ESSAYS ON MICROECONOMICS

A Dissertation
Presented to
The Academic Faculty

By

Catalina Ravizza
B.A. and M.A., Pontificia Universidad Católica de Chile, 2012
M.A., Brown University, 2015

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy
in the Department of Economics at Brown University

Providence, Rhode Island
May 2020

This dissertation by Catalina Ravizza is accepted in its present form

by the Department of Economics as satisfying the

dissertation requirements for the degree of Doctor of Philosophy.

Date ____________ ______________________________

Brian Knight, Advisor

Recommended to the Graduate Council

Date ____________ ______________________________

Emily Oster, Reader

Date ____________ ______________________________

Anna Aizer, Reader

Approved by the Graduate Council

Date ____________ ______________________________

Andrew G. Campbell, Dean of the Graduate School

VITA

Catalina Ravizza graduated with a Bachelor of Arts in Business and Economics with honors and a Master of Arts in Economics from the Pontificia Universidad Católica de Chile in 2012. Her M.A. Thesis was approved with honors. In 2015, Catalina received his Master's in Economics from Brown University. After receiving his Ph.D. from Brown in 2020, Catalina became an Economist at Keystone Strategy.

To my mom and my Alejandro.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Brian Knight, whose patient guidance and continuous support have been invaluable to accomplish this dissertation and this doctorate. I am incredibly thankful for his kindness and generosity with his time. He constantly gave me constructive and insightful feedback on my projects, regardless of my research interests moving towards topics outside of his main focus. I am lucky to have him as an advisor.

I am also deeply grateful to my committee members, Anna Aizer and Emily Oster. Every meeting with them and with Brian has been a learning experience and I am amazed by the speed and high quality of their comments. Their advice and support during the job market process has been vital. Crucially, Emily organized a reading group about health economics at the perfect time in the evolution of my research interests. She made innumerable suggestions that significantly improved my job market paper and focused it towards the right direction. Her blunt and practical advice has driven me to become a more efficient and focused researcher. Anna always gave me sharp and timely advice and her feedback was invaluable for the development and progress of my job market paper. She went above and beyond academic advising and her care and encouragement was key for me to keep always moving forward.

I also want to express my gratitude to John Friedman, whose helpful feedback and constant support throughout my third year in the doctorate pushed me to complete my first research project and led me to explore new research ideas.

I am thankful of my coauthor, Diego Gentile, whose curiosity and resolution was key to transform a discussion that began at a casual lunch into a chapter of this dissertation.

My research has also benefited from discussions with Andrew Foster, Jesse Shapiro, Neil Thakral, Matthew Turner, and participants in the Brown Applied Microeconomics Seminar, the Applied Breakfast and the Health Breakfast; I am grateful for their helpful comments and suggestions. I owe special thanks to Angelica Spertini who from the very first day has been a strong and positive force that has shaped my life at Brown. I will always be grateful for her immense and far-ranging support. I am incredibly grateful to my friends, office mates, econ girls, gym partners, basketball teammates, Latin and Italian representatives who shared this road with me, especially to Lorenzo Aldeco, José Belmar, Samuele Giambra, Violeta Gutkowski, Alessandro Lin, Amanda Loyola, Margarita Machelett, Hui Wen Ng, Carolina Siegel, and Ella Getz Wold. They were part of productive discussions, numerous coffee breaks, unforgettable parties, and more importantly, they made my stay in Providence much more lively with their companionship, laughs, and support. I especially want to thank Lorenzo Aldeco (again) for being the best work husband I could have wished for these last six years.

I thank David Debrott and Eduardo Salazar at the Superintendence of Health, Chile, for their kind help procuring data. I gratefully acknowledge the financial support from the Center for Latin American and Caribbean Studies at Brown University and from Conicyt Chile.

Last but not least, I thank my husband Alejandro Pacheco for supporting and encouraging me unconditionally, patiently, and lovingly through this journey. I dedicate this dissertation to Alejandro, my mom Angela, my nonna Maria Angelina, and my brother Diego. My family has been a source of endless support in my life and I thank them to be here.

PREFACE

This dissertation studies the incentives and consequences of public policies and behaviors through the lens and with the tools of frontier applied microeconomics. In the first chapter, I investigate the effects of information shocks on health care use before, during, and after the population's exposure to the information shock. I use high-dimensional health claims data and exploit a nationwide government information campaign promoting HIV testing. I find an increase of 30% in HIV testing and in diagnosis reports. Marginal testers, who responds to the information shock, are slightly more likely to be younger and single, suggesting that marginal testers are similar to individuals testing in periods without an information campaign. With a contagious and transmittable disease such as HIV, the timing of detection is key for both managing the disease and to contain its spreading. In the second chapter, my coauthor Diego Gentile and I examine the impacts of implementing policies fostering safe abortion practices in Uruguay on women's outcomes. We use survey data and a difference-in-differences approach and synthetic control methods. We show that the policy decrease maternal mortality, affects reproductive behavior, and ultimately increases women's employment. In the third chapter, I analyze following financial advice and return chasing as drivers of portfolio choice and their implications on wealth accumulation and retirement policies. I use fund rebalances data from retirement savings accounts in the Chilean pension system. I construct and estimate statistics that capture these drivers. I find evidence of return chasing and of following advice. High income men are more likely to follow financial advice and they switch after larger return differences. Also, low income return chasers earn significantly lower cumulative returns than higher income individuals, when compared to a benchmark, suggesting that active choices can potentially affect income inequality at the retirement age. Overall, the findings of this dissertation provide useful inputs for discussions about health screening policies, abortion policies, and social security reform.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

HEALTH INFORMATION AND HEALTH CARE USE: EVIDENCE FROM HIV TESTING

1.1 Introduction

Information campaigns are a common tool in national preventive screening interventions, but there is little research on their impact (Buchmueller & Goldzahl (2018)). Moreover, their public health consequences are driven by selection of marginal screeners into acting on the campaign. This is an often overlooked challenge (Einav et al. (2019)). Individuals acting on the campaign may not be similar to individuals acting on their own and may not be the most at risk. As a result, campaigns may not encourage the use of health care services in patients who will benefit the most.

Selection determines the ultimate impact of a policy that involve individuals' choice, such as with the endogenous choice of testing following a screening intervention. Even when a campaign increases testing, it can be deemed as a negative result from a public policy perspective. It is possible that marginal testers exhibit generally positive health behaviors or are frequently screening for diseases. Hence, the campaign may be only moving forward the timing of testing instead of encouraging unscreened or at-risk individuals. Nevertheless, encouraging repeated testers would be a positive result for contagious diseases where the timing of screening is important for early detection and to avoid spread of the disease.

An evident problem is that marginal testers cannot be identified individually, and therefore cannot be distinguished from testers that would have acted even in absence of the campaign. Nevertheless, it is possible to learn about the marginal testers' characteristics from the variation in the campaign exposure. The underlying assumption is that any change in characteristics between individuals exposed and not exposed to the campaign is driven by the marginal testers or compliers. To this end, it is key to have detailed and high-frequency individual level data on characteristics and health services use.

I explore the relevance of the selection driver in the context of Human Immunodeficiency Virus (HIV) testing and a government information campaign in Chile. I study the impact of the campaign on individuals' screening and health services use, focusing on how selection drives this impact. The HIV information campaign effectively increases HIV testing. The evidence shows that marginal testers that act on the campaign are similar in their health care use and demographics characteristics to people who self-select into testing regardless of exposure to the campaign. Back of the envelope calculations show that marginal testers are more likely to be young, single, and repeated testers, which are characteristics associated with higher riskiness. Furthermore, they have similar diagnosis rates. These similarities suggest that the campaign attracts people that is comparable to those who would have gotten an HIV test anyway in the future, or that are more in contact with health care services. In this situation, the campaign may be moving forward in time the HIV testing of some individuals, and hence, it also be moving forward their diagnosis. With a contagious and transmittable disease such as HIV, timing of diagnosis is key for both managing the disease and to contain its spreading, so the results still point to the campaign having a positive impact in terms of public health.

To examine this driver empirically, I use administrative health claims data for all privately insured individuals in Chile between 2012 and 2017. The privately insured are 14% of the Chilean population, a sample that is of higher income than the rest of the population. Private health insurance often provides better, more timely attention than the public one, and the plans offered are generally more expensive. These data includes a rich set of demographics, including, age, gender, marital status, and income, and a detailed account of individuals' health services use. Importantly, although test results are not included in the service use data, I use data from a registry of self-reported diagnosis for a list of diseases with expensive treatment, where HIV is included. Individuals must voluntarily register with their health insurer to receive benefits related to their diagnosed disease.

The Chilean government promotes preventive health care use mainly through nationwide health information campaigns. These respond to a generally low use of preventive health services. Although HIV infection remains incurable, early detection and treatment reduce the risk of transmission, and hence the government routinely launches HIV information campaigns. Both the content and timing of these campaigns vary, providing an unexpected information shock. Roughly one campaign is launched per year and they usually last two months, consisting of mostly commercials on national TV, radio, out of home advertising (eg. bus shelter posters), internet, and social media.

The analysis is organized into three main parts. First, I begin by studying the impact of the information campaign on HIV testing. I construct a balanced panel of individuals and estimate an event study around the HIV information campaign, controlling for seasonality. I find a large increase in HIV testing after the HIV information campaign is launched. The number of testers increases by 30% with respect to the weeks prior to the campaign. To further support my findings, I test for unknown structural break dates in the trend of HIV testing. The estimated break date is the campaign launch date or a week before. Raw data do not show any demographic groups distinctively driving the increase in testing.

Second, I study whether testers at the time of the campaign are different in terms of observables from testers at other times when there was no campaign. I find that individuals tested around the campaign are more likely to be young (ages 18 to 24), single, and of high income. Some of these characteristics are associated with risky sexual behaviors. Moreover, there are no differences in the health services use at the time of the HIV test event. Lab work use is similar, with most testers bundling tests for sexually transmitted diseases (STDs) and general check-ups. Testers at the time of the information campaign are also more likely to have used any health service in the year prior and to have taken an HIV test. Taking another approach to compare the testers, I find that the predicted probability of HIV testing is very similar for individuals that got an HIV test regardless of the occurrence of the information campaign, although the predictive power of this exercise is low given that there are very few HIV testers.

Third, I investigate the intended and unintended consequences of getting an HIV test. The campaign pushes individuals to contact health services, presenting an opportunity to also undergo preventive health screening tests. I study the impact on an intended consequence, the diagnosis rates, and on an unintended consequence, the use of health services after taking the HIV test, which indicates usage of treatments and follow-ups of detected health problems. Again, I compare the groups of testers exposed and not exposed to the campaign. Diagnosis rates in the first 12 weeks after the test are statistically indistinguishable between these groups. However, since the number of testers increased as a result of the campaign this translated into 11 more diagnosis with respect to the year without campaign. This roughly constitutes one additional diagnosis per week. Using a difference-in-differences design, I find that health services usage after the HIV test does not show any statistically significant differences.

These results demonstrate that the campaign was effective in increasing testing. Furthermore, marginal testers are very similar to other testers, in terms of demographics, health care utilization before, around, and after testing, as well as in terms of diagnosis rates. This provides indirect evidence that marginal testers may be moving forward their testing and also their detection of HIV infection. Nevertheless, the number of tests does not decline right after the campaign ends and therefore I do not observe direct evidence of this channel in the short term[1]. Testing could alternatively decrease in the longer term. The increase in the number of reported diagnosis, although with a similar rate, is a key element to assess the effectiveness of the campaign in light of selection into testing. This suggests that marginal testers are similarly risky and are moving forward their detection. Therefore, they may initiate treatment earlier and become less likely to further transmit the infection.

This paper relates to three main literatures. First, it relates to a literature on the role of information on health behaviors and preventive screening. Broadly, findings support small impacts from information shocks coming from different sources, such as organized screenings (Buchmueller & Goldzahl (2018)), celebrity promotional campaigns (Cram et al. (2003)), social networks (Deri (2005), Fadlon & Nielsen (2017)), psychological interventions (Haushofer et al. (2019)), diagnosis (Oster (2018b)). Mass media information campaigns, very common tools in national screening interventions, have been studied mostly in the health literature showing small and short-lived impacts (Snyder et al. (2004), Snyder (2007), Wang et al. (2012), Wakefield et al. (2010))[2]. The majority of the literature has focused on non-contagious and chronic health problems. I confirm the significant role that campaigns have on screening in the setting of a contagious disease, where higher rates of screening affect not only individual health, but also have public health consequences.

Second, it relates on a more recent literature studying the role of selection in the context of health-related choices. The importance of selection stems from its impact on the effectiveness of policies such as screening interventions or recommendations. Einav et al. (2019) uses an oncology model to show that selection into screening overestimates the benefits of recommending early screening, compared to the case where responders were representative of covered individuals. Oster (2018a) shows that selection makes difficult learning about causal effects of recommendations. I provide empirical evidence on the selection driver in the context of HIV testing by characterizing marginal testers using a rich set of individuals' characteristics. This allows me to better assess the effectiveness of the campaign. The role of selection has also been documented in other health settings, such as health insurance choice (Bundorf et al. (2019)). Moreover, Myerson et al. (2018) discuss that economists have paid scant attention to these complexities and their implications for

[1]Due to data limitations, I can only observe a short period after the campaign was launched and hence cannot explore if testing decreases in the longer term.

[2]Short-lived impacts of information campaigns are also observed in other settings relevant for public policy, such as social security systems (Finseraas et al. (2017)).

evaluating screening programs. Screening interventions in general may produce different outcomes depending on the reasons why patients went unscreened in the first place, which can also impact treatment take-up.

Third, it relates to a broader literature on the incentives and benefits of preventive screening for a potential disease. There is a vast literature documenting the benefits of preventive screening (CDC (2009), Maciosek et al. (2010), McMorrow et al. (2014)). More particularly, a rich set of papers studies the case of HIV testing (Baggaley et al. (2017), Montoy et al. (2016), Montoy et al. (2018), Nichols & Meyer-Rath (2017), Nakagawa et al. (2012), Reitsema et al. (2019), Zah & Toumi (2016)). They document that early detection of HIV reduces morbidity, mortality, the probability of onward transmission, and their associated costs. Moreover, they find that screening for HIV in primary care is cost-effective, despite the cost of earlier initiation of antiretroviral treatment. I present indirect evidence of a policy that moves forward HIV detection and highlight the role of selection in its assessment.

The remainder of this paper is organized as follows. Section 1.2 describes the background information for the information campaign and the data sources used. Section 1.3 shows evidence of the direct impact of the campaign on HIV testing. Section 1.4 investigates the role of selection as a driver of the increase in testing. Section 1.5 studies whether the increase in testing had impacts in outcomes after the HIV test took place. Section 1.6 concludes.

1.2 Background and Data

1.2.1 HIV information campaign in Chile

In Chile, HIV infection has been a relentless public health challenge. The HIV epidemic has been characterized by a rapid increase in the estimated new infections since the early 2000s, (see figure 1.13a in Appendix 1.7.1) despite having the highest access to ART among Latin American countries. Sexual transmission accounts for 99% of the infections.

The Chilean government estimates that 65,000 individuals live with HIV by 2016 and 5,212 infections were diagnosed in 2017.

The medical advances in the last decades have transformed HIV infection into a chronic disease. Preventive screening is a standard and powerful policy to deal with chronic diseases. Since HIV remains incurable, early detection and treatment have an important role in reducing the risk of developing Acquired Immune Deficiency Syndrome (AIDS) and of transmission. Moreover, preventive health services use in Chile is generally low (Rotarou & Sakellariou (2018)), therefore the government promotes preventive health care use mainly through nationwide health information campaigns. The Chilean government routinely launches HIV information campaigns to foster preventive screening of HIV infection. Both the content and timing of these campaigns varies, providing an unexpected information shock to the whole Chilean population. Roughly one campaign is launched per year and they usually last two months, consisting mostly of commercials on national TV, radio, out of home advertising (eg. bus shelter posters), internet and social media.

In early August of 2017, the government launched an HIV information campaign was launched under the slogan "The more we ignore it, the stronger it becomes. Use a condom and take the test". Hence, it explicitly promoted HIV testing and condom use. Table 1.1 shows the timeline of events surrounding the campaign. The launch of the campaign occurred just two weeks after a report from the Joint United Nations Programme on HIV/AIDS (UNAIDS) was released, describing Chile as the Latinamerican country with the highest increase in new HIV infections between 2010 and 2016. The findings of this report had wide newspaper coverage in Chile.

Google searches surrounding the campaign launch suggest awareness among the Chilean population of the HIV information campaign. Figure 1.1 shows the trends of Google searches for two terms related to sexually transmitted diseases (STDs). I observe a large spike for the search term "HIV" around the 2017 campaign but not for "syphilis", suggesting that the spike may be prompted by the information campaign. The increase in

TABLE 1.1: Timeline of HIV information campaign in 2017

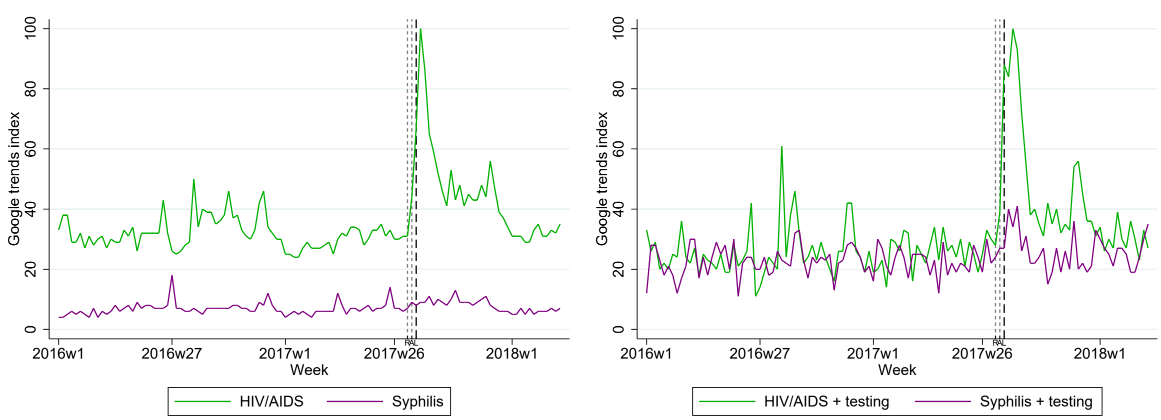
Date	Event	Description
July 20th (week 29)	UNAIDS report	"Ending Aids 2017" shows Chile is the Latin American country with largest increase of new cases between 2010 and 2016
July 28th (week 30)	Announcement of campaign	Ministry of health announces HIV campaign will be released the following week
August 3rd (week 31)	Launch of campaign	*Slogan:* The more we ignore it, the stronger it becomes. Use a condom and take the test. *Media:* TV, Radio, Internet, Social media

Notes: See poster in figure 1.13b in Appendix 1.7.1.

