Pi_h, \Gamma_h), \Sigma_m(\Pi_m, \Gamma_m),\) and \(\Sigma_l(\Pi_l, \Gamma_l)\) are known, I can determine the path to the
high steady state by evaluating

\[
\Pi_{t-1}^J = \Pi_t^J - \beta G(\Gamma_t^J) \quad \text{and} \quad (B.3)
\]
\[
\Gamma_{t-1}^J = (1 - \beta - r)\Gamma_t^J + \gamma(\Pi_{t-1}^J) \quad (B.4)
\]
at \(\Pi_T = \Pi_h - \epsilon\) and \(\Gamma_T = \Gamma_h - \epsilon\). This process yields \(\Pi_{T-1}\) and \(\Gamma_{T-1}\), which can then be used
to find \(\Pi_{T-2}\) and \(\Gamma_{T-2}\). I repeat this process until the middle steady state is reached. The
result is a series, \(\{\Gamma_t, \Pi_t\}_{t=n}^T\), that maps out the path from the middle steady state to the high
steady state.

I can determine the path to the low steady state by evaluating equations (B.3) and (B.4)
at \(\Pi_T = \Pi_l + \epsilon\) and \(\Gamma_T = \Gamma_l + \epsilon\). This process yields \(\Pi_{T-1}\) and \(\Gamma_{T-1}\), which can then be used
to find \(\Pi_{T-2}\) and \(\Gamma_{T-2}\). I repeat this process until the middle steady state is reached. The
result is a series, \(\{\Gamma_t, \Pi_t\}_{t=n}^T\), that maps out the path from the middle steady state to the low
steady state.
For the parameterization described in Section 3.4 the steady states are the solutions to the following system of equations:

\[ \Pi^* = \Gamma^* + 0.10 \]  \hspace{1cm} (B.5)
\[ \Gamma^* = 4\gamma(\Pi^*) \]  \hspace{1cm} (B.6)

where

\[ \gamma(\Pi^J_t) = \begin{cases} 
0 & \text{if } \Pi^J_t \leq 1/3 \\
\frac{1}{3}(\frac{3\Pi^J_t-1}{\Pi^J_t+1}) & \text{if } 1/3 < \Pi^J_t < 2/3 \\
\frac{2}{3}(\frac{2\Pi^J_t}{\Pi^J_t+1} - \frac{\Pi^J_t}{3-\Pi^J_t}) & \text{if } \Pi^J_t \geq 2/3.
\end{cases} \]  \hspace{1cm} (B.7)

Solving this yields the following three steady states: \( \Sigma_h(\Gamma^h = 0.6631, \Pi^h = 0.7631) \), \( \Sigma_m(\Gamma^m = 0.3687, \Pi^m = 0.4687) \) and \( \Sigma_l(\Gamma^l = 0.0, \Pi^l = 0.1000) \).

In order to generate the optimistic path pictured in Figure 3.5, I induct backwards from \( \Pi_T = \Pi_h - 0.0001 \) and \( \Gamma_T = \Gamma_h - 0.0001 \) using the following equations:

\[ \Pi^J_{t-1} = 1.25\Pi^J_t - 0.25\Gamma^J_t - 0.025 \]  \hspace{1cm} (B.8)
\[ \Gamma^J_{t-1} = 0.75\Gamma^J_t + \gamma(\Pi^J_{t-1}). \]  \hspace{1cm} (B.9)

The resulting series \( \{\Gamma_t, \Pi_t\}_{t=n}^T \) describes the optimistic path. Similarly, I generate the pessimistic path by inducting backwards from \( \Pi_T = \Pi_l + 0.0001 \) and \( \Gamma_T = \Gamma_l + 0.0001 \). The resulting series, \( \{\Gamma_t, \Pi_t\}_{t=n}^T \), describes the pessimistic path.

### B.2 Welfare calculations

#### B.2.1 Welfare in the base model

In order to understand the welfare effects of the three policies, it is necessary to determine the average expected welfare in both the high and low steady states of the base model for the parameterization described in Section 3.4. The general equation for the average expected welfare...
welfare in one period is

\[ W(\Pi_t^I) = \Pi_t^I \left[ \int_0^1 \omega_\Delta(\theta, \Pi_t^I) f_q(\theta) d\theta - \beta \left( \frac{f_r^G(\Pi_t^I) c_g(c) dc}{G(\Gamma(\Pi_t^I))} \right) \right] + (1 - \Pi_t^I) \int_0^1 \omega_\Delta(\theta, \Pi_t^I) f_u(\theta) d\theta. \] (B.10)

