THE IMPACT OF NEW TECHNOLOGIES ON HEALTH KNOWLEDGE AND BEHAVOR: EVIDENCE FROM KENYA, UGANDA, AND THE UNITED STATES

Samuel Halvor Masters

A dissertation submitted to the faculty at 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 Health Policy and Management in the Gillings School of Global Public Health.

Chapel Hill
2017

Approved by:

Harsha Thirumurthy

Sally Stearns

Deborah Tate

Justin Trogdon

John E. Paul

ABSTRACT

Samuel Halvor Masters: The impact of new technologies on health knowledge, behavior and outcomes: Evidence from Kenya, Uganda and the United States
(Under the direction of Harsha Thirumurthy)

New technologies have the potential to dramatically change health behavior in both high- and low-income countries. Increasingly, researchers and practitioners have begun to design interventions that leverage the benefits of these new technologies to improve health. These programs attempt to help individuals overcome a number of different barriers to desirable health behaviors. I explore three technologies and their impact on health behavior: oral HIV self-tests, web-based weight loss programs, and mobile phones. Oral HIV self-tests are HIV tests that can be conducted by the tester themselves, confidentially, and only requires swabbing the mouth for oral fluid. The web-based weight loss program is an interactive web portal that provides users with weight loss tips and allows users to input metrics to benchmark progress. Mobile phones are now ubiquitous, but little is understood about how their rapid expansion over the past 20 years has impacted health. I examine the effect of each of these technologies on specific health behaviors and explore heterogeneity in their utilization. I use a different dataset to examine each technology and utilize advanced econometric modeling to account for endogeneity of technology adoption.

ACKNOWLEDGEMENTS

I am grateful for the patients who participated in the self-testing study and the health facility and research staff that made the study possible. I am thankful for my committee, John E. Paul, Sally Stearns, Deborah Tate, Justin Trogdon, and especially my Chair, Harsha Thirumurthy, for all their patience with me as I worked through various versions of these papers. I wish to acknowledge my former mentors, Lou Garrison, Srikanth Kadiyala, Emmanuela Gakidou, Santosh Kumar, and Mike Hanlon for helping me to develop as a researcher and urging me to pursue a PhD. I am thankful for the strong support of my family and friends throughout the entire PhD program.

PREFACE

Chapter 2 was previously published in *PLOS Medicine* on November 8th, 2016. The full citation is included in a footnote at the beginning of the chapter.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANC	Antenatal Care
ART	Antiretroviral therapy
BMI	Body mass index
CACE	Complier average causal effect
CCA	Complete case analysis
CI	Confidence interval
ENV	Study group in WAY weight loss trial that received environmental change, including healthy food options in school cafeterias
FP	Family planning
HIV	Human immunodeficiency virus
HIVST	HIV self-testing
IPV	Intimate partner violence
IQR	Interquartile range
IV	Instrumental variables
MAR	Missing at random
MCAR	Missing completely at random

mHealth	Mobile health
MI	Multiple imputation
NACE	Non-complier average causal effect
PMTCT	Prevention of mother-to-child transmission of HIV
PPC	Postpartum care
PS	Propensity scores
RR	Risk ratio
SD	Standard deviation
SSA	Sub-Saharan Africa
SUTVA	Stable unit treatment value assumption
UNPS	Uganda National Panel Survey
WAY	Worksites Activities for You weight loss trial
WEB	Study group in WAY weight loss trial that received access a website for weight loss
WPI	Study group in WAY weight loss trial that received access a website for weight loss and financial incentives

CHAPTER 1: INTRODUCTION

1.1 Specific Aims:

Policymakers, industry, and researchers' desire to improve health behavior leads to significant innovation in health technology. Technological innovation for health behavior is driven by identifying barriers that prevent optimal individual level behavior and intervening on them. Examples of these technologies are numerous, from pedometers in mobile phones which help people keep track of weight loss goals, to diagnostic tools such as personal blood pressure cuffs which make potentially costly clinic visits for routine monitoring unnecessary. In addition to health-focused approaches, technologies designed for a different primary purpose other than health can have a direct impact on health knowledge and behavior. For example, the primary purpose of mobile phones was not health improvement; however, many innovative interventions utilize phones for population health gain [1, 2].

The overall objective of this dissertation was to determine if specific technological advances, such as mobile phones and new diagnostics, are capable of transforming health behaviors. Identifying these impacts required empirical methods to address the endogenous adoption of new technologies, with people "selecting" into technology use and ownership. Unobserved factors can lead to adoption and may be correlated positively or negatively with certain health behaviors, thus leading to spurious conclusions about the true impact of technology. The long-term goal of this research is to better understand how technology can be effectively used to improve not only health behavior but ultimately population health.

I explored three different ways in which technology may influence health behavior by overcoming barriers faced by the consumer. First, new easy-to-use diagnostics have potentially reduced the costs associated with practicing healthy behavior. For example, HIV self-tests were designed to provide confidential, immediate results to users, overcoming barriers related to clinic visits and stigma[3, 4]. Second, some technologies attempt to overcome psychological biases by employing behavioral economic principles to improve behavior[5-7]. For example, web-based interventions can nudge individuals towards healthier habits by sending salient, tailored messages[8]. These technologies can be paired with traditional economic incentives to maximize their potential for health gain. Finally, new technologies can change health behavior indirectly, such as increasing communication between peers using mobile phones, which could lead to greater awareness of health behaviors.

I tested the above hypothesized relationships using three different technologies. Using data from multiple countries I aimed to:

- Aim 1: Determine whether distribution of HIV self-tests to women receiving antenatal or postnatal care leads to higher rates of HIV testing among their primary partners than distribution of referral cards for clinic testing.
- Aim 2.1: Understand utilization of a web-based weight loss intervention (WWI). Aim 2.2: Determine if utilization differed among employees in colleges randomized to receive the WWI alone compared with employees in colleges randomized to receive the WWI plus financial incentives for weight loss. Aim 2.3: Determine if utilization of the WWI was associated with weight loss.

- Aim 3: Determine if owning a mobile phone leads women to know and use family planning (FP) methods, use health facilities for delivery, and give birth with medical personnel present.

For Aims 1 and 3 I examined whether the impacts of technology were heterogeneous based on characteristics of users such as rurality, education, and age. This heterogeneity analysis enabled a greater understanding of which subpopulations benefit most from technology available today, and which subpopulations may need more targeted technological interventions in the future. Understanding how technology can remove barriers and influence behavior helps identify the most effective technological means to change health outcomes. This research is particularly important today as increasingly more resources are invested in technology to improve population health.

1.2 Significance

Technological advances such as new drugs and machines have historically played a vital role in improving curative health; in public health, the same is true of sanitation and clean water [9, 10]. Although morbidity and mortality have improved due to technology, a large proportion of the population has not benefited from these technologies as one would predict. Behavioral factors have been a major reason why maximum health gains have not been achieved. People can know that something benefits them but still not act on that information. This gap has generated an innovative area for health researchers: health behavior change using technology. One of the main goals of my research is to help design technologies that reduce the barriers people face to behave "health consciously." The barriers that people face in practicing good health behavior are numerous; the barriers depend on which behavior someone is trying to change, and also individual idiosyncrasies.

For my dissertation I aimed to determine if three different new technologies improved health behavior. Each of the technologies I analyzed attempts to reduce barriers in one of two ways, either by removing barriers that are inherent to an existing health product, or reducing barriers that are intrinsic to the individual. The example of HIV testing (Aim 1) can be used to illustrate these issues. One of the main contributors to the high burden of HIV is the low rates of HIV testing, especially in low-resource settings. Periodic HIV testing, especially among those at high risk of acquiring the disease, is essential for reducing HIV burden and improving population health [11]. Currently, HIV testing is done in clinics by medical staff across sub-Saharan Africa. For someone to get tested they have to invest considerable time, including time related to travel, waiting for testing, and waiting for their result [12]. Testing is also associated with stigma, which induces a psychological cost [13, 14]. HIV self-testing is a new technology that aims to reduce stigma, as it can be conducted alone, confidentially. Self-tests have lower opportunity cost—there is no need to travel to the clinic or wait for testing, and whereas traditional blood testing takes hours for a result, HIV self-testing only takes 30 minutes. Interventions that have similar approaches to removing barriers—such as mobile testing clinics that reduce travel time [12] and rapid result tests that reduce wait time—are effective at increasing HIV testing [15].

Often intention does not equate to action in health behavior. For example, many people desire and intend to quit smoking or lose weight, but for various reasons fail to realize those goals. Behavioral economics has helped identify some of the psychological biases we are prone to in our everyday lives that make our intended goals so difficult to reach. For Aim 2, I focused on the biases that prevent people from reaching their weight loss goals. These biases include present-biased preferences, self-control, and salience [5, 16, 17]. Today, technology is increasingly focused on helping people overcome their psychological biases to reach their goals. For example, weight loss programs can send tailored reminders for exercise and dieting, making goals and the steps required to reach them more salient [18, 19]. Some include timed messages that attempt to boost self-control, such as sending a motivational message when a person regularly exercises [20].

In an attempt to maximize the effectiveness of behavioral economic interventions for overcoming psychological biases, researchers are now pairing interventions with more traditional economic interventions for behavior change, such as incentives [21-23]. Incentive structures can take many forms. For example, incentives can be placed on outputs, such as weight loss [24, 25] or on an input, e.g. a behavior, such as modifying diet and exercise [26, 27]. In addition, other incentives leverage behavioral responses to maximize their effectiveness, such as loss aversion and overweighing of small probabilities, e.g. lotteries [28, 29]. Studies on incentives related to weight loss have found mixed results [21], however their effectiveness in tandem with a new technology—such as a web-based weight loss intervention—remains unknown.

While Aims 1 and 2 focused on specific technologies for health and their impact on health behavior, Aim 3 focused on mobile phones, which were not created for health, but nonetheless are used for health specific purposes. The worldwide explosion in mobile phone usage has driven health professionals to embrace it as a platform for interventions, called “mobile health” or mHealth interventions. Researchers and practitioners in developing countries are already using mobile phones to send treatment reminders for medication adherence [30]. New innovative companies are using mobile phones as a platform to distribute health messages [31]. The programs vary in sophistication—from direct messages informing people about national health days to complex interactive messages built to educate people on health issues over time.

Mobile phones may also impact health in indirect ways. More generally, mobile phones change the way people communicate and obtain information. Two studies showed that introducing cellular network coverage into areas that previously did not have network coverage changed how businesses obtained information on prices, which led to welfare improvements for both businesses and consumers [32, 33]. Household consumers are likely to have similar gains for health. For example, women would be able to call health workers who live far away to get information about FP methods and their availability, or community health workers would be able to send informative text messages to women regarding when family health days are held, during which women could obtain FP. Mobile phones may also affect educational and income opportunities which could subsequently impact health. My reduced form modeling approach allowed me to estimate the total effect of mobile phones on women's health behaviors. In other words, I was able to estimate the impact mobile phones have due to specific mHealth programs and due to other means, such as through changes in poverty level or education.

This study is significant because I determined the extent to which each technology—HIV self-tests, websites for weight loss, and mobile phones—improves health behavior. Beyond its significance for existing technologies, this contribution will be increasingly important in the future as new technologies emerge that provide new opportunities for people to change their behavior. Understanding the impact of these technologies on health behavior can help priority setting, targeting, and investment moving forward.

1.3 Innovation

These studies are among the first to rigorously evaluate how different technological innovations have impacted health behaviors. Regarding each aim separately: The HIV self-testing study is the first to compare distribution of HIV self-testing kits for partner testing to standard clinic referrals for partner testing. Secondary distribution of HIV self-tests is highly innovative since it utilizes women visiting the clinic to distribute tests to their partners, something that has not been examined before. Second, although a large literature surrounds technology and weight loss, fewer studies have looked at how utilization of these technologies influences behavior. To my knowledge, no studies have examined the impact of financial incentives on utilization of web-based technology for weight loss. Finally, no identified studies have looked at the total effects of mobile phone ownership on knowledge and use of FP. As the world becomes increasingly more dependent on technology, it is important to understand how these innovations can be leveraged to improve the health of populations, especially disadvantaged populations.

The unique datasets and methods I used allowed me to determine causal effects. Randomized controlled trial data for Aim 1 will allow me to determine the causal impact of HIV self-tests provision on men's HIV testing. The Worksite Activities for You (WAY) to Health study (Aim 2) was also a randomized controlled trial, and its random assignment allowed me to use traditional trial analysis methods and other, new innovative techniques to determine causal effects of interest. The UNPS dataset (Aim 3) is a longitudinal panel datasets that allowed me to employ econometric techniques which aid in the identification of causal effects of interest. Te

UNPS has a rich set of covariates that allowed me to control for various time-varying factors that may confound the relationships of interest. In addition, the panel dataset allowed me to use fixed effects models which control for time-invariant differences between people. These robust data are a major methodological advantage over cross-sectional data which typically cannot be used to determine causal effects.

1.4 Theory

Intended health behavior often differs from realized health behavior. Depending on the person, the reason for this disconnect can vary. Using HIV testing as an example, a person may desire to test for HIV, but various factors prevent them from doing so. The reason could be cost of testing, opportunity cost of time, or stigma associated with testing.

In the standard economics framework of rational choice, all decisions are reduced to utility tradeoffs, and we assume that a person is rational and that their failure to test is because the utility associated with testing does not outweigh the utility associated with not testing. The rational framework can be adapted to incorporate other personal preferences, such as time preferences, e.g. preferring consumption today vs. in the future.

However, research in behavioral economics has shown that people are not rational and that they do not act upon simple utility calculations for each behavior. Individual psychological biases cause people to err. These biases include present-biased preferences, that is, placing too much emphasis on present costs as opposed to future benefits. Present-bias causes people to overweight today at the expense of tomorrow and helps explain why people overeat or smoke today at the expense of a potentially shorter life in the future [17]. People also have psychological biases with respect to the information they choose to act on. In a rational choice framework, holding truth constant, all information, regardless of the source, is treated equally.

However, research suggests that this is not the case; instead, the messenger, the person or device or method that delivers information to a person regarding a specific health behavior, can have a large effect [5]. In addition, goal setting and reminders can have a powerful impact on behavior [34]. Messages about the importance of weight loss can increase the salience of goals for weight loss.

Many health behaviors require commitment over time rather than a single decision, such as weight loss. For example, a person may desire to lose weight, but never do so because they systematically lack commitment over time to achieve their goal. Psychological biases can also help explain this disconnect. We are psychologically inclined to go with the status quo, especially in regards to behavior [35, 36]. People become accustomed to certain behaviors, such as eating large portions and being physically inactive to the point that it becomes difficult to deviate from this default behavior.

New technologies can improve the probability of healthy behavior by decreasing traditional costs and by utilizing insights from behavioral economics to reduce or remove the barriers associated with traditional technologies. In the case of HIV testing, self-tests potentially increase HIV testing rates by reducing traditional economic costs associated with testing, such as time. Self-tests can potentially lead to higher usage because they can be distributed by peers, such as friends and spouses, thus leveraging messenger biases, which may lead to higher uptake than through professional health staff. For weight loss, cash compensation and providing users with a website to track progress may help people overcome status quo bias and self-control issues, potentially leading to eating less and/or exercising more [21, 24]. Finally, owning a mobile phone may increase access to salient messages about community “health days” that could benefit the family. These messages may encourage an individual to create a plan for visiting a health day event. Table 1 provides a list of barriers to health behavior that may be affected by the technologies studied in this dissertation.

1.5 Guide to the Dissertation

The remainder of this dissertation proceeds as follows. Chapter 2 provides the results from the study on HIV self-test provision to women attending antenatal or postnatal clinics for male partner HIV testing. Chapter 3 explores the relationship between financial incentives and utilization of a website for weight loss, as well as the ability of the website to promote weight loss. Chapter 4 examines the role of mobile phone ownership and FP in Uganda. Chapter 5 concludes.

TABLE 1: Barriers to healthy behavior and the technologies examined in this research that address them

Barrier	Technologies that address the barrier
Access	HIV self-tests; Mobile phones
Commitment	HIV self-tests; Website for weight loss
Convenience	HIV self-tests; Mobile phones
Goal setting	Website for weight loss
Knowledge	HIV self-tests; Website for weight loss; Mobile phones
Messenger	HIV self-tests; Mobile phones
Motivation	Website for weight loss; Mobile phones
Present-bias	HIV self-tests; Website for weight loss
Status quo bias	HIV self-tests; Website for weight loss
Stigma	HIV self-tests; Website for weight loss

REFERENCES

1. Chib A, Velthoven MHv, Car J. mHealth Adoption in Low-Resource Environments: A Review of the Use of Mobile Healthcare in Developing Countries. Journal of Health Communication. 2015;20(1):4-34. doi: 10.1080/10810730.2013.864735.

2. Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low- and middle-income countries – what has been shown to work? Glob Health Action. 2014;7. doi: 10.3402/gha.v7.25606.

3. Napierala Mavedzenge S, Baggaley R, Corbett EL. A Review of Self-Testing for HIV: Research and Policy Priorities in a New Era of HIV Prevention. Clinical Infectious Diseases. 2013;57(1):126-38. doi: 10.1093/cid/cit156.

4. Spielberg F, Levine RO, Weaver M. Self-testing for HIV: a new option for HIV prevention? The Lancet Infectious Diseases. 4(10):640-6. doi: 10.1016/S1473-3099(04)01150-8.

5. Dolan P, Hallsworth M, Halpern D, King D, Metcalfe R, Vlaev I. Influencing behaviour: The mindspace way. Journal of Economic Psychology. 2012;33(1):264-77. doi: 10.1016/j.joep.2011.10.009.

6. Loewenstein G, Brennan T, Volpp KG. Asymmetric paternalism to improve health behaviors. JAMA. 2007;298(20):2415-7.

7. Loewenstein G, Asch DA, Volpp KG. Behavioral Economics Holds Potential To Deliver Better Results For Patients, Insurers, And Employers. Health Affairs. 2013;32(7):1244-50. doi: 10.1377/hlthaff.2012.1163.

8. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Translational behavioral medicine. 2011;1(1):53-71. doi: 10.1007/s13142-011-0021-7. PubMed PMID: PMC3142960.

9. Cutler D, Miller G. The role of public health improvements in health advances: the twentieth-century United States. Demography. 2005;42(1):1-22.

10. Cutler D, Deaton A, Lleras-Muney A. The Determinants of Mortality. Journal of Economic Perspectives. 2006;20(3):97-120. doi: doi: 10.1257/jep.20.3.97.

11. Mugglin C, Estill J, Wandeler G, Bender N, Egger M, Gsponer T, et al. Loss to programme between HIV diagnosis and initiation of antiretroviral therapy in sub-Saharan Africa: systematic review and meta-analysis. Tropical Medicine & International Health. 2012;17(12):1509-20. doi: 10.1111/j.1365-3156.2012.03089.x.

12. Morin SF, Khumalo-Sakutukwa G, Charlebois ED, Routh J, Fritz K, Lane T, et al. Removing Barriers to Knowing HIV Status: Same-Day Mobile HIV Testing in Zimbabwe. JAIDS

Journal of Acquired Immune Deficiency Syndromes. 2006;41(2):218-24. doi: 10.1097/01.qai.0000179455.01068.ab. PubMed PMID: 00126334-200602010-00014.

13. Turan JM, Bukusi EA, Onono M, Holzemer WL, Miller S, Cohen CR. HIV/AIDS Stigma and Refusal of HIV Testing Among Pregnant Women in Rural Kenya: Results from the MAMAS Study. AIDS and Behavior. 2010;15(6):1111-20. doi: 10.1007/s10461-010-9798-5.

14. Kalichman SC, Simbayi LC. HIV testing attitudes, AIDS stigma, and voluntary HIV counselling and testing in a black township in Cape Town, South Africa. Sexually Transmitted Infections. 2003;79(6):442-7. doi: 10.1136/sti.79.6.442.

15. Kassler WJ, Alwano-Edyegu MG, Marum E, Biryahwaho B, Kataaha P, Dillon B. Rapid HIV testing with same-day results: a field trial in Uganda. International Journal of STD & AIDS. 1998;9(3):134-8. doi: 10.1258/0956462981921882.

16. Hoch SJ, Loewenstein GF. Time-Inconsistent Preferences and Consumer Self-Control. Journal of Consumer Research. 1991;17(4):492-507.

17. O'Donoghue T, Rabin M. Doing It Now or Later. American Economic Review. 1999;89(1):103-24.

18. Liu F, Kong X, Cao J, Chen S, Li C, Huang J, et al. Mobile Phone Intervention and Weight Loss Among Overweight and Obese Adults: A Meta-Analysis of Randomized Controlled Trials. American journal of epidemiology. 2015;181(5):337-48. doi: 10.1093/aje/kwu260.

19. Servick K. Mind the phone. Science. 2015;350(6266):1306-9. doi: 10.1126/science.350.6266.1306.

20. Patrick K, Griswold WG, Raab F, Intille SS. Health and the Mobile Phone. American Journal of Preventive Medicine. 2008;35(2):177-81. doi: http://dx.doi.org/10.1016/j.amepre.2008.05.001.

21. Paloyo AR, Reichert AR, Reinermann H, Tauchmann H. The Causal Link between Financial Incentives and Weight Loss: An Evidence-Based Survey of the Literature. Journal of Economic Surveys. 2014;28(3):401-20. doi: 10.1111/joes.12010.