"HIV" searches slowly begins two weeks prior to the campaign, corresponding to the time at which the UNAIDS report was released. The observed trends for the combined search of each STD and terms related to testing also show a spike for HIV, but not for syphilis, although the trends are much noisier.

FIGURE 1.1: Google trends of term searches, sexually transmitted diseases

(a) STDs searches (b) STDs testing searches



Notes: This figure shows the Google trend searches for selected terms. The black dashed vertical line corresponds to the campaign launch and the gray dashed vertical lines correspond to the release of the UNAIDS report and the campaign announcement, respectively. Panel (a) corresponds to searches of the terms "HIV/AIDS", and "syphilis". Panel (b) corresponds to the same searches as above combined with testing related terms; for HIV/AIDS I consider "test" and "ELISA", while for syphilis I consider "test" and "VDRL".

In Chile, HIV testing is widely entangled with doctor visits. The clinical practice guidelines for HIV detection follow the most widely used approach for diagnosis of this infection and instruct using the Enzyme Linked Immunosorbent Assay (ELISA) test that requires sophisticated equipment and skilled technicians. This test is labor intensive and time consuming, since it requires the collection of a blood sample at a laboratory, which then is sent to the National Laboratory that performs the ELISA test. Recently, resource-limited countries have introduced the rapid diagnostic tests (RDTs), which has the benefit of allowing counseling and results during a single encounter, although some studies find slightly worse diagnostic performance. The Chilean government has not followed this trend. Only on December 2017 the Ministry of Health launched a short-lived program to make available the RDTs for HIV for individuals enrolled in a few public primary care clinics in the country. The program has not been scaled since.

1.2.2 Health claims data from Chilean private health insurers

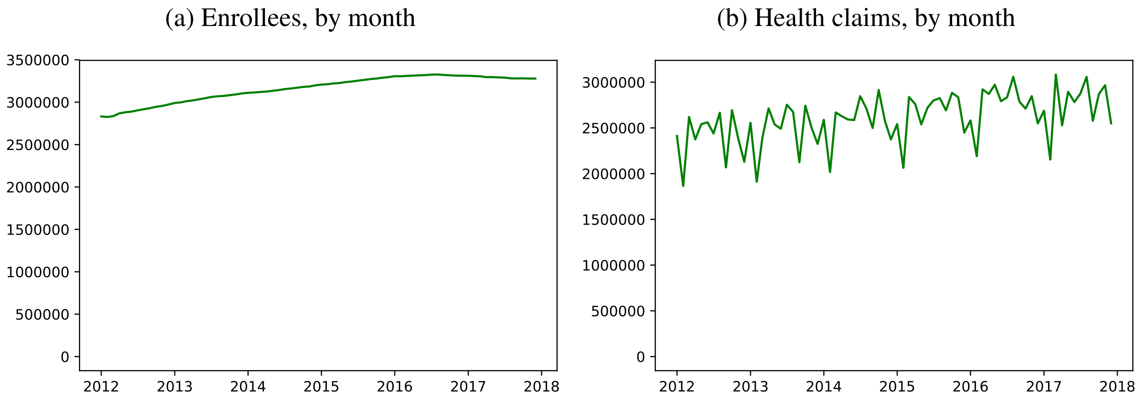
I use administrative data for all privately insured individuals between 2012 and 2017. Figure 1.2 shows the trends for all enrollees and the health services use for a balanced panel of enrollees, showing an increase of 15% over the full period. The privately insured are 14% of the Chilean population, a sample that is higher income than the rest of the Chilean population[3]. Private health insurance often provides better, more timely attention than the public one. For instance, wait times in the public sector are 6 times larger than those in the private sector and are concentrated in specialist doctor visits and surgeries. The in-network providers for the private health insurance are not capacity constrained for general doctor visits. Also, the private plans offered are generally more expensive. Plan choice is not associated with the employer, as individuals can switch insurer at any time.

The dataset I use is composed by three main pieces; beneficiaries characteristics, health service claims, and the registry of a subset specific diagnosis. I use each of the three data sources in turn to construct individual level characteristics, health service use variables, and HIV diagnosis reports, respectively.

First, I determine enrollment status and construct individual level characteristics from

[3]78% of the population has public health insurance, 3% have military health insurance and the rest are not covered.

FIGURE 1.2: Enrollees and health claims, raw data



Notes: The enrollees correspond to 2,575,393 families and 4,399,932 individuals, of which 54.31% are men. The median number of months enrolled is 58 and 40% of the individuals are enrolled the full period. The trend oh health services correspond to a balanced panel of individuals. It includes 187,568,004 observations, 1,039,812 families, and 1,774,224 individuals of which 43.92% are men.

the list of beneficiaries with demographics information that I observe each month. I observe whether individuals are enrolled or not each month and determine enrollment spells, which allows me to accurately identify when there are no health services use. For each individual, I observe their gender, date of birth and death, and municipality of residence, as well as an indicator for being the main insured in the health insurance plan and their relationship with the main insured (self, spouse, child, parent, other). For the main insured only, I observe marital status and taxable income. Marital status is a relevant characteristic when studying risky sexual behavior, therefore I infer it as best as possible in certain cases. I assign marital status of *married* to spouses, and *single* to children. Further, I assign marital status to some of those with a relationship labeled as *other* who are in a group plan with a family structure given the age differences or because of the presence of children. For the rest of the sample I assign marital status as *unknown*.

Second, I construct a comprehensive list of types of health care services which capture a broad range of health services use from the health claims data. These data are high frequency, including daily dates, health service codes, co-payment, and plan and provider characteristics. The large majority of the health service codes are defined by the Chilean government, which I use to construct eight types of services. For each type I construct an indicator for *any* service use and a *count* for the number of a particular type of services.

I consider the following types of services: doctor visits with a general physician; specialist visits, such as urology, traumatology and orthopedic, ophthalmology, cardiology, or in some cases no specialization is provided; preventive care services, from service codes included in a national list of services such as fasting blood glucose (diabetes), syphilis blood test, smoking questionnaire, measurement of weight and height and waist circumference (obesity), blood pressure measurement (hypertension); diagnosis and therapy services, such as ECG, physical therapy, vitreoretinal exam, endoscopy; surgeries, such as circumcision, appendectomy, meniscectomy, tonsillectomy, nose surgery, corneal surgery; hospitalizations coded as any day of hospitalization; lab tests, such as blood test (CBC, lipid profile, biochemical profile), or urine test(complete urinalysis, urine sediment); and mental health services, such as visits to the psychologist or psychiatrist.

Third, I use the reports of diagnosis for a list of health problems, where HIV infection and AIDS are included, to construct an indicator for HIV diagnosis. The reports are freely given by individuals to the health insurer and allow them to gain access to benefits related to their health problem. In the case of HIV, individuals will have access to examinations and lab tests, to the antiretroviral therapy treatment, and to monthly follow-up examinations and lab tests, at a reduced co-payment that varies between zero and 20 percent of the price determined by the Ministry of Health. Although the prices have slightly increased over time, the battery of benefits has been the same.

1.3 Impact on HIV testing

A first question is whether the information campaign increased HIV testing. I construct a balanced panel of individuals and estimate an event study at the weekly level for the number of HIV tests around the information campaign. I find a large increase in testing in the weeks following the campaign launch. To further support my findings, I test for unknown structural break dates in the trend of HIV tests. The estimated break date is statistically significative and coincides with the campaign launch date or a week before. No single demographic group appears to distinctively drive the testing increase.

1.3.1 Event study design

I use a balanced sample of individuals enrolled with any health insurer between the years 2012 and 2017. I focus on individuals that *voluntarily* take the test; therefore I exclude women likely to be pregnant since they have mandatory HIV testing. Each HIV test observed for the sample of individuals has a daily date, which I aggregate at the weekly level to construct the outcome of interest, y_t, the weekly number of HIV tests. The raw trends of this measure are shown in figure 1.3. I observe a large spike in the number of HIV tests around the launch of the 2017 campaign, that stays at a higher level than the prior trend, although with plenty of noise. The number of tests increased by 39% in the first 15 weeks after the campaign is launched with respect to the prior 15 weeks[4]. A similar pattern is observed for the trends by gender, as shown in figures 1.13c and 1.13d in Appendix 1.7.1.

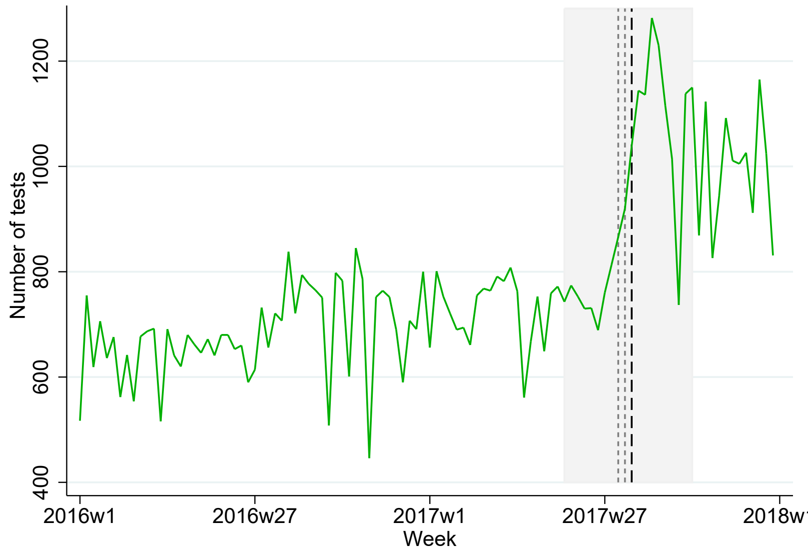
I quantify the effects of the unexpected launch of the information campaign using an event study design, relying on a time trend analysis. For non-emergency related health services use, seasonality is an important factor that can bias the results. For instance, health services use is lower around holidays, but more use may be observed during winter. I use a balanced panel of individuals enrolled during the full period of 2012 to 2017, which allows me to include calendar time trends to control for the seasonality concern. Also, using a balanced panel fixates the individuals considered in the analysis, ensuring that testing trends do not change because the universe of enrolled individuals changes.

$$y_t = \alpha + \gamma(t - t_0) + \sum_{\tau=-15}^{15} \beta_\tau \mathbb{1}(t - t_{2017} = \tau) + \theta^W_{w(t)} + \theta^A_{a(t)} + \varepsilon_t \qquad (1.1)$$

Equation 1.1 shows the event study model. I estimate the campaign's impact from

[4]This also represents a 50% increase in the first 15 weeks after the campaign is launched with respect to the average since 2016.

FIGURE 1.3: Trend of HIV tests, weekly balanced sample

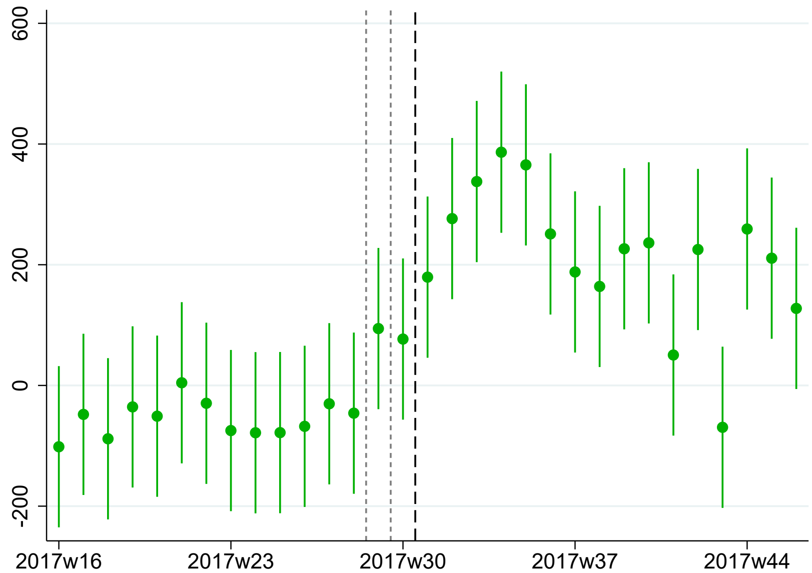


Notes: This figure shows the raw weekly trend for the number of HIV tests. The sample includes individuals enrolled for the full period, which yields 132,340 individuals, ages 18 to 50, 57% of which are men. The black dashed vertical line corresponds to the campaign launch and the gray dashed vertical lines correspond to the release of the UNAIDS report and the campaign announcement, respectively. The shaded gray area corresponds to the period used in the event study design.

dummies for each week around the campaign launch, represented by t_{2017}, and where τ is the relative time with respect to t_{2017} in an event window of 15 periods. $\widehat{\beta_\tau}$ is the estimated additional number of tests in week τ. The model implicitly restricts the coefficients $\widehat{\beta_\tau}$ to be zero for the weeks outside of the event window. I include a time trend, γ, where t_0 corresponds to week 1 of 2012. To deal with seasonality, I control for a set of calendar week and year fixed effects, $\theta^W_{w(t)}$ and $\theta^A_{a(t)}$, respectively, where $w(t)$ and $a(t)$ yield calendar week and year from weekly date t. I estimate the model using an OLS regression.

Figure 1.4 shows the estimated coefficients $\widehat{\beta_\tau}$. I observe a large increase in the weeeks following the campaign launch, suggesting that the campaign increased the number of tests, after controlling by seasonal trends. Note that the estimated coefficients of the two weeks just before the campaign launch are positive, although not statistically different from zero. These weeks coincide with the timing of the release of the UNAIDS report and the campaign announcement, shown as shown gray dashed lines in figure 1.4. In all the weeks prior to the campaign, the estimated coefficients are not statistically different from zero, indicating that there was no pre-trend and that they do not differ from the coefficients outside the event window. Figures 1.13e and 1.13f in Appendix 1.7.1 show similar estimated trends for separate regressions by gender. Table 1.4 in Appendix 1.7.2 shows the coefficients of the estimation of equation 1.1. Results are similar when excluding the time trend or year fixed effects.

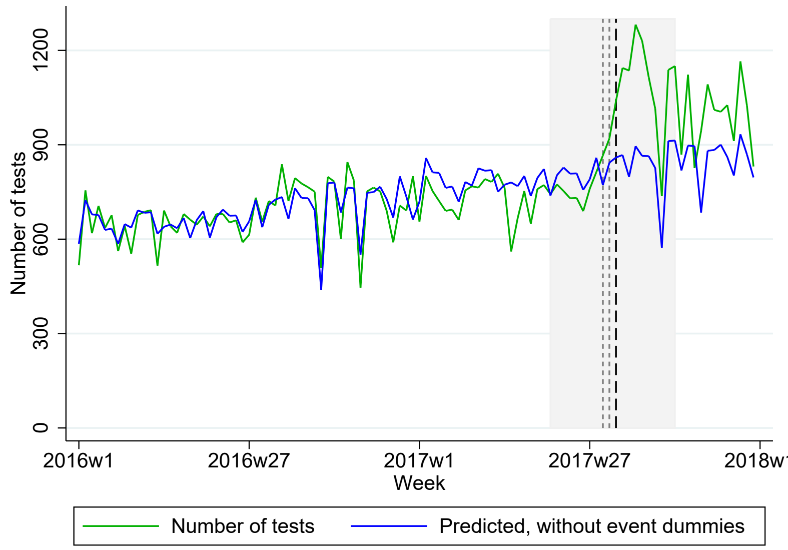
FIGURE 1.4: Event study of HIV tests



Notes: This figure shows the estimated $\widehat{\beta_\tau}$ coefficients of equation 1.1 estimated using data from 2012 to 2017. The coefficients represent the number of additional HIV tests in the weeks around the 2017 campaign launch, pictured as a black dashed vertical line. The gray dashed vertical lines correspond to the release of the UNAIDS report and the campaign announcement, respectively. The model is estimated using OLS and includes a time trend and a set of calendar week and year fixed effects. The average number of tests during the weeks prior to the campaign launch is 759.

The increase in testing can be better appreciated by comparing the testing trend with the predicted trend of tests excluding the dummies around the event study, $\widehat{\beta_\tau}$. I show this in figure 1.5, in green and blue lines, respectively. I plot the period 2016 to 2017, although the estimation used data between 2012 and 2017. The predicted pre-period trend closely follows the actual trend, showing the relevance of seasonality modeled through the calendar week fixed effects. After the campaign launch, the predicted trend does not increase, showing a large gap between the actual and predicted trends. Two weeks after the campaigns show a plunge in the number of tests. They corresponds to weeks in which there was a national holiday and therefore had fewer business days.

FIGURE 1.5: Predicted HIV tests and raw data



Notes: The green line shows the actual testing trend. The blue line shows the predicted testing trend using the model in equation 1.1 and the data from 2012 to 2017, but excluding the dummies $\widehat{\beta_\tau}$ in the event window, which is shown as the shaded gray area. The coefficients represent the number of additional HIV tests in the weeks around the 2017 campaign launch, pictured as a black dashed vertical line. The gray dashed vertical lines correspond to the release of the UNAIDS report and the campaign announcement, respectively.

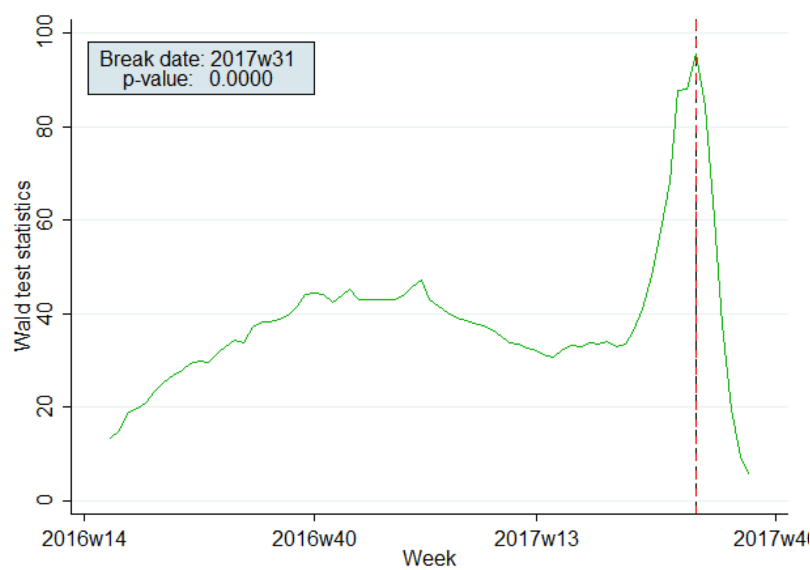
1.3.2 Test for unknown structural breaks

A well-known limitation of event studies is that it is difficult to identify the exact date of a specific event. In the context of the information campaign I study, the timeline of events is well known (see table 1.1). Furthermore, I can take advantage of the high-frequency of the data to test for a structural break date without imposing a known break date. Intuitively, this approach provides an objective measure to investigate whether the raw data shows evidence of the impact of the campaign by comparing the maximum sample test with what could be expected under the null hypothesis of no break.