For the parameterization discussed in Section 3.3 the average expected welfare can be simplified to

\[ W(\Pi_t^I) = \Pi_t^I \left[ E_t[\omega_\Delta|q] - \beta \left( \frac{\Gamma(\Pi_t^I) - 0.10}{2} \right) \right] + (1 - \Pi_t^I) E_t[\omega_\Delta|u], \] (B.11)

where

\[ E_t[\omega_\Delta|u] = \begin{cases} 1 & \text{if } \Pi_t^I \leq 1/3 \\ \frac{2}{3} \left( \frac{3\Pi_t^I + 1}{\Pi_t^I + 1} \right) & \text{if } 1/3 < \Pi_t^I < 2/3 \\ \frac{2}{3} \left( \frac{2\Pi_t^I}{2 - \Pi_t^I} + \frac{2\Pi_t^I}{1 + \Pi_t^I} \right) & \text{if } \Pi_t^I \geq 2/3 \end{cases} \] (B.12)

and

\[ E_t[\omega_\Delta|q] = \begin{cases} 1 & \text{if } \Pi_t^I \leq 1/3 \\ \frac{1}{3} \left( \frac{9\Pi_t^I + 1}{\Pi_t^I + 1} \right) & \text{if } 1/3 < \Pi_t^I < 2/3 \\ \frac{2}{3} \left( \frac{\Pi_t^I}{2 - \Pi_t^I} + \frac{4\Pi_t^I}{1 + \Pi_t^I} \right) & \text{if } \Pi_t^I \geq 2/3 \end{cases} \] (B.13)

Evaluating (B.11) at the high and low steady states provides the baseline welfare level for workers in the advantaged and disadvantaged groups. In this case, \( W(\Pi_h) = 1.4833 \) and \( W(\Pi_l) = 1.0010 \).

### B.2.2 Welfare under a hiring subsidy

The first step to determining the welfare effects of the hiring subsidy policy is to calculate the path that the disadvantaged group will take when the policy is put in place as well as the path that they will take to the high steady state when the policy is lifted. Once these paths are
established, I can calculate the welfare of each group for each period of interest.

I begin by determining the single steady state of the dynamic system under the policy. This steady state is the solution to the following system of equations:

\[ \Pi^\xi = G(\Gamma^\xi) \]  
\[ \Gamma^\xi = \frac{\gamma^\xi(\Pi^\xi)}{\beta + r}. \]  

I refer to the single steady state as \( \Sigma^\xi(\Gamma^\xi, \Pi^\xi) \). Once this is determined I find the disadvantaged group’s path to the steady state by inducting backwards from \( \Gamma_T = \Gamma^\xi - \epsilon \) and \( \Pi_T = \Pi^\xi - \epsilon \) using the following equations:

\[ \Pi^d_{t-1} = \frac{\Pi^d_t - \beta G(\Gamma^d_t)}{1 - \beta} \]  
\[ \Gamma^d_{t-1} = (1 - \beta - r)\Gamma^d_t + \gamma^\xi(\Pi^d_{t-1}). \]  

To determine the path taken after the policy is removed I use (B.3) and (B.4) and induct backwards from \( \Pi_T = \Pi_h + \epsilon \) and \( \Gamma_T = \Gamma_h + \epsilon \).

For the parameterization of interest the steady state is \( \Sigma^\xi(\Gamma^\xi = 0.8159, \Pi^\xi = 0.9159) \). This is determined by finding the solution to the following system of equations:

\[ \Pi^\xi = \Gamma^\xi + 0.10 \]  
\[ \Gamma^\xi = 4\gamma^\xi(\Pi^\xi), \]  

where

\[ \gamma^\xi(\Pi^d_t) = \begin{cases} 
0 & \text{if } \Pi^d_{t-1} \leq 0.10 \\
\frac{1}{3} \left( \frac{10\Pi^d_t - 1}{\Pi^d_{t+1}} \right) & \text{if } 0.10 < \Pi^d_t < 0.31 \\
\frac{11}{6} \left( \frac{2\Pi^d_t}{\Pi^d_{t+1}} - \frac{\Pi^d_t}{2 - \Pi^d_t} \right) & \text{if } \Pi^d_t \geq 0.31.
\end{cases} \]  

In order to find the path that the disadvantaged group will take to this steady state, I
induct backwards from $\Gamma_T = \Gamma^\xi - \epsilon$ and $\Pi_T = \Pi^\xi - \epsilon$ using the following equations:

\begin{align*}
\Pi^d_{t-1} &= 1.25\Pi^d_t - 0.25\Gamma^d_t - 0.025 \quad (B.21) \\
\Gamma^d_{t-1} &= 0.75\Gamma^d_t + \gamma^\xi(\Pi^d_{t-1}) \quad (B.22)
\end{align*}

where $\epsilon = 0.0000001$.