22. Cawley J, Price JA. A case study of a workplace wellness program that offers financial incentives for weight loss. J Health Econ. 2013;32(5):794-803. doi: 10.1016/j.jhealeco.2013.04.005. PubMed PMID: 23787373.

23. Marteau TM, Ashcroft RE, Oliver A. Using financial incentives to achieve healthy behaviour. BMJ. 2009;338. doi: 10.1136/bmj.b1415.

24. Finkelstein EA, Linnan LA, Tate DF, Birken BE. A pilot study testing the effect of different levels of financial incentives on weight loss among overweight employees. Journal of

occupational and environmental medicine / American College of Occupational and Environmental Medicine. 2007;49(9):981-9. Epub 2007/09/13. doi: 10.1097/JOM.0b013e31813c6dcb. PubMed PMID: 17848854.

25. Paul-Ebhohimhen V, Avenell A. Systematic review of the use of financial incentives in treatments for obesity and overweight. Obesity Reviews. 2008;9(4):355-67. doi: 10.1111/j.1467-789X.2007.00409.x.

26. Mitchell MS, Goodman JM, Alter DA, John LK, Oh PI, Pakosh MT, et al. Financial Incentives for Exercise Adherence in Adults: Systematic Review and Meta-Analysis. American Journal of Preventive Medicine. 2013;45(5):658-67. doi: http://dx.doi.org/10.1016/j.amepre.2013.06.017.

27. Wall J, Mhurchu CN, Blakely T, Rodgers A, Wilton J. Effectiveness of Monetary Incentives in Modifying Dietary Behavior: a Review of Randomized, Controlled Trials. Nutrition Reviews. 2006;64(12):518-31. doi: 10.1111/j.1753-4887.2006.tb00185.x.

28. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica. 1979;47(2):263-91.

29. Volpp KG, John LK, Troxel AB, Norton L, Fassbender J, Loewenstein G. Financial incentive–based approaches for weight loss: A randomized trial. JAMA. 2008;300(22):2631-7. doi: 10.1001/jama.2008.804.

30. Mbuagbaw L, Kop MLvd, Lester RT, Thirumurthy H, Pop-Eleches C, Ye C, et al. Mobile phone text messages for improving adherence to antiretroviral therapy (ART): an individual patient data meta-analysis of randomised trials. BMJ Open. 2013;3(12):e003950. doi: 10.1136/bmjopen-2013-003950.

31. Chib A, Wilkin H, Ling LX, Hoefman B, Biejma HV. You Have an Important Message! Evaluating the Effectiveness of a Text Message HIV/AIDS Campaign in Northwest Uganda. Journal of Health Communication. 2012;17(sup1):146-57. doi: 10.1080/10810730.2011.649104.

32. Jensen R. The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. The Quarterly Journal of Economics. 2007;122(3):879-924. doi: 10.1162/qjec.122.3.879.

33. Aker JC. Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger. American Economic Journal: Applied Economics. 2010;2(3):46-59. doi: 10.1257/app.2.3.46.

34. Bagozzi RP, Dholakia U. Goal setting and goal striving in consumer behavior. The Journal of Marketing. 1999:19-32.

35. Madrian BC, Shea DF. The power of suggestion: inertia in 401(k) participation and savings behavior. Quarterly Journal of Economics. 2001;116(4):1149-87.

36. Samuelson W, Zeckhauser R. Status quo bias in decision making. Journal of Risk and Uncertainty. 1(1):7-59. doi: 10.1007/bf00055564.

CHAPTER 2: PROMOTING PARTNER TESTING AND COUPLES TESTING THROUGH SECONDARY DISTRIBUTION OF HIV SELF-TESTS: A RANDOMIZED CLINICAL TRIAL

2.1 Overview

Background: Achieving higher rates of partner HIV testing and couples testing among pregnant and postpartum women in sub-Saharan Africa is essential for the success of combination HIV prevention, including the prevention of mother-to-child transmission. We aimed to determine whether providing multiple HIV[1] self-tests to pregnant and postpartum women for secondary distribution is more effective at promoting partner testing and couples testing than conventional strategies based on invitations to clinic-based testing.

Methods and Findings: We conducted a randomized trial in Kisumu, Kenya, between June 11, 2015, and January 15, 2016. Six hundred antenatal and postpartum women aged 18–39 y were randomized to an HIV self-testing (HIVST) group or a comparison group. Participants in the HIVST group were given two oral-fluid-based HIV test kits, instructed on how to use them, and encouraged to distribute a test kit to their male partner or use both kits for testing as a couple. Participants in the comparison group were given an invitation card for clinic-based HIV

---

[1] This work was previously published in *PLOS Medicine*. Full citation:

Masters, S. H., Agot, K., Obonyo, B., Napierala Mavedzenge, S., Maman, S., & Thirumurthy, H. (2016). Promoting Partner Testing and Couples Testing through Secondary Distribution of HIV Self-Tests: A Randomized Clinical Trial. *PLoS Med*, 13(11), e1002166. doi:10.1371/journal.pmed.1002166

testing and encouraged to distribute the card to their male partner, a routine practice in many health clinics. The primary outcome was partner testing within 3 mo of enrollment. Among 570 participants analyzed, partner HIV testing was more likely in the HIVST group (90.8%, 258/284) than the comparison group (51.7%, 148/286; difference = 39.1%, 95% CI 32.4% to 45.8%, $p < 0.001$). Couples testing was also more likely in the HIVST group than the comparison group (75.4% versus 33.2%, difference = 42.1%, 95% CI 34.7% to 49.6%, $p < 0.001$). No participants reported intimate partner violence due to HIV testing. This study was limited by self-reported outcomes, a common limitation in many studies involving HIVST due to the private manner in which self-tests are meant to be used.

Conclusions: Provision of multiple HIV self-tests to women seeking antenatal and postpartum care was successful in promoting partner testing and couples testing. This approach warrants further consideration as countries develop HIVST policies and seek new ways to increase awareness of HIV status among men and to promote couples testing.

2.2 Introduction

Low uptake of HIV testing services in sub-Saharan Africa (SSA) is among the key barriers to meeting the 90-90-90 targets established by UNAIDS and to improving the effectiveness of HIV treatment as prevention. HIV testing among men in particular remains low in many countries, as does knowledge of HIV status among HIV-infected persons [1]. Door-to-door testing and mobile testing strategies have moved testing services out of health facilities and into communities, overcoming barriers related to clinic-based testing and, subsequently, increasing testing coverage. However, despite these advancements, there remains a need for novel interventions that can promote testing among men and other hard-to-reach populations [2,3].

In addition to increasing HIV testing uptake among men, achieving higher rates of couples testing can also contribute to HIV prevention efforts. Low uptake of couples testing is particularly concerning in light of data indicating that four out of every ten new HIV infections occur within stable heterosexual partnerships and that the majority of persons in sero-discordant relationships are unaware of their HIV status [4]. The benefits of couples testing may include safer sexual behavior in couples [5], higher uptake of interventions such as antiretroviral therapy (ART) for HIV-positive partners [6], and pre-exposure prophylaxis (PrEP) among HIV-negative partners in sero-discordant relationships, as well as increased uptake of and adherence to prevention of mother-to-child transmission (PMTCT) interventions [7–9]. Given the need to achieve better PMTCT outcomes and prevent new infections in couples, a number of countries have sought to promote partner testing and couples testing among pregnant and postpartum women [10]. However, efforts to encourage pregnant and postpartum women to refer their male partners for HIV testing have had limited success [11,12]. The barriers to testing among male partners have included stigma, fear of prognosis, lack of awareness of HIV risk, inconvenience, fear of disclosure, transportation costs, opportunity costs such as time off from work, and behavioral factors such as a tendency to delay behaviors with immediate costs and delayed benefits [13,14].

HIV self-testing (HIVST) is a promising approach that addresses many barriers associated with clinic-based HIV testing and has had high acceptability in SSA [15–17]. Self-tests can enable individuals to test themselves for HIV privately and at their own convenience.

Simple oral-fluid-based tests have achieved high sensitivity and specificity, with some studies also having shown that the tests can be used accurately by individuals [18]. A number of countries in SSA have developed policies for implementation and support of HIVST [19,20], with Kenya being the first country in SSA to include HIVST in its national testing guidelines [21]. Recent research in Kenya has also demonstrated the acceptability and feasibility of a novel "secondary distribution" strategy that seeks to promote HIV testing among men and in couples through provision of multiple self-tests to women seeking health services [22].

We conducted a randomized trial in Kenya among women receiving antenatal care (ANC) or postpartum care (PPC) services to test whether the provision of multiple self-tests to women for distribution to their partners can increase uptake of male partner testing and couples testing.

2.3 Methods

2.3.1 Ethics Statement

The study received approval from the Scientific and Ethics Review Unit at the Kenya Medical Research Institute and the Office of Human Research Ethics at the University of North Carolina at Chapel Hill.

2.3.2 Study Setting

The study was conducted in urban and peri-urban areas within Kisumu County, Kenya. Adult HIV prevalence in Kisumu is 19.3% [23], the third highest among the counties in the country. Women visiting ANC and PPC clinics were recruited from three health facilities in Kisumu.

2.3.3 Study Design and Participants

Trained research assistants screened and enrolled women seeking ANC or PPC at the three facilities, in a private location away from regular clinic activities. Women were given the opportunity to enroll in the study if they met the following eligibility criteria: were 18–39 y of age, reported that their primary partner was not known to be HIV-positive or had not tested in the past 6 mo, resided in or around Kisumu County, and had no intention of leaving the area within 3 mo. In addition, at the ANC clinic eligibility was limited to women with gestation age ≤ 20 wk, and at the PPC clinic eligibility was limited to women who had given birth in the past 6 wk to 12 mo. Following the provision of written informed consent, participants were administered a baseline questionnaire that measured demographic characteristics, sexual behavior, HIV testing history, and partner characteristics. All study staff received ethical training on research with human participants.

2.3.4 Randomization Procedures

Participants were randomized in a 1:1 ratio using balanced block randomization (block size 20) to an HIVST group or a comparison group. Sealed randomization envelopes were offered to participants sequentially, and these revealed the study group assignment to the participant and study staff simultaneously.

2.3.5 Intervention

Participants in the HIVST group received two oral-fluid-based rapid HIV tests (OraQuick Rapid HIV-1/2 Antibody Test, OraSure Technologies). Each test was accompanied with an instruction sheet that described step-by-step self-testing procedures in multiple languages. Study staff also provided the participants with a brief demonstration of how to use the test. Participants were encouraged to distribute a test kit to their male partner or to use both test kits to undertake couples testing if they felt comfortable doing so; they were also counseled on how to talk to their partners about HIV testing, the possibility of adverse reactions associated with suggesting HIV testing to their partner, learning their partner’s HIV status, and disclosing their own HIV status. Following Kenya’s 2015 HIV testing services guidelines [19], participants were informed about the need to seek clinic-based confirmatory testing if a positive (reactive) self-test result was obtained, and an invitation card for confirmatory testing at a clinic in the study area was included with each test.

Participants in the comparison group were counseled on the importance of partner testing and provided with an invitation card to give to their partner for HIV testing at a study clinic. The use of invitation cards to promote male partner testing is currently standard practice in many facilities. The cards mentioned the importance of testing, listed the health facility where the participant was enrolled, and encouraged the male partner to get tested at the study facility.

Both groups received information on where to seek advice and assistance for clinical, counseling, and legal support in case of intimate partner violence (IPV). They also were given a study phone number to call in case they had questions or needed advice about clinic-based testing or self-testing, or IPV or other adverse events.

2.3.6 Follow-Up Assessments

Follow-up data collection occurred over a 3-mo period. Participants were contacted each month to determine if they had distributed a self-test kit to their sexual partner (HIVST group) or if their partner had sought HIV testing at a clinic (comparison group). Research assistants scheduled and conducted an in-person follow-up interview with participants who reported having distributed a test to their partner or who reported that their partner sought clinic-based testing, while participants who had not done so or were not reached at 1 and 2 mo were interviewed at 3 mo. If participants were unable to meet with research assistants, a follow-up phone interview was conducted. Participants in both groups were asked whether their partner had been tested for HIV since study enrollment.

2.3.7 Statistical Analyses

The unit of analysis was the study participant. All outcomes were self-reported by study participants. The primary, prespecified outcome was whether the primary partner of the participant had an HIV test within 3 mo of enrollment, which was determined from the follow-up survey question: “Has your partner had an HIV test since you were enrolled in the study?” The primary analysis compared this outcome in the HIVST and comparison groups using an unadjusted modified Poisson regression with robust standard errors [24]. Our original analysis plan proposed estimation of a logistic regression model, but ultimately we selected a modified Poisson model because risk ratios can be easier to interpret than odds ratios. We chose to present both the absolute risk differences between the two study groups and the risk ratios from modified Poisson regressions. Participants who were not successfully followed up were not included in the analyses as it was not possible to determine the primary and secondary outcomes for them.

In secondary analyses we examined the impact of the intervention on the following six outcomes reported by participants: (1) discussion of HIV testing with partner, (2) couples testing, (3) couples testing among participants whose partner tested for HIV, (4) awareness of partner’s HIV test result, (5) awareness of partner’s HIV test result among participants whose partner tested for HIV, and (6) partner’s HIV test result. Discussion of HIV testing was defined as having occurred if the participant reported that she and her partner had talked about HIV testing since enrollment in the study. Couples testing was defined as having occurred when a participant reported that she had tested together with her partner at the same time. Awareness of partner's HIV test result was defined as the participant having learned her partner's HIV status. Additionally, we examined whether partners of participants in the HIVST group who tested positive sought confirmatory testing and whether partners in both groups who received a positive result were reported to be in care at the time of follow-up. We also assessed IPV at baseline and follow-up using questions adapted from the Kenya Demographic and Health Survey [25] that asked whether participants experienced physical, emotional, verbal, or sexual violence from their partner. Participants were coded as having experienced IPV if they responded affirmatively to any of the IPV questions. Survey questions used to measure study outcomes are reported in S1 Table.

In order to determine whether there were differences in intervention effectiveness in certain populations, we estimated modified Poisson regression models among participants who were enrolled at each of the three health facilities, among those whose primary partner had tested for HIV in the 12 mo prior to enrollment or not, and among those who had experienced IPV in the 12 mo prior to enrollment or not. All statistical tests were two-sided, and significance level was set at $p < 0.05$. No adjustment was made for multiple testing since the secondary analyses were considered exploratory. Statistical analyses were performed using Stata 14.1.

The planned sample size for the study was 600, with 300 participants in each study group. Power calculations assuming a two-sided unadjusted independent proportions test indicated that with a sample size of 300 per study group and 20% uptake of partner testing in the comparison group, there would be 80% power to detect a difference in partner testing as small as 10%.

2.4 Results

2.4.1 Participant Recruitment and Flow

Between June 11, 2015, and October 16, 2015, a total of 1,929 women were screened for participation. Among those, 614 (32%) were determined to be ineligible, 715 declined to participate (37%), and 600 (31%) were enrolled and randomized (Fig. 1). Reasons for ineligibility included no primary partner (28%), partner HIV-positive (22%), intention of leaving study area during follow-up period (15%), age of participant (8%), age of child (8%), and fear of IPV due to discussing HIV testing with partner (5%). Common reasons for refusal included women reporting they were “in a hurry” or “too busy” (384/715, 53.7%), needing permission from partner to enroll in a study (54/715, 7.6%), and reporting their partner had tested recently and therefore did not have interest in participating in the study (111/715, 15.5%). Follow-up interviews were conducted until January 15, 2016. One person from the comparison group withdrew from the study during the follow-up period. Of the 600 participants who were enrolled, follow-up was completed for 570 (95%), 286 (94.4%) in the comparison group and 284 (95.6%) in the HIVST group.

2.4.2 Participant Characteristics

Participants in the two study groups had largely similar characteristics at baseline (Table 1). Their mean age was 24 y, and the vast majority were married. Median monthly earnings was US$0 since the majority did not report any engagement in income-earning activities during or after pregnancy. Participants’ self-reported sexual behavior and their reports of their partner’s

HIV testing history were similar in both groups (Table 2). Nearly 4% of all participants self-reported being HIV-positive. The majority of participants reported that their partner had tested for HIV in the past 12 mo (56%), and only a small percentage of participants (14%) had heard of HIVST prior to the study. Nearly 30% of participants reported experiencing IPV in the past 12 mo.

2.4.3 Male Partner Testing

Male partner testing within 3 mo of enrollment was higher in the HIVST group (258/284, 90.8%) than the comparison group (148/286, 51.7%), as shown in Table 3. The difference of 39.1% between the two groups was statistically significant (95% CI 32.4% to 45.8%, $p < 0.001$). Among participants in the HIVST group whose partners used a self-test, 76% and 17% reported that their partner found it “very easy” or “somewhat easy,” respectively, to use the self-test, while 6% reported that their partner found it “somewhat difficult” or “very difficult.” In the comparison group, 45% (67/148) of partners who tested were reported to have done so outside of the three study facilities.

2.4.4 Secondary Outcomes

Over 95% of participants in both groups reported discussing HIV testing with their partner since enrollment, and there was no significant difference between the two groups (difference = −1.1%, 95% CI −4.3% to 2.2%, $p = 0.512$). Participants in the HIVST group were more likely to test as a couple than participants in the comparison group (difference = 42.1%, 95% CI 34.7% to 49.6%, $p < 0.001$). In addition, among participants whose partner tested for HIV during the follow-up period, couples testing was more likely in the HIVST group than the comparison group (difference = 18.8%, 95% CI 9.8% to 27.8%, $p < 0.001$).

At follow-up, participants in the HIVST group were more likely to know their partner's HIV status than those in the comparison group (difference = 39.1%, 95% CI 32.4% to 45.8%, $p < 0.001$). However, among participants whose partner tested for HIV during the follow-up period, participants' awareness of their partner's HIV status did not differ significantly between the two groups (difference = 0.9%, 95% CI −1.8% to 3.5%, $p < 0.519$), suggesting that the increase in awareness of partner HIV status in the HIVST group was driven by the greater likelihood of partner testing having occurred rather than a greater likelihood of becoming aware if a partner did get tested. Among participants whose partner tested for HIV, almost all were aware of their partner's HIV test result (98.0% in comparison group, 98.8% in HIVST group). A small number of participants in both groups reported that their partner tested HIV-positive (1.4% in comparison group, 2.8% in HIVST group). Among the eight partners who tested positive in the HIVST group, two went for confirmatory testing, were confirmed positive, and were linked to care. Among the four partners who tested positive in the comparison group, three were reported to have sought HIV care at the time of the 3-mo interview. No participants in either group reported IPV due to HIV testing.

2.4.5 Heterogeneity of Intervention Effectiveness

Participants in the HIVST group reported higher partner testing rates than participants in the comparison group in all subgroups examined (Table 4). While partner testing was significantly more likely in the HIVST group than the comparison group in all three study sites, the HIVST intervention was more effective in promoting partner testing in the hospital setting as compared to the urban health clinic setting ($p < 0.001$). There was no difference in intervention effectiveness by partner testing status in the past 12 mo ($p = 0.172$). Similarly, we found no difference in intervention effectiveness between participants who had experienced IPV at baseline and those who had not (p = 0.111).

2.5 Discussion

Provision of multiple self-tests to women led to secondary distribution of the self-tests to their male partners and ultimately achieved higher HIV testing among their male partners and higher couples testing than a more conventional approach of giving women invitation cards for their male partners to test at health facilities. In the group that received multiple self-tests, partner testing was reported by 91% of participants who were followed up, and 75% of participants followed up tested together with their partner. To our knowledge, this is the first randomized trial to test whether provision of multiple self-tests to women promotes partner and couples testing. In subgroup analyses, the intervention was more effective than the partner invitation approach even among women who reported a history of IPV at baseline and among women whose partners had not gone for HIV testing in the past 12 mo.

Male partner testing was nearly universal among women who received multiple self-tests. This striking result is consistent with findings from a pilot study we previously conducted in the study region, in which male partner testing was reported to have occurred for 91% of women seeking ANC and 86% of women receiving PPC [22]. The study results are also consistent with the high acceptability of HIVST that has been documented throughout SSA and elsewhere [15–17].

Uptake of partner testing and couples testing in the comparison group that received invitation cards for their male partner was largely similar to what has been reported in two other recent studies. One study conducted in the same region of Kenya reported that couples testing occurred among 36% of pregnant women who received clinic invitation cards for their partner [26]. Further, a study conducted among HIV-positive pregnant women in Malawi reported that couples testing occurred among 52% of women who received invitation cards for their partner [27]. The similarity in male partner and couples testing levels in the comparison group of our study with those reported in these other studies of the partner invitation approach provide further support for the validity of the self-reported measures obtained in our study. In addition, it is notable that the couples testing rate in the HIVST group of our study was similar to or exceeded the rates achieved by the interventions tested in the two other studies: home visits and invitations followed by home tracing. While formal cost-effectiveness analyses are necessary, it is plausible that interventions relying on secondary distribution of self-tests would ultimately require fewer resources in total and therefore would have greater sustainability.

While prior HIV testing in this urban and peri-urban study setting was fairly high, we found no difference in the effectiveness of the HIVST intervention based on whether partners had tested for HIV in the past 12 mo. This result is encouraging since it suggests that the strategy of giving multiple self-tests to women can effectively increase access to HIV testing in hard-to-reach populations such as men who do not test regularly, and perhaps more generally in settings where testing rates are not as high as they were in our study setting. In addition, the large differences in partner testing between the HIVST and comparison groups was observed in all population subgroups, which suggests broader applicability of this intervention among various subgroups of pregnant and postpartum women.