I begin this analysis by deseasonalizing the trend of the number of HIV tests using the balanced sample of individuals for the full period between 2012 and 2017 aggregated at the weekly level. Then, I take the deseasonalized subset between 2016 and 2017 and estimate a simple time trend regression, $y_t = a + b(t - t_0) + \varepsilon_t$, where I test for unknown structural break dates on the coefficients. I perform this test for the intercept and the slope simultaneously, as well as separately for each coefficient. The null hypothesis in the test is that there is no structural break date on the coefficient of interest.

For each possible break date, figure 1.6 plots the Wald statistic of the test. The red vertical dashed line shows the estimated break date and the black vertical dashed line shows the actual campaign launch date. Both lines coincide at the campaign launch date, and I find evidence of statistically significative structural break in the intercept and slope at that exact date. When I test for unknown structural breaks in a subsample excluding the weeks after the campaign launch, I do not find any statistically significative structural break. Figures 1.13g and 1.13h in Appendix 1.7.1 show similar results when testing for unknown structural break date in the slope and intercept separately. The estimated break date is two weeks before the campaign launch and corresponds to the date of the release of the UNAIDS report. In each case I find evidence of a statistically significative structural break.

FIGURE 1.6: Structural break test of HIV tests' trend, for slope and intercept



Notes: This figure plots the Wald statistics for a structural break test at each possible break date. The red vertical dashed line shows the estimated break date, displayed in the top-left box along with the p-value for the test under the null hypothesis of no break date. The black vertical dashed line shows the actual campaign launch date.

1.3.3 Testing and demographics groups

Having shown that the information campaign increased HIV testing, the public health consequences of this result would be different if the increases were concentrated in low versus high-risk groups. According to general trends of sexual and preventive behaviors by demographic groups, higher-risk groups would be more exposed to HIV infection. Therefore increasing testing among such groups would hinder the spread of HIV. Usually, individuals that are younger, single, and lower income, are regarded as higher-risk towards HIV infection.

From the raw data, I do not observe a sharp change in the pattern of testers by demographic groups around the time of the campaign. Figure 1.13 in Appendix 1.7.1 shows that the number of testers increased thoroughly around the campaign for every demographic group, although the share of single and young individuals slightly increased. This evidence proves to be suggestive at best, but further analysis needs to be done to draw any conclusion.

I investigate whether the campaign is encouraging people to get their first HIV test or if they are recurring testers. I use two different categorizations. First, I define an initiation status for either being a first-time tester or a recurring testers. Second, I define the recent testers as individuals that got an HIV tests in he year prior. Figure 1.13 in Appendix 1.7.1 plots the trends showing no marked change around the campaign. About 40% of testers have had a prior HIV test and 20% of testers had the prior test within a year.

1.4 Selection into testing

Screening choices are inherently an endogenous choice, even when individuals are exposed to an exogenous and unexpected encouragement to get screened. Researchers studying such an intervention will identify its impact from individuals who comply with the encouragement. Nevertheless, it is not possible to individually identify the compliers. Observed testers after the encouragement can do so motivated by the encouragement or would have done so regardless of it. In this context, an increase in testing can be considered more or less effective from a public health perspective depending on who are the people that act on the encouragement. In this section, I take a deep dive to characterize the test takers and understand the role of selection into responding to the HIV information campaign.

I study the selection driver by drawing a parallel between characterizing the compliers in an instrumental variables' framework and characterizing the test takers that act encouraged by the campaign. The underlying assumption is that any change in characteristics between individuals exposed and not exposed to the campaign is driven by marginal testers or compliers. I investigate the marginal testers' distribution of characteristics from the variation in the campaign exposure by comparing two groups of testers. I compare testers periods with and without an information campaign. Marginal testers that act on the information campaign do not differ greatly in the health care use from other testers and, as a result, do not have different health care use after the test and have similar diagnosis rates. The information campaign may be moving forward in time the HIV testing of individuals that may be more in contact with health care services or may have gotten tested anyway.

1.4.1 Study groups and health services use event

I take advantage of the detailed and high-frequency individual level data to construct two groups of testers, differentiated by their exposure to the campaign when taking an HIV test[5]. The group of testers exposed to the campaign, which I refer to as the treatment group, consists of individuals observed taking an HIV test at the time of the information campaign, between weeks 29 and 36 of 2017. This group comprises 11,452 testers, of which 46% are men. The control group consists of individuals who took an HIV test in a period without an information campaign, between weeks 29 and 36 of 2016. This group consists of 7,935 testers, of which 50% are men. I use data for the same weeks in each years as a way to consider the seasonality concern (discussed in section 1.3) between the two groups of testers.

Comparing individuals according to their exposure to the campaign does not guarantee that testers in the treatment group are compliers or marginal testers in the sense that they chose to screen solely motivated by the encouragement of the campaign. The study of selection in this context is challenging since it is not possible to differentiate between always takers and compliers to the campaign. This situation is analogous to describing the distribution of compliers' characteristics in an instrumental variables' framework, where the effects are identified from those who comply with the encouragement. Nevertheless, I can compare the average of the two groups under the assumption that differences in these averages will be driven by the campaign.

HIV tests are performed following a doctors' prescription order. As such, it requires a visit to a doctor and introduces the possibility of getting it along with other tests. For instance, individuals may be taking only an HIV test, or may be bundling it with only STDs tests, or with a more general health check up. I define a health services event as the set of health services that occurs with an interval of 10 days or less. Note that this means that the event may last longer than 10 days. For instance, health services on October 3rd, 10th and 18th will be considered part of the same event spanning 15 days. Figure 1.13i in Appendix 1.7.1 shows that most events occur within just a few days. This allows me to identify any bundling of health services presumably originating from the same testing event and to separately study the health services use surrounding the HIV test event and in

[5] Analogous to the previous section, I use a balanced sample for the period between 2015 to 2017. I restrict individuals to ages 18 to 50 on December 31st on 2017, both male and non-pregnant female.

the period following the HIV test (in section 1.5).

1.4.2 Marginal testers' distribution of characteristics

Marginal testers may differ from other testers in terms of demographic characteristics or their health services use. I compare the average characteristics of testers in both groups under the assumption that differences in these averages will be driven by the campaign. I use a proportion test to test whether the share of individuals with each characteristic varies between treatment and control. Figure 1.7 shows bar graphs for each characteristic, along with the p-values in the square brackets, for the test of equality of proportions. The null hypothesis is that the two populations have equal proportion of the characteristic.

The bar graphs on the left side of figure 1.7 show the results for a set of demographic variables. I find that testers exposed to the campaign are slightly more likely to be young, single, and of high income than testers in the year without a campaign. These differences are statistically significant but very small in magnitude. The bar graphs on the right side of figure 1.7 show the results for a set of health service use variables. The health services use at the HIV health event does not differ by groups. Marginal testers are slightly more likely to have used any health service in the previous year and less likely to be first time testers, as shown in figures 1.7d and 1.7f. Taken together, the results suggest that marginal testers are more likely to be part of higher risk groups and that they may be more in contact with health care services.

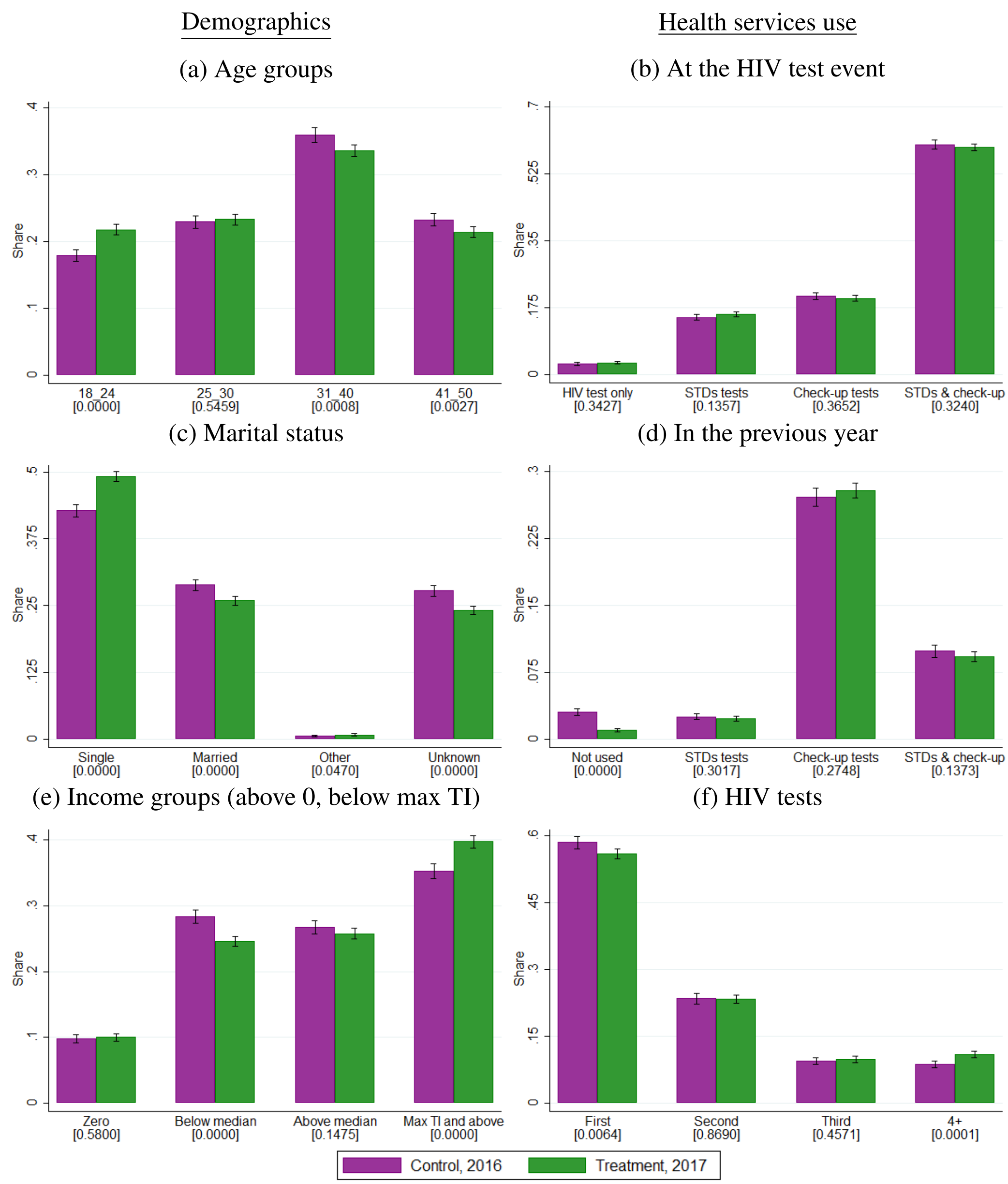
I take a deeper dive at two higher risk groups, single and young individuals ages 18 to 24. The results are shown in figure 1.8. I observe that among young individuals, the group exposed to the campaign is slightly less likely to undertake comprehensive health check-ups and STDs tests, both at the time of the HIV test event and in the previous year. Similarly to the results for the full sample, marginal testers are more likely to have used any health service in the previous year and to be repeated testers. This evidence indicates that the campaign brings marginal testers that are in contact with health care services and that have taken HIV tests before, suggesting that some of these testers may be moving forward in time their testing decision encouraged by the information campaign.

All things considered, the evidence shows that most people are getting tested for HIV when they also test for general check-ups and for other STDs. From the results in the previous section, I quantified an increase in the number of testers of 30% that is statistically significative. This points to a situation in which some individuals go to the doctor anyway and decide to get an HIV test, while the increase of 30% of individuals act on the campaign and end up going to the doctor when they would not have gone otherwise. The latter group is what I call marginal testers and, under the assumption that changes are driven by the campaign, fully accounts for the 30% increase in testers exposed to the campaign.

I use a back of the envelope calculation to compare the characteristics of marginal testers and regular testers exposed to the campaign. I consider two assumptions. First, marginal testers fully account for the 30% increase in testers exposed to the campaign, which corresponds to a 23% of the testers in 2017[6]. Second, regular testers' characteristics are equivalent to those of testers not exposed to the campaign. With these assumptions in hand and since I observe the total share of testers exposed to the campaign for each characteristics, I back out the shares for the marginal testers. Figure 1.9 shows bar graphs that compare the average characteristics of these groups.

Let's consider figure 1.9a. From the previous results we know that the share of young individuals, ages 18 to 24, from 18% to 22% and that marginal testers comprise 23% of all testers exposed to the campaign. Testers exposed to the campaign comprise regular and marginal testers. I assume that the share of young individuals among the regular testers is the same as among those not exposed to the campaign, 18%, and back out the share of young individuals among the marginal testers as 35%, such that the total share is the observed value of 22%. Using this approach, I observe that marginal testers are 16% more likely to be young. They are also 18% more likely to be single and 10% less likely to be first time testers.

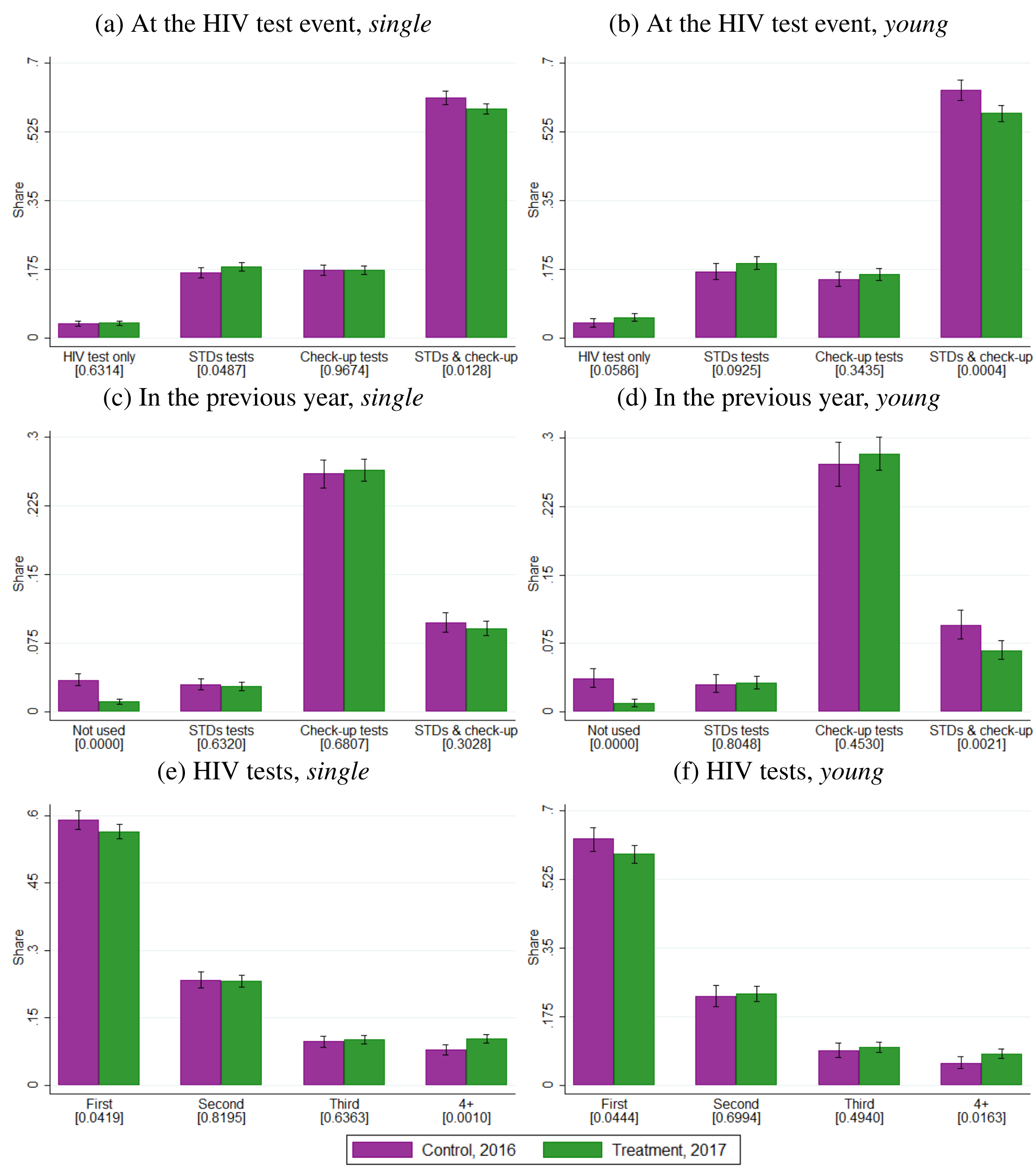
FIGURE 1.7: Demographics and health services use, all testers



Notes: These figures show bar graphs for the distribution of characteristics for testers exposed (Treatment, 2017) and not exposed (Control, 2016) to the campaign. I use an equality of proportion test to compare the share of individuals with each characteristic between treatment and control, under the assumption that differences in these shares will be driven by the campaign. p-values for the test are shown in square brackets below each characteristic. The null hypothesis is that the two population have equal proportion of the characteristic.