I use (B.8) and (B.9) and induct backwards from $\Pi_T = \Pi_h + \epsilon$ and $\Gamma_T = \Gamma_h + \epsilon$ to find the path taken after the policy is removed.

To determine the welfare effects while the subsidy is in place, I look at the $\Gamma_t, \Pi_t$ pairs on the path to $\Sigma^\xi$ starting from $\Pi_t$, the disadvantaged group starting point, through the steady state values under the policy. To find how welfare changes after the policy is lifted I look at the $\Gamma_t, \Pi_t$ pairs starting from $\Pi^\xi$ until the high steady state is reached. With these values I am able to calculate each group’s welfare during each period that the policy is in place as well as during the periods after it is removed.

The policy is applied in a general equilibrium framework so that the workers in the advantaged group are taxed in order to fund the policy. As such, the welfare of both groups will be impacted by the policy, even though the behavior of the advantaged group is not affected. The welfare of the advantaged group while the policy is in place is the following

\[ W^a(\Pi^a_t, \Pi^d_t) = W(\Pi^a_h) - \xi + \frac{\lambda^d}{\lambda^a} \Pi^d_t + \int_{\theta(\Pi^d_h)}^{1} f_q(\theta) d\theta. \]

(B.23)

This is simply the average welfare in the high steady state minus the average expected transfer. Note that the transfer each worker in the advantaged group provides depends on the relative sizes of the two groups. Clearly, if more of the population is disadvantaged, then the loss in welfare to those in the advantaged group is higher.

The average expected welfare of the disadvantaged group while the policy is in place is the
following

\[
W^d(\Pi^d_t) = \Pi^d_t \left[ \int_0^1 \omega^\xi(\theta, \Pi^d_t) f_q(\theta)d\theta - \beta \left( \frac{\int \Gamma(\Pi^d_t) cg(c) dc}{G(\Gamma(\Pi^d_t))} \right) \right] + (1 - \Pi^d_t) \int_0^1 \omega^\xi(\theta, \Pi^d_t) f_u(\theta)d\theta. \tag{B.24}
\]

Expected wage of qualified workers under the subsidy policy

Expected wage of unqualified workers under the subsidy policy

The only way that this differs from (B.10) is that it depends on the expected wage under the subsidy policy, \( \omega^\xi(\theta, \Pi^d_t) \); all other components are the same.

For the parameterization, the advantaged group’s average expected welfare while the policy is in place is

\[
W^a(\Pi^a_t, \Pi^d_t) = 1.4833 - \frac{7\Pi^d_t}{6} P[\theta_t > \hat{\theta}(\Pi^d_t)|q], \tag{B.25}
\]

where

\[
P[\theta_t > \hat{\theta}(\Pi^d_t)|q] = \begin{cases} 
0 & \text{if } \Pi^d_t \leq 0.10 \\
\frac{2}{3} & \text{if } 0.10 < \Pi^d_t < 0.31 \\
1 & \text{if } \Pi^d_t \geq 0.31.
\end{cases} \tag{B.26}
\]

The average expected welfare for the disadvantaged group is

\[
W^d_t(\Pi^d_t) = \Pi^d_t \left( E_t[\omega^\xi|u] - \beta \left( \frac{\Gamma(\Pi^d_t) - 0.10}{2} \right) \right) + (1 - \Pi^d_t) E_t[\omega^\xi|u], \tag{B.27}
\]

where

\[
E_t[\omega^\xi|u] = \begin{cases} 
1 & \text{if } \Pi^d_t \leq 0.10 \\
\frac{13\Pi^d_t + 2}{3(\Pi^d_t + 1)} & \text{if } 0.10 < \Pi^d_t < 0.31 \\
\frac{11}{3} \left( \frac{\Pi^d_t}{2 - \Pi^d_t} + \frac{\Pi^d_t}{1 + \Pi^d_t} \right) & \text{if } \Pi^d_t \geq 0.31
\end{cases} \tag{B.28}
\]
and

\[
E_t[\omega^\xi_q] = \begin{cases} 
1 & \text{if } \Pi^d_t \leq 0.10 \\
\frac{1}{3}(23\Pi^d_t + 1) & \text{if } 0.10 < \Pi^d_t < 0.31 \\
\frac{11}{6}(\frac{\Pi^d_t}{1-\Pi^d_t} + \frac{4\Pi^d_t}{1+\Pi^d_t}) & \text{if } \Pi^d_t \geq 0.31.
\end{cases}
\]  