From a policy standpoint, providing self-tests to women in clinic settings has substantial appeal not only because it promotes male partner testing but also because it helps women learn their partner's HIV status. The intervention's feasibility is enhanced by the fact that pregnant and postpartum women represent an easier-to-reach segment of the population by virtue of their higher utilization of health services. Couples testing, which is recommended by the World Health Organization and the Kenyan Ministry of Health, is another important benefit of the intervention. Individuals who test as a couple and mutually disclose their HIV status are more likely than those testing alone to adopt a range of HIV prevention and care behaviors [5]. Despite these benefits, only 37.2% of people who have tested for HIV in Kenya reported ever testing together with a sexual partner [28]. Notably, the uptake of couples testing observed among women given multiple self-tests in this study (75%) was higher than the uptake reported in the recent pilot study we conducted in the study area, in which women receiving ANC and PPC tested as couples 47% and 58% of the time, respectively [22].

This study has several limitations that warrant discussion. First, we relied on self-reported data for the main outcomes. This is a common limitation in many studies involving HIVST due to the private manner in which self-tests are meant to be used. Despite the potential for self-reporting to be associated with reporting bias, we believe reporting bias was minimal given the above-mentioned consistency of our results for partner testing in both study groups with other studies conducted in SSA [15,18,22,26,27] and given the lack of material incentives tied to participants' responses. In addition, any bias in reporting of testing uptake is unlikely to be differential by study group. Male partners in the comparison group were able to test at multiple facilities in the study area, and it was as difficult in practice to verify their clinic-based testing as it was to verify self-test usage by partners in the HIVST group. These factors are likely to strengthen the validity of comparing self-reported partner testing in the two study groups. Since objective verification of self-test use will remain a challenge, there is a need for larger-scale studies that examine downstream outcomes such as the proportion of partners linking to HIV prevention and treatment. Second, our study did not include women who knew their partner was HIV-positive because we believed that a partner testing intervention would have little additional benefit to them. This feature of the study design, coupled with high rates of HIV testing in the urban and peri-urban study setting [29], likely led to relatively few HIV-positive partners being identified in this study. This limited our ability to make statistical inferences with respect to confirmatory testing and linkage to care. More research is needed to rigorously assess levels of confirmatory testing and linkage to care following HIVST, as well as to understand the decision-making process of whether or not to seek these services.

Finally, the third limitation stems from the fact that roughly one-third of women seeking ANC or PPC declined to participate in the study, and some were ineligible because they reported a fear that violence would result from offering a self-test to their partner. Among women declining participation, the most commonly reported reason was a lack of adequate time to enroll in the study, but other reasons such as a lack of interest in partner testing likely played a role. While these two reasons for declining to participate in the study do not impact the internal validity of the study results, they do limit the generalizability of the findings to all pregnant and postpartum women. Refusal also reinforces the feasibility and safety of offering multiple self-tests because women demonstrated considerable agency and ability to decide themselves whether to accept self-tests and offer them to their partner. Prior work has documented the high acceptability of this intervention among women receiving multiple self-tests [22], and ongoing qualitative research with women receiving multiple self-tests shows that women have a strong sense of agency when deciding whether to offer self-tests to others and appreciate the opportunity to learn their partner's status. Additional qualitative research will provide insights and lessons for wider implementation. Given the novelty of HIVST and this particular strategy for promoting partner testing (i.e., secondary distribution of self-tests by women receiving ANC and PPC), it is also likely that the broader acceptability of secondary distribution strategies will grow as HIVST becomes more common. Additional research is necessary to assess the effectiveness of the intervention in other populations and settings outside western Kenya. However, to the extent that men experience similar barriers to clinic-based HIV testing elsewhere, the results from this study could be applicable to other settings and populations.

One concern about providing multiple self-tests to women for distribution to partners has been the possibility of IPV. Despite women reporting high rates of IPV in the past 12 mo at baseline (27%), it is noteworthy that the intervention was highly effective even among women who reported a history of IPV at baseline, and there were no cases of IPV due to HIV testing reported in either study group during the follow-up period. Few male partners had a reactive self-test result in the study, which may have contributed to the lack of reported IPV due to testing. However, prior research we have conducted with women receiving multiple self-tests—including female sex workers who identified a greater proportion of HIV-positive partners than ANC or PPC women in our study—also suggests IPV is rare [22]. The fact that there were no cases of IPV also suggests that women have the agency and discretion to decide whether to accept self-tests and whether to offer self-tests to their partner.

This study provides key insights on a strategy—secondary distribution of self-tests to sexual partners—that may become common in many populations in SSA and elsewhere as HIV self-tests become more widely available, whether formally endorsed or not. For example, the feasibility of this approach is also being explored among key populations such as men who have sex with men [30,31]. The promising results from this study suggest that secondary distribution of self-tests warrants further consideration as countries develop HIVST policies and seek new ways to promote partner testing. Implementing this intervention at scale is feasible as the primary requirements are that clinic staff be trained on how to explain self-test use and to offer self-tests to women. However, there are potential challenges to programmatic implementation of the intervention, such as ensuring adequate counseling when self-tests are offered to women, making counseling available post-test, and including interventions to achieve high linkage to appropriate services. Ongoing and planned implementation research will assess these issues and further develop strategies for maximizing the potential for HIVST in achieving HIV prevention and care objectives.

FIGURE 1: Assessment of eligibility, randomization, and follow-up

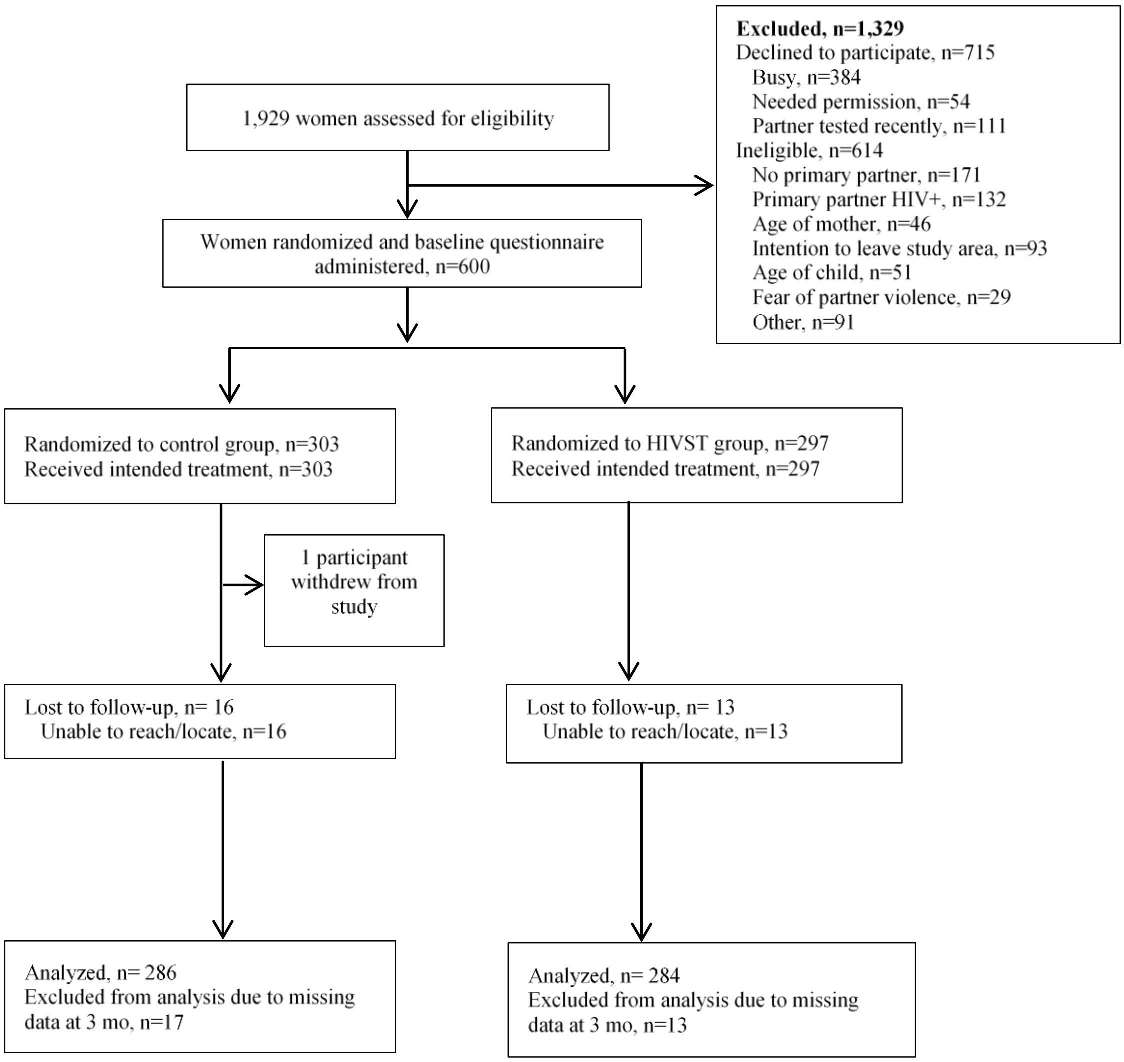


TABLE 2: Baseline characteristics of study participants

Characteristic	Comparison Group (*n* = 286)	HIV Self-Testing Group (*n* = 284)	Total (*n* = 570)
Age (years), mean (SD)	24.2 (4.3)	24.2 (4.5)	24.2 (4.4)
Monthly earnings (US dollars), median (IQR)	0 (0–30)	0 (0–40)	0 (0–36)
Ethnic group			
Luo	221 (77)	219 (77)	440 (77)
Luhya	33 (12)	43 (15)	76 (13)
Other	32 (11)	22 (8)	54 (9)
Education			
Some or completed primary	138 (48)	143 (50)	281 (49)
Some secondary	133 (47)	120 (42)	253 (44)
Completed secondary or greater	15 (5)	21 (7)	36 (6)
Married	266 (93)	266 (94)	532 (93)
Occupation			
Non-manual	74 (26)	83 (29)	157 (28)
Manual	19 (7)	28 (10)	47 (8)
Housewife/unemployed	193 (67)	173 (61)	366 (64)

For all variables frequencies are presented, with percentages in parentheses, except where otherwise noted.

IQR, interquartile range; SD, standard deviation.

TABLE 3: Self-reported sexual behavior and HIV testing history

Behavior or HIV Testing History	**Comparison Group (*n* = 286)**	**HIV Self-Testing Group (*n* = 284)**	**Total, (*n* = 570)**
Age at first intercourse (years), mean (SD)	17.7 (2.8)	17.9 (2.5)	17.8 (2.7)
Condom used during last sex	54 (19)	46 (16)	100 (18)
Had at least one other sexual partner in the past 12 mo	4 (1)	5 (2)	9 (2)
Number of times been tested for HIV in the past 12 mo, mean (SD)	2.8 (1.4)	2.8 (1.5)	2.8 (1.4)
Self-reported HIV-positive	10 (3.5)	13 (4.6)	23 (4.1)
Heard of HIV self-testing prior to study	39 (14)	41 (14)	80 (14)
Primary partner ever tested for HIV			
Yes	220 (77)	216 (76)	436 (76)
No	19 (7)	21 (7)	40 (7)
Don't know	47 (16)	47 (17)	94 (16)
Primary partner tested for HIV in the past 12 mo			
Yes	173 (60)	149 (52)	322 (56)
No	35 (12)	42 (15)	77 (14)
Don't know	78 (27)	93 (33)	171 (30)
Know partner's status	192 (67)	194 (68)	386 (68)
Experienced intimate partner violence in the past 12 mo	76 (27)	78 (27)	154 (27)

For all variables frequencies are presented, with percentages in parentheses, except where otherwise noted.

SD, standard deviation.

TABLE 4: Effects of HIV self-testing intervention within 3 mo.

Outcome	Comparison Group, Number (Percent) (*n* = 286)	HIV Self-Testing Group, Number (Percent) (*n* = 284)	Absolute Difference, Percentage Points (95% CI)*	Risk Ratio (95% CI)**	*p*-Value*
Primary outcome					
Male partner HIV testing	148 (51.7)	258 (90.8)	39.1 (32.4 to 45.8)	1.76 (1.56–1.98)	<0.001
Secondary outcomes					
Discussed HIV testing with partner	276 (96.5)	271 (95.4)	−1.1 (−4.3 to 2.2)	0.99 (0.96–1.02)	0.512
Couples testing for HIV	95 (33.2)	214 (75.4)	42.1 (34.7 to 49.6)	2.27 (1.90–2.71)	<0.001
Couples testing conditional on partner HIV testing***	95 (64.2)	214 (82.9)	18.8 (9.8 to 27.8)	1.29 (1.13–1.48)	<0.001
Aware of partner's HIV test result	145 (50.7)	255 (89.8)	39.1 (32.3 to 45.9)	1.77 (1.57–2.00)	<0.001
Aware of partner's HIV test result conditional on partner HIV testing***	145 (98.0)	255 (98.8)	0.9 (−1.8 to 3.5)	1.01 (0.98–1.04)	0.519
Partner tested HIV-positive	4 (1.4)	8 (2.8)	1.4 (−0.9 to 3.8)	2.01 (0.61–6.62)	0.239

*Estimates and confidence intervals are marginal effects from unadjusted modified Poisson regression.

**Estimates and confidence intervals are risk ratios from unadjusted modified Poisson regression.

***Model includes the subset of participants whose partner tested for HIV.

TABLE 5: Comparison of intervention effectiveness in participant subgroups

Subgroup	HIV Testing Uptake, Number/Total Number (Percent)		Effect of Self-Testing		*P*-Value for Interaction**
	Comparison Group	HIV Self-Testing Group	Absolute Difference, Percentage Points (95% CI)*	*P*-Value for Subgroup*	
Study site					
Urban health clinic	80/120 (66.7)	117/129 (90.7)	24.0 (14.2 to 33.9)	<0.001	—
Hospital	47/122 (38.5)	97/105 (92.4)	53.9 (43.8 to 63.9)	<0.001	<0.001
Peri-urban health clinic	21/44 (47.7)	44/50 (88.0)	40.3 (22.9 to 57.7)	<0.001	0.093
Partner tested for HIV in 12 mo prior to enrollment					
Tested ≥1 time	102/173 (59.0)	142/149 (95.3)	36.3 (28.3 to 44.4)	<0.001	—
Did not test	16/35 (45.7)	37/42 (88.1)	42.4 (23.1 to 61.7)	<0.001	0.389
Do not know if tested	30/73 (38.5)	79/93 (84.9)	46.5 (33.5 to 59.5)	<0.001	0.057
Participants reported intimate partner violence in past 12 mo at baseline					
No	114/210 (54.3)	185/206 (89.8)	35.5 (27.6 to 43.4)	<0.001	—
Yes	34/76 (44.7)	73/78 (93.6)	48.9 (36.4 to 61.3)	<0.001	0.111

*Estimates and confidence intervals are marginal effects from a modified Poisson regression of outcome on study group for the subgroup described.

***P*-Value for interaction coefficient between subgroup and first category (urban health clinic, tested ≥1 time in past 12 mo, and no IPV).

REFERENCES

1. UNAIDS. 90-90-90: an ambitious treatment target to help end the AIDS epidemic. Geneva: UNAIDS; 2014.

2. Suthar AB, Ford N, Bachanas PJ, Wong VJ, Rajan JS, Saltzman AK, et al. Towards universal voluntary HIV testing and counselling: a systematic review and meta-analysis of community-based approaches. PLoS Med. 2013;10(8):e1001496. Epub 2013/08/24. doi: 10.1371/journal.pmed.1001496. PubMed PMID: 23966838; PubMed Central PMCID: PMC3742447.

3. Sharma M, Ying R, Tarr G, Barnabas R. Systematic review and meta-analysis of community and facility-based HIV testing to address linkage to care gaps in sub-Saharan Africa. Nature. 2015;528(7580):S77-85. doi: 10.1038/nature16044. PubMed PMID: 26633769.

4. Gelmon L, Kenya P, Oguya F, Cheluget B, Haile G. Kenya HIV prevention response and modes of transmission analysis. Nairobi: Kenya National AIDS Control Council; 2009.

5. Allen S, Meinzen-Derr J, Kautzman M, Zulu I, Trask S, Fideli U, et al. Sexual behavior of HIV discordant couples after HIV counseling and testing. AIDS. 2003;17(5):733-40. PubMed PMID: 00002030-200303280-00012.

6. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med. 2011;365(6):493-505. doi: 10.1056/NEJMoa1105243.

7. Aluisio A, Richardson BA, Bosire R, John-Stewart G, Mbori-Ngacha D, Farquhar C. Male antenatal attendance and HIV testing are associated with decreased infant hiv infection and increased HIV free survival. J Acquir Immune Defic Syndr. 2011;56(1):76-82. doi: 10.1097/QAI.0b013e3181fdb4c4. PubMed PMID: PMC3005193.

8. Drake AL, Wagner A, Richardson B, John-Stewart G. Incident HIV during pregnancy and postpartum and risk of mother-to-child HIV transmission: a systematic review and meta-analysis. PLoS Med. 2014;11(2):e1001608. doi: 10.1371/journal.pmed.1001608.

9. Medley A, Baggaley R, Bachanas P, Cohen M, Shaffer N, Lo Y-R. Maximizing the impact of HIV prevention efforts: interventions for couples. AIDS Care. 2013;25(12):1569-80. doi: 10.1080/09540121.2013.793269. PubMed PMID: PMC4664148.

10. World Health Organization. Guidance on couples HIV testing and counselling including antiretroviral therapy for treatment and prevention in serodiscordant couples: recommendations for a public health approach. Geneva: World Health Organization; 2012.

11. Farquhar C, Kiarie JN, Richardson BA, Kabura MN, John FN, Nduati RW, et al. Antenatal couple counseling increases uptake of interventions to prevent HIV-1 transmission. J

Acquir Immune Defic Syndr. 2004;37(5):1620-6. Epub 2004/12/04. PubMed PMID: 15577420; PubMed Central PMCID: PMCPmc3384734.

12. Msuya SE, Mbizvo EM, Hussain A, Uriyo J, Sam NE, Stray-Pedersen B. Low male partner participation in antenatal HIV counselling and testing in northern Tanzania: implications for preventive programs. AIDS Care. 2008;20(6):700-9. Epub 2008/06/26. doi: 10.1080/09540120701687059. PubMed PMID: 18576172.

13. Obermeyer CM, Osborn M. The utilization of testing and counseling for HIV: a review of the social and behavioral evidence. Am J Public Health. 2007;97(10):1762-74. Epub 2007/09/01. doi: 10.2105/ajph.2006.096263. PubMed PMID: 17761565; PubMed Central PMCID: PMCPMC1994175.

14. Hutchinson AB, Corbie-Smith G, Thomas SB, Mohanan S, del Rio C. Understanding the patient's perspective on rapid and routine HIV testing in an inner-city urgent care center. AIDS Educ Prev. 2004;16(2):101-14. Epub 2004/05/12. PubMed PMID: 15134119.

15. Choko AT, Desmond N, Webb EL, Chavula K, Napierala-Mavedzenge S, Gaydos CA, et al. The uptake and accuracy of oral kits for HIV self-testing in high hiv prevalence setting: a cross-sectional feasibility study in Blantyre, Malawi. PLoS Med. 2011;8(10):e1001102. doi: 10.1371/journal.pmed.1001102.

16. Figueroa C, Johnson C, Verster A, Baggaley R. Attitudes and acceptability on HIV self-testing among key populations: a literature review. AIDS Behav. 2015;19(11):1949-65. doi: 10.1007/s10461-015-1097-8. PubMed PMID: 26054390; PubMed Central PMCID: PMC4598350.

17. Napierala Mavedzenge S, Baggaley R, Corbett EL. A review of self-testing for HIV: research and policy priorities in a new era of HIV prevention. Clin Infect Dis. 2013;57(1):126-38. doi: 10.1093/cid/cit156.

18. Choko AT, MacPherson P, Webb EL, Willey BA, Feasy H, Sambakunsi R, et al. Uptake, accuracy, safety, and linkage into care over two years of promoting annual self-testing for HIV in Blantyre, Malawi: a community-based prospective study. PLoS Med. 2015;12(9):e1001873. Epub 2015/09/09. doi: 10.1371/journal.pmed.1001873. PubMed PMID: 26348035; PubMed Central PMCID: PMCPmc4562710.

19. National AIDS and STI Control Programme. Guidelines for HIV testing services in Kenya. Nairobi: National AIDS and STI Control Programme; 2015.

20. World Health Organization. Consolidated guidelines on HIV testing services. Geneva: World Health Organization; 2015.

21. National AIDS and STI Control Programme, Ministry of Public Health and Sanitation. Guidelines for HIV testing and counselling in Kenya. Nairobi: National AIDS and STI Control Programme; 2010.

22. Thirumurthy H, Masters S, Napierala Mavedzenge S, Maman S, Omanga E, Agot A. Promoting male partner HIV testing and safer sexual decision making through secondary distribution of self-tests by HIV-negative female sex workers and women receiving antenatal and post-partum care in Kenya: a cohort study. Lancet HIV. 2016;3(6):e266-74.

23. National AIDS and STI Control Programme. Kenya HIV county profiles. Nairobi: Ministry of Health; 2014.

24. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol. 2004;159(7):702-6. Epub 2004/03/23. PubMed PMID: 15033648.