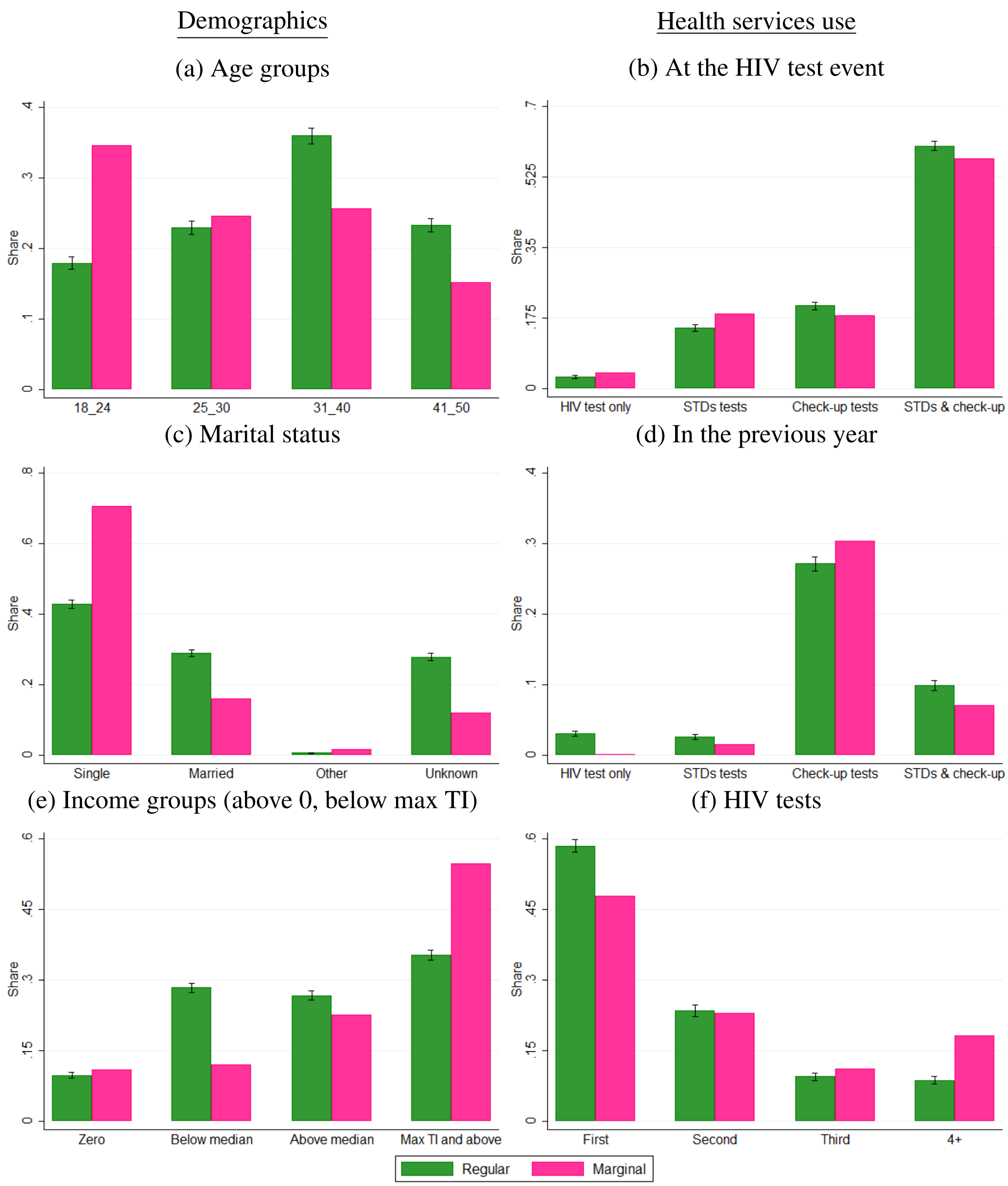
[6] This implies that if is there are 130 testers, 100 belong to the regular testers' group and 30 belong to the marginal testers' group. Therefore, 30 over 130 yields constitutes 23% of marginal testers.

FIGURE 1.8: Health services use, higher-risk groups



Notes: These figures show bar graphs for the distribution of characteristics for testers exposed (Treatment, 2017) and not exposed (Control, 2016) to the campaign. I use an equality of proportion test to compare the share of individuals with each characteristic between treatment and control, under the assumption that differences in these shares will be driven by the campaign. p-values for the test are shown in square brackets below each characteristic. The null hypothesis is that the two population have equal proportion of the characteristic.

FIGURE 1.9: Demographics and health services use, marginal and regular testers



Notes: These figures show a back of the envelope calculation to compare the groups of characteristics of regular and marginal testers, among testers exposed to the campaign. I use two assumptions. First, marginal testers fully account for the 30% increase in testers exposed to the campaign. Second, regular testers' characteristics are as if they would behave as the testers not exposed to the campaign. With these assumptions I back out the share of marginal testers needed for to get the total share of the exposed testers.

1.4.3 Probability of HIV testing

Previous results show that marginal testers have statistically different characteristics with respect to other testers, but that these differences are quite small in magnitude. Therefore, individuals that select into testing after exposure to the campaign are similar to those that select into testing in absence of the campaign. I this section, I use a complementary approach to further contrast the testers exposed and not exposed to the campaign with the goal of comparing the predicted likelihood of getting an HIV test between them.

I construct a model to predict the likelihood of getting an HIV test using a rich set of observables without the nudge of the information campaign. To this end, I estimate the model using a sample of every individual around weeks 29 to 36 in 2016, when there was no campaign. I pool all individuals in the mentioned period in a cross-section format. This sample includes the testers in the control group defined above as well as every non-tester during that period. Note that very few take the test (about 0.86 %) as shown in table 1.2.

TABLE 1.2: Sample sizes of study groups by campaign exposure

	2016 (no campaign)	2017 (campaign)	Total
Not tester	1,049,534	0	1,049,534
Tester	9,066	11,111	20,177
Total	1,058,600	11,111	1,069,711

I estimate a logit regression where the left hand size variable is an indicator for getting an HIV test at any moment during weeks 29 to 36 in 2016. The right hand size variables include observables about demographic characteristics and past health service use. Table 1.3 shows the estimated coefficients. Most of the observables are statistically significant, nevertheless the pseudo-R^2 is very low signaling that the predictive power of the model is not strong. This may be due to the fact that only a very small share of individuals in the full sample take an HIV test.

I use the estimated model from table 1.3 to compute the predicted probability of the

TABLE 1.3: Prediction of HIV test taking, without campaign exposure

Probability of taking HIV test	Coef.	Std. Err.	p-value
Last year - Any specialist visit	0.416	0.036	0.000
Last year - Any panel tests	-0.267	0.024	0.000
Last year - Any gyn or proc visit	0.216	0.031	0.000
Last year - Any doctor visit	0.483	0.023	0.000
Last year - Any psychologist visit	0.277	0.027	0.000
Last year - Any service	0.103	0.120	0.388
Last year - Any surgery	0.169	0.042	0.000
Last year - Any preventive screening	0.448	0.052	0.000
Last year - Any STDs tests	0.650	0.035	0.000
Last year - Any HIV tests	0.758	0.039	0.000
Last year - Any hospitalization	-0.136	0.049	0.006
Marital status - Married	-0.296	0.030	0.000
Marital status - Unknown	0.073	0.029	0.011
TI - Below median	-0.161	0.039	0.000
TI - Above median	-0.138	0.038	0.000
TI - Max and above	-0.165	0.038	0.000
Age - 25_30	0.215	0.035	0.000
Age - 31_40	0.130	0.035	0.000
Age - 41_50	-0.300	0.039	0.000
Female	-0.079	0.025	0.002
Constant	-5.528	0.137	0.000

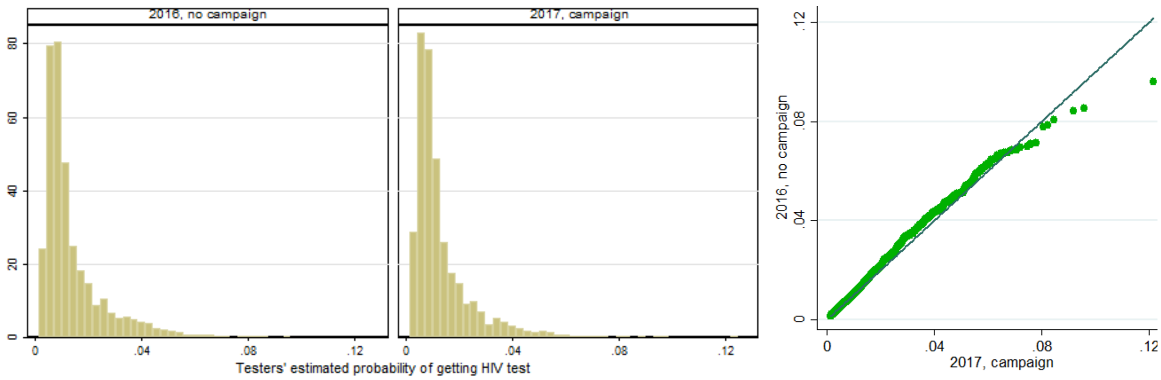
Notes: This table shows the estimated coefficients of demographic and health service use variable of the logit model that predicts the likelihood of getting an HIV test. I use data for weeks 29 to 36 of 2016, when there was no campaign. Region dummies coefficients are ommitted from the table. Number of observations: 1,058,596. Pseudo R2: 0.038.

testers in each group and compare their distributions. As expected, the probabilities are mostly close to zero, due the low predictive power of the model. The results suggest, similarly to the previous section, that marginal testers are similar in terms of observables to testers taking an HIV test without exposure to the campaign. The two-sample Kolmogorov-Smirnov test for equality of distribution functions has a p-value of 0.0076, rejecting the null of equality of distributions, nevertheless, the largest difference between the distribution functions in any direction is 0.0237, a small magnitude. Figure 1.10b shows the findings from a different perspective. Considering that 99 percent of the estimated probabilities fall below 5%, the distributions are quite similar as the quantile-quantile plot falls mostly over the 45 degree line.

FIGURE 1.10: *Testers'* estimated probability of HIV testing

(a) Distributions, density

(b) Quantile-quantile plot



Notes: The figures on the left and center show an histogram of the estimated probabilities of getting an HIV test for the groups of testers in 2016 and 2017, respectively. The estimations come from the logit model. The figure on the right shows a quantile-quantile plot that compares distributions in the previous two figures. Note that 99% of the estimated probabilities fall below 5%, for both groups of testers.

1.5 Impact after the HIV test

The HIV information campaign fosters contact with health care services that is less likely to be driven by individuals' health status or by preference toward health screening. This push to contact health services presents an opportunity for testers to not only screen for HIV infection, but to undergo other preventive health screening tests. This can have intended and unintended consequences on health outcomes and health services use. An intentional result would be to find an increase in HIV detection, depending on whether marginal testers are more or less at risk with respect to other testers. An unintentional result would be the early detection of other health problems unrelated to HIV. In the data I cannot observe test results, nevertheless I observe the use of health services which can shed light on the follow-up and treatment of health problems. I find an increase in the number of reported HIV diagnosis. Although individuals bundle their HIV screening with other check-up services, I find no differences in the use of other health services such as specialist visits or hospitalizations. The latter suggests that marginal testers are not more likely to undergo diagnosis of other health problems.

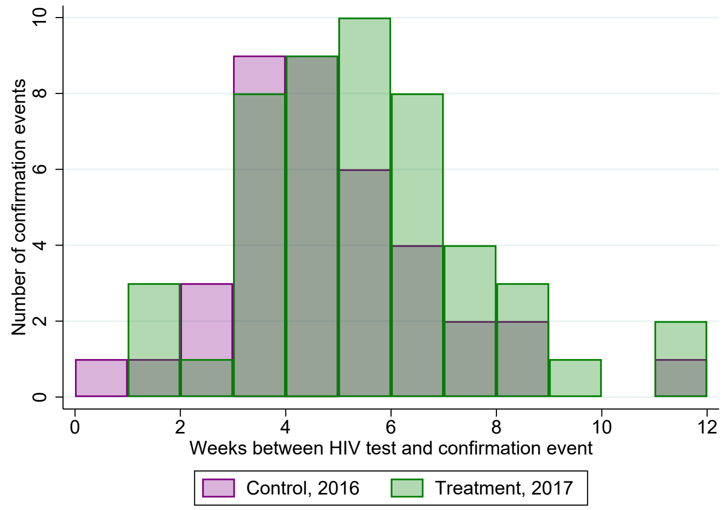
1.5.1 Self-reported diagnosis

Two issues difficult the assessment of the campaign impact on outcomes after the testing event. First, the health services data does not include test results. Nevertheless, I observe the full universe of self-reported diagnosis for the private insured population, where insurees can self-report it with the health insurers to receive health care benefits, such as reduced-price treatment. The self-reporting nature of these records poses two main concerns; individuals may be changing their reporting behavior and they may take too long to report it. Reassuringly for the former concern, during the period under study there was no change in the benefits program that may have directly affected the incentives to registering. For the latter, I identify the HIV test that is closest to the date of the diagnosis report. About 60% of all diagnosed individuals in the period 2012 to 2017 (2,460 over 4,034) take an HIV test at most 60 days prior to registering their diagnosis, suggesting that reporting is spread over time but that a large share of them occur soon after receiving the result.

Second, the information campaign occurs by the end of the period of time under study, which factually restricts the timespan to detect any potential impact. To make a congruent comparison of diagnosis reports, and given the self-reporting nature of the data, I take advantage of the high-frequency data and use only diagnosis reports occurring within 12 weeks of the observed HIV test.

I compare the share and number of diagnosis reports between testers in the treatment and control groups. Diagnosis rates are very small in magnitude, 0.48% for the control and 0.43% for the treatment, and they are not statistically different; the test of proportions yields a p-value of 0.60. Figure 1.11 shows the number of reported diagnosis by weeks relative to the HIV test. In 2016, 38 diagnosis were reported in the first 12 weeks out of 7,892 testers, while in 2017 there were 49 reported diagnosis out of 11,411 testers. This yields an increase of 11 diagnosis in the first 12 weeks after the HIV test.

FIGURE 1.11: Interval between HIV test and diagnosis report, by weeks



Notes: This figure shows the overlapped distribution of the weekly number of reported diagnosis for testers in the treatment and control groups. I include reported diagnosis occurring only within the first twelve weeks after the HIV test took place (week zero). For most of the weeks, the number of reported diagnosis of testers exposed to the campaign (green bars) is higher than those of testers not exposed to the campaign (purple bars).

The increase in reported diagnosis, from 38 to 49, implies an increase of 30% and yields roughly one additional reported diagnosis per week after the campaign. This can be very valuable in the context of a contagious disease such as HIV. The evidence suggests that marginal testers are similar to regular testers, and hence the increase in testing may be pushing forward the diagnosis for a few marginal testers. Although small in magnitude, this increase is large in proportion. Hence, it has a potential great impact, since in the health literature, delays in diagnosis pose the greatest risk of excess mortality for people with HIV. Nakagawa et al. (2012) perform a series of simulations varying the rate of diagnosis. Decreasing the diagnosis rate mechanically increases the number of late diagnosis, this means diagnosis at later stages of the HIV infection. They show that late diagnoses not only reduce the life expectancy of the HIV-positive person but also impact the probability of onward transmission because treatment, which reduces infectivity, is also delayed. Therefore, the campaign may be moving forward in time the diagnosis and can result in decreased risk of death by 10 years from infection and decreasing transmission.

1.5.2 Health services use results

Section 1.4 showed evidence that many testers bundle their HIV test with other health check-up tests. Therefore, an unintentional result of the campaign would be the early detection of other health problems. Individuals undergoing a battery of lab work would increase the knowledge of their health status, which would allow them to receive appropriate health care treatment. This can be investigated through the use of health care services after the HIV test. I consider health services such as doctor visit, specialist visit, preventive care, diagnosis/therapy, surgery, hospitalization, lab test, and mental health services, which are measured as indicators for *any* service use and as the *count* of a type of services, as detailed in section 1.2.2.

I use a difference-in-differences methodology to investigate whether testers encouraged by the campaign use more health services after the HIV test event. I define an event window of 24 weeks before and 12 weeks after the test. I pool all the lags and leads outside of the event window into two dummies: farther lags and farther leads. The main event date corresponds to the date of the individual's HIV test, provided that it occurred between weeks 29 to 36. This takes into account that the encouragement to get an HIV test slowly began when the UNAIDS report was released. Moreover, the event window considers that the campaign occurs very close to the end of the sample period. The model is shown in equation 1.2.

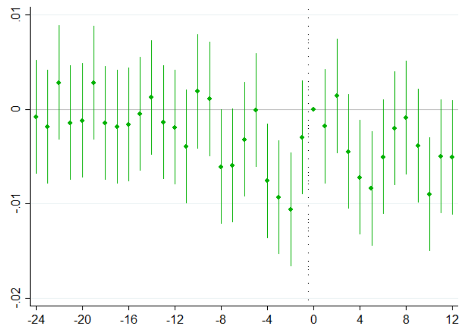
$$y_{it} = \alpha + \gamma Treat_i + \sum_{\tau=T_{pre}}^{T_{post}} \delta_\tau D_{it}^\tau + \sum_{\tau=T_{pre}}^{T_{post}} \beta_\tau D_{it}^\tau Treat_i + \beta X_i + \mu_G + \mu_{h(t)} + \varepsilon_{it} \quad (1.2)$$

D_{it}^τ is an indicator for the relative time with respect to the date of the test. T_{pre} and T_{post} corresponds to the bounds of the interval between 24 weeks prior and 12 weeks a posteriori. $\mu_{h(t)}$ corresponds to calendar time fixed effects. μ_G corresponds to geographic area indicators. The model includes controls for individual level characteristics, X_i, such as age group, civil status, region, income. I also include controls for health services use in the previous year such as any HIV or STD test, and doctor visits, among others.

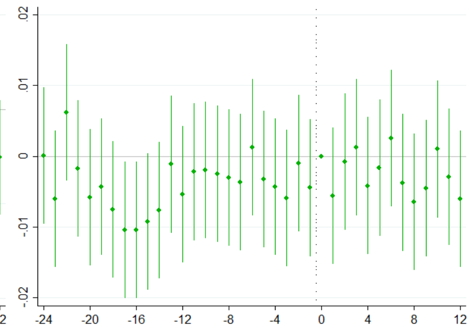
Presumably, marginal testers end up going to the doctor when they would not have gone otherwise. Figure 1.12 shows that testers at the time of the campaign use similar types or levels of health services after the HIV testing with respect to testers not exposed to the campaign. This supports the finding that marginal testers have fairly similar characteristics to regular testers, hence they also have similar use of health services after the campaign. Note that this comparison implies that the contact with health services has no impact for marginal testers relative to the control group. When studying the treated testers under an event study design, results show an increase in certain types of services such as hospitalizations and surgeries after the HIV test. Therefore, these increases also occur in the control group, either by seasonality of health services use or because the two groups share similar characteristics or risks.