(B.29)

To generate Figure 3.19 I evaluate (B.25) and (B.27) at all \(\Gamma_t, \Pi_t\) pairs along the path starting from \(\Pi_t = 0.10\) till the group is within 0.0001 of the single steady state of the dynamic system under the hiring subsidy. To calculate the welfare after the policy is removed I evaluate (B.11) at all \(\Gamma_t, \Pi_t\) pairs along the path from \(\Pi^\xi\) till the group is within 0.0001 of the high steady state.

**B.2.3 Welfare under an investment voucher**

In order to find the welfare effects of the investment voucher policy, I must determine the path that the disadvantaged group will take once the policy is put in place as well as the path that they will take to the high steady state once the policy is lifted. Once these paths are established, I can calculate the welfare of each group for each period that the policy is in place.

The first step to calculating the path that the disadvantaged group will take once the investment voucher policy is implemented is to determine the single steady state of the dynamic system under the policy. This steady state is the solution to the following system of equations:

\[
\Pi^v = G(\Gamma^v + v) \\
\Gamma^v = \frac{\gamma(\Pi^v)}{\beta + r}.
\]  

(B.30, B.31)

I refer to it as \(\Sigma^v(\Gamma^v, \Pi^v)\). Once this is determined, I can find the disadvantaged group’s path to the steady state by inducting backwards from \(\Gamma_T = \Gamma^v - \epsilon\) and \(\Pi_T = \Pi^v - \epsilon\) using the following equations:

\[
\Pi^d_{t-1} = \frac{\Pi^d_t - \beta G(\Gamma^d_t + v)}{1 - \beta}
\]  

(B.32)

\[
\Gamma^d_{t-1} = (1 - \beta - r)\Gamma^d_t + \gamma(\Pi^d_{t-1}).
\]  

(B.33)
I then use (B.3) and (B.4) and induct backwards from $\Pi_T = \Pi_h + \epsilon$ and $\Gamma_T = \Gamma_h + \epsilon$ to find the path taken after the policy is removed.

For the simple parameterization, the steady state is $\Sigma^u(\Gamma^u = 0.5058, \Pi^u = 0.8391)$. This is determined by finding the solution to the following system of equations:

\[
\begin{align*}
\Pi^u &= \Gamma^u + \frac{1}{3} \\
\Gamma^u &= 4\gamma(\Pi^u),
\end{align*}
\]

where $\gamma(\cdot)$ is as in (B.7).

In order to find the path that the disadvantaged group will take to this steady state, I induct backwards from $\Gamma_T = \Gamma^u - \epsilon$ and $\Pi_T = \Pi^u - \epsilon$ using the following equations:

\[
\begin{align*}
\Pi^d_{t-1} &= 1.25\Pi^d_t - 0.25\Gamma^d_t - 0.0833 \\
\Gamma^d_{t-1} &= 0.75\Gamma^d_t + \gamma(\Pi^d_{t-1}),
\end{align*}
\]

where $\epsilon = 0.00000001$.

To determine the path taken after the policy is removed I use (B.8) and (B.9) and induct backwards from $\Pi_T = \Pi_h + \epsilon$ and $\Gamma_T = \Gamma_h + \epsilon$.

I can calculate the welfare effects while the voucher is in place by looking at all the $\Gamma_t, \Pi_t$ pairs along the path to $\Sigma^u$ starting from $\Pi_l$, the disadvantaged group starting point. To find how welfare changes after the policy is lifted, I look at all the $\Gamma_t, \Pi_t$ pairs along the path to $\Sigma_h$ starting from the $\Pi^u$. With these values, I can to calculate each group’s welfare during each period that the policy is in place as well as during the periods after it is removed.

The policy is applied in a general equilibrium framework so that the workers in the advantaged group are taxed in order to fund the policy, as such the welfare of both groups will be impacted by the policy, even though the behavior of the advantaged group is not affected.
The welfare of the advantaged group while the policy is in place is

\[ W^a(\Pi^a_t, \Pi^d_t) = W(\Pi^a_h) - \lambda^a \frac{\beta}{\lambda^a} G(\Gamma(\Pi^d_t) + \nu) \]  

(B.38)

Note that the transfer that each worker in the advantaged group provides depends on the relative sizes of the two groups. Clearly if more of the population is disadvantaged, then the loss in welfare to those in the advantaged group is higher.