25. Kenya National Bureau of Statistics, ICF Macro. Kenya Demographic and Health Survey 2008-09. Calverton (Maryland): ICF Macro; 2010.

26. Osoti AO, John-Stewart G, Kiarie J, Richardson B, Kinuthia J, Krakowiak D, et al. Home visits during pregnancy enhance male partner HIV counselling and testing in Kenya: a randomized clinical trial. AIDS. 2014;28(1):95-103. Epub 2013/08/15. doi: 10.1097/qad.0000000000000023. PubMed PMID: 23942059; PubMed Central PMCID: PMCPmc4337402.

27. Rosenberg NE, Mtande TK, Saidi F, Stanley C, Jere E, Paile L, et al. Recruiting male partners for couple HIV testing and counselling in Malawi's option B+ programme: an unblinded randomised controlled trial. Lancet HIV. 2015;2(11):e483-91. doi: http://dx.doi.org/10.1016/S2352-3018(15)00182-4.

28. Ng'ang'a A, Waruiru W, Ngare C, Ssempijja V, Gachuki T, Njoroge I, et al. The status of HIV testing and counseling in Kenya: results from a nationally representative population-based survey. J Acquir Immune Defic Syndr. 2014;66(Suppl 1):S27-36. Epub 2014/04/16. doi: 10.1097/QAI.0000000000000102. PubMed PMID: 24732818.

29. Kimanga DO, Ogola S, Umuro M, Ng'ang'a A, Kimondo L, Murithi P, et al. Prevalence and incidence of HIV infection, trends, and risk factors among persons aged 15-64 years in Kenya: results from a nationally representative study. J Acquir Immune Defic Syndr. 2014;66(Suppl 1):S13-26. doi: 10.1097/QAI.0000000000000124. PubMed PMID: 24445338; PubMed Central PMCID: PMCPMC4794992.

30. Carballo-Dieguez A, Frasca T, Balan I, Ibitoye M, Dolezal C. Use of a rapid HIV home test prevents HIV exposure in a high risk sample of men who have sex with men. AIDS Behav. 2012;16(7):1753-60. Epub 2012/08/16. doi: 10.1007/s10461-012-0274-2. PubMed PMID: 22893194; PubMed Central PMCID: PMC3458207.

31. Carballo-Dieguez A, Frasca T, Dolezal C, Balan I. Will gay and bisexually active men at high risk of infection use over-the-counter rapid HIV tests to screen sexual partners? J Sex Res. 2012;49(4):379-87. Epub 2012/02/02. doi: 10.1080/00224499.2011.647117. PubMed PMID: 22293029.

CHAPTER 3: THE IMPACT OF FINANCIAL INCENTIVES ON UTILIZATION OF A WEB-BASED WEIGHT LOSS INTERVENTION AMONG NORTH CAROLINA COMMUNITY COLLEGE EMPLOYEES

3.1 Overview

Background: Overweight and obesity are pressing public health issues in the United States that are often addressed in worksite-based wellness programs. Various interventions for weight loss, including web-based behavioral programs and incentives, have been tested, but it is not well understood how multiple interventions interact.

Objective: To assess whether financial incentives for weight loss (1) increased use of a web-based application designed to assist people with their weight loss goals, and (2) if utilization of the web-based application led to weight loss.

Methods: Secondary analysis of data from a cluster randomized trial of 17 community colleges in North Carolina that took place between 2005 and 2008. Colleges and their employees were randomized to one of three groups: 1) environmental change in which healthy food options were offered in the school cafeterias (ENV), 2) a web-based weight loss application plus ENV (WEB), or 3) WEB plus financial incentives for weight loss, plus ENV (WPI). Participants were followed up at 3, 6, and 12 months. We compared website utilization between the WPI and WEB groups using modified Poisson regression. We estimated the complier average causal effect (CACE) of website utilization on weight loss using two different methods, instrumental variables and propensity scores, to address possible bias from self-selection of website use. We used the study group, which was randomly assigned, to instrument for web utilization.

Results: Among 610 participants analyzed in the WEB and WPI groups, utilization of the website was more common among WPI participants (13.9%, 38/273) than WEB participants (8.6%, 29/337) at all follow-up points (Pooled Risk Ratio=1.885, 95% CI 1.552-2.289, P=<0.001). Among 693 participants analyzed in the WEB and ENV groups, propensity score analyses of the CACE found that logging into the website led to weight loss, but the amount of weight loss declined over time (3-month CACE=-6.76lbs, 95% CI -10.27 to -3.30 vs. 12-month CACE=-4.16lbs, 95% CI -9.30 to 1.04).

Conclusions: Positive financial incentives for weight loss increased use of a web-based weight loss application. Utilization of the web-based application led to weight loss among members of the WEB group early in the study period (0-6 months) but not later (6-12 months). Future research needs to focus on the sustainability of weight loss effects. Without an intervention that promotes weight loss and helps people keep weight off, the impact of population based web interventions is likely to be minimal.

3.2 Introduction

Overweight and obesity affect 68% of adults in the United States [1] and significantly contribute to higher disease burden and health expenditures [2, 3]. Employees and employers bear the cost associated with medical care, productivity loss, and insurance for overweight employees [4, 5]. Given such financial externalities, companies are now investing in employee wellness programs with the hope that it becomes a win-win that saves money for both the company (reduced health care spending and less absenteeism and presenteeism due to health problems) and employee (better health and financial gain through reduced premiums or direct incentives) [6].

One potential way to increase weight loss among employees is to provide them incentives for meeting weight loss goals. Incentives can take different forms, both positive (e.g. financial reward for weight loss) or negative (e.g. loss of deposit if no weight loss), and can be placed on different aspects of weight loss, such as inputs (e.g. exercise) or outputs (e.g. weight loss). Evidence of the impact of positive incentives on weight loss outside of worksites is mixed [7-9]. One pilot randomized controlled trial implemented positive incentives in the worksite and found no statistically significant weight loss difference at 6 months between those randomized to receive financial incentives early in the study, late in the study, or throughout the study [10]. The mixed results associated with positive incentives led to new research on combining incentives with other weight loss strategies, such as web-based weight loss aids, in the hope that these combination strategies lead to sustained weight loss.

The evidence on the effectiveness of web interventions on weight loss is generally positive, with one review concluding that more research is needed and others suggesting that type of web intervention may be important [11-13]. The lack of clear evidence is in part due to the vast array of possible interventions that can be administered with web technology [14]. Further, although numerous studies have assessed the intent-to-treat effect of web interventions on weight loss, fewer have looked at why the web interventions are effective, specifically, how individuals' behavior is altered by actively utilizing web-based applications [15-17]. Furthermore, the studies that examined the impact of web-based interventions on weight loss have been unable to accurately estimate the causal impact of utilization due to the self-selection of usage inherent in web-based interventions. This study addresses these prior limitations by utilizing multiple econometric methods, including instrumental variables and propensity scores, to account for self-selection and estimate the causal impact of website utilization on weight loss among those who use the website.

We conducted a secondary analysis of data from a cluster randomized trial that combined a web-based weight loss intervention with positive incentives for weight loss to better understand how incentives and web-based weight loss interventions interact when paired together. Two questions were addressed: 1) how web utilization differed as a result of a financial incentive and 2) if web utilization had an impact on weight loss for those who utilized the website.

3.3 Methods

3.3.1 Study setting and design

The Worksite Activities for You (WAY) to Health cluster randomized controlled trial took place in North Carolina between 2005 and 2008 (ClinicalTrials.gov identifier: NCT01282775). The study goal was to increase weight loss among overweight and obese employees and the main study details are described in depth elsewhere (Crane, Tate, Finkelstein, & Linnan, 2012; Linnan et al., 2012). Prevalence of overweight in the community college employee population (64%) is similar to the rate found nationwide (68%) (Flegal et al., 2010; Linnan et al., 2012). Employees of each selected college were eligible to participate if they met the following criteria: overweight or obese (BMI>25), aged 18+, and employed full-time or part-time. Certain employees were excluded due to various factors, including: low BMI (<25), pregnancy, and type 1 diabetes. Details of exclusion criteria have been discussed elsewhere (Linnan et al., 2012). Once enrollment was completed at all schools, 17 college campuses were randomized to one of three study groups: 1) environmental change, in which community colleges offered healthy food options in the school cafeteria and vending machines (ENV), 2) environmental change plus a web-based weight loss program (WEB), and 3) environmental change plus a web-based weight loss program plus incentives for weight loss (WPI). All enrolled employees within a given campus received the same intervention. The intervention period lasted 12 months. Participants were reweighed at 3, 6, and 12 months and received $5, $10, and $20 contingent upon being reweighed at those times, respectively.

The website intervention was modeled after a weight loss intervention described in detail elsewhere (Tate, Jackvony, & Wing, 2003). In brief, the website was primarily self-directed and included information on worksite-based nutrition- and exercise-related events on campus, weekly online weight loss lessons, and a study progress tracking system in which people were able to input their weight, diet, and exercise information periodically. The website did not contain tailored information based on participant input; rather, lessons were structured so that participants could self-tailor by clicking on a statement that best described their progress, and then receive suggestions for next steps. The web intervention also included a weekly email sent to participants alerting them of a new lesson and physical diaries to record information for inputting into the web application later. Participants could login as frequently as they chose. The WPI group was identical to the WEB group except that they also received financial incentives for weight loss at each of the follow-up visits and the website displayed a personalized table of the incentives the participant would earn for various levels of weight. At each of the 3 follow-up visits, participants in the WPI group received $5 for each 1% decrease in weight from their baseline measurement, up to a maximum of 10% ($50). They were eligible to receive the cash incentive at each official weigh-in monitored by study staff (at 3, 6, and 12 months) for a total maximum incentive of $150.

3.3.2 Measures and descriptive analyses

The primary outcome was whether a study participant logged into the website over the each follow-up period: 0-3 months, 3-6 months, and 6-12 months. The key explanatory variable was a binary variable indicating if the study participant was in the WPI group or the WEB group. There was no missing participation data given that website usage was electronically monitored for the entire period. We conducted bivariate analyses comparing study participant characteristics among those who utilized the website and those who did not using Pearson's Chi-squared tests for binary variables and Mann Whitney tests for continuous variables. Covariates analyzed included baseline weight and BMI, age, self-reported health (poor/fair/good, very good/excellent, missing), sex (male, female, missing), race (white, black, other, missing), marital status (married, unmarried, missing), income (>$50,000, $30,000 - $49,999, $0 - $29,999, missing), education (post-graduate degree, associate or bachelor's degree, no college degree, missing) and job type (faculty, non-faculty, missing). The secondary outcome was weight change in pounds at the end of the study (12 months). The main explanatory variable of the secondary analysis was website utilization. Website utilization was instrumented using an indicator for the WEB group. There was missing weight change data due to study participants failing to get weighed at the various time points.

3.3.3 Impact of financial incentive on utilization

The unit of analysis was the study participant. The primary analysis compared utilization of the website between the WEB and WPI groups using modified Poisson regression (Zou, 2004). We analyzed utilization at each follow-up point, and we also ran a pooled model in which we included all observations over the entire follow-up period (with variables added to control for follow-up points). We included community college fixed effects to account for potential differences in intervention fidelity by study site. We accounted for potential autocorrelation among participants within the same community college by clustering our standard errors using White's general correction at the community college level. Similarly, for the overall estimate with all follow-up data, multiple measurements were used from the same individuals over time, and the clustered standard errors at the community college level account for potential autocorrelation between individuals in the same college and across the same individual over time. The model included covariates for baseline weight and BMI, age, sex, race, marital status, education, job type, income, and self-reported health. In a separate model we analyzed the absolute number of logins to the website as the dependent variable, rather than any logins. We used a pooled ordinary least squares regression with errors clustered at the community college level to analyze the number of logins.

3.3.4 Impact of utilization on weight loss

In secondary analyses, we estimated the causal effect of website utilization among users. We conducted these analyses for two primary reasons: first, to understand the true impact of website utilization on weight loss; and second, to understand if the financial incentive would have extraneous impacts on weight loss through utilization, rather than solely through its direct impact on weight loss. In order to understand the impact of website utilization on weight loss we limited our analyses to those who did not have access to weight loss incentives because their access to incentives potentially biased their usage of the website, and subsequently the causal impact of website utilization on weight loss. Thus, our secondary analyses are limited to participants in the ENV and WEB groups.

One method for estimating the efficacy of an intervention is a “per protocol” analysis, which compares the outcomes of study participants in the WEB group who utilize the website with outcomes of participants in the ENV group. Another is an “as treated” analysis, in which the utilizers in the WEB group are compared with all participants randomized to the ENV group as well as participants randomized to the WEB group who did not utilize the web service. Both the “as treated” and “per protocol” estimates are likely to be biased because they group participants based on their behavior (comply or not), and the participants who comply may differ from those who do not in unobservable ways (BARNARD, DU, HILL, & RUBIN, 1998; Sheiner & Rubin, 1995). The potential for bias is high, given that uptake of the intervention in the WEB group was very low—only 8.6% of eligible people logged into the website by 12 months. Therefore, we used two methods, instrumental variables (IV) (Angrist, Imbens, & Rubin, 1996; Newhouse & McClellan, 1998; Sussman & Hayward, 2010) and propensity scores (PS) (Jo & Stuart, 2009), to account for self-selected compliance in the WEB group. These methods are used to estimate a complier average causal effect (CACE) (Dunn, 2011), which is an unbiased estimator of the efficacy of the web-based intervention. The goal of the IV and PS methods is to accurately compare the group of individuals who utilized in the WEB group with a group of individuals in the ENV group who potentially would have complied (i.e. used the website) if they had been randomized to the WEB group. We describe the methods used to generate the CACE and their assumptions below.

3.3.5 Instrumental variables

In order to estimate the CACE using IV we instrument website utilization using the study group. We used two stage least squares in which we predicted utilization using the study group in the first stage:

$$(1)\ UTIL_{it} = \alpha_i + \rho WEB_i + \varepsilon_{it}$$

where *UTIL* is a binary variable indicating if a participant utilized the website during the 12 month study period. *WEB* is the instrument; it is an indicator for if a participant belonged to the WEB study group and not the ENV group. Then in the second stage we included the fitted values for utilization from the first stage in the regression of weight loss on utilization.

$$(2)\ WL_i = \alpha_i + \delta\widehat{UTIL}_i + \varepsilon_i$$

where WL is a continuous measure of weight change at 12 months and $\widehat{UTIL}$ is predicted values of utilization from the first stage regression. The effect, δ, is the estimated CACE using instrumental variables. The IV CACE is an unbiased estimator of the causal impact of utilization on weight loss. However, IV CACE estimate requires several key assumptions for validity, which we explore below.

3.3.6 Assumptions of the IV

In short, participants in the trial can be categorized into four distinct groups: 1) compliers—those who complied with their assigned treatment, i.e. those who logged into the website in the WEB group and those who did not utilize in the ENV group, 2) always takers—those who would always comply, no matter what group they were randomized to, i.e. those randomized to either the WEB or ENV group who would always log into the website; 3) never takers—those who would never comply, no matter what group they are randomized to, i.e. those randomized to either the WEB or ENV group who would never log into the website; and 4) defiers—those who would always do the opposite of the group they were randomized to, i.e. those randomized to the WEB who would not utilize the website and those randomized to the ENV group who would use the website (Angrist et al., 1996). The WAY trial was designed so that those randomized to the ENV control group did not have access to the weight loss website to which the WEB group had access. This trial design, therefore, allows for only compliers and never takers.

Five assumptions must be met for the IV method to be an unbiased estimator of the CACE for randomized trials with noncompliance (Angrist et al., 1996). The first assumption is the stable unit treatment value assumption (SUTVA), which assumes that the potential outcomes for study participants are unrelated to the treatment status of other study participants. This assumption may be violated, for example, if a participant discussed their strategies for weight loss with other participants. We do not have data on the extent of participant conversations with each other in the WAY dataset, and therefore SUTVA is impossible to test empirically. Although conversations between participants are possible, it seems unlikely that these discussions, if they did happen, would significantly impact individual behavior and therefore we assume SUTVA holds. The second assumption is random assignment, which is inherent in the randomized trial design of the WAY study. Third is the exclusion restriction, which assumes that the impact of the instrument on outcomes, is entirely through the effect of compliance on outcomes. Fourth is that the instrument and compliance are positively correlated. The first stage regression from two stage least squares of website utilization on study group yielded a F-statistic of 30.47, which confirms that study group is strongly correlated with website utilization. Fifth is monotonicity, which states there are no defiers; there are no defiers in the WAY trial given the design.

3.3.7 Bias correction of IV

In the WAY trial, a strong case can be made that assumptions 1,2,4 and 5 were met. However, it is less clear if assumption 3, the exclusion restriction, was met. Fortunately, the WAY data were collected in such a way that we could test to see if the outcomes for those who used and did not used the website in the WEB group differed, thus testing the exclusion restriction assumption. In the WAY data, mean 12-month weight loss in the ENV group was -0.5 pounds (Table 6). In the WEB group, mean weight loss was -2.8 pounds. Weight loss among compliers in the web group was -4.9 pounds. Weight loss among non-compliers (never-takers) in the WEB group was -2.6 pounds. Given that weight loss in the ENV group differed from weight loss among non-compliers in the WEB group (-0.5 in ENV group vs. -2.6 among non-compliers in WEB group), we violate the exclusion restriction because those randomized to the WEB group and not utilizing the website lost more weight than ENV participants. Participants in the WEB group had an unobservable weight loss benefit due to being randomized to the WEB group besides use of the website. For example, participants in the WEB study group may have read weekly weight loss emails or utilized the paper self-monitoring records or calorie books. Failing to meet the exclusion restriction assumptions means that the IV estimates of website efficacy could be biased. To account for the bias we "corrected" the IV estimates by subtracting off the bias (assuming the bias is additive) (Angrist et al., 1996). The bias in the WAY trial was simply the difference between the weight loss among ENV group participants and weight loss among non-compliant WEB group participants. This correction yielded a bias corrected CACE, which is the causal impact of the website on weight loss among participants who utilized the website.

Thus, the estimation of the CACE using instrumental variables involved four steps. First, we estimated the biased CACE using two stage least squares. Second, we generated the magnitude of the bias by dividing the impact of randomization among never-takers by the proportion of compliers. Third, we assumed additive bias and subtracted the bias from the CACE estimate. Last, we bootstrapped standard errors for the CACE estimate with 1000 replications.

3.3.8 Propensity scores

A different approach for estimating complier average treatment effects comes from principal stratification (Frangakis & Rubin, 2002). In principal stratification study participants are stratified by their compliance, either utilizers or not. It was straightforward to stratify by compliance in the WEB group, where compliance was observed; however, compliance was not observed in the ENV group since they were never given the opportunity to use the website and therefore compliance needed to be estimated. One way to estimate compliance, and subsequently the CACE, is to use propensity scores (Jo & Stuart, 2009; Joffe, Ten Have, & Brensinger, 2003; Stuart & Jo, 2015). The PS method relies on concept of principal ignorability (Jo & Stuart, 2009), which assumes no differences in potential outcomes across different principal strata (never takers or compliers) given the observed pretreatment variables (baseline covariates). In other words, strata membership is identifiable using only observed variables and not driven by unobserved differences. Since we observed outcomes and compliance in the WEB group, this assumption only applies to the ENV group. Estimation of the CACE requires the same SUTVA and randomization assumptions of IV analysis, but does not require the exclusion restriction. Thus, estimation of the CACE using PS did not need to be corrected for the bias of the exclusion restriction as estimation using IV did.

The PS method for CACE estimation required two steps. First, a logistic model was used to determine probability of compliance among WEB group recipients only using baseline covariates, and propensity scores were predicted for participants in the ENV group:

$$(3)\quad logit(PR_i) = \alpha_i + X_i'\beta + \varepsilon_{it}$$

where *PR* is the probability that individual *i* is a complier and *X* represents a vector of baseline covariates used to predict compliance. The logistic model included the following variables in *X*: baseline weight and BMI, age, education, income, and self-reported health.

Second, the propensity scores were used as weights in an ordinary least squares regression model of weight loss on website utilization in order to estimate the CACE:

$$(4)\quad WL_i = \alpha_i + \rho WEB_i + X_i'\beta + \varepsilon_i$$

where *WL* represents weight change at 12 months, *WEB* is an indicator for if the participant belonged in the WEB study group, and *X* represents the same vector of covariates as in Equation (3). The estimated effect, ρ, is the CACE using PS. We chose to weight on the propensity score because it has better performance in CACE estimation than matching on the propensity score (Jo & Stuart, 2009). Participants in the WEB group who utilized the website were given a weight of 1, participants in the WEB group who did not utilize the website were given a weight of 0, and participants in the ENV group were given a weight of PR/(1-PR). One unique advantage of this method is that it allowed for estimation of the CACE and also the never-taker average causal effect (NACE). The NACE is a measure of the weight loss among those who did not utilize the website in the WEB group and can be used to assess violations of the exclusion restriction. If the NACE is different from zero, the randomization itself may have affected outcomes for those in the treatment group and therefore the CACE found in the IV analysis would be biased. We bootstrapped the standard errors with 1000 replications.