FIGURE 1.12: Dynamic event study, weekly

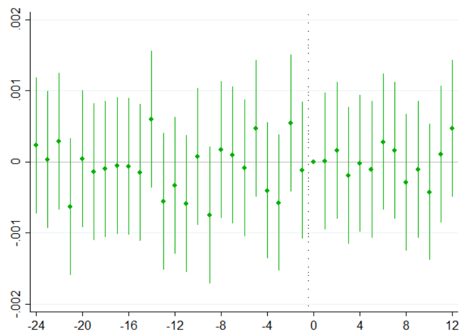
(a) Doctor visit



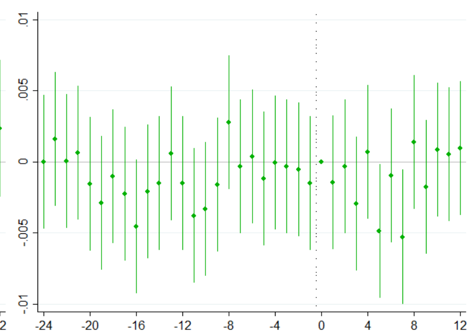
(b) Specialist visit



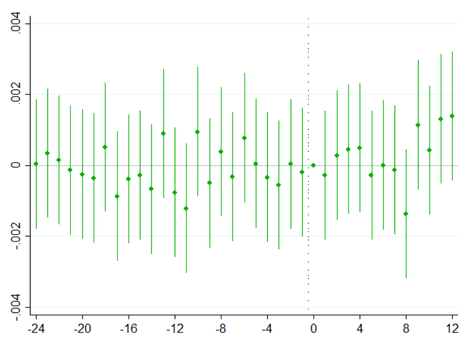
(c) Preventive screening



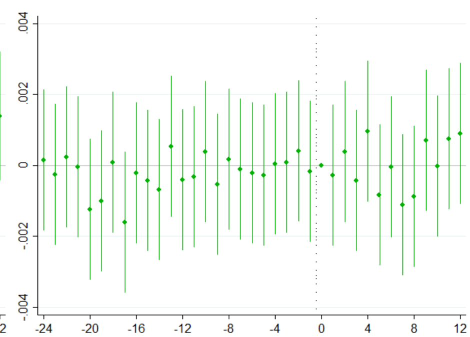
(d) Diagnosis or therapy proc.



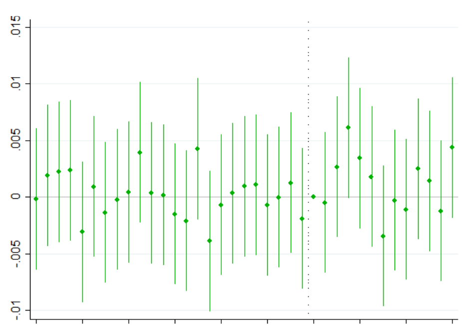
(e) Hospitalization



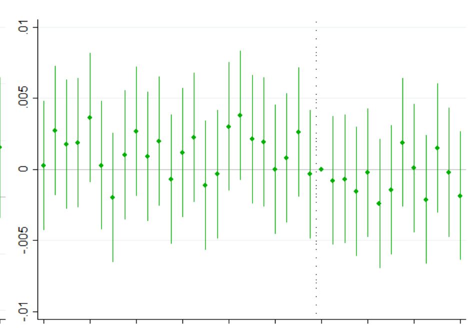
(f) Surgery



(g) Lab tests



(h) Mental health



Notes: The figures above show the $\widehat{\beta_\tau}$ coefficients of equation 1.2 estimated separately for outcomes of types of health services. τ is the relative time with respect to the date of the test for each individual, considering an interval between 24 weeks prior and 12 weeks a posteriori. The model is estimated using OLS and includes calendar time fixed effects, geographic area indicators, and controls for individual level characteristics.

1.6 Conclusion

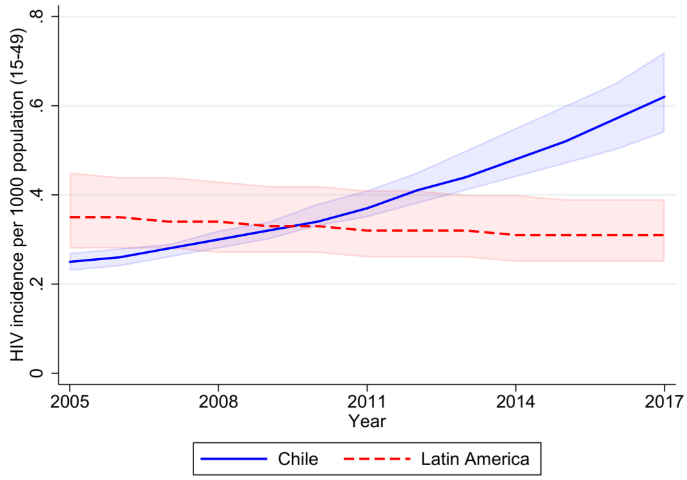
The campaign is effective in its intended goal of increasing HIV testing. Nevertheless, selection has implications for policy-making and to assessing the overall effectiveness of the campaign. The characterization of marginal testers shows that they are quite similar to other testers. Marginal testers who select into testing after being exposed to the campaign have statistically significative differences in terms of demographics and health services use before, at the time of testing, and after the test with respect to other testers. Although the differences are very small in magnitude, they point out to marginal testers belonging to groups usually related with riskier sexual behaviors. Marginal testers are slightly more likely to be repeated testers and to have used any health care in the year before the test. Hence, they may be just moving forward their testing and may benefit from early detection of HIV infection. The latter is important in the context of contagious and transmittable diseases such as HIV, where timing of detection is key for both managing the disease and to contain its spreading.

1.7 Appendix

1.7.1 Figures

FIGURE 1.13: Appendix figures

(a) Trends of HIV incidence



Notes: Incidence is defined as the number of infections in the population and therefore includes both diagnosed and undiagnosed cases. The trends corresponds to UNAIDS estimations.

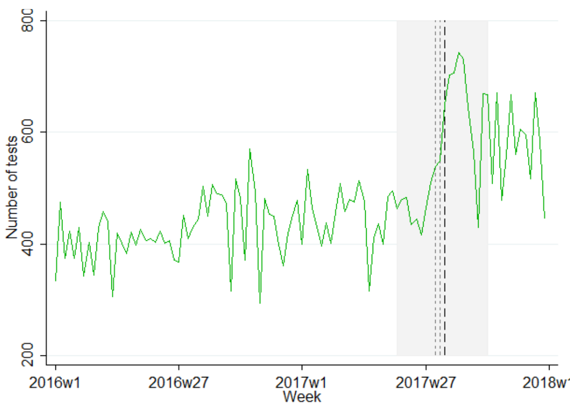
(b) Poster, August 2017

Notes: The slogan's translation reads "The more we ignore it, the stronger it becomes. Use a condom and take the test".

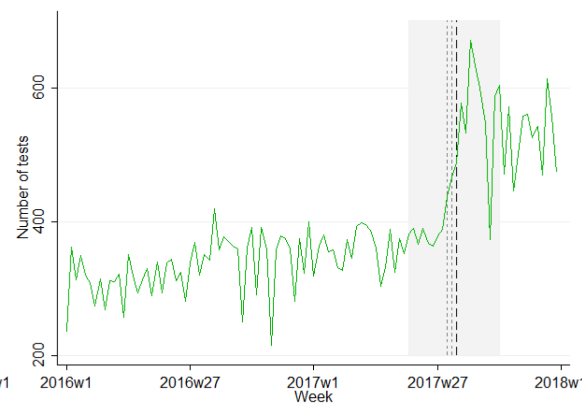
FIGURE 1.13: Appendix figures (continued)

Trends of HIV tests, weekly balanced sample by gender

(c) Men



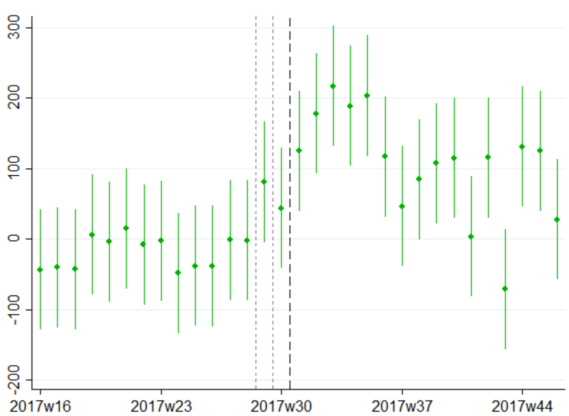
(d) Women



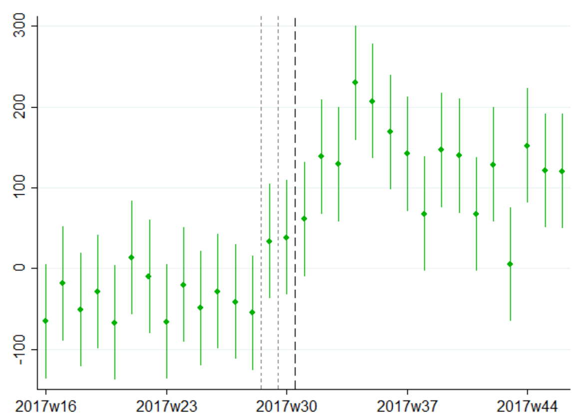
Notes: These figures show the weekly raw trends for the number of HIV tests, by gender. The full sample includes 132,340 individuals, ages 18 to 50, 57% of which are men. The black dashed vertical line corresponds to the campaign launch and the gray dashed vertical lines correspond to the release of the UNAIDS report and the campaign announcement, respectively. The shaded gray area corresponds to the period used in the event study design.

Event study of HIV tests, by gender

(e) Men



(f) Women

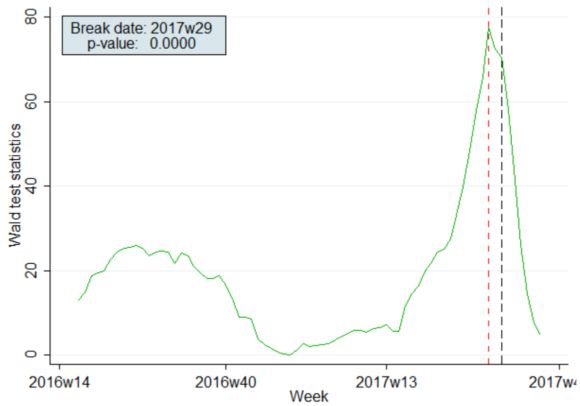


Notes: Each figure corresponds to a separate event study regression by gender and presents the $\widehat{\beta_\tau}$ coefficients of equation 1.1. The coefficients represent the number of additional HIV tests in the weeks around the 2017 campaign launch, pictured as a black dashed vertical line. The gray dashed vertical lines correspond to the release of the UNAIDS report and the campaign announcement. The model is estimated using OLS and includes a time trend and a set of calendar week and year fixed effects.

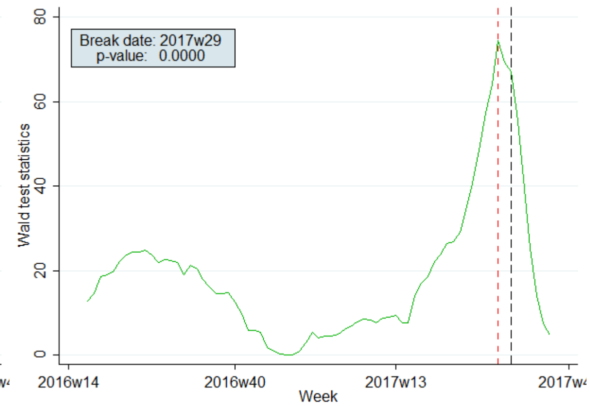
FIGURE 1.13: Appendix figures (continued)

Structural break test of HIV tests' trend

(g) Test for intercept only



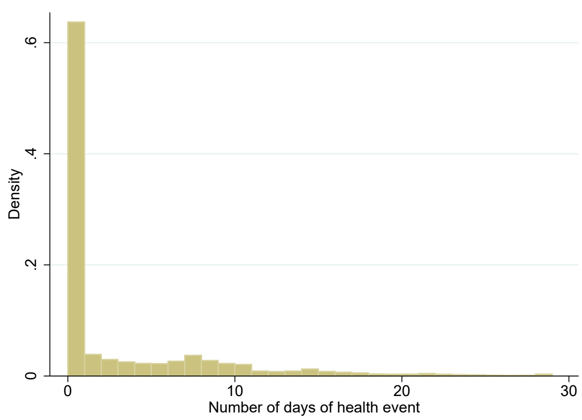
(h) Test for slope only



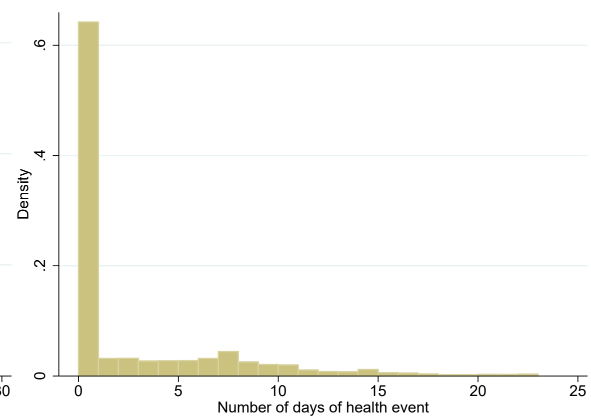
Notes: These figures plot the Wald statistics for a structural break test at each possible break date. The red vertical dashed line shows the estimated break date, displayed in the top-left box along with the p-value for the test under the null hypothesis of no break date. The black vertical dashed line shows the actual campaign launch date.

Number of days within health event

(i) All health services



(j) Health services around HIV test

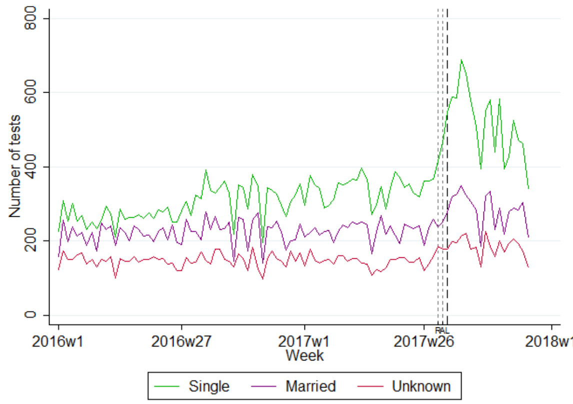


Notes: These figures show the distribution of the health events' length in days. The figure on the left shows the distribution for all health services, while the right figure shows the distribution of a subset of health events where an HIV test occurred.

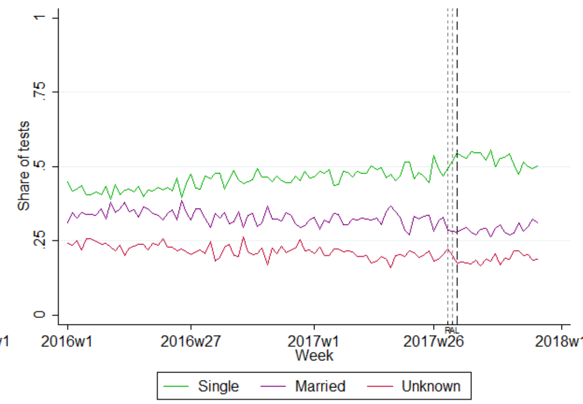
FIGURE 1.13: Appendix figures (continued)

Trends by demographics, level and share

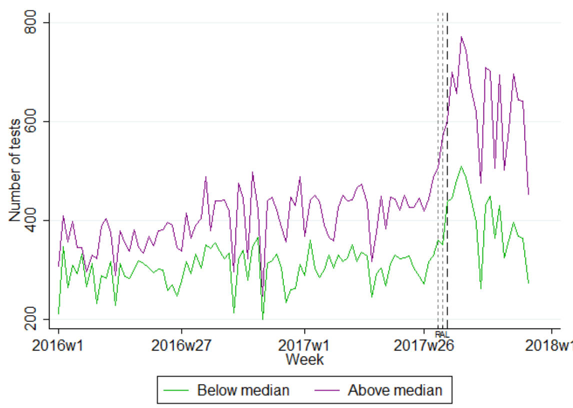
(k) Marital status, level



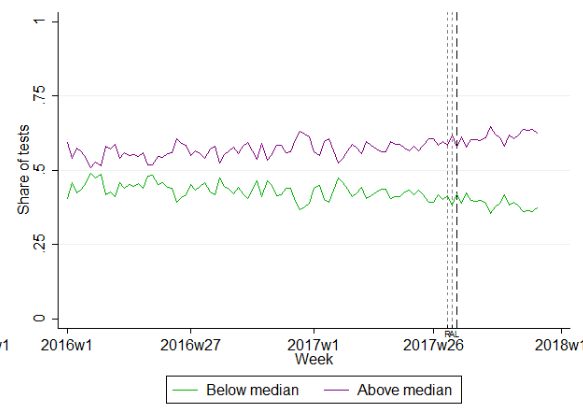
(l) Marital status, share



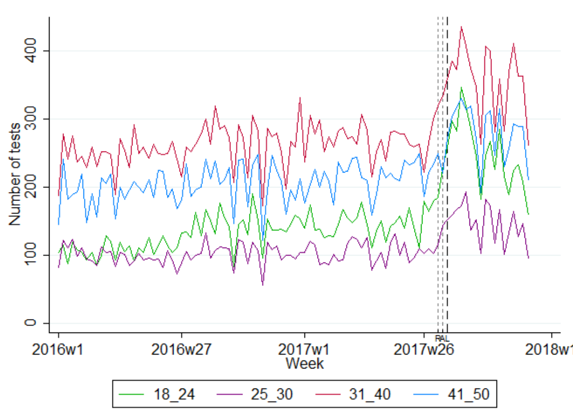
(m) Income, level



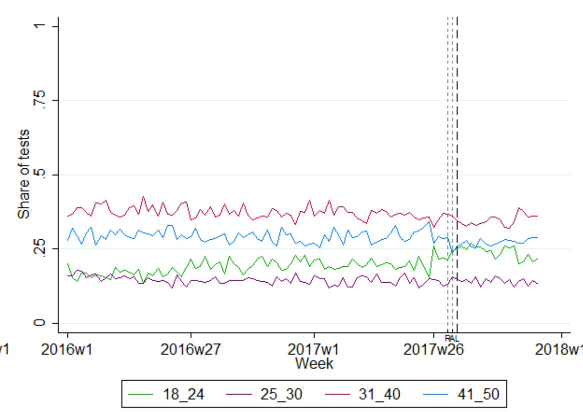
(n) Income, share



(o) Age groups, level



(p) Age groups, share

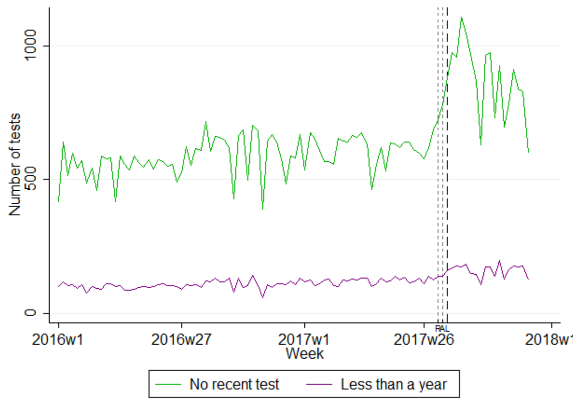


Notes: These figures show weekly raw trends for the number of HIV tests, by demographics. The black dashed vertical line corresponds to the campaign launch and the gray dashed vertical lines correspond to the release of the UNAIDS report and the campaign announcement, respectively. I use a balanced sample of individuals for the period 2015 to 2017. Figures on the left show the trends in levels and those on the right show the trend in shares over the full sample. Figures (k) and (l) break down the trend by marital status, single or married, and a third category is unknown. Figures (m) and (n) break down the sample by income of the head of household. Figures (o) and (p) break down the sample by age groups.