The welfare of the disadvantaged group while the policy is in place is

\[ W^d(\Pi^d_t) = \Pi^d_t \left[ \int_0^1 \omega_\Delta(\theta, \Pi^d_t) f_q(\theta) d\theta - \beta \left( \frac{\int g(c \mid q) dc}{G(\Gamma(\Pi^d_t) + \nu)} - \nu \right) \right] + (1 - \Pi^d_t) \int_0^1 \omega_\Delta(\theta, \Pi^d_t) f_u(\theta) d\theta. \]  

(B.39)

The only way that this differs from (B.10) is that workers are more likely to invest (as a result of the voucher) and the costs associated with this investment are reduced by \( \nu \).

For the parameterization, the average expected welfare for the advantaged group while the policy is in place is

\[ W^a(\Pi^a_t, \Pi^d_t) = 1.4833 - 0.0156 \left( \Gamma(\Pi^d_t) + 0.1333 \right). \]  

(B.40)

The average expected welfare for the disadvantaged group while the policy is in place is

\[ W^d(\Pi^d_t) = \Pi^d_t \left( E_t[\omega_\Delta | q] - \beta \left( \frac{\Gamma(\Pi^d_t) - 0.20}{2} \right) \right) + (1 - \Pi^d_t) E_t[\omega_\Delta | u], \]  

(B.41)

where \( E_t[\omega_\Delta | u] \) and \( E_t[\omega_\Delta | q] \) are as in (B.12) and (B.13), respectively.

To generate Figure 3.20, I evaluate (B.40) and (B.41) at all \( \Gamma_t, \Pi_t \) pairs along the path to \( \Sigma^\nu \) starting from \( \Pi_t \), the disadvantaged groups starting point, till they are within 0.0001 of
Σ\(^v\). To calculate the welfare after the policy is removed I evaluate (B.11) at all \(\Gamma_t, \Pi_t\) pairs along the path to \(\Sigma_h\) starting from \(\Pi^v\) till they are within 0.0001 of the high steady state.

### B.2.4 Welfare under the equal treatment policy

Finding the welfare effects of the equal treatment policy is slightly more involved than the process for the previous policies. Two steps are necessary. First, I must find the path that the entire population takes to the high steady state once the policy is put in place. Using this path I can then find the individual paths of the two groups. It is important to consider the fact that the population as a whole will reach the steady state before the two groups do, as such I must account for several periods where the population investment rate is not changing but the group investment rates still are. Once the two groups reach the steady state, then the policy can be safely removed.

The dynamic system is the same under the equal treatment policy as it is without the policy, but it depends on the firms beliefs about the average investment rate of the entire population. As such the first step is to calculate the path that the population as a whole will take once the policy is put in place. I assume that the population is optimistic so that they will take the path to the high steady state. I determine the population’s path to the high steady state by inducting backwards from \(\Gamma^{pop}_T = \Gamma_h - \epsilon\) and \(\Pi^{pop}_T = \Pi_h - \epsilon\) using the following equations:

\[
\Pi^{pop}_{t-1} = \frac{\Pi^{pop}_t - \beta G(\Gamma^{pop}_t)}{1 - \beta} \quad \text{(B.42)}
\]
\[
\Gamma^{pop}_{t-1} = (1 - \beta - r)\Gamma^{pop}_t + \gamma(\Pi^{pop}_{t-1}) \quad \text{(B.43)}
\]

After this process I have all the \(\Gamma^{pop}_t, \Pi^{pop}_t\) pairs that form the path to the high steady state. I am only interested in the path starting at \(\Pi^{pop} = \lambda^a \Pi_h + \lambda^d \Pi_l\). Once I calculate the population path, I can find the paths that the individual groups will take by forward inducting...
from $\Pi^d_0 = \Pi_t$ and $\Pi^a_0 = \Pi_h$ using the following equations:

\[
\begin{align*}
\Pi^d_t &= (1 - \beta)\Pi^d_{t-1} + \beta G(\Gamma^\text{pop}_t) \quad (B.44) \\
\Pi^a_t &= (1 - \beta)\Pi^a_{t-1} + \beta G(\Gamma^\text{pop}_t). \quad (B.45)
\end{align*}
\]

Note that I know the relevant $\Gamma^\text{pop}_t$ values from the previous calculation of the population’s path the high steady state.