3.3.9 Missing weight loss data

The estimation of the CACE is complicated by the fact that there are missing weight loss (outcome) data. Approximately 32% of participants in the WEB and WPI groups were missing a 12-month weight measurement. When missing data are present, the central question is whether or not the missing data are random, and if not, are they random conditional upon observable covariates. We performed two analyses using different assumptions for missing data. In the first, we assumed the outcome data were missing completely at random (MCAR). MCAR requires missing data to be uncorrelated with both observed and unobserved data. This strong assumption allows for unbiased complete case analysis (CCA) of the data. Second, we assumed data are missing at random (MAR). Unlike MCAR, MAR assumes that the missing weight loss data can be correlated with observable data as long as it is included in the regression model, and that missing weight loss data are unrelated to their values of weight loss. We used multiple imputation (MI) to account for MAR data. Missing weight loss data were imputed using all covariates available (see Table 6). Within each bootstrapped dataset that includes missing data, we created 10 multiple imputed datasets (Schomaker & Heumann, 2016). We then generated the MI CACE within each bootstrapped dataset. We did this for both the IV and PS estimators. It is important to note that neither of these two assumptions can be directly tested, and therefore we present both as sensitivity analyses.

We generated 12 month CACE estimates using each of the above methods: per protocol, IV with CCA, IV with MI, PS with CCA and PS with MI. Additionally, we present CACE estimates using our preferred method, PS with MI, for each time period (3, 6, and 12 months) and overall. We included these measurements to examine how the CACE changed over time. We chose to use the PS with MI approach for examining the CACE over time because we believe is has the most defensible assumptions, predominantly its lack of reliance on the exclusion restriction. We used nonparametric percentile based confidence intervals for all bootstrapped estimates. All analyses were conducted in Stata v14.1 (StataCorp, College Station, TX, USA). All hypothesis tests were two-sided with α=0.05, and no adjustment was made for multiple comparisons.

3.4 Results

In total for the original study, 7 colleges and 375 employees were randomized to the EC group, 5 colleges and 350 employees to the WEB group, and 5 colleges and 279 employees to the WPI group. The majority of the sample was female and white (Table 6). Over half of the participants were married. Approximately 39% of the sample was faculty at the community colleges. Mean baseline weight at enrollment was 204 pounds, and mean BMI was 34. Among those randomized to the WEB or WPI group, utilization of the website was relatively low, as only 67 of 610 participants (11%) logged into the website during the follow-up period. The proportion of participants who logged into the website decreased during each follow-up period for both the WPI and WEB groups (Figure 2). Similarly, the overall mean number of logins per participant decreased during each follow-up period (Figure 3).

3.4.1 Impact of incentive on utilization of website

The incentive increased the number of people who logged into the web-based weight loss application (Table 7) [Pooled RR=1.885, 95% CI, 1.552-2.289, *p*=<0.001]. The impact of the incentive did not vary significantly between follow-up periods. There was no significant difference in the number of logins between WPI and WEB participants [Coef=0.329, 95% CI, -0.206 to 0.865, *p*=0.198], suggesting that the incentive was effective at increasing initial usage of the website but not sustained usage over time.

3.4.2 Impact of website utilization on weight loss

Weight loss was higher for those who utilized the website than those who did not in the WEB group. In the WPI group the opposite was true: weight loss was higher among those who did not utilize the website (Table 6). Participants who used the website in the WEB and WPI groups and those who did not were generally similar. In bivariate analyses, baseline BMI was significantly higher in the group who utilized than the group who did not (*P*=0.049). Those missing a self-reported health measurement were significantly less likely to utilize the website (*P*=0.023) and participants who reported an income between $0 and $29,999 were more likely to use the website (*P*=0.021). Age, sex, race, education, marital status, and job type were not significantly related to website utilization in bivariate analyses.

Per protocol effects for the web-based weight loss application suggest that the intervention was successful at reducing weight among those who used it [Coef=-4.34, 95% CI, -8.49 to -0.20] (Table 8). Results from the CACE estimation using IV are presented in columns 2 and 3 in Table 8. Estimation of the CACE using IV with CCA, without accounting for the violation of the exclusion restriction, resulted in very large estimates of the impact of the incentive [naïve CACE = -21.74 pounds]. However, once accounting for the bias, estimates were lower than the per protocol estimates [CACE = -2.25, 95% CI, -7.71 to 2.85]. Results from the IV with both CCA and MI were insignificant and imprecisely estimated.

CACE estimation using PS yielded an estimate that was qualitatively similar to that found with the per protocol estimate [CACE using PS = -4.89, 95% CI, -10.89 to 1.19]. The magnitude of the CACE estimates from PS were similar between the CCA and MI approaches to missing data. None of the IV or PS results were statistically significant, suggesting that utilization of the website did not lead to weight loss among the group of compliers at 12 months.

Table 9 presents the CACE results at each study point individually, using the PS with MI approach. At 3 and 6 months, the CACE was qualitatively similar [0-3 month CACE = -6.76, 95% CI -10.27 to -3.30 and 3-6 month CACE = -6.92, 95% CI -12.24 to -1.01]. The 6-12 month CACE estimate was lower than the estimate from the previous two periods [6-12 month CACE = -4.16, 95% CI -9.30 to 1.04], suggesting that website usage in the 6 to 12 month period was associated with less weight loss than from enrollment to 6 months. The pooled CACE estimate [pooled CACE = -6.03, 95% CI -8.39 to -3.82] was statistically significant.

3.5 Discussion

Employees who received a financial incentive for weight loss were more likely to utilize a web-based weight loss application than employees who did not receive a financial incentive. Absolute usage of the website among participants in the WEB group was low (29/337, 8.6%), but receipt of the incentive increased the probability of logging onto the website. The results of the estimation of the CACE using IV and PS suggest that utilization of the website was ineffective at increasing weight loss among members of the WEB group at 12 months. However, subsample analyses at each time point showed that website utilization led to weight loss early in the study period (0-6 months) but was ineffective at promoting weight loss later in the study period (6-12 months).

Given that the incentive was tied to weight loss and not website utilization, the result of higher website usage in the WPI group is somewhat surprising. The finding suggests that employees in the WPI group believed that use of the web application would help them achieve more weight loss and therefore receive more incentive. The effect of the incentive on utilization was remarkably consistent over time, which further supports the hypothesis that participants viewed the website as a helpful way to achieve weight loss goals over the entire study period. The positive impact of the website on weight loss among those who used the website in the WEB group, the CACE in our study, appeared to decline over time. The declining effect is likely due to regaining weight over time and is a common finding in studies related to weight loss (Arem & Irwin, 2011; Wing , Tate , Gorin , Raynor , & Fava 2006). The declining impact of the CACE over time suggests that the website was effective at helping compliers lose weight early, soon after enrollment, but ineffective at helping them maintain weight loss over the study period.

Although we were unable to isolate the causal impact of web utilization on weight loss in the WPI group because the WPI group received incentives which may motivate weight loss directly, descriptive analyses showed that participants in the WPI group who logged into the website lost less weight than those who did not log in. This finding is contradictory to the hypothesized direction and the CACE analysis, in which we found that utilization of the website for WEB group participants did not significantly impact 12 month weight change. One possible explanation for this difference is that the most attractive feature of each intervention was utilized by the highly motivated. For the WEB group, highly motivated individuals utilized the website for weight loss. For the WPI group, highly motivated people focused on the incentive. Subsequently, participants who ended up using the website in the WPI group had relatively low motivation compared to those who did not and therefore did not lose as much weight as the highly motivated participants who did not utilize the website. Thus, the finding that positive financial incentives led to website usage should not be viewed as another possible way to promote weight loss among study participants, and that the web effect and incentive effect acted independently. Future studies should test pairing financial incentives for weight loss with a variety of evidence-based weight loss interventions, such as more tailored web-based interventions, to potentially leverage synergies and help promote weight loss. Access to these resources could lead to greater weight loss than financial incentives alone.

We are not aware of other weight loss studies that accounted for noncompliance using an instrumental variables or propensity score estimation strategy to determine the CACE. The CACE allows for precise estimation of the impact of utilization of the website for the group of participants who used the website in the WEB group, accounting for noncompliance. Despite the unbiasedness of the CACE, careful attention must be paid when interpreting these effects. The CACE is estimated from a group of individuals who chose to log into the weight loss website. People who used the website may have unobserved characteristics correlated with weight loss and utilization; for example, they may be highly motivated and therefore their outcomes may differ from the outcomes that would be observed if other participants who chose not to log in, were "forced" to log into the website (Sussman & Hayward, 2010). Thus the CACE is potentially not generalizable to the nonusers. Although the estimate is generated from a self-selected group, it nonetheless represents the weight loss we could expect to see among compliers if a website for weight loss was introduced to a community college employee population.

This study has other limitations besides the generalizability of the CACE. First, significant attrition occurred in the study with respect to follow-up weight loss measurements, which could impact our estimate of the CACE. Although we used multiple imputation for dealing with attrition, inherently we are relying on the assumption that data are missing at random conditional on observed covariates, which may not be true. If this is not true because participants who did not lose weight were less likely to get weighed, the estimated CACE would be an overestimate of true weight loss due to the website. Second, those who did not comply in the WEB group had different outcomes from those in the ENV group. The weight loss differences between those who did not utilize the website in the WEB group and the ENV group were significant. Participants in the WEB group who did not use the website likely lost more weight due to other aspects of the intervention (e.g., weekly emails, paper-based calorie diaries and a comprehensive book with the calories of foods). However, we were unable to monitor usage of these aspects. Correcting for the bias in IV analysis required assumptions including that the bias was equal to the difference in weight loss between the non-compliers in the WEB group and weight loss in the ENV group and that the bias was additive. We believe that these assumptions are accurate and valid based on the observable data; however, other approaches could be taken. Because of the IV assumptions, we also used a different estimation strategy, PS, to generate unbiased CACE estimates. The PS analysis has its own assumptions, which we noted earlier, but we view the PS method as the most robust estimation strategy for the CACE because it does not rely on the exclusion restriction assumption—as IV does—which is violated in our study. We present both estimates so that readers can understand the differences in assumptions and results between the methods.

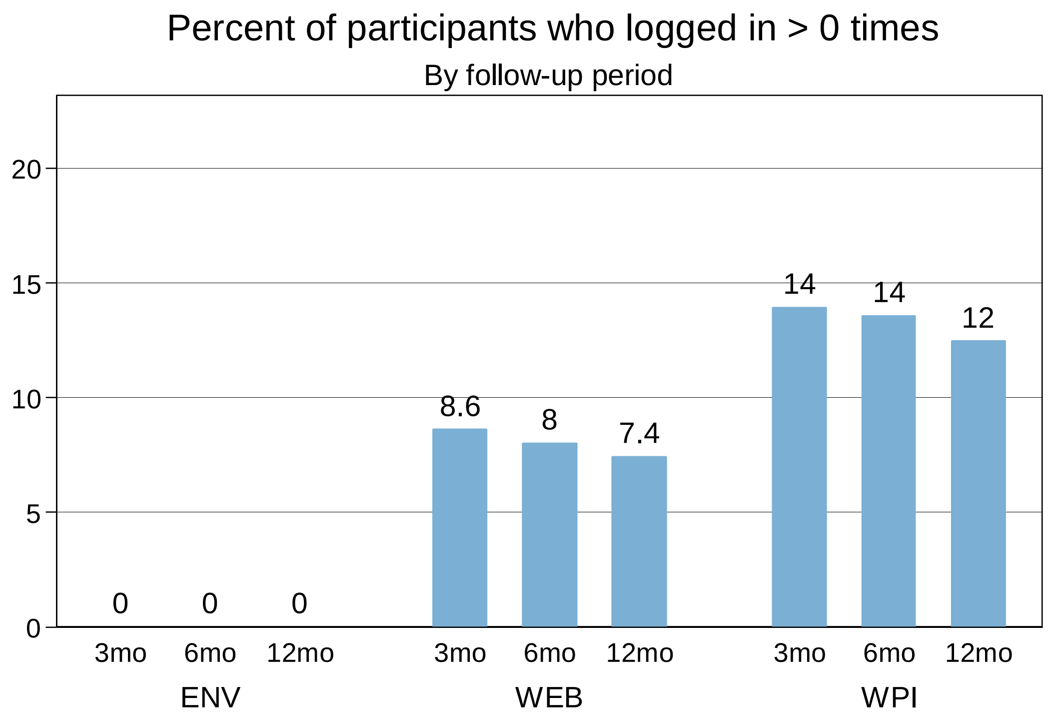
The CACE estimates from this study represent the impact that a researcher or policy maker could expect if a similar weight loss website were rolled out at other worksites. It represents the impact for utilizers of the website, which were a small minority in this study. Since the WAY trial websites have become more sophisticated and are potentially more likely to impact users looking to lose weight. One interesting area of research would be to determine how CACE estimates of website utilization on weight loss are changing with updated websites.

Another interesting new area of research would be to determine what impact positive incentives paired with an input, such as website utilization, would have. It is unclear whether incentives placed on "process" inputs (e,g., website use) are less, equally, or more effective at reducing weight than incentives placed on outputs.

3.6 Conclusions

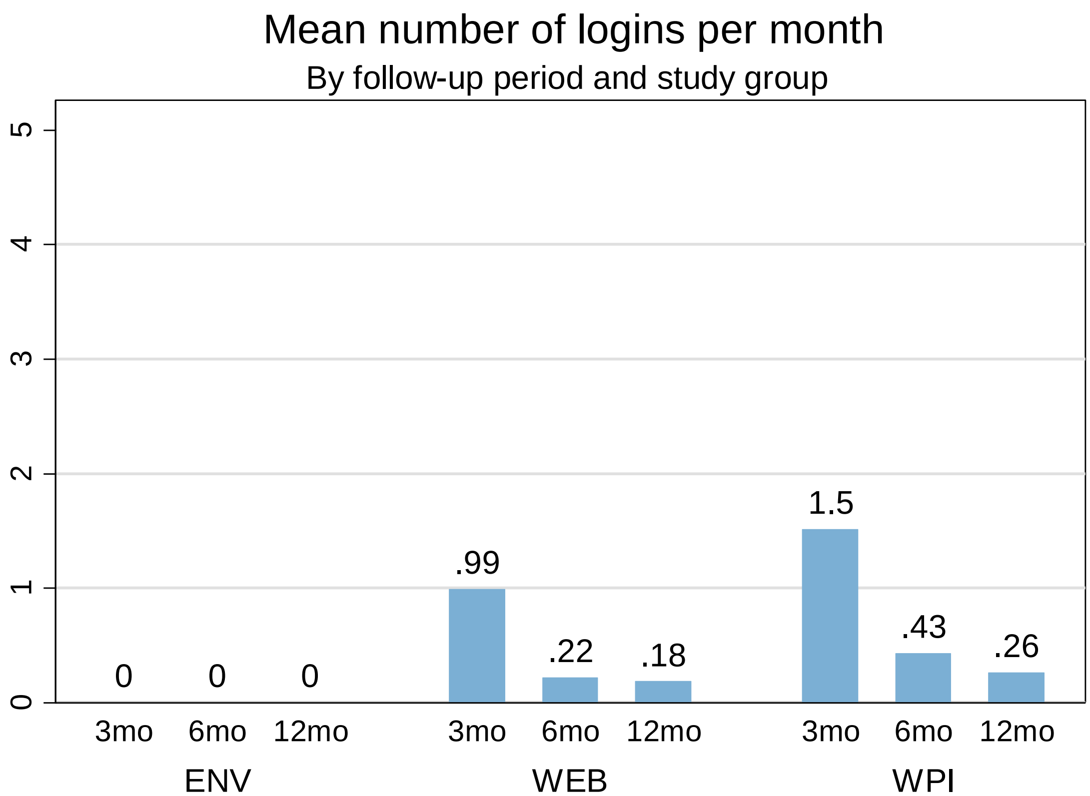
Financial incentives tied to weight loss led to greater utilization of a website designed to assist with weight loss. For participants who were not randomized to receive incentives, utilization of the website led to weight loss during the first 6 months of use, but not later in the study period. From a policy standpoint, websites for weight loss are relatively inexpensive to produce and may significantly impact a subset of people. Thus, it may be beneficial to roll them out to large populations given the benefit associated with even small amounts of weight loss. However, attention needs to be paid to the sustainability of weight loss effects. Without an intervention that promotes weight loss and helps people keep weight off, the impact of population based web interventions is likely to be minimal.

FIGURE 2: Percent of participants who used the website at least once



Abbreviations: ENV – study group that received environmental change only; WEB – study group that was able to access a website for weight loss, as well as environmental change; WPI – study group that received incentives for weight loss, access to a website for weight loss and environmental change.

FIGURE 3: Logins over time for all participants by study group



Abbreviations: ENV – study group that received environmental change only; WEB – study group that was able to access a website for weight loss, as well as environmental change; WPI – study group that received incentives for weight loss, access to a website for weight loss and environmental change.

TABLE 6: Study participant characteristics of those in ENV group and WEB or WPI group, grouped by whether or not they utilized the study website during the study period

	ENV	WEB		WPI			
	Did not utilize, (n=356)	**Did not utilize, (n=308)**	**Utilized, (n=29)**	**Did not utilize, (n=235)**	**Utilized, (n=38)**	**Total, (n=965)**	***P*-value***
12-month weight change, mean (SD)	-0.5 (9.4)	-2.6 (11.8)	-4.9 (12.2)	-4.1 (11.3)	-2.8 (11.1)	-2.2 (10.9)	0.768
Missing 12 month weight measurement	100 (28)	117 (38)	6 (21)	70 (30)	6 (16)	299 (30)	0.007
Baseline weight, mean (SD)	204.9 (44.2)	205.6 (45.2)	212.9 (54.2)	199.7 (41)	214.8 (54.6)	204.5 (44.6)	0.225
Baseline BMI, mean (SD)	33.8 (6.76)	33.3 (6.6)	34.6 (6.9)	33.4 (6.5)	35.3 (7.2)	33.6 (6.7)	0.049
Age, mean (SD)	46.3 (9.6)	47.3 (10.7)	48 (10.5)	47.9 (9.3)	46.5 (10)	47.1 (9.9)	0.722
Self-reported health							
Poor, fair, good	170 (48)	138 (45)	12 (41)	108 (46)	23 (61)	451 (47)	0.283
Very good, excellent	125 (35)	112 (36)	16 (55)	84 (36)	11 (29)	348 (36)	0.500
Missing	61 (17)	58 (19)	1 (3)	43 (18)	4 (11)	167 (17)	0.023
Sex							
Female	272 (76)	218 (71)	21 (72)	196 (83)	30 (79)	737 (76)	0.982
Male	63 (18)	68 (22)	6 (21)	24 (10)	5 (13)	166 (17)	0.914
Missing	21 (6)	22 (7)	2 (7)	15 (6)	3 (8)	63 (7)	0.843
Race							
White	271 (76)	244 (79)	22 (76)	189 (80)	29 (76)	755 (78)	0.490
Black	56 (16)	31 (10)	3 (10)	23 (10)	4 (11)	117 (12)	0.897
Other	8 (2)	9 (3)	2 (7)	8 (3)	1 (3)	28 (3)	0.559
Missing	21 (6)	24 (8)	2 (7)	15 (6)	4 (11)	66 (7)	0.600
Income							
>$50,000	75 (21)	81 (26)	4 (14)	58 (25)	7 (18)	225 (23)	0.100

$30,000 - $49,999	134 (38)	109 (35)	9 (31)	93 (40)	12 (32)	357 (37)	0.348
$0 - $29,999	95 (27)	71 (23)	9 (31)	47 (20)	14 (37)	236 (24)	0.021
Missing	52 (15)	47 (15)	7 (24)	37 (16)	5 (13)	148 (15)	0.605
Education							
Post-graduate degree	145 (41)	112 (36)	11 (38)	97 (41)	14 (37)	379 (39)	0.852
Associate or bachelor's degree	155 (44)	132 (43)	14 (48)	100 (43)	15 (39)	416 (43)	0.931
No college degree	35 (10)	40 (13)	2 (7)	24 (10)	6 (16)	107 (11)	0.971
Missing	21 (6)	24 (8)	2 (7)	14 (6)	3 (8)	64 (7)	0.889
Marital status							
Married	239 (67)	208 (68)	22 (76)	159 (68)	24 (63)	652 (67)	0.860
Unmarried	96 (27)	77 (25)	5 (17)	61 (26)	11 (29)	250 (26)	0.785
Missing	21 (6)	23 (7)	2 (7)	15 (6)	3 (8)	64 (7)	0.889
Job category							
Faculty	129 (36)	122 (40)	14 (48)	98 (42)	9 (24)	372 (39)	0.329
Non-faculty	208 (58)	170 (55)	14 (48)	123 (52)	26 (68)	541 (56)	0.373
Missing	19 (5)	16 (5)	1 (3)	14 (6)	3 (8)	53 (5)	0.881

Notes: All presented statistics are No. (%) unless otherwise noted.
*-*P*-value from Mann Whitney tests for continuous variables and Pearson Chi-squared tests for binary variables. Tests compare those who utilized the web intervention vs. those who did not across both WEB and WPI groups.

TABLE 7: Impact of incentive on utilization of study website

	Follow-up period				
	0-3 months*	**3-6 months***	**6-12 months***	**Pooled***	**Pooled****
Outcome	**Login or not**	**Login or not**	**Login or not**	**Login or not**	**# of logins**
Incentive, RR	2.015	1.900	1.738	1.885	0.329
95% CI	(1.592 to 2.550)	(1.579 to 2.287)	(1.470 to 2.054)	(1.552 to 2.289)	(-0.206 to 0.865)
P-value	<0.001	<0.001	<0.001	<0.001	0.198
Number of individuals	610	610	610	610	610
Total observations	610	610	610	1830	1830

Abbreviations: RR, risk ratio; CI, confidence interval
*-Estimates and CI are risk ratios from modified Poisson regression. Covariates include baseline weight and BMI, age, sex, race, marital status, education, job type, income, self-reported health and study site.
**-Estimates and CI are from ordinary least squares regression. Covariates included were same as those in the modified Poisson regressions.