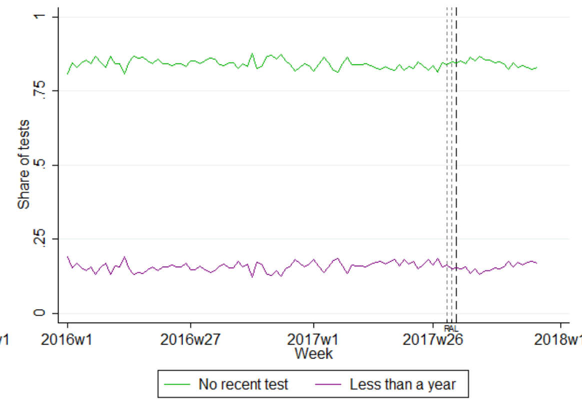
FIGURE 1.13: Appendix figures (continued)

Trends by testing behavior, level and share

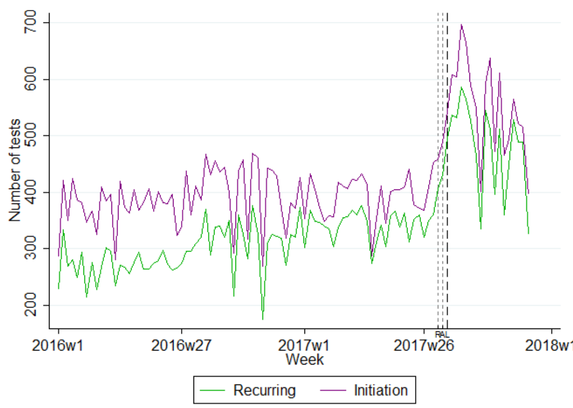
(q) Recent test, level



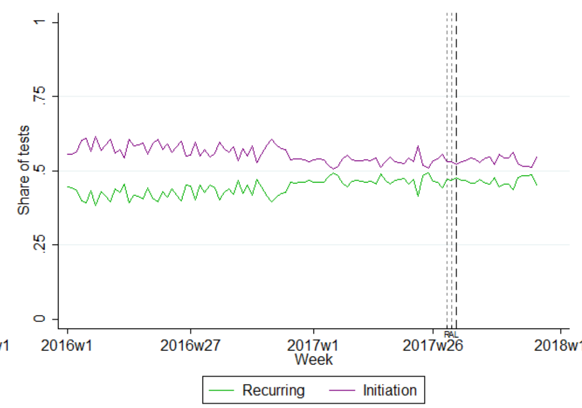
(r) Recent test, share



(s) Initiation, level



(t) Initiation, share



Notes: These figures show weekly raw trends for the number of HIV tests, by testing behavior. The black dashed vertical line corresponds to the campaign launch and the gray dashed vertical lines correspond to the release of the UNAIDS report and the campaign announcement, respectively. I use a balanced sample of individuals for the period 2015 to 2017. Figures on the left show the trends in levels and those on the right show the trend in shares over the full sample. Figures (q) and (r) break down the sample according to whether the individuals had a recent HIV test, that is to say within one year or more. Figures (s) and (t) break down the sample according to whether the individuals had any prior HIV test, namely by recurring testers and first time testers (initiation status).

1.7.2 Tables

TABLE 1.4: Event study of HIV tests

	(1) All		(2) Men		(3) Women	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Constant	329.19	(0.000)	207.72	(0.000)	120.43	(0.000)
Trend	1.50	(0.000)	0.90	(0.000)	0.90	(0.000)
β_{-15}: 2017w16	-101.54	(0.135)	-44.00	(0.311)	-65.50	(0.069)
β_{-14}: 2017w17	-47.94	(0.480)	-40.80	(0.348)	-18.90	(0.599)
β_{-13}: 2017w18	-88.34	(0.194)	-43.00	(0.322)	-51.10	(0.156)
β_{-12}: 2017w19	-35.54	(0.600)	5.80	(0.894)	-28.90	(0.422)
β_{-11}: 2017w20	-50.94	(0.453)	-4.00	(0.927)	-67.10	(0.063)
β_{-10}: 2017w21	4.46	(0.948)	14.60	(0.737)	13.10	(0.716)
β_{-9}: 2017w22	-29.54	(0.663)	-8.20	(0.850)	-10.30	(0.774)
β_{-8}: 2017w23	-74.74	(0.271)	-2.80	(0.949)	-66.10	(0.067)
β_{-7}: 2017w24	-78.34	(0.249)	-48.40	(0.265)	-20.30	(0.572)
β_{-6}: 2017w25	-78.14	(0.250)	-38.20	(0.379)	-49.10	(0.173)
β_{-5}: 2017w26	-67.74	(0.319)	-38.80	(0.372)	-28.50	(0.428)
β_{-4}: 2017w27	-30.34	(0.655)	-1.40	(0.974)	-41.30	(0.251)
β_{-3}: 2017w28	-45.94	(0.499)	-2.40	(0.956)	-55.30	(0.125)
β_{-2}: 2017w29	94.26	(0.166)	80.80	(0.064)	33.70	(0.349)
β_{-1}: 2017w30	76.86	(0.258)	43.80	(0.313)	38.30	(0.287)
β_{0}: 2017w31	179.46	(0.009)	124.80	(0.004)	60.90	(0.091)
β_{1}: 2017w32	276.46	(0.000)	178.00	(0.000)	138.50	(0.000)
β_{2}: 2017w33	337.86	(0.000)	217.20	(0.000)	128.90	(0.000)
β_{3}: 2017w34	386.46	(0.000)	188.60	(0.000)	229.70	(0.000)
β_{4}: 2017w35	365.46	(0.000)	203.20	(0.000)	207.10	(0.000)
β_{5}: 2017w36	251.06	(0.000)	117.00	(0.007)	168.90	(0.000)
β_{6}: 2017w37	188.06	(0.006)	46.40	(0.286)	141.70	(0.000)
β_{7}: 2017w38	164.06	(0.016)	84.40	(0.053)	67.70	(0.061)
β_{8}: 2017w39	226.46	(0.001)	107.60	(0.014)	146.50	(0.000)
β_{9}: 2017w40	236.26	(0.001)	114.60	(0.009)	139.50	(0.000)
β_{10}: 2017w41	50.46	(0.457)	3.00	(0.945)	67.10	(0.063)
β_{11}: 2017w42	225.26	(0.001)	115.60	(0.008)	128.70	(0.000)
β_{12}: 2017w43	-69.34	(0.307)	-71.60	(0.100)	4.90	(0.892)
β_{13}: 2017w44	259.26	(0.000)	131.00	(0.003)	152.10	(0.000)
β_{14}: 2017w45	210.86	(0.002)	124.60	(0.004)	121.10	(0.001)
β_{15}: 2017w46	127.66	(0.061)	27.40	(0.528)	120.10	(0.001)

Notes: This table shows selected coefficients of equation 1.1. The $\widehat{\beta_\tau}$ coefficients represent the number of additional HIV tests in the weeks around the 2017 campaign launch (β_0). The model is estimated using OLS and includes a time trend and a set of calendar week and year fixed effects.

CHAPTER 2

THE EFFECTS OF ACCESS TO SAFE ABORTION METHODS ON WOMEN'S MORTALITY, FERTILITY, EMPLOYMENT, AND BIRTH OUTCOMES

2.1 Introduction

In many developing countries abortion remains illegal or highly legally restricted and the fertility transition is still in the early stages, which reinforces the dichotomy between safe and legal abortion versus unsafe and illegal abortion. A harm-reduction approach to break with this dichotomy is the provision of access to safe abortion methods, within the legal framework. Such an approach has the potential to have broader impacts on women's life choices through an increased certainty about reproduction. Nevertheless, these impacts remain scarcely studied, particularly due to data challenges driven by the illegality of the activity.

A group of doctors in Uruguay united in 2001, when abortion was highly legally restricted, to develop a harm reduction initiative to reduce abortion-related mortality and morbidity using the best medical practices available and without the need of changing the ruling laws. This initiative was known as "Iniciativas Sanitarias" (IS program), translated as "Health Initiatives". Broadly, the IS program had the goal of providing care before and after the abortions to women that had an unplanned pregnancy and were considering abortion as an alternative. The strategy involved a multidisciplinary team who provided information on safe pregnancy termination practices, mainly using misoprostol, conducted health checks before and after the abortion, and offered medical, psychological, and social counseling. They also recommended the use of contraceptives after the abortions were performed. Even though access to misoprostol was restricted, it was sold in pharmacies only when prescribed by a gastroenterologist, Fernandez et al. (2009) report that misoprostol sales in Uruguay rose 261% from 2002 to 2007 compared to a decline of 17% in Latin America. This suggests that women actually changed their behavior following the doctors' protocol.

The IS program might seem like a small impact intervention, however, the IS model has been featured for its success in the New York Times[1], and in partnership with international organizations the model has been replicated in a small scale in nine other countries. In Uruguay, the IS program had a successful initial implementation in 2004 that established the structure and laid the grounds for the path to the legalization of voluntary termination of pregnancy (VTP law) in 2012, hence the VTP law simply legitimized a behavior that was already occurring before. It has been well documented that the IS program decreased maternal mortality substantially (Briozzo et al. (2016)) and that when the VTP law was implemented there were no major effects on women's mortality - it was already low. Similarly, Antón et al. (2016) studied the effect of the VTP law on low income teen mothers fertility[2] and find no effects, attributing this result to the fact that most likely the reproductive behavior changed before the law due to the IS program and a 2010 reform making contraceptives free and widely available for this population group.

This paper studies the impact of access to a safe and comprehensive abortion method on women's life choices, using the Uruguayan case. Our hypothesis is that when access to the IS program is made available, fertility would be reduced, women might increase their labor supply, and birth outcomes might improve. These impacts might occur through several channels. First, women that would otherwise have a baby may have more time to commit to a job related activities in absence of a baby. Second, going over the IS program protocol would likely increase their probability of taking contraceptives after an abortion, therefore making the likelihood of facing an unplanned pregnancy much lower than before.

[1]"From Uruguay a model to make abortion safer", Patrick Adams, New York Times, June 2016.

[2]They exploit the existence of a special protocol for pregnant teenagers that asked and recorded whether the pregnancy was planned or not. Their main assumption is that the VTP law affected only unplanned pregnancies.

Third, women would face fewer negative health shocks, such as hospitalizations, which might have lasting effects on labor and education outcomes, especially if the shocks occur in early adulthood. Fourth, for women that could not risk the consequences of an unsafe abortion, the IS program might introduce a safe alternative that is now feasible. Fifth, the IS program can foster adequate care during pregnancy since it provides an early first contact with health care services for pregnant women, and since only women with wanted pregnancies (which may be more likely to commit to a healthy pregnancy) continue it. Finally, the IS protocol might have a behavioral response through empowering women and through reducing social intolerance towards abortion.

We exploit the timing and geographic variation from the expansion of the IS program in our empirical exercise, using a difference-in-difference design. We begin by showing a 20% decrease in maternal mortality, and a slight decrease in two fertility measures around the time of the implementation of the program. Then, we analyze the effects on birth outcomes and employment. We find that access to safe abortion promotes prenatal care and increases women's employment. This effect was statistically significant for women that are not single. We find no significant effects on hours worked or part-time work. Our findings suggest that even in absence of legalization, access to safe abortion methods can impact women's life choices.

A wide range of literature studies the effects of reproductive policies on women's life choices, mostly focused in contraception and abortion policy changes in the developed world. The widely cited paper from Goldin & Katz (2002) show that access to contraceptive pills lowers the cost of pursuing a career and raises the age at first marriage. More recently, Fischer et al. (2017) show that the closure of abortion clinics in Texas reduced abortions by 20% and increased births by 3%. Currently, many developing countries have recently legalized abortion or are starting to discuss about doing so, and In most cases the discussion is loaded of ethical content. This paper contributes to the literature by analyzing a developing country path to the legalization of abortion and its effects on maternal mortality and women's life choices, since we believe that there is scope for contributing to the debate from a scientific point of view.

The paper is organized as follows. Section 2.2 describes the institutional and legal background in Uruguay, the program implementation, some documented early impacts, and the data sources. Section 2.3 describes the empirical strategy, presents the study groups, and highlights our main assumptions. Section 2.4 presents the impacts on maternal mortality, women's fertility, labor, and birth outcomes. Section 2.5 concludes.

2.2 Background and Data

2.2.1 Unsafe abortion in Latin America and Uruguay

According to the World Health Organization (WHO), unsafe abortion is defined as a procedure for terminating an unintended pregnancy that is carried out lacking one or more of the requirements for medical environment, equipment, and trained health workers. Developing countries have a high prevalence of unsafe abortions, which results in a widespread public health problem from medical complications. In fact, about seven million women were treated for complications of unsafe abortion in the developing world, with around 760,000 women from Latin America and the Caribbean (Singh & Maddow-Zimet (2016)).

Compared to developed countries, developing countries have two characteristics that build up to the high prevalence of unsafe abortion: their legal framework and the fertility transition. First, abortion remains illegal or highly restricted in most developing countries[3], which makes access to safe abortion scarce or inexistent. In Latin America and the Caribbean, more than 97% of women of childbearing age live in countries where abortion is restricted or banned altogether[4] and 95% of abortions are unsafe (Sedgh et al. (2012)). Second, the fertility transition is still an ongoing process in developing countries. In the

[3]"Fact Sheet. Abortion in Latin America And the Caribbean". Guttmacher Institute, 2017 (accessed on March 2020).

[4]"Status of the world's 193 countries and six territories/nonstates, by six abortion-legality categories and three additional legal grounds under which abortion is allowed". Guttmacher Institute, 2017 (accessed on March 2020).

early stage of the transition there is a growing demand for contraceptive methods, while at the same time contraceptive use and availability remains insufficient leading to an increase in unplanned pregnancies (Ahman & Shah (2011)).

The case of Uruguay in the last decades is representative of the developing world, while also becoming differentiate by presenting a process to reduce unsafe abortion. Abortion was illegal since 1938 with a two exceptions: either the lives of the mother or the baby were endangered, or in the case of rape. To this day, the neighboring countries of Argentina and Brazil have that same policy. During the 90's the doctors' association considered unsafe abortions a public health issue since they were the primary cause of maternal mortality nationally. They proceeded to declare that health care providers were obliged to intervene in the periods "before" and "after" the abortion. Unsafe abortions accounted for 28% of the maternal deaths, compared to an average 13% worldwide (Briozzo et al. (2006)), with almost half of these maternal deaths registered in Montevideo. Then, in 2004 the "Health Initiatives" (IS) program was first implemented in Montevideo providing access to safe abortion methods. Soon after it was expanded to the Montevideo metropolitan area, and after the initial success, it was expanded to the other states. This process culminated with the legalization of abortion with the voluntary termination of pregnancy (VTP) law in late 2012, making Uruguay the first country in South America to do so.

2.2.2 "Health Iniciatives" (IS) program

The IS program was developed in 2004 with the goal of reducing the risk and harm of illegal and dangerous abortions by providing information, counseling, and medical supervision to women seeking abortion, within the legal framework. The program was executed by a trained interdisciplinary team of doctors, midwives, psychologists, and nurses. The recommended procedure for pregnancy termination consisted in the use of misoprostol pills, a non-invasive procedure that greatly simplifies the requirements of place, equipment, and health worker skills. Nevertheless, it is still critical for safe abortions to have medical supervision along the process, particularly for dosing, possible side effects, and follow-up health care.

A key feature of the IS program was its provision of health care "before" and "after" the abortion using a three-visit method to achieve its goal, while also deterring future unwanted pregnancies that could end in unsafe abortions. In an initial visit, women undergo pregnancy tests and psychological evaluation. After pregnancy confirmation, they receive confidential advice about a delineated set of available options, including both pregnancy continuation and termination. A few days later, after deciding to abort, women attend the second visit to receive counseling on the avoidable risks and safe methods for the procedure of pregnancy termination. The recommended method is the use of misoprostol pills with detailed directions for use and dosage, although doctors were legally unable to prescribe the pills. Women would just receive information for their safety, not advice nor prescriptions nor publicity, which allowed the IS model be lawful. At this point and taking into account the planned date of abortion, women schedule a third visit for after the abortion. Post-abortion evaluation is a key feature of a safe abortion method. The final third visit consists of an abortion follow-up that allows caregivers to detect complications and to provide contraceptive guidance to promote safe reproductive behaviors.

The initial success of the IS program set in motion almost a decade of program expansions, as shown in table 2.1[5], and regulatory milestones that laid the groundwork for abortion legalization. In early 2004 the IS program was first implemented in the state of Montevideo and resulted, soon after, in reduced maternal mortality at Uruguay's main public maternity hospital (CHPR) and increased awareness among peers and patients, while also gaining support from the main stakeholders: Universidad de la República and the Uruguayan Medical Union. In August 2004 the Ministry of Health recognized unsafe abortion as a Public Health problem under Normative 369, which formalized the IS program as the official protocol to comprehensively treat women with an unwanted pregnancy. In 2006 the IS model expanded to neighboring hospitals within the city, and later on to the whole metropolitan area, including neighboring states Canelones and San José. Then, it expanded to other states and finally, in late 2012 the Congress legalized abortion (VTP law) following the IS model, and also recognizing the right of doctors to conscientious objection. The only changes coming from the VTP law was that misoprostol can now be legally prescribed by doctors to women voluntarily deciding to abort, and that a formal protocol for these cases was enforced. Before, women could get the drug on the black market, or they could get it if the abortion was required for a few legally justified reasons.

TABLE 2.1: Implementation timeline of the IS program

Project	**Health Center**	**State**	**Date**
FIGO/IS	Hospital de la Mujer del CHPR	Montevideo	Apr-04
FIGO/IS	Centro R. Misurraco	Montevideo	Mar-06
FIGO/IS	Centro de Salud de la Costa	Canelones	Apr-06
FIGO/IS	Centro de Salud Jardines del Hipódromo	Montevideo	Apr-08
FIGO/IS	Centro de Salud Giordano	Montevideo	Apr-08
FIGO/IS	Hospital Las Piedras	Canelones	May-08
FIGO/IS	Hospital de San José	San Jose	May-08
FIGO/IS	Hospital de Canelones	Canelones	Aug-08
FIGO/IS	Hospital Departamental de Florida	Florida	Aug-08
OMS/IS	Centro de Salud Jose Royol	Rivera	Jun-10

2.2.3 Documented impacts of the IS program

The IS program, by promoting safe abortion methods, could have induced a transition from unsafe to safe abortions. Provided that such transition occurred, we would expect an increase in the sales of misoprostol and a decrease in maternal mortality. The subsections below present a review of descriptive evidence related to availability of misoprostol and maternal mortality.

Misoprostol sales

An important characteristic of the IS program is the promotion of safe abortion through the prescription of misoprostol pills to be used at home. Even though obtaining this medication

[5] At least three other implementations occurred in other states of Uruguay between December 2010 and 2012, but we do not have the exact dates.

legally was virtually impossible, women could still get it. Before 2012, when abortion was still illegal, doctors were not allowed to prescribe, provide, nor say where to get the drug. Pharmacists were only allowed to sell misoprostol if the drug was prescribed by a gastroenterologist, since the drug was originally created as an anti-ulcer treatment. On the other hand, anecdotal evidence[6] shows that the drug was available in the black market, and could also be purchased in some pharmacies.

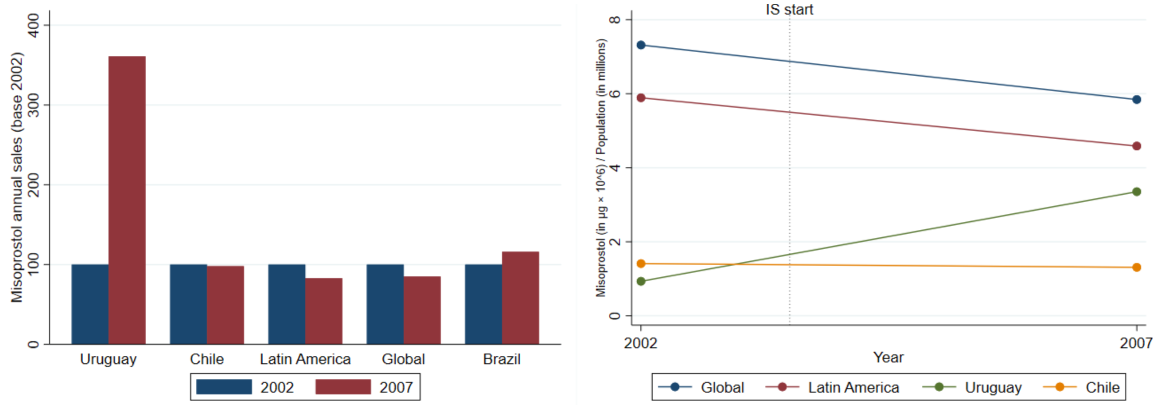
Despite the legal restrictions, misoprostol was shown to become much more available in Uruguay during the implementation of the IS program, possibly enabling women to change their behavior following the doctors' protocol. Fernandez et al. (2009) report that misoprostol sales experienced an enormous increase in Uruguay compared to other Latin American countries between 2002 and 2007, around the time of the IS implementation (see figure 2.1). While misoprostol sales experienced a 17% decline in Latin America, misoprostol sales rose 261% in Uruguay, reaching 3.3 μg/population in 2007. In Brazil they increased just 16% with about 0.3 μg/population, and in Chile there was a 2% decline with about 1.3 μg/population. There is no data for misoprostol-only sales in Argentina, but misoprostol-NSAID sales (drugs that contain misoprostol along with NSAIDs) decreased in the same period by 34%.

More black market activity to obtain the drug would be inferred from an increase in misoprostol sales accompanied by an increase in related web searches. In figure 2.2 we show relative search interest for the terms "Abortion" and "Misoprostol" between 2004 and 2013 in Uruguay, since there is not sufficient data at the state level. Although we do not observe clear spikes in the searches for "Misoprostol", we do observe several spikes for the "Abortion" searches that coincide with several milestones in the abortion legalization process. The last three peaks correspond to November 2008 when the Senate approved the legalization of voluntary termination of pregnancy, but it was immediately vetoed by President Vázquez, to September 2012 when the Voluntary Termination of Pregnancy law was

[6]This comes mostly from newspaper reports about specific cases.

FIGURE 2.1: Misoprostol sales trends in Uruguay and Latin America

(a) Misoprostol sales (b) Misoprostol sales per capita



Notes: This figure plots the misoprostol sales in different regions and countries. In figure (a) we use data from Table 2 of Fernandez et al. (2009) and we use 2002 as a base year for comparison. In figure (b) we plot the ratio of the misoprostol sales and the population, using population data from the World Bank.

enacted, and to June 2013 at the time of a popular consultation to call a referendum against the decriminalization of abortion. This suggests that the abortion legalization process and discussion was quite salient for Uruguayans.

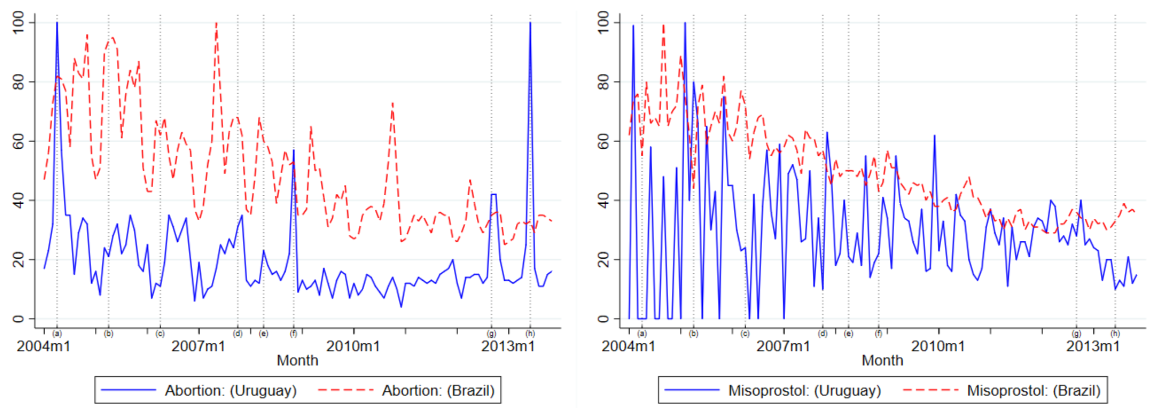
Contraceptives

To prevent the probability that women would face an unwanted pregnancy, the IS program provided information about contraceptive methods at the third visit. Additionally, starting on 2008 contraceptives were distributed for free in Uruguay, which might have increased its use and affected women's reproductive behavior. Currently we do not have data to test this hypothesis. Nevertheless, it is documented that teen pregnancy was still considered a public health issue after this policy change. The authorities argued that this was because there were barriers to the correct use of contraceptives, as traditional methods depend on a strict intake to be effective[7]. Also, as documented in Fiol et al. (2012) for the case of Rivera, the introduction of IS clinics could cause higher awareness among inhabitants, possibly affecting women's reproductive behavior.

[7]El Observador, September 10th, 2016 newspaper.

FIGURE 2.2: Search interest over time in Uruguay: abortion and misoprostol

(a) Abortion (b) Misoprostol



Notes: we use Google Trends data (link) on the relative search interest over time in Uruguay for two terms: Abortion and Misoprostol. The numbers represent relative search interest for a region and time, with 100 being the peak popularity for the term in the period. The vertical dotted lines represent abortion-related milestones: (a) April 2004: rejection of initiative to decriminalize abortion and Montevideo IS implementation; (b) April 2005: President Vázquez announces that if the abortion law is approved, he will veto it; (c) April 2006: Canelones IS implementation (March in Montevideo); (d) October 2007: the Senate discussed a new project to decriminalize abortion; (e) April 2008: Montevideo IS implementation, (f) November 2008: the Senate approved the legalization of voluntary termination of pregnancy, however, it was vetoed by President Vázquez; (g) September 2012: Voluntary Termination of Pregnancy law; (h) June 2013: popular consultation to call a referendum against the decriminalization of abortion that failed.

Hospitalizations

Another possibility is that the IS program could have resulted in fewer hospitalizations associated to abortion complications - as opposed to deaths. This would imply that women would now be facing fewer health complications that could interfere with other activities such as work or education. At this time, we do not have data to test this hypothesis.

2.2.4 Data sources

We use two main data sources. Our primary source of data is a pooled cross-section at the individual level that comes from the Encuesta Continua de Hogares (ECH). This is a quarterly CPS-like survey, nationally representative in both urban and rural areas, and with quarterly frequency from 1998 to 2016. In each year, the sample size is of around 120,000 observations after 2000, and 60,000 in the prior years. The variables included in the ECH comprise individual level information about employment, education, and health, and also household level information, such as total income and characteristics of the house (materials, water supply, sewage, etc). Because many variables have suffered modifications over time, we homogenize them to obtain a balanced panel with the relevant variables.

Our second main data source consists of administrative birth records, at the individual level, from 1999 to 2015 in Uruguay. The geographic unit in this dataset is at the state level. Each birth record includes several variables for birth and pregnancy outcomes, such as birth weight, APGAR score, prenatal visits, and state of birth. It also includes information about the parents such as age, relationship status, and state of residence.

We complement our main data sources with data from other sources. A key piece of information in our empirical strategy is a compilation of a timeline for the timing and locations of states where the IS program expansion occurred over time. We obtained this directly from health workers involved in the IS program as well as from the Uruguayan Ministry of Health. To compute fertility measures we use estimates of the population by state, gender, and 5-year age brackets from the National Institute of Statistics of Uruguay (INE, ine.gub.uy). Finally, to descriptively complement our results, we've also gathered data on mortality by state from the Ministry of Health.

2.3 Research design and outcomes

We use the process of IS program implementation over time to study whether the access to safe abortion practices by the most vulnerable women in Uruguay led to better outcomes. We begin by documenting possible spillovers in health care related to pregnancy and births. We exploit the IS program time and geographic variation to define our study groups and treatment dates, accounting for spillovers in the program's implementation. We pool all the units and perform a difference-in-differences analysis focused on women of childbearing age. Then, we present raw evidence on the decline in the fertility trends from a pooled analysis.

The IS program is a policy impacting reproductive behavior and hence impact women's life choices, by giving women greater certainty about reproduction, even in the absence of legalization. This program can also impact fertility rates by facilitating access to safe abortion methods. Therefore, we expect a decline in the fertility after the implementation of the IS program. Also, with access to a safe abortion method, women who interrupt their pregnancies would have more time to dedicate to work or to school related activities, due to several complementary reasons: improved family planning, higher probability of contraceptives use, fewer hospitalizations, and behavioral empowerment.

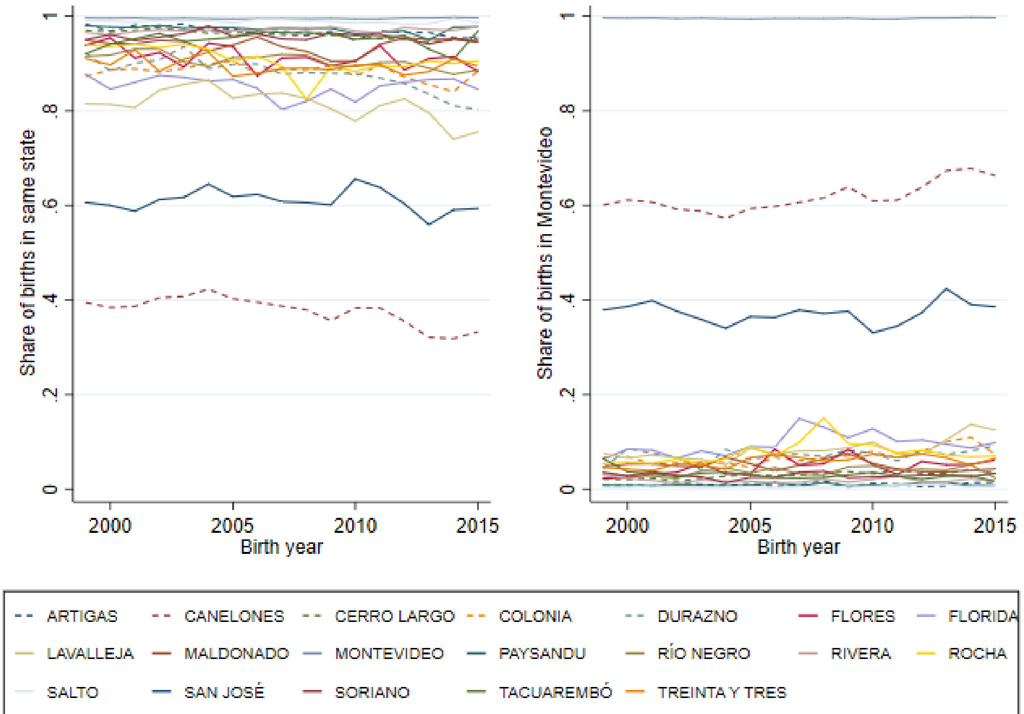
2.3.1 Spillover impacts of the IS program

In order to analyze the overall impacts of the IS program implementation, we need to take into account that Uruguay is a small country and that the largest infant hospital (CHPR) is located in Montevideo. This suggests that implementations in one state, particularly in Montevideo, could have spillover effects in other states. Ideally, we would like to observe where women receive their health care, but we do not have this information. Instead, we use birth level data to study whether women give birth where they reside or not, which can proxy for pregnancy-related health care.

We compute the share of total births in the mother's residence state and in Montevideo, as shown in Figure 2.3, to detect possible spillover effects from Montevideo. In most states the vast majority of births occur in the mother's home state, and the shares remain fairly constant over time, as shown in the left panel with most shares above 80%. Canelones and San José are the exception, with 40% and 60% of births occurring out of state, which exactly corresponds with the share of births that occur in the neighboring state of Montevideo, as shown in the right panel. We use this evidence to make the assumption that the *actual* implementation date of Canelones and San José corresponds to the implementation date of Montevideo[8].

[8] We state this assumption in table 2.2.

FIGURE 2.3: Share of births in the mother's residence state and in Montevideo



Notes: This figure shows the share of births that occur the mother's state of residence (left) or in Montevideo (right). We use administrative birth records data that includes both the mother's residence state and the state where the birth took place.

2.3.2 Difference-in-differences approach

We exploit the temporal and geographic variation stemming from the expansion of the IS program, as shown in table 2.1, to study the impacts of access to safe abortion on the lives of the most vulnerable women in Uruguay. We assign control states to each treatment unit, using in some cases synthetic control methods to select appropriate control units.[9]. We pool the study groups, comprised of treatment and control units, to perform a pooled difference-in-differences analysis.

For each study group we set a treatment date that identifies the first time the program had an effect on the state. We do this in two different ways. On one side, we assign the implementation date of the first occurrence to states with no suspected spillovers from

[9]See Figure 2.11c in Appendix 2.6 for a map of Uruguay and its states. The case studies are available upon request.

other states (Rivera and Florida). On the other side, we assign the implementation date of Montevideo to states in which we suspect, given the evidence on the previous section, that Montevideo's implementation had a spillover effect (San Jose and Canelones). Table 2.2 shows a summary of the study groups and the treatment dates. Note that in our analysis we exclude Montevideo, the largest urban area that comprises 40% of the Uruguayan population, since we do not have a control unit for this state.

TABLE 2.2: Study groups and assigned implementation date

Treatment state	**Control states**	**Treatment date**
Montevideo	–	Apr-04
San Jose	Lavalleja	Apr-04
Canelones	Maldonado	Apr-04
Florida	Colonia, Flores	Aug-08
Rivera	Artigas, Cerro Largo	Jun-10

Our baseline difference-in-differences analysis focuses on women of childbearing age, namely from 15 to 44 years old. The ECH data allows us to, additionally, study a placebo group of older women ages, 46 to 60 years old, that are unlikely to be affected by changes in reproductive policies. As it has been documented in other studies (e.g. Alzúa & Velázquez (2018)) reproductive policies can have heterogeneous impacts across different groups of women. We focus on groups of women by marital status or having previous children, so we split each age group into four subsamples: single, not single, have kids, or do not have kids. Note that the birth records data does not include an exact variable for having kids, but instead we use an indicator for having previous pregnancies.

We study three groups of outcomes that we hypothesize might be affected by the IS program: fertility, labor, and birth outcomes. First, the IS program can impact whether women terminate their pregnancy or not by introducing a safe abortion method, which may impact women that otherwise would not risk to have complications from an unsafe abortion. Therefore, we study fertility outcomes that are computed at the state level using data from administrative birth records and population estimates: the average number of births and the general fertility rate[10]. The general fertility rate corresponds to the number of births per 1,000 women of childbearing age and is computed as the ratio between the number of births of women of childbearing age and the number of women of childbearing age. We use the ECH to compute the shares of women in the subsamples single, not single, have kids, and do not have kids[11]. Note that the latter two subsamples are the best available proxies for the indicator of first time pregnant and had previous pregnancy subsamples in the birth records data.

Second, women in the absence of a baby would have more time to allocate to different activities, such as working. Therefore, we study labor outcomes that include employment, hours worked, and an indicator for part-time work (below 32 hours per week). Finally, the IS program can result in the case that only women with wanted pregnancies end up giving birth and, hence, might be more likely to commit to a healthy pregnancy. or alternatively, can provide an early first contact with health care services for pregnant women. We study four birth outcomes that consist of indicator variables for low birth weight (below 2500 grams), low APGAR score (below 7), adequate prenatal care (at least 5 prenatal visits), and pre-term birth (under 37 weeks).

We begin our analysis with a graphical examination which we implement using a restricted dynamic difference-in-differences model as shown in equation 2.1.

$$y_{its} = \alpha + \lambda_t + \mu_s + \delta Treat_s + \sum_{l=-T}^{T} \tau_l D^l_{ts} + \sum_{l=-T}^{T} \beta_l Treat_s D^l_{ts} + z_{its}\gamma_{ts} + u_{its} \quad (2.1)$$

The unit of observation is woman i at time t in state s, as in the ECH data and the birth records. Note that we use pooled cross-section data. y_{its} is the outcome (birth or labor outcomes), we include fixed effects for calendar time (λ_t) and for geographic areas (μ_s), other individual controls (z_{its}), indicators for the relative time with respect to the

[10]Note that because of varying population across states, the fertility rate is a scaled version of the number of births for the women of childbearing age.

[11]The shares are computed yearly using the ECH and are smoothed over time.

implementation date (D^{l}_{its}), an indicator for treated states ($Treat_s$), and individual-specific errors (u_{its}). The coefficients of interest are the set of β_l, where $l = 0$ is the treatment date. For the fertility outcomes we run a version of this model where the unit of observation is state s at time t, and hence we do not include individual controls. The restricted version of the dynamic difference-in-differences model we estimate allows us to use all the available data while focusing on the periods around the treatment date. We set an event window of 6 periods before and after the implementation date, and we pool all the lags and leads outside of the event window into two dummies: farther lags and farther leads. We only plot the event window coefficients, and we normalize $\tau_{l=-1}$ to zero. From equation 2.1 the relative time set is $T = \{T_{lags}, -6, -5, .., 0, .., 5, 6, T_{leads}\}$. We run the model for the different subsamples by age groups and demographics that were described above.

Additionally, we run a canonical static difference-in-differences model as shown in equation 2.2.

$$y_{its} = \alpha + \lambda_t + \mu_s + \delta Treat_s + \tau Post_{ts} + \beta Treat_s Post_{ts} + z_{its}\gamma_{ts} + u_{its} \qquad (2.2)$$

Equation 2.2 is a simplified version to equation 2.1. The main difference between the two models is that we now assume that there are no pre-trends and that the treatment effect is constant for every periods relative to the implementation date, l. We include calendar time and unit fixed effects, individual controls, and individual-specific errors. We set the lags to zero, imposing no pre-trends, and the leads are replaced by a dummy variable for the period after the implementation, $Post_{ts}$. The coefficient of interest in this model is the interaction term β, and we run the model for the same subsamples.

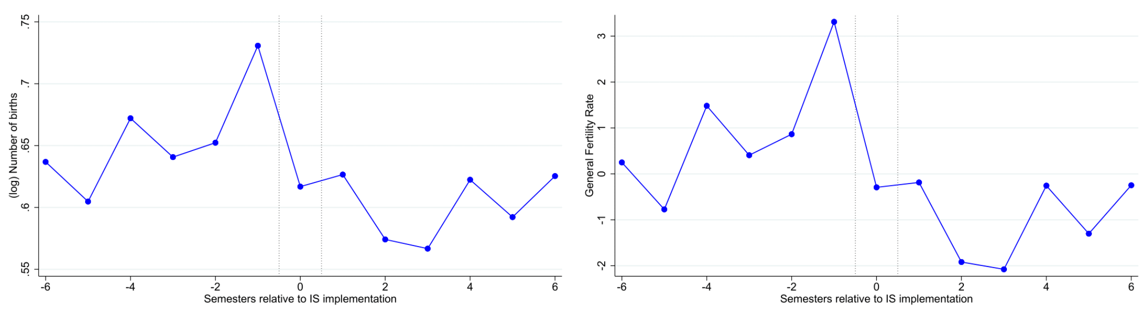
2.3.3 Descriptive statistics: fertility trends

We plot the difference of the averages of the fertility measures between treated and control areas. Figure 2.4 shows the trends relative to the implementation dates with two vertical lines that mark the implementation date and one year after the implementation date. We observe a decline in the average difference for both fertility measures around the treatment date. The general fertility rate shows an increasing pre-trend, and a decline after the implementation.

FIGURE 2.4: Average differences, fertility outcomes

(a) (log) Number of births

(b) General fertility rate



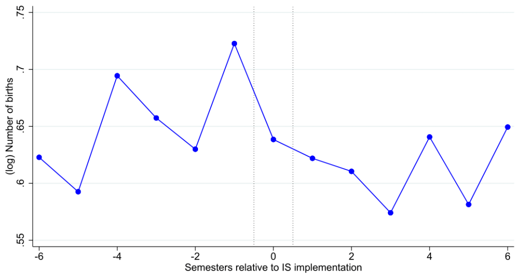
Notes: This figure plots the difference between the average number of births and fertility rate for the treated areas and untreated areas from 1999 to 2015 in Uruguay. We pooled the trends such that time zero corresponds to the implementation date of each state. The two vertical lines that mark the implementation date and one year after it.

Underlying the aggregated trends in figure 2.4 there could be heterogeneous effects for different groups of women. Figure 2.5 shows the trends relative to the implementation dates by subsamples for the number of births and the general fertility rate. For each subsample, we observe similar trends for the two fertility measures. For women that had a previous pregnancy, or for whom it was their first pregnancy or for those that are not single, we observe a decline around the implementation of the IS program. Only for the subsample of single women we detect an increasing trend (most of the time) for both fertility measures.

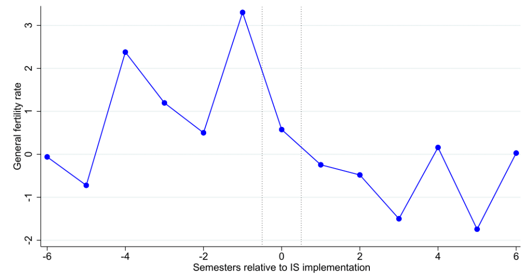
FIGURE 2.5: Average differences, fertility outcomes, by subsamples

Had previous pregnancy

(a) (log) Number of births

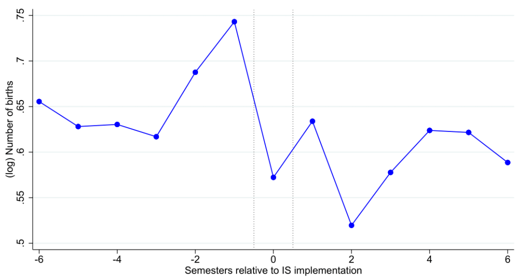


(b) General fertility rate

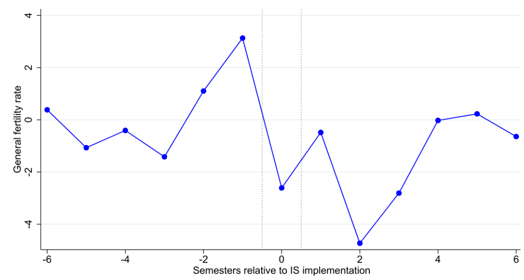


First pregnancy

(c) (log) Number of births

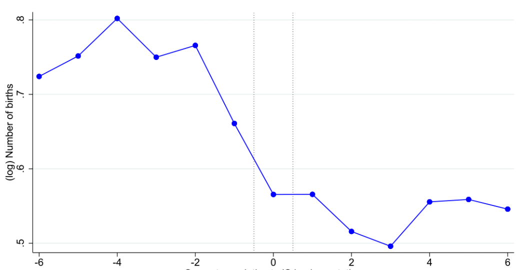


(d) General fertility rate

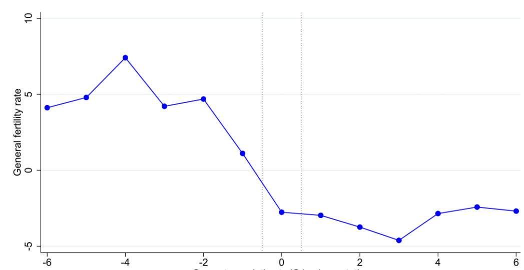


Not single

(e) (log) Number of births

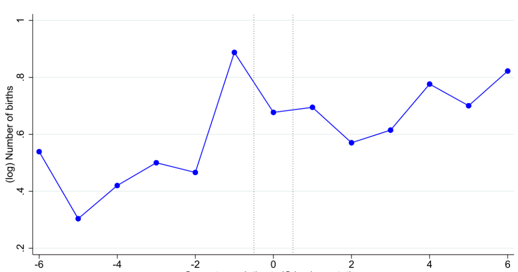


(f) General fertility rate

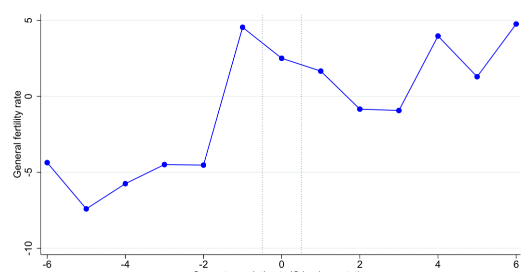


Single

(g) (log) Number of births



(h) General fertility rate

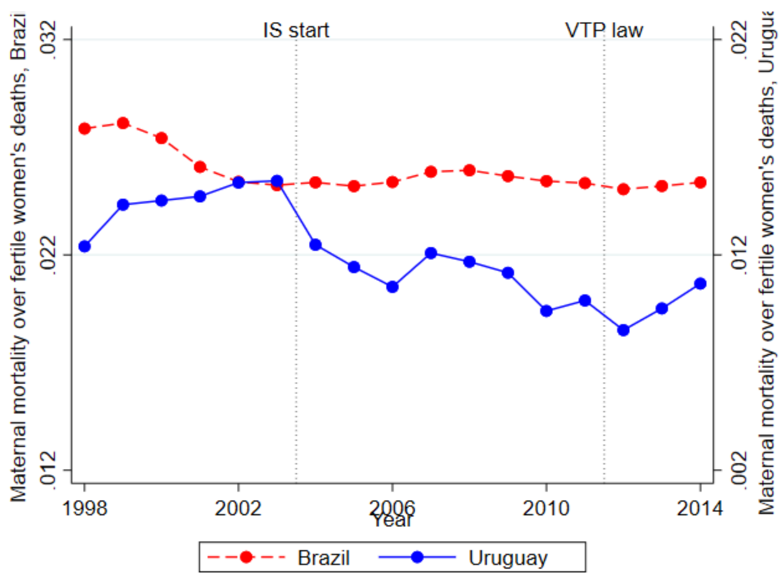


Notes: This figure plots the difference of the average general fertility rate in treated and untreated states, relative to the date of the implementation of the IS program. We use birth records data from 1999 to 2014 in Uruguay. Period zero represents the treatment date. The two vertical lines that mark the implementation date and one year after it.

2.3.4 First stage: maternal mortality

According to Briozzo et al. (2016), the first implementation of the IS program in 2004 caused a decrease in maternal mortality, and given that, when the VTP law was implemented in 2012 there were no major effects on women's mortality - as it was already low. Using uruguayan and brazilian administrative data we produced figure 2.6 that shows the 5-year (two years before, the contemporaneous period, and two periods after) moving averages, using Brazil, a neighbouring country as a control. We can see how maternal mortality, as a share of women of childbearing age mortality, decreased strongly after IS implementation.

FIGURE 2.6: Ratio of maternal mortality over deaths, women of childbearing age



Notes: Own elaboration with administrative data from the Ministry of Public Health, Uruguay, and the Ministry of Health, Brazil. The figure shows 5-years moving averages. The dotted line corresponds to the left axis, and the solid line to the right axis.

Running a difference in differences exercise with the data in figure 2.6, we get the results shown in table 2.3. Taken at face value, and assuming that the exclusion restriction holds, our results show that the ratio of maternal deaths as a share of women of childbearing age decrease by 0.004, which is over 20% of the mean.

TABLE 2.3: Difference in differences estimates, mortality outcomes

	Maternal mortality ratio
Mean	0.019
Coefficient	-0.004
Std. error	(0.001)
Observations	52

Notes: This table shows the estimated coefficient from a OLS regression for the interaction term between a treatment and post dummies, using the data displayed in figure 2.6.

2.4 Impacts of access to safe abortion on women's life choices

In this section, we present the results for our difference-in-differences analysis. As shown in table 2.2, we pooled the treated areas along with their corresponding controls, where the control areas are assigned with the event date of the matched treated areas. We run our analysis on a baseline sample of women of childbearing age, a placebo group of women (only when using ECH data), and four subsamples defined by demographics. Our results are organized in three subsections, each for a group of outcomes: fertility, labor, and birth.

In the subsections below, we begin with a graphical examination from the restricted dynamic difference-in-differences estimation (equation 2.1). We include two vertical lines in each graph that denote the implementation of the IS program and one year after the implementation. Then, we run the regressions for the static difference-in-differences model, which assumes a constant treatment effect and no pre-trends (equation 2.2). We estimate the models using an OLS regression, even for the outcomes that are indicator variables since it simplifies the interpretation and works well for outcomes with a moderate range of probabilities. In our research design, treatment is assigned at the state level on different dates, therefore we cluster standard errors by state and calendar time.

We argue that the impact of the IS program might have a gradual onset between the date of implementation and one year after on the basis of two observations. First, women that got pregnant a few months after the implementation would be impacted by this policy, but we would not observe the impacts on their life choices right away. Second, women that were more than 12 weeks (about 3 months) pregnant at that time would not be affected by this policy change, and hence they wouldn't be able to change their life choices right after the IS program implementation.

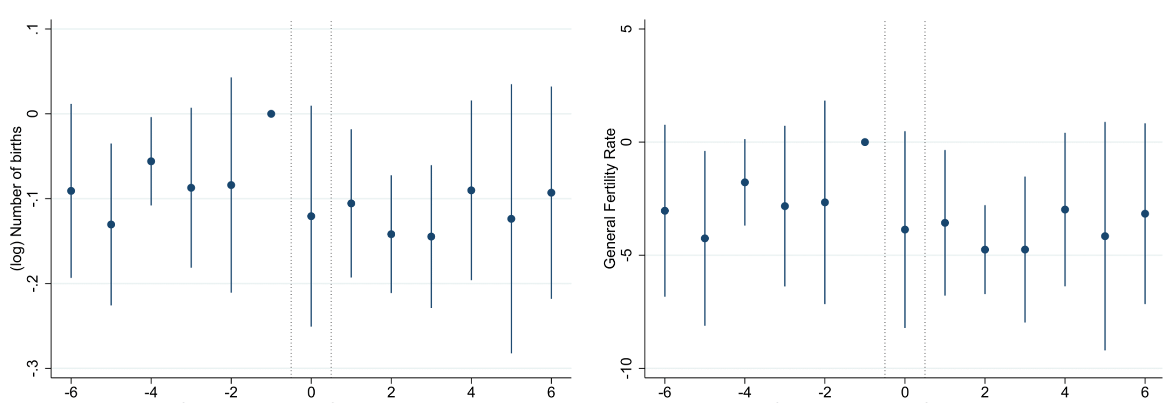
2.4.1 Fertility outcomes

We use the birth records data aggregated at the state level to construct our fertility measures. Figure 2.7 shows the estimated interaction terms in the event window from equation 2.1 for the number of births and the general fertility rate. We observe a decline in both fertility measures around the implementation date. These changes are not statistically significative, but the point estimates suggest that the implementation of the IS program affected the fertility in the treated areas with respect to the untreated areas. We take a closer look at these changes in fertility by looking at four subsamples in figure 2.8. We observe that the impact in figure 2.7 is driven by the decrease in fertility for women that are not single or that had their first pregnancy. The impact is less clear for women that had a previous pregnancy, and it seems to be slightly increasing for single women.

FIGURE 2.7: Difference in differences estimates, fertility outcomes

(a) (log) Number of births

(b) General fertility rate

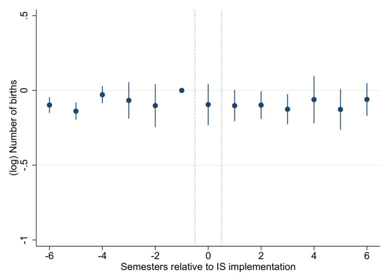


Notes: The figures show the estimated coefficients for the lags and leads of the relative time interacted with a treatment dummy for the fertility outcomes as in equation 2.1, using ordinary least square regressions. Standard errors are clustered by state and calendar time. Period 0 represents the treatment date and we normalize the coefficient for period -1 to zero. We use births records data from 1999 to 2014 in Uruguay.

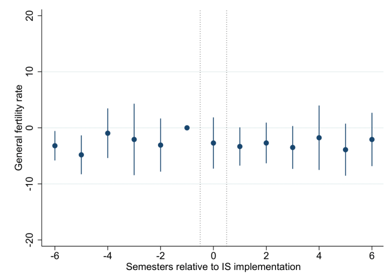
FIGURE 2.8: Difference in differences estimates, fertility outcomes, by subsamples

Had previous pregnancy

(a) Number of births

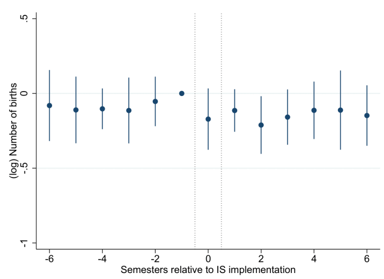


(b) General fertility rate

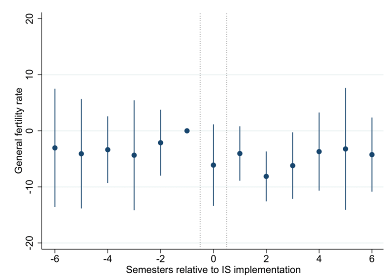


First pregnancy

(c) Number of births

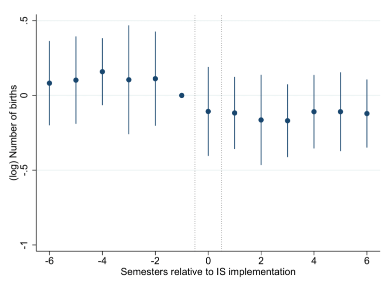


(d) General fertility rate

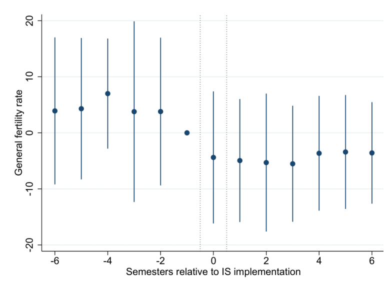


Not single

(e) Number of births

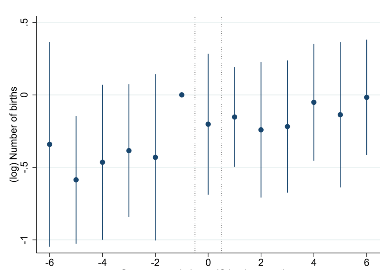


(f) General fertility rate

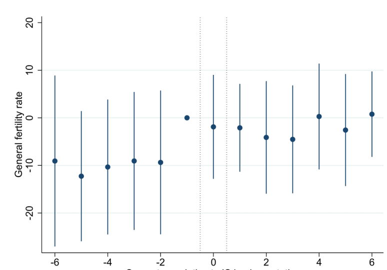


Single

(g) Number of births



(h) General fertility rate



Notes: The figures show the estimated coefficients for the lags and leads of the relative time interacted with a treatment dummy for the fertility outcomes as in equation 2.1, using ordinary least square regressions. Standard errors are clustered by state and calendar time. Period 0 represents the treatment date and we normalize the coefficient for period -1 to zero. We use births records data from 1999 to 2014 in Uruguay.

Table 2.4 shows the estimated coefficient for the interaction term in equation 2.2, where we assume no pre-trends and a constant treatment effect. We observe a decrease of 2% for the number of births and of 1 unit for the fertility rate after the implementation of the IS program for our baseline estimates, although they are not statistically significative. The implementation of the IS program implementation. Similarly to the graphs by subsamples, the impact in table 2.4 is driven by the decrease