For the parameterization of interest the high steady state is $\Sigma_h(\Gamma_h = 0.6631, \Pi_h = 0.7631)$. In order to find the path that the population will take to this steady state I induct backwards from $\Gamma^\text{pop}_T = \Gamma_h - \epsilon$ and $\Pi^\text{pop}_T = \Pi_h - \epsilon$ where $\epsilon = 0.00000001$ using the following equations:

\[
\begin{align*}
\Pi^\text{pop}_{t-1} &= 1.25\Pi^\text{pop}_t - 0.25\Gamma^\text{pop}_t - 0.025 \quad (B.46) \\
\Gamma^\text{pop}_{t-1} &= .75\Gamma^\text{pop}_t + \gamma^\text{pop}(\lambda^a\Pi^a_{t-1} + \lambda^d\Pi^d_{t-1}). \quad (B.47)
\end{align*}
\]

After this process I know all the $\Gamma^\text{pop}_t, \Pi^\text{pop}_t$ pairs that form the path to the high steady state. I am only interested in the path starting at $\Pi^\text{pop} = 0.5973$. With this path I am able to determine the paths that the two groups will take by forward inducting from $\Pi^d_0 = 0.1000$ and $\Pi^a_0 = 0.7631$ using the following:

\[
\begin{align*}
\Pi^d_t &= 0.80\Pi^d_{t-1} + 0.20\Gamma^\text{pop}_t + 0.02 \quad (B.48) \\
\Pi^a_t &= 0.80\Pi^a_{t-1} + 0.20\Gamma^\text{pop}_t + 0.02. \quad (B.49)
\end{align*}
\]

In this parameterization it is also necessary to include ten periods where $\Pi^\text{pop}$ is at the steady state level but the individual groups are still approaching it.

In order to determine the average expected welfare of the disadvantaged group while the equal treatment policy is in place, I look at both the $\Gamma^\text{pop}_t, \Pi^\text{pop}_t$ pairs that form the population path and the $\Gamma^\text{pop}_t, \Pi^d_t$ pairs that form the disadvantaged group’s path starting from $\Pi^d = \Pi_t$ and $\Pi^\text{pop} = \lambda^a\Pi_h + \lambda^d\Pi_t$ until $\Sigma^\text{pop}$ is reached. For the advantaged group I consider both the $\Gamma^\text{pop}_t, \Pi^\text{pop}_t$ pairs that form the population path and the $\Gamma^\text{pop}_t, \Pi^a_t$ pairs that form the advantaged
group’s path starting from \( \Pi^a = \Pi_h \) and \( \Pi^{pop} = \lambda^a \Pi_h + \lambda^d \Pi_l \) until \( \Sigma^{pop} \) is reached. With these values I can then calculate each group’s welfare during each period that the policy is in place.

The welfare of the each group while the policy is in place is

\[
W^J(\Pi^J_t, \Pi_t^{pop}) = \Pi^J_t \left[ \int_0^1 \omega_\Delta(\theta, \Pi_t^{pop}) f_q(\theta) d\theta \right] \frac{\beta}{\Gamma(\Pi_t^{pop})} \left[ \int G(\Gamma(\Pi_t^{pop})) c g(c) dc \right]
\]

\[
+ (1 - \Pi^J_t) \int_0^1 \omega_\Delta(\theta, \Pi_t^{pop}) f_u(\theta) d\theta \quad \text{for } J = a, d.
\]

(B.50)

Note that these two welfare values depend on both the population investment rate and the relevant group investment rate.

For the parameterization the average expected welfare of each group while the policy is in place is

\[
W^J(\Pi^J_t, \Pi_t^{pop}) = \Pi^J_t \left( E_t[\omega_\Delta^{pop} | q] - \beta \left( \frac{\Gamma(\Pi_t^{pop})}{2} - 0.10 \right) \right) + (1 - \Pi^J_t) E_t[\omega_\Delta^{pop} | u],
\]

(B.51)

where \( E_t[\omega_\Delta^{pop} | u] \) and \( E_t[\omega_\Delta^{pop} | q] \) are as in (B.12) and (B.13), respectively. They are simply evaluated at \( \Pi_t^{pop} \) rather than \( \Pi^J_t \).

To generate Figure 3.21 I evaluate (B.51) at all relevant pairs along the paths from the starting point \( \Pi^{pop} = 0.5973, \Pi^d = 0.1000, \Pi^a = 0.7631 \) till both groups are within .0001 of the high steady state.
Bibliography