TABLE 8: Impact of website utilization on weight loss for website utilizers at 12 months

	Per Protocol*	Instrumental Variables **		Propensity Scores ^^	
		Complete case	Imputation^	Complete case	Imputation^
Biased CACE, weight change (lbs)	-4.34	-21.74	-31.08		
95% CI	(-8.49 to -0.20)	(-45.09 to -3.24)	(-61.62 to -4.40)		
Bias, weight change (lbs)		-19.49	-34.34		
95% CI		(-44.10 to 0.30)	(-63.37 to -12.94)		
NACE, weight change (lbs)				-2.23	-2.24
95% CI				(-4.71 to 0.08)	(-4.65 to -0.31)
CACE, weight change (lbs)		-2.25	3.27	-4.89	-3.98
95% CI		(-7.71 to 2.85)	(-9.84 to 20.77)	(-10.89 to 1.19)	(-8.97 to 0.36)
Number of individuals	279	470	693	470	693

Abbreviations: CACE, complier average causal effect of website utilization on weight loss; CI, confidence interval; NACE, Never-taker average causal effect

Notes: *-Estimates and CI are from ordinary least squares linear regression

**-Estimates are from instrumental variables two stage least squares regression. Study group instrumented for website utilization. CACE estimate for IV corrected by subtracting bias from biased CACE estimate

^-Multiple imputation of weight loss predicted by study group, baseline BMI, weight, age, self-reported health, education, income, job type, race, sex and marital status.

^^-Propensity score estimated using baseline weight and BMI, age, self-reported health, education and income as predictors

TABLE 9: Impact of web application on weight loss by study follow-up period

	Follow-up period			
	0-3-months	**3-6-months**	**6-12-months**	**Pooled**
CACE, weight change (lbs)	-6.76	-6.92	-4.16	-6.03
95% CI	(-10.27 to -3.30)	(-12.24 to -1.01)	(-9.30 to 1.04)	(-8.93 to -3.49)
NACE, weight change (lbs)	-2.43	-1.80	-2.33	-2.21
95% CI	(-3.47 to -1.37)	(-3.62 to -0.38)	(-4.51 to -0.31)	(-3.08 to -1.23)
N	693	693	693	2079

Abbreviations: CACE, complier average causal effect; CI, confidence interval; NACE, Never taker average causal effect
Notes: All models used the propensity score estimator with multiple imputation for weight loss.

REFERENCES

1. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999-2008. JAMA. 2010;303(3):235-41. Epub 2010/01/15. doi: 10.1001/jama.2009.2014. PubMed PMID: 20071471.

2. Finkelstein EA, Trogdon JG, Cohen JW, Dietz W. Annual Medical Spending Attributable To Obesity: Payer-And Service-Specific Estimates. Health Affairs. 2009;28(5):w822-w31. doi: 10.1377/hlthaff.28.5.w822.

3. Flegal KM, Graubard BI, Williamson DF, Gail MH. EXcess deaths associated with underweight, overweight, and obesity. JAMA. 2005;293(15):1861-7. doi: 10.1001/jama.293.15.1861.

4. Bhattacharya J, Bundorf MK. The Incidence of the Healthcare Costs of Obesity. J Health Econ. 2009;28(3):649-58. doi: 10.1016/j.jhealeco.2009.02.009. PubMed PMID: PMC4224588.

5. Hammond RA, Levine R. The economic impact of obesity in the United States. Diabetes, metabolic syndrome and obesity : targets and therapy. 2010;3:285-95. doi: 10.2147/DMSOTT.S7384. PubMed PMID: PMC3047996.

6. Benedict MA, Arterburn D. Worksite-Based Weight Loss Programs: A Systematic Review of Recent Literature. American Journal of Health Promotion. 2008;22(6):408-16. doi: doi:10.4278/ajhp.22.6.408. PubMed PMID: 18677881.

7. Paloyo AR, Reichert AR, Reinermann H, Tauchmann H. The Causal Link between Financial Incentives and Weight Loss: An Evidence-Based Survey of the Literature. Journal of Economic Surveys. 2014;28(3):401-20. doi: 10.1111/joes.12010.

8. Patel MS, Asch DA, Troxel AB, Fletcher M, Osman-Koss R, Brady J, et al. Premium-Based Financial Incentives Did Not Promote Workplace Weight Loss In A 2013-15 Study. Health Aff (Millwood). 2016;35(1):71-9. Epub 2016/01/07. doi: 10.1377/hlthaff.2015.0945. PubMed PMID: 26733703.

9. Cawley J, Price JA. A case study of a workplace wellness program that offers financial incentives for weight loss. J Health Econ. 2013;32(5):794-803. doi: 10.1016/j.jhealeco.2013.04.005. PubMed PMID: 23787373.

10. Finkelstein EA, Linnan LA, Tate DF, Birken BE. A pilot study testing the effect of different levels of financial incentives on weight loss among overweight employees. Journal of occupational and environmental medicine / American College of Occupational and Environmental Medicine. 2007;49(9):981-9. Epub 2007/09/13. doi: 10.1097/JOM.0b013e31813c6dcb. PubMed PMID: 17848854.

11. Seo D-C, Niu J. Evaluation of Internet-Based Interventions on Waist Circumference Reduction: A Meta-Analysis. Journal of Medical Internet Research. 2015;17(7):e181. doi: 10.2196/jmir.3921. PubMed PMID: 26199208; PubMed Central PMCID: PMCPMC4527011.

12. Arem H, Irwin M. A review of web-based weight loss interventions in adults. Obesity Reviews. 2011;12(5):e236-e43. doi: 10.1111/j.1467-789X.2010.00787.x. PubMed PMID: 20804523; PubMed Central PMCID: PMCPMC3086798.

13. Tang J, Abraham C, Greaves C, Yates T. Self-Directed Interventions to Promote Weight Loss: A Systematic Review of Reviews. J Med Internet Res. 2014;16(2):e58. doi: 10.2196/jmir.2857.

14. Sherrington A, Newham JJ, Bell R, Adamson A, McColl E, Araujo-Soares V. Systematic review and meta-analysis of internet-delivered interventions providing personalized feedback for weight loss in overweight and obese adults. Obesity Reviews. 2016;17(6):541-51. doi: 10.1111/obr.12396. PubMed PMID: PMC4999041.

15. Krukowski RA, Harvey-Berino J, Ashikaga T, Thomas CS, Micco N. Internet-Based Weight Control: The Relationship Between Web Features and Weight Loss. Telemedicine Journal and e-Health. 2008;14(8):775-82. doi: 10.1089/tmj.2007.0132. PubMed PMID: 18954247; PubMed Central PMCID: PMCPMC2998280.

16. Watson S, Woodside JV, Ware LJ, Hunter SJ, McGrath A, Cardwell CR, et al. Effect of a Web-Based Behavior Change Program on Weight Loss and Cardiovascular Risk Factors in Overweight and Obese Adults at High Risk of Developing Cardiovascular Disease: Randomized Controlled Trial. Journal of Medical Internet Research. 2015;17(7):e177. doi: 10.2196/jmir.3828. PubMed PMID: 26183659; PubMed Central PMCID: PMCPMC4527000.

17. Bennett GG, Herring SJ, Puleo E, Stein EK, Emmons KM, Gillman MW. Web-based Weight Loss in Primary Care: A Randomized Controlled Trial. Obesity. 2010;18(2):308-13. doi: 10.1038/oby.2009.242. PubMed PMID: 19696764; PubMed Central PMCID: PMCPMC3710670.

18. Crane MM, Tate DF, Finkelstein EA, Linnan LA. Motivation for participating in a weight loss program and financial incentives: an analysis from a randomized trial. Journal of obesity. 2012;2012:290589. Epub 2012/05/12. doi: 10.1155/2012/290589. PubMed PMID: 22577524; PubMed Central PMCID: PMCPmc3345232.

19. Linnan L, Tate DF, Harrington CB, Brooks-Russell A, Finkelstein E, Bangdiwala S, et al. Organizational- and employee-level recruitment into a worksite-based weight loss study. Clinical Trials. 2012;9(2):215-25. doi: 10.1177/1740774511432554. PubMed PMID: 22273588.

20. Tate DF, Jackvony EH, Wing RR. Effects of Internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. JAMA. 2003;289(14):1833-6. Epub 2003/04/10. doi: 10.1001/jama.289.14.1833. PubMed PMID: 12684363.

21. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol. 2004;159(7):702-6. Epub 2004/03/23. PubMed PMID: 15033648.

22. BARNARD J, DU J, HILL JL, RUBIN DB. A Broader Template for Analyzing Broken Randomized Experiments. Sociological Methods & Research. 1998;27(2):285-317. doi: 10.1177/0049124198027002005.

23. Sheiner LB, Rubin DB. Intention-to-treat analysis and the goals of clinical trials. Clin Pharmacol Ther. 1995;57(1):6-15. Epub 1995/01/01. doi: 10.1016/0009-9236(95)90260-0. PubMed PMID: 7828382.

24. Newhouse JP, McClellan M. Econometrics in outcomes research: the use of instrumental variables. Annual Review of Public Health. 1998;19(1):17-34. doi: 10.1146/annurev.publhealth.19.1.17. PubMed PMID: 9611610.

25. Angrist JD, Imbens GW, Rubin DB. Identification of Causal Effects Using Instrumental Variables. Journal of the American Statistical Association. 1996;91(434):444-55. doi: 10.2307/2291629.

26. Sussman JB, Hayward RA. An IV for the RCT: using instrumental variables to adjust for treatment contamination in randomised controlled trials. BMJ. 2010;340:c2073. doi: 10.1136/bmj.c2073. PubMed PMID: 20442226; PubMed Central PMCID: PMCPMC3230230.

27. Jo B, Stuart EA. On the use of propensity scores in principal causal effect estimation. Stat Med. 2009;28(23):2857-75. Epub 2009/07/18. doi: 10.1002/sim.3669. PubMed PMID: 19610131; PubMed Central PMCID: PMCPMC2757143.

28. Dunn G. Complier-Average Causal Effect (CACE) Estimation. In: Lovric M, editor. International Encyclopedia of Statistical Science. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011. p. 273-4.

29. Frangakis CE, Rubin DB. Principal stratification in causal inference. Biometrics. 2002;58(1):21-9. Epub 2002/03/14. PubMed PMID: 11890317; PubMed Central PMCID: PMCPMC4137767.

30. Stuart EA, Jo B. Assessing the sensitivity of methods for estimating principal causal effects. Statistical Methods in Medical Research. 2015;24(6):657-74. doi: 10.1177/0962280211421840.

31. Joffe MM, Ten Have TR, Brensinger C. The compliance score as a regressor in randomized trials. Biostatistics. 2003;4(3):327-40. Epub 2003/08/20. doi: 10.1093/biostatistics/4.3.327. PubMed PMID: 12925501.

32. Schomaker M, Heumann C. Bootstrap Inference when Using Multiple Imputation. arXiv preprint arXiv:160207933. 2016.

33. Wing RR, Tate DF, Gorin AA, Raynor HA, Fava JL. A Self-Regulation Program for Maintenance of Weight Loss. New England Journal of Medicine. 2006;355(15):1563-71. doi: doi:10.1056/NEJMoa061883. PubMed PMID: 17035649.

CHAPTER 4: HAS MOBILE PHONE OWNERSHIP IMPROVED HEALTH KNOWLEDGE AND BEHAVIOR? EVIDENCE FROM A LONGITUDINAL STUDY IN UGANDA

4.1 Overview

Mobile phone ownership in developing countries increased exponentially in the past two decades. Since the introduction of mobile phones, health practitioners and researchers have heralded them as a new tool to improve the health and welfare of the world's poorest. In this study we used three waves of a panel survey from Uganda to determine the impact of mobile phones on knowledge of family planning (FP) methods and use of FP. We exploited the panel dataset to estimate linear fixed effects models. The two primary outcomes of interest were the number of FP methods women were aware of and usage of FP. Mobile phone ownership rose from 59% to 72% over the three study years (2009 to 2011). We find that women who belonged to a household that obtained a mobile phone over the study period learned 0.192 [95% CI=0.015 to 0.368, *P*-value=0.033] more FP methods than women who did not. This was equivalent to a 3.1 percent increase from the sample mean number of methods known. However, obtaining a mobile phone did not make women more likely to use FP [Coef=-0.013, 95% CI=-0.052 to 0.027, *P*-value=0.528]. The impact of mobile phone ownership on knowledge of FP was more pronounced among rural women, who experienced a statistically significant impact [Coef=0.275, 95% CI=0.079 to 0.471, *P*-value=0.006], whereas urban women were not significantly affected [Coef=-0.165, 95% CI=-0.588 to 0.259, *P*-value=0.445]. Although the impact on FP knowledge was significant, the impact was small, suggesting that future mobile health campaigns need to be paired with specific FP interventions, such as sending text messages about FP distribution at health days, in order to generate large changes in FP knowledge and usage. It appears that alternative approaches to increasing awareness and use of FP methods may be needed.

4.2 Introduction

Over the past two decades, people’s ability to communicate increased rapidly due to exponential increases in affordability, access, and usage of mobile phone technology. Access increased due to two key factors: plummeting costs of cell phones and widespread rollout of cellular network coverage. In the vast majority of households in Africa, mobile phones effectively leapfrogged landline-based phone systems and are now widely used. From 2002 to 2014, mobile phone ownership among individuals in Uganda rose from 10% to 65% [1]. While texting and calling are the most common activities for cell phones, a survey in Uganda found that 20% of respondents reported using their mobile phone to access health information [2].

This expansion has led to growing enthusiasm about the prospects for expanded phone access to deliver information to the poor in developing countries. The potential for large impacts is especially true for women, who continue to face significant inequalities in access to information in the developing world. The widespread uptake of mobile phones could have a direct impact on women’s health through three main channels: 1) consumption of health related information, 2) increased communication, and 3) improvement of the health system through the supply side.

First, mobile phones can increase people’s health knowledge through direct consumption of health information. Recent surveys from developing countries show that people use their mobile phones to access health information in some manner [2]. People can use their phones to receive health related messages or call specific hotlines set up to provide information.

Additionally, people may receive health related information inadvertently, such as through governments sending text messages to all mobile subscribers nationwide regarding upcoming national health days [3]. Second, decreasing communication barriers between peers and health workers from different areas could potentially decrease the search cost of finding new health information. Increased communication using mobile phones can lead to people learning from their peers and social network. Phones also allow for easier access to health professionals, such as community health workers being able to send informative text messages to women regarding when family health days are held, during which women could obtain FP. Third, phones may impact health seeking through the supply side. They can make health systems more efficient, such as improving health workers' ability to manage pharmaceutical stockouts [4], thus potentially increasing people's perceptions of the health system and subsequently improve health seeking behavior.

The explosion of mobile phone usage in developing countries has driven health professionals to embrace them as a platform for interventions. Researchers and practitioners in developing countries historically focused on creating formal mHealth programs to influence consumption of health information and the supply of health care. For example, various programs have been tested that target health knowledge, such as text messages for promoting medication adherence [5-10]. The programs vary in sophistication—from direct messages informing people about national health days to complex interactive messages built to educate people on health issues. Additionally, some studies examined the potential for mobiles phones to impact the supply side of health care. These programs have a wide range of goals, from providing a mobile messaging platform to rural health workers so that they can ask questions and receive timely feedback to questions and concerns for patient safely [11] to sending care guideline reminders to ensure that patients seeking services receive high quality care [12].

Despite widespread interest in expanding mHealth programs, they are still limited in coverage and scope although numerous studies have been published on mHealth, evidence of their effectiveness is lacking [13-15]. One significant drawback of the existing studies on mobile phones and health—which may potentially explain their lack of effectiveness–is that they generally only examine a specific mHealth program and do not take into account the larger, more systemic effects that owning a mobile phone may have. In this study we estimate the total effect of mobile phones on family planning (FP) through a reduced form approach. This approach allows us examines the aggregate impact of mobile phones on health indicators, including those that may be affected by both supply and demand sides.

The reduced form approach also allows us to capture the effect of more “informal” means, such as people taking initiative and seeking out health information among family, friends and health workers [16]. These informal ways of utilizing mobile phones for health care are potentially more important than formal mHealth programs since change associated with new technologies is often driven by individual consumers. A study in Uganda found that a majority of women reported friends as a trusted source of FP related information [17], thus suggesting that increased communication between peers may increase knowledge of FP. To date, less attention has been paid to the role of peer communication as it is not straightforward how a mHealth practitioner would intervene; however, peer communication has the potential to have a substantial effect on health knowledge and should not be overlooked.

Although mobile phones are a broad intervention that likely impacts numerous aspects of health, we limited our focus to a persistently difficult problem in Uganda, FP. Over the past decade, substantial progress has been made in increasing knowledge of FP, with almost all women of reproductive age knowing at least one modern method, but the level of unmet need remains persistently high (34% of married women not wishing to become pregnant are not using FP), showing no trend in decreasing over time [18]. Few studies examine mHealth and FP in Uganda. Previous studies examining the role of mHealth and FP found mHealth programs for FP are beneficial, although the studies were predominantly conducted in the United States [19]. One recent study on a text messaging based program for FP in Kenya found improvements in knowledge of FP but no significant impact on use [20]. A pilot study in Tanzania showed that the general public accessed a text messaging-based service for FP information, but evidence of knowledge and use was not captured [21]. Existing mHealth and FP studies have focused on specific, targeted programs and their impact on FP, rather than examining the aggregate impacts of mobile phone ownership.

In this study we estimate the impact of mobile phone ownership on women's knowledge of FP methods and FP use. In practice, estimation requires careful attention to issues of endogeneity due to selection; women who own phones may be different than those who do not. We account for potential selection using a robust fixed effects model to control for unobserved confounding factors. In secondary analyses we explore the impact of mobile phone ownership among different subgroups. Additionally, we examine difference exposure definitions, including the number of phones owned by a household and the amount spent on usage of mobile phones.

4.3 Background

In addition to this study's direct relevance to mHealth in developing countries, this study is also related to three other literatures: 1) information dissemination, 2) the social determinants of health, and 3) the cross-sector impacts of development. First, this research is tied to the limited literature on how mobile phones change the way obtain information. Three studies showed that introducing cellular network coverage in an area changed how businesses obtained information on prices [22-24]. Businesses from these studies were able to use their phones to get information so that prices were more uniform across markets, reducing asymmetries. This information led to welfare improvements for both businesses and consumers. Women are likely to have similar gains for health knowledge. For example, community health workers would be able to send informative text messages to women regarding when family health days are held, during which women could obtain FP; or women would be able to call health workers who live far away to get information about FP methods and their availability.

This study adds to the large literature on population health improvement due to investment in other sectors. Investment in the social determinants of health can have a profound impact on the health of individuals and families [25, 26]. For example, investment in women's education reduces child mortality [27, 28], and studies show the important role of poverty in influencing the health of people [29, 30]. Despite a growing literature on the social determinants of health, little attention has been paid to the impact of advancements in the IT industry on health. Innovations in IT can help reduce inequalities in health by reducing disparities related to health information. For example, dropping mobile phone prices may enable women from rural areas who previously did not have access to FP information to afford mobile phones and subsequently access high quality FP information using the phones.

Finally, this study pertains to research on the role of technology in low-income countries and its impact on development broadly. Several studies have identified the effect of technology on non-health sectors, such as rural electrification on labor outcomes [31] and television on women's status [32], but we are not aware of any studies that have looked at health outcomes. In this study we add to this literature by examining how a new technology impacts a key aspect of development, health.

4.4 Methods

4.4.1 Study setting and data

We used data from three waves of the nationally representative Uganda National Panel Survey (UNPS) which was administered in 2009/2010, 2010/2011, and 2011/2012. The UNPS survey is part of the World Bank's Living Standards Measurement Survey, Integrated Survey on Agriculture (LSMS-ISA) series. The survey was designed to capture detailed information on demographics, socio-economic status, health, education and agriculture, among other topics.

The primary outcomes were the number of FP methods known and FP usage. In the UNPS survey, women aged 15-49 are asked to list all FP methods that they are aware of. After women listed FP methods they were aware of without prompting, surveyors asked if they knew of each method not previously listed. In total there were 8 possible modern methods, including: female sterilization, male sterilization, pill, intrauterine device, injectable, implant, condom, or female condom. We generated a knowledge score that ranged from 0 to 8 which was the sum of the total methods known. Participants who knew all methods were assigned a knowledge score of 8. Women were defined as using FP if they were using at least one method at the time of survey.

One unique advantage of using the UNPS over other longitudinal datasets is the highly detailed consumption and asset information that was collected as part of the survey. The survey was specifically designed to capture accurate measures of economic welfare. Measures of consumption and assets are necessary because they are likely correlated with both mobile phone ownership and health outcomes. This correlation could lead to confounding and therefore it is imperative that any analyses account for it. The consumption variable in the UNPS is an average household monthly consumption value equal to monetary expenditures, as well as consumption through barter and kind, all smoothed over time. Extreme consumption outliers (>US$1,600 per month) were recoded as missing (<0.01%). In addition to expenditure we included information on total assets, which can be a better measure for overall wealth in low-resource settings. For each asset we assigned a code of 1 if a household owned it and 0 if it did not. We generated a total asset score by summing up all assets owned by each household. The list of assets can be found in Table 18.

Mobile phone ownership was included as a binary variable indicating whether a household owned at least one phone at the time of survey. The UNPS does not differentiate between smartphone ownership and basic (talk and text only) phone ownership. Smartphone ownership was very low during the survey years [2] and therefore all mobile phones recorded in the dataset were likely basic call or text only phones.

4.4.2 Primary empirical analysis

We estimated the effect of mobile phone ownership on each outcome using a person level fixed effects ordinary least squares regression model shown in Equation (1) below:

$$FP_{ikt} = \alpha_i + \delta MP_{kt} + X'_{ikt}\beta + \gamma_t + \varepsilon_{ikt} \tag{1}$$

where *FP* was one of two outcome variables for individual *i* in household *k* in time period *t*, including: 1) number of FP methods known, and 2) use of FP. The binary variable *MP* indicated whether a woman lived in a household with a mobile phone or not. The vector *X* included time-varying individual and household characteristics that varied over individual *i* or household *k* and time *t*; including age, rural status, education, total consumption, and total assets. α_i were individual level fixed effects and γ_t were fixed effects for each survey year. We included time-varying income and consumption variables to control for time-varying income shocks. We included interaction terms between age, consumption and assets and the year fixed effects to account for differential trends over time. Errors were clustered at the household level to account for multiple women living in the same household and thus having the same mobile phone ownership.

The coefficient of interest, δ, captured the fixed effects estimate of the effect of mobile phones on FP outcomes. Under the assumption of no confounding time-varying omitted variables, either from the survey participant or phone companies, the estimated effect, δ, can be viewed as the causal estimate. However, this is a strong assumption. Time-varying unobserved factors may exist that confound the relationship between mobile phones and FP outcomes; such factors would cause the estimate of δ to be biased. We explored these assumptions in robustness analyses below.

4.4.3 Secondary empirical analyses

Heterogeneity

In order to understand how the impact of mobile phones differed between urban and rural populations, we estimated the above equation on each subgroup (urban/rural, education categories) separately. We hypothesized that the impact of mobile phones would be more pronounced among rural women, since their ability to access information regarding FP was likely more constrained than women residing in urban settings. We also examined the impact of mobile phones among women of different education levels: none, some primary, some secondary and post-secondary. A priori we did not have a strong hypothesis regarding education and phone ownership. On one hand, women with less education may be able to use phones to learn more information due to their relatively low starting point compared to more educated women. However, conversely, poorly educated women may not know how to use technology and therefore it would not benefit them in gaining new knowledge.

Different exposure measurements

Mobile phones users in Uganda generally buy airtime (credit) from vendors that can be used for calling, texting, or data. We used the amount spent on airtime as a proxy for household mobile phone usage. We converted the amount spent in Ugandan Shilling to US dollars and censored outliers above the 99th percentile to the 99th percentile. The expenditure on airtime measure is able to capture more heterogeneity in phone usage than the binary phone ownership measure. For example, households could have owned a phone but not used it, or households may not have owned a phone, but instead bought airtime to use with phones they share with relatives or neighbors—a practice which has been frequently reported in previous studies [16]. We hypothesized that greater household airtime expenditure would lead to more FP methods known and higher probability of FP usage.

Second, we examined the number of phones the household owned. We used a categorical variable indicating whether the household owned no phones, 1 phone, or ≥2 phones. In many households in Uganda, men own and use the mobile phone and women are not allowed to use the phone. The disparity in access to mobile phones has been widely noted [16, 33]. Thus the number of phones a household owned may provide another indication of exposure to mobile phones for female residents that accounts for men monopolizing use of a single household phone.

We were also interested in understanding if there was a community effect associated with mobile phone ownership. That is to say, did communities with higher proportions of mobile phone ownership have higher rates of FP knowledge and usage. To test this hypothesis, we added a variable for the percent of households in the community that owned a mobile phone during a given survey year. We calculated this variable as the percent of sampled households in each primary sampling unit that owned a mobile phone. Inherently we are relying on the sampling of a cluster to be a good representation of local neighborhood characteristics. Including the proportion who owned phones also controls for spillovers of mobile phone ownership, such as a household sharing their mobile phone with neighbors.

Additional outcomes

In addition to FP outcomes, mobile phone ownership may impact other health behaviors. We examined a number of secondary outcomes including: 1) birth in a health facility, 2) birth attendance by medical personnel, 3) child received 3 doses of diphtheria, pertussis and tetanus (DPT) vaccine, 4) child received one dose of measles vaccine, 5) child ever received vitamin A supplements. All secondary outcomes were binary. Despite the theoretical possibility that phone ownership could impact these outcomes, it remains difficult to test with this dataset. The limitation with analyzing these outcomes with this short three-year panel is that there were very few individuals for which we observe outcomes at least twice. Therefore, the fixed effect linear model was not appropriate given its reliance on within variation. Given this limitation, we decided to use a linear random effect models to analyze the association between mobile phone ownership and these outcomes of interest.

Robustness analyses

The most significant concern of the fixed-effects analysis was that a time-varying unobserved variable might have driven both changes in FP outcomes and mobile phone ownership. For example, attitudes towards modernity were not observed and could have led to selection bias—people who were more open to modern things may have been more likely to own a mobile phone and may have had higher health knowledge of FP. We tested for selection effects by estimating a linear probability model (Equation (2)) in which knowledge of FP was regressed on a binary indicator for whether or not a survey participant obtained a mobile phone during the subsequent survey round.

$$(2) \qquad KOFP_{ikt} = \alpha_i + \lambda GAIN\_MP\_NEXT\_WAVE_{ikt} + \delta MP_{kt} + X'_{ikt}\beta + \varepsilon_{ikt}$$

We estimated this equation to determine whether FP knowledge and usage were predictable from future mobile phone ownership. We included the same set of controls as included in Equation (1). If there were unobserved factors influencing both FP outcomes and mobile phone ownership, we may see changes in FP outcomes anticipating changes in mobile phone ownership because the unobserved variable would be driving both. Thus, we were interested in the significance of λ, insignificance would provide suggestive evidence that unobserved variables were not driving simultaneous changes over time.

Given that a proportion of the sample was not observed all three years, we were concerned that our results may be biased due to compositional effects of the dataset, such as mobile phone owners being more likely to be followed-up. To test if results were robust to compositional effects we estimated Equation (1) but only included observations for which we observed them in the dataset all three years. All results from robustness checks are included in the supplementary appendix. All analyses were conducted using Stata 14.1.

4.5 Results

4.5.1 Results from descriptive analysis

Household mobile phone ownership rose from 59% among women in the 2009/2010 UNPS survey to 72% in the 2011/2012 survey (Table 10). Characteristics of women were similar over time. Approximately 75% of the sample resided in rural areas and majority had some primary schooling. The mean age of women in the sample was 29.5 and mean household consumption per month was US$77.50. In pooled analysis including all three waves of the data, pronounced differences occurred in participant characteristics between those who did and did not own mobile phones. Rural residents were much less likely to own a phone than urban. Women that lived in a household with a phone had nearly two times the median consumption of women that did not own (US$51 vs. US$100, respectively). More educated women were more likely to own phones. FP usage was almost twice as likely among women who resided in households that owned a phone (27% vs. 15%) and they also knew more contraceptive methods (6.3 methods vs. 5.5 methods).

4.5.2 Results from primary analysis

The results from estimating Equation (1) are presented in Table (12). We find that women who belonged to a household that obtained a mobile phone over the study period learned 0.192 [95% CI=0.015 to 0.368, *P*-value=0.033] more FP methods than women who did not. The estimated effect is equivalent to a 3.1 percent increase in knowledge of FP methods from the overall mean of FP knowledge (6.1 methods known). Although women in the sample did exhibit greater knowledge, we found that obtaining a mobile phone did not lead to greater use of FP [Coef=-0.013, 95% CI=-0.052 to 0.027, *P*-value=0.528].

4.5.3 Results from secondary analyses

Heterogeneity

We limited our subgroup analyses to the knowledge of FP methods outcome. The effect of mobile phone ownership on knowledge was more pronounced in rural areas than urban. In rural areas obtaining a mobile phone led to 0.275 [95% CI=0.079 to 0.471, *P*-value=0.006] more FP methods known (Table 13). The estimated effect is equivalent to a 4.6 percent increase in knowledge of FP methods. In urban areas the estimated effect was negative, but the result was insignificant [Coef=-0.165, 95% CI=-0.588 to 0.259, *P*-value=0.445]. In an interacted model the difference of the impact of mobile phones between urban and rural areas was significant [*P*-value=0.028]. The impact of mobile phones did not differ by education—among all education subgroups, none, primary, secondary, or greater than secondary, the estimated effects were similar. Mobile phone ownership did not lead to significant increases in knowledge of FP methods for any of the education subgroups.

Different exposure measures

Expenditure on airtime did not lead to more FP methods known or used (Table 14). For each additional US$1 households spent on airtime, women knew 0.005 more FP methods, suggesting that the impact, if any, was extremely small.

The number of mobile phones owned did not seem to impact knowledge or use significantly (Table 15). Whether a household owned 1 or 2+ mobile phones led to similar gains in knowledge of modern methods [0.188 and 0.182 for 1 and 2+ respectively]. The number of phones a household owned did not significantly affect FP use, similar to the primary analysis of binary mobile phone ownership.

The percent of households in the community that owned a mobile phone had a small positive impact on the number of FP methods known, however it was insignificant (Table 16). Including the percent ownership variable slightly decreased the statistical significance of the coefficient for household mobile phone ownership on knowledge of FP. However, inclusion did not significantly decrease the coefficient estimate, suggesting that some of the household mobile phone ownership benefit on FP knowledge is explained by community ownership. One explanation for this small effect is that members of a community may share phones and can therefore learn new information through use of neighbor's phones; another is that women who learn about new methods using their phones diffuse this new information in the community to those who do not own phones. The percent of households owning a mobile phone was statistically significantly associated with reduced FP usage, however the coefficient suggested that a 10 percent increase in households owning a mobile phone in the community would decrease the probability of a woman using family planning by 0.01. Thus, the magnitude of this effect was so small that the result has little policy significance.

Different outcome measures

Mobile phone ownership was significantly associated with probability of delivery in a health facility and skilled birth attendance [Coef=0.111 and 0.108, respectively] using a linear random effects model (Table 17). The results were equivalent to an 18.4 percent increase in probability from the mean of in-facility delivery and 17.6 percent increase in probability from the mean of skilled birth attendance. Ownership of mobile phones was not significantly associated with child vaccination or vitamin A supplementation.

Robustness checks

Future phone ownership—i.e. obtaining a phone in the next survey wave—was not significantly predictive of FP knowledge or usage (Table 19). This provides suggestive evidence that there were no unobserved variables simultaneously driving change in both FP outcomes and mobile phone ownership. This result increases our confidence that the fixed effects estimates are unbiased. Despite this result, we cannot completely rule out the possibility that there was some important unobserved variable that impacted ownership of mobile phones and FP knowledge/usage simultaneously.

We found similar effects for mobile phone ownership on knowledge of FP methods and FP usage using the balanced panel of only those observed in all three surveys. This suggests that the compositional effects from people entering and leaving the panel did not influence the main results (Table 20).

4.6 Discussion

Mobile phone ownership did not have a large impact on women's knowledge of FP methods. Even where results were statistically significant, the magnitude of the effect was modest. We found that mobile phone ownership led to women knowing 3.1 percent more FP methods. This finding does not support the hypothesis that extending mobile phone access will significantly improve knowledge of FP methods. Similarly, we found no relationship between phone ownership and usage, suggesting that even though ownership may increase knowledge of FP methods, a woman's decision on whether or not to use FP is complex and cannot be explained by household mobile phone ownership.

Mobile phones had a more pronounced impact on knowledge of FP methods for rural women than urban women. This result supports initial hypotheses that rural women had less access to information prior to mobile phone ownership and obtaining a mobile phone helped increase access to information regarding FP. Accessing information in rural areas can be costly, either due to long distances or poor services. Mobile phones can essentially reduce the cost of accessing information for rural areas by making communication to previously remote areas much easier [11].

We found that mobile phone ownership was unrelated to FP usage. One potential explanation for this finding is that the availability of FP in Uganda is generally inadequate, especially in rural areas [34]. Women may have knowledge of FP, but they are unable to obtain safe and effective FP in health facilities. Thus, supply side issues present a bottleneck for usage and must be addressed in order for mobile phones to make a large impact. If supply improves, mobile phones may impact usage in a positive way, but given the current environment, it is difficult to determine if the null finding represents a lack of an effect for phone ownership, or a supply side bottleneck.

One finding that encourages further research is the impact of mobile phone ownership on in-facility delivery and skilled birth attendance. Although we were unable to run the individual level fixed effects model with these outcomes due to their rarity, the results from the random effects model suggested that mobile phone ownership is significantly associated with these outcomes. Further, the estimated effects were quite large, 18.4 and 17.6 percent for in-facility and skilled birth attendance, respectively. This strong relationship is not surprising given that phones could increase the probability of communication between health workers and expecting mothers, thus making it more likely for women to set up the necessary steps for in-facility childbirth. Clinic appointment reminders in developed countries have shown similar findings [35] but robust evidence regarding mobile phone ownership and maternal health is lacking [36]. Further, we found that child vaccination was unrelated to mobile phone ownership, perhaps due to the large supplementary immunization activities that now take place as a part of child health days. These health days bring immunization services directly to villages and therefore could reduce any impact mobile phones may have on immunization.

In this study we provide the first evidence of the aggregate impact of mobile phone ownership on FP outcomes using a large national dataset from a developing country where phone ownership has risen dramatically. Previous studies have shown positive impact of mobile phones on various economic development indicators [37], but none have examined ownership and FP, specifically. Similarly, we are not aware of any other studies that have examined mobile phone ownership and its impact on maternal and child health indicators using a large nationally representative sample. This analysis is an important contribution to the literature as researchers and policy makers attempt to understand how mobile phones impact health.

This study has limitations. First, mobile phone ownership is measured at the household level and not the individual level. Although we assumed that women in a household have access to a mobile phone if there is one within the household, this may not necessarily be true. During this period in Uganda mobile phones were widespread but still not quite ubiquitous and in many cases a household would have one phone for all residents. Given that we may be incorrectly assuming that women residing in households that own a phone are able to use the phone—and do use—we have measurement error in our treatment variable. This measurement error could bias our estimated effect. However, given that women are likely receiving less treatment than we assume (assume everyone has access and uses while in reality that may not be true) our estimated coefficient is likely biased downward and may represent a lower bound of the true effect.

Second, as in any observational study, selection into mobile phone ownership may be due to unobserved factors that were also correlated with FP. Unobserved factors could lead to biased coefficient estimates for mobile phone ownership. Selection could be occurring from either the demand, individual side, or the supply, phone company side. While we attempted to reduce the possibility of time-varying confounders by including a rich set of control variables, we acknowledge the possibility of other factors, such as attitudes towards modernity, that may affect estimates. On the supply side, selection may be determined by the mobile phone companies and where they choose to provide service. Mobile network providers may target areas for network expansion that are associated with positive changes in FP knowledge over time. Although this is important to consider, for the period we analyzed it is less relevant because mobile phones were already popular and coverage was high. In the UNPS survey there were no surveyed villages that lacked mobile network coverage, even in 2009. Thus, selection by mobile phone companies is unlikely to impact estimates.

Third, we only had access to a three-year panel of consecutive years. While this short panel works well for FP outcomes, which are measured each year, it makes identification for outcomes that are rare, such as births, very difficult. The panel limits our ability to detect causal effects of mobile phones on maternal and child health outcomes, such as in-facility delivery and skilled birth attendance. Future research should make use of longer panels to explore birth outcomes associated with mobile phone ownership.

This study relied on a large, nationally representative dataset to explore the relationship between mobile phones and FP. The use of this dataset makes it hard to unpack the specific mechanisms through which mobile phone ownership may impact FP, whether it be through peer communication, direct contact with providers, or receiving text messages for health days. More focused studies, such as randomized controlled trials, are needed to understand the main mechanisms at work. Although the study design inhibits our ability to identify the most salient factors associated with change, it does provide greater generalizability than a small study examining one mechanism would.

Mobile phone ownership in Uganda has flourished in recent years, yet 35% of the population still does not own a mobile phone [2]. Uganda is fairly representative of other developing countries in sub-Saharan Africa (SSA) in terms of uptake of mobile phones. The results from this observational study using a nationally representative dataset from Uganda are likely generalizable to other settings given Uganda’s similarities with other countries in terms of mobile infrastructure development. The results from this study suggest that increasing mobile phone ownership, without including strong evidence based mHealth programs in tandem, would not lead to better FP outcomes. However, suggestive evidence from the in-facility delivery and skilled birth attendance analyses point to the potential for mobile phones to dramatically alter uptake of health services. More research is needed on mobile phone ownership and other health indicators, especially with regards to specific mHealth programs.

Ownership of smartphones, capable of more advanced health interventions, such as running mHealth applications, is very low in Uganda and their effect was not addressed in this analysis—only 5% of people owned one in 2014 [2]. Access to smartphones is bound to increase in the coming years as they become cheaper and more accessible. Smartphones may expedite learning, especially in rural areas that did not previously have easy access to public health information. As governments and the international health community plan for the future of health development in SSA and other poor regions of the world, they should pay close attention to findings from smartphone trials in developed countries to help guide their interventions for at-scale adoption.

Although we did not find large effects on FP, mobile phones have the potential for impacts on many development sectors. Mobile phones have led to the widespread adoption of mobile banking in many countries, which has had positive impacts on savings and social protection, such as the ability to better able to cope with catastrophic expenditures associated with health shocks [38]. Additionally, it has changed the way that farmers and traders get price information, leading to welfare improvements [23]. More research is needed to understand the diverse impact of mobile phones across different aspects of development.

4.7 Conclusions

Widespread adoption of mobile phones without expanding targeted programs for health using mobile phones is unlikely to significantly impact FP. Health researchers need to develop mHealth interventions for FP and implement them at scale. However, even in the absence of such programs, small benefits to FP and other health outcomes due to mobile phones have the potential for large population health impacts, given how ubiquitous mobile phone ownership has become. More research about how women acquire FP information and make FP usage decisions in an increasingly changing technological environment is needed for mHealth interventions to be truly effective and, subsequently, population health to be significantly impacted.

TABLE 10: Descriptive statistics of study sample during each wave

	2009	2010	2011	Total
Total number of women interviewed	2476	1911	2172	6559
Mobile phone ownership	1459 (59)	1240 (65)	1573 (72)	4272 (65)
Reside in rural area	1848 (75)	1479 (77)	1678 (77)	5005 (76)
Highest education level				
No education	347 (14)	204 (11)	272 (13)	823 (13)
Some primary	1453 (59)	1171 (61)	1313 (60)	3937 (60)
Some secondary	590 (24)	449 (23)	488 (22)	1527 (23)
Post secondary	86 (3)	87 (5)	99 (5)	272 (4)
Age in years, mean (SD)	29.1 (9.8)	29.6 (9.8)	30 (9.6)	29.5 (9.7)
Asset count, mean (SD)	6.1 (2.3)	5.9 (2.2)	5.7 (2.1)	5.9 (2.2)
Total consumption past month (US$), median (IQR)	89.6 (56.1 - 160.8)	70 (42.8 - 124.5)	70.8 (43.5 - 121.9)	77.5 (47.2 - 135.9)
Total amount spent on airtime past month (US$), median (IQR)	0.8 (0 - 4)	1.6 (0 - 4)	1.6 (0 - 4.8)	1.6 (0 - 4)
Using FP	529 (21)	449 (23)	493 (23)	1471 (22)
FP methods known, mean (SD)	6 (1.8)	6.1 (1.7)	6.1 (1.9)	6.1 (1.8)

Abbreviations: SD, standard deviation; IQR, interquartile range; FP, family planning

Notes: For all variables frequencies are presented with percentages in parentheses except where otherwise noted

Results are not weighted so they represent sample rather than population characteristics

TABLE 11: Participant characteristics by mobile phone ownership for full sample

	Do not own a mobile phone	Own a mobile phone	Total	P-Value*
Total number of women interviewed	2411	4927	7338	
Mobile phone ownership	0 (0)	4927 (100)	4927 (67)	
Reside in rural area	2195 (91)	3357 (68)	5552 (76)	<0.001
Highest education level				
No education	522 (22)	386 (8)	908 (12)	<0.001
Some primary	1659 (69)	2714 (55)	4373 (60)	<0.001
Some secondary	223 (9)	1504 (31)	1727 (24)	<0.001
Post secondary	7 (0)	323 (7)	330 (4)	<0.001
Age in years, mean (SD)	30.5 (10)	29.4 (9.3)	29.7 (9.6)	0.823
Asset count, mean (SD)	4.3 (1.5)	6.6 (2.1)	5.8 (2.2)	<0.001
Total consumption past month (US$), median (IQR)	50.9 (34.1 - 76.1)	97.8 (62.5 - 168.6)	78.2 (48.1 - 133.8)	<0.001
Total amount spent on airtime past month (US$), median (IQR)	0 (0 - 0)	2 (0.8 - 6)	0.8 (0 - 4)	<0.001
Using FP	460 (19)	1790 (36)	2250 (31)	<0.001
FP methods known, mean (SD)	5.3 (2.1)	5.6 (2.4)	5.5 (2.3)	<0.001

Abbreviations: SD, standard deviation; IQR, interquartile range; FP, family planning

Notes: For all variables frequencies are presented with percentages in parentheses except where otherwise noted. Results are not weighted so they represent sample rather than population characteristics. *-P-value from Chi2 tests for dichotomous variables and Mann Whitney tests for continuous variables.

TABLE 12: Impact of mobile phones on knowledge and use of modern family planning methods

Outcome	Improved FP methods known	Using improved FP
Household owns a mobile phone	0.192	-0.013
95% CI	[0.015 to 0.368]	[-0.052 to 0.027]
P-value	0.033	0.528
Mean of dependent variable	6.06	0.22
Observations	6,559	6,739
R-squared	0.025	0.010
Number of id	3,722	3,795

Abbreviations: FP, family planning

Notes: Results are from linear fixed effects regression with standard errors clustered at the household level.

TABLE 13: Impact of mobile phones on knowledge of modern family planning methods in specific subgroups

Subgroup	Rural	Urban	No primary education	Primary education	Secondary education	> Secondary education
Household owns a mobile phone	0.275	-0.165	0.163	0.180	0.154	0.143
95% CI	[0.079 to 0.471]	[-0.588 to 0.259]	[-0.337 to 0.664]	[-0.032 to 0.392]	[-0.286 to 0.595]	[-0.621 to 0.906]
P-value	0.006	0.445	0.522	0.097	0.491	0.713
Mean of dependent variable	5.93	6.48	5.41	5.92	6.54	7.36
Observations	5,005	1,554	823	3,937	1,527	272
R-squared	0.022	0.080	0.032	0.034	0.042	0.128
Number of women	2,807	979	530	2,291	1,024	192

Abbreviations: FP, family planning

Notes: Results are from linear fixed effects regression of number of modern family planning methods known on each subgroup. All regressions include standard errors clustered at the household level.

TABLE 14: Amount spent on mobile telephone airtime as explanatory variable of interest

Dependent variable	FP methods known	Using FP
Total amount spent on airtime (US$)	0.005	0.0003
95% CI	[-0.006 to 0.017]	[-0.003 to 0.003]
P-value	0.368	0.879
Mean of dependent variable	6.33	0.26
Observations	6,559	6,739
R-squared	0.024	0.01
Number of women	3,722	3,795

Abbreviations: FP, family planning; CI, confidence interval

Notes: Results are from linear fixed effects regression with standard errors clustered at the household level.

TABLE 15: Impact of the number of mobile phones a household owns on FP methods known and FP usage

Outcome	Improved FP methods known	Using improved FP
Household does not own a mobile phone	Reference	Reference
95% CI		
P-value		
Household owns 1 mobile phone	0.188	-0.0129
95% CI	[0.009 to 0.367]	[-0.053 to 0.027]
P-value	0.0397	0.523
Household owns 2+ mobile phones	0.182	-0.023
95% CI	[0.182 to -0.038]	[-0.023 to -0.081]
P-value	0.106	0.453
Mean of dependent variable	6.06	0.22
Observations	6,559	6,739
R-squared	0.025	0.010
Number of women	3,722	3,795

Abbreviations: FP, family planning

Notes: Results are from linear fixed effects regression with standard errors clustered at the household level.

TABLE 16: Impact of community mobile phone ownership on FP methods known and FP usage

Outcome	Improved FP methods known	Using improved FP
Household owns a mobile phone	0.166	0.001
95% CI	[-0.013 to 0.344]	[-0.041 to 0.043]
P-value	0.069	0.955
Percent of household communities that own a phone	0.002	-0.001
95% CI	[-0.003 to 0.007]	[-0.002 to 0]
P-value	0.377	0.049
Mean of dependent variable	6.06	0.22
Observations	6,559	6,739
R-squared	0.025	0.012
Number of women	3,722	3,795

Abbreviations: FP, family planning

Notes: Results are from linear fixed effects regression with standard errors clustered at the household level.

TABLE 17: Impact of mobile phone ownership on select maternal and child health outcomes using random effects linear regression

Outcome	In-facility delivery	Skilled birth attendance	Measles vaccine	DPT vaccine	Vitamin A
Household owns a mobile phone	0.111	0.108	0.041	0.010	0.043
95% CI	[0.067 to 0.154]	[0.064 to 0.152]	[-0.003 to 0.085]	[-0.036 to 0.055]	[-0.006 to 0.092]
P-value	<0.001	<0.001	0.0701	0.678	0.0825
Mean of dependent variable	0.60	0.61	0.65	0.73	0.67
Observations	2,508	2,469	2,326	2,309	2,256
Number of women	1,675	1,662	1,934	1,926	1,891

Abbreviations: DPT, diphtheria, pertussis and tetanus

Notes: Results are from linear random effects regression with standard errors clustered at the household level.

TABLE 18: List of assets in Uganda National Panel Survey

List of assets
Jewelry and Watches
Generators
Mobile phone
Television
Furniture/furnishings
Radio/Cassette
Bicycle
Land
Boat
Other buildings
Motor vehicle
Household appliances
Internet Access
Other household assets e.g. lawn mowers, etc.
Other electronic equipment
Other Transport equipment
Computer
Solar panel/electric inverters
Motor cycle
House

TABLE 19: Testing for preexisting trends: Examining if FP outcomes are predictable from future phone ownership

	FP methods known	Using FP
Obtained phone in next period	-0.110	0.004
95% CI	[-0.351 to 0.132]	[-0.057 to 0.064]
P-value	0.374	0.906
Household owns a mobile phone	0.124	-0.011
95% CI	[-0.095 to 0.344]	[-0.065 to 0.044]
P-value	0.266	0.707
Mean of dependent variable	6.06	0.22
Observations	6,559	6,739
R-squared	0.025	0.010
Number of id	3,722	3,795

Abbreviations: FP, family planning

Notes: Results are from linear fixed effects regression with robust standard errors.

TABLE 20: Testing for compositional effects: Primary model specification but limited to those observed all three waves of the survey

	FP methods known	Using FP
Household owns a mobile phone	0.273	-0.012
95% CI	[0.064 to 0.482]	[-0.062 to 0.038]
P-value	0.011	0.639
Mean of dependent variable	6.35	0.25
Observations	2,890	2,890
R-squared	0.027	0.015
Number of id	988	988

Abbreviations: FP, family planning
Notes: Results are from linear fixed effects regression with standard errors clustered at the household level.

REFERENCES

1. Pew Research Center. Emerging Nations Embrace Internet, Mobile Technology. 2014 February 2014. Report No.

2. Pew Research Center. Cell Phones in Africa: Communication Lifeline. 2015.

3. World Health Organization. mHealth: New horizons for health through mobile technologies. Geneva: World Health Organization,, 2011.

4. Barrington J, Wereko-Brobby O, Ward P, Mwafongo W, Kungulwe S. SMS for Life: a pilot project to improve anti-malarial drug supply management in rural Tanzania using standard technology. Malaria journal. 2010;9:298-. doi: 10.1186/1475-2875-9-298. PubMed PMID: PMC2978233.

5. Mbuagbaw L, Kop MLvd, Lester RT, Thirumurthy H, Pop-Eleches C, Ye C, et al. Mobile phone text messages for improving adherence to antiretroviral therapy (ART): an individual patient data meta-analysis of randomised trials. BMJ Open. 2013;3(12):e003950. doi: 10.1136/bmjopen-2013-003950.

6. Aranda-Jan CB, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC public health. 2014;14:188. doi: 10.1186/1471-2458-14-188.

7. Nglazi MD, Bekker LG, Wood R, Hussey GD, Wiysonge CS. Mobile phone text messaging for promoting adherence to anti-tuberculosis treatment: a systematic review. BMC Infect Dis. 2013;13(1):566. Epub 2013/12/04. doi: 10.1186/1471-2334-13-566. PubMed PMID: 24295439.

8. Obasola O, Mabawonku I, Lagunju I. A Review of e-Health Interventions for Maternal and Child Health in Sub-Sahara Africa. Matern Child Health J. 2015;19(8):1813-24. doi: 10.1007/s10995-015-1695-0.

9. Bloomfield GS, Vedanthan R, Vasudevan L, Kithei A, Were M, Velazquez EJ. Mobile health for non-communicable diseases in Sub-Saharan Africa: a systematic review of the literature and strategic framework for research. Global Health. 2014;10:49. Epub 2014/06/15. doi: 10.1186/1744-8603-10-49. PubMed PMID: 24927745; PubMed Central PMCID: PMCPMC4064106.

10. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet. 2010;376(9755):1838-45. Epub 2010/11/13. doi: 10.1016/S0140-6736(10)61997-6. PubMed PMID: 21071074.

11. Lemay NV, Sullivan T, Jumbe B, Perry CP. Reaching remote health workers in Malawi: baseline assessment of a pilot mHealth intervention. J Health Commun. 2012;17 Suppl

1:105-17. Epub 2012/05/11. doi: 10.1080/10810730.2011.649106. PubMed PMID: 22548604.

12. Zurovac D, Sudoi RK, Akhwale WS, Ndiritu M, Hamer DH, Rowe AK, et al. The effect of mobile phone text-message reminders on Kenyan health workers' adherence to malaria treatment guidelines: a cluster randomised trial. The Lancet. 2011;378(9793):795-803. doi: 10.1016/S0140-6736(11)60783-6.

13. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling Up mHealth: Where Is the Evidence? PLoS medicine. 2013;10(2):e1001382. doi: 10.1371/journal.pmed.1001382.

14. Chib A, van Velthoven MH, Car J. mHealth adoption in low-resource environments: a review of the use of mobile healthcare in developing countries. J Health Commun. 2015;20(1):4-34. Epub 2014/03/29. doi: 10.1080/10810730.2013.864735. PubMed PMID: 24673171.

15. Déglise C, Suggs SL, Odermatt P. Short Message Service (SMS) Applications for Disease Prevention in Developing Countries. J Med Internet Res. 2012;14(1):e3. doi: 10.2196/jmir.1823.

16. Hampshire K, Porter G, Owusu SA, Mariwah S, Abane A, Robson E, et al. Informal m-health: How are young people using mobile phones to bridge healthcare gaps in Sub-Saharan Africa? Social Science & Medicine. 2015;142:90-9. doi: http://dx.doi.org/10.1016/j.socscimed.2015.07.033.

17. Alege SG, Matovu JK, Ssensalire S, Nabiwemba E. Knowledge, sources and use of family planning methods among women aged 15-49 years in Uganda: a cross-sectional study. Pan African Medical Journal. 2016;24(39).

18. Uganda Bureau of Statistics (UBOS), ICF International Inc. Uganda Demographic and Health Survey 2011. Kampala, Uganda: UBOS and Calverton, Maryland: ICF International Inc.: 2012.

19. Smith C, Gold J, Ngo TD, Sumpter C, Free C. Mobile phone-based interventions for improving contraception use. Cochrane Database Syst Rev. 2015;(6). doi: 10.1002/14651858.CD011159.pub2. PubMed PMID: CD011159.

20. Johnson D, Juras R, Riley P, Chatterji M, Sloane P, Choi SK, et al. A randomized controlled trial of the impact of a family planning mHealth service on knowledge and use of contraception. Contraception. 2016. doi: http://dx.doi.org/10.1016/j.contraception.2016.07.009.

21. L'Engle KL, Vahdat HL, Ndakidemi E, Lasway C, Zan T. Evaluating feasibility, reach and potential impact of a text message family planning information service in Tanzania. Contraception. 2013;87(2):251-6. doi: http://dx.doi.org/10.1016/j.contraception.2012.07.009.

22. Jensen R. The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. The Quarterly Journal of Economics. 2007;122(3):879-924. doi: 10.1162/qjec.122.3.879.

23. Aker JC. Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger. American Economic Journal: Applied Economics. 2010;2(3):46-59. doi: 10.1257/app.2.3.46.

24. Muto M, Yamano T. The Impact of Mobile Phone Coverage Expansion on Market Participation: Panel Data Evidence from Uganda. World Development. 2009;37(12):1887-96. doi: 10.1016/j.worlddev.2009.05.004.

25. Marmot M. Social determinants of health inequalities. Lancet. 2005;365(9464):1099-104. doi: 10.1016/S0140-6736(05)71146-6. PubMed PMID: 15781105.

26. Marmot M, Friel S, Bell R, Houweling TAJ, Taylor S, Hlt CSD. Closing the gap in a generation: health equity through action on the social determinants of health. Lancet. 2008;372(9650):1661-9. doi: Doi 10.1016/S0140-6736(08)61690-6. PubMed PMID: WOS:000260899900028.

27. Desai S, Alva S. Maternal education and child health: Is there a strong causal relationship? Demography. 1998;35(1):71-81. doi: 10.2307/3004028.

28. Gakidou E, Cowling K, Lozano R, Murray CJL. Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: a systematic analysis. The Lancet. 2010;376(9745):959-74. doi: 10.1016/S0140-6736(10)61257-3.

29. Pritchett L, Summers LH. Wealthier is healthier. Journal of Human resources. 1996:841-68.

30. Wagstaff A. Poverty and health sector inequalities. Bulletin of the world health organization. 2002;80(2):97-105.

31. Dinkelman T. The Effects of Rural Electrification on Employment: New Evidence from South Africa. The American Economic Review. 2011;101(7):3078-108.

32. Jensen R, Oster E. The Power of TV: Cable Television and Women's Status in India. The Quarterly Journal of Economics. 2009;124(3):1057-94. doi: 10.1162/qjec.2009.124.3.1057.

33. Wesolowski A, Eagle N, Noor AM, Snow RW, Buckee CO. Heterogeneous Mobile Phone Ownership and Usage Patterns in Kenya. PLoS One. 2012;7(4):e35319. doi: 10.1371/journal.pone.0035319. PubMed PMID: PMC3338828.

34. Wang W, Wang S, Pullum T, Ametepi P. How family planning supply and the service environment affect contraceptive use: Findings from four East African countries. Calverton, Maryland, USA: ICF International, 2012.

35. Guy R, Hocking J, Wand H, Stott S, Ali H, Kaldor J. How Effective Are Short Message Service Reminders at Increasing Clinic Attendance? A Meta-Analysis and Systematic Review. Health Services Research. 2012;47(2):614-32. doi: 10.1111/j.1475-6773.2011.01342.x.

36. Noordam AC, Kuepper BM, Stekelenburg J, Milen A. Improvement of maternal health services through the use of mobile phones. Tropical Medicine & International Health. 2011;16(5):622-6. doi: 10.1111/j.1365-3156.2011.02747.x.

37. Aker JC, Mbiti IM. Mobile Phones and Economic Development in Africa. Journal of Economic Perspectives. 2010;24(3):207-32. doi: 10.1257/jep.24.3.207.

38. Jack W, Suri T. Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution. The American Economic Review. 2014;104(1):183-223. doi: 10.1257/aer.104.1.183.

CHAPTER 5: CONCLUSIONS

New technologies are changing health knowledge and behavior. In this dissertation I have shown this using three different technologies—HIV self-tests, a web-based weight loss application, and mobile phones—in three different settings—Kenya, the United States and Uganda. The positive impact of technologies on health is true of technologies designed specifically for health behavior change, such as HIV self-tests and a website for weight loss, and those designed for non-health purposes, such as mobile phones.

In Chapter 2 I presented results from a randomized trial in Kenya among women receiving antenatal care (ANC) or postpartum care (PPC) services to test whether the provision of multiple self-tests to women for distribution to their partners can increase uptake of male partner testing and couples testing. I found that, among the participants analyzed, partner testing within 3 months of study enrollment was significantly higher for the study group receiving self-tests than a comparison group receiving referral cards for clinic-based testing. In the group that received multiple self-tests, partner testing was reported by 91% of participants who were followed up, and 75% of participants followed up tested together with their partner. Additionally, those randomized to the self-testing group were more likely to test with their partner and know their partner's HIV status. There were no cases of violence or abuse associated with use of the new technology.

In Chapter 3 I used data from a cluster randomized controlled trial of worksites in North Carolina to determine the impact of financial incentives for weight loss on utilization of a website for weight loss. Additionally, I aimed to determine the impact of utilization of the website for weight loss on weight loss, accounting for noncompliance by a majority of study participants. I found that overall utilization of the website was extremely low, however the financial incentive significantly increased utilization of the website. Further, I found that utilization of the website was only effective at increasing weight loss among people that used it early in the study period, and ineffective over the entire 12-month follow-up period.

In Chapter 4 I used data from three waves of a nationally representative longitudinal survey in Uganda to examine the impact of mobile phones on knowledge and use of family planning. I used a robust individual fixed effects design that allowed us to control for unobserved individual level factors. I found that ownership of mobile phones led to women knowing significantly more family planning methods. However, it did not lead to a change in use of family planning. Subgroup analyses showed that the effect on knowledge was more pronounced in rural populations, which have historically less access to quality information compared to people living in urban areas.

The magnitude of the effect of technology on health behavior and health knowledge differed in each study. Availability of HIV self-tests led to large improvements in HIV testing for males, whereas use of a website for weight loss did not significantly improved weight loss for those that utilized it at 12 months. Despite the positive impact of technology on health behavior in the HIV self-testing study and early in the weight loss trial, mobile phone ownership was not associated with use of family planning in Uganda. Knowledge measurably improved with the new technologies: couples were more likely to know each other's HIV status if they were given self-tests and those who owned a mobile phone in Uganda knew of more modern family planning methods than those who did not.

There were large differences in the effect of each technology on health behavior and knowledge. Given the heterogeneity of impact, it is important to identify what could potentially be driving the differences. HIV self-tests are an intervention that has high utility for the user in that it provides a known service, HIV testing, but removes many of the existing barriers associated with testing, such as stigma and cost. We provided self-tests directly to the study participants as part of the study, and there was very little required of the study participants to use the test. Similarly, the website in the worksite based weight loss trial in NC was designed specifically to help people lose weight. Although usage was lower, the website did help people lose weight early in the study period. Both of these studies were provided to study participants and focused on an intervention that targeted a specific health issue. Conversely, in the mobile phone study we found that mobile phone ownership was not related to family planning use. In the mobile phone study there was no specific intervention that was tested. The impact observed is from owning mobile phones generally, not due to a specific mobile health program. This difference in outcomes points to the fact that in order to significantly impact health, new technologies need to be available and distributed to people and must focus on a specific health behavior for maximum effectiveness.

Given the findings from the three studies, it is important to think about what policy implications these may have. Generally, these results are positive and suggest that technology should be viewed as a tool to help improve health through targeted interventions, as well as something that may improve health on its own, such as just owning a mobile phone. The positive results highlight the need for more funding and greater adoption of new technologies. More widespread and rapid adoption could be extremely beneficial given the positive effects new technologies have. One significant finding from this research is that investment in information technology can impact health and that technology should be viewed as a social determinant of health. Health improvements due to technology will be especially strong if mHealth programs are paired with the general take-up of new technologies already occurring. For example, generating evidence-based new applications for smartphones that target a specific behavior, such as diet.

Another important finding is that technology may help reduce health inequalities, either in terms of health behavior or knowledge that currently exist. For example, men testing for HIV or rural women learning more FP methods. Overall, we showed no negative effects due to new technology. This is extremely encouraging and suggests that people use discretion when using new technology and that they understand the risks and benefits of using that new technology. We acknowledge that the risk is high for new technologies such as new diagnostic tests, which, when paired with secondary distribution, could lead to harm.

This dissertation has numerous strengths. First, I approached the question of how new technologies impact health behavior using different interventions and outcomes. The disparate nature of the interventions allows for a more complete picture of how new technologies are impacting health. In this dissertation I examined a new diagnostic technology, a website for weight loss and ownership of mobile phones. Each of these technologies could independently impact health. Second, I utilized rigorous study designs and complex survey datasets to examine the impacts of technology. Randomized trials and longitudinal datasets allow for more robust analyses. Finally, I used advanced econometric methods to determine the effect of these technologies. Estimation of the impact of new technologies was not straightforward due to issues of selection. To combat issues of selection I used three strategies, a randomized controlled trial design, instrumental variables and propensity scores, and a difference-in-difference model. These strategies account for selection and help ensure that the results found in these analyses are not driven by selection.

Despite its strengths, this dissertation also has limitations. The randomized controlled trial design used in Chapters 2 and 3 ensures high internal validity of study findings, but it makes generalizing study findings difficult. We had strict exclusion criteria in each study, and it is unlikely that the results from this study would be broadly generalizable to other populations and geographies. More studies are needed to determine the impact of HIV self-tests and websites for weight loss in other study settings and with different populations. Although RCTs have low external validity, they do have high internal validity, and the two studies allowed a nuanced understanding of the interventions. For example, we were able to determine each step along the testing cascade for male partners, from discussion, to use, to disclosure. The data collected in the RCTs allows us to understand the mechanisms driving change, something that we were unable to do in the mobile phones study. The mobile phones study, conversely, was much more generalizable, since we used a large nationally representative survey sample, however it is extremely difficult to unpack which mechanisms drive change. For example, we found that ownership led to knowledge of more family planning methods, but we were unable to test whether or not it was it through increased communication with peers, health workers, texting from the government or something else. Future studies should identify the specific factors driving change associated with mobile phone use.

These studies show that technology improves health knowledge, even without specific programs for health information dissemination, but linkages to behavior change are not as robust. Technology will become increasingly important for public health as new innovations continue to be developed and people all over the world continue to utilize them. The pace at which new technology develops is increasing exponentially, and public health practitioners need to adapt quickly to ensure that new technologies can be used for health improvement. More research is needed to confirm findings from RCTs in other settings and with different populations. Additionally, health practitioners and governments should consider pairing together multiple new technologies, such as self-testing and mobile phones, to maximize the population health impact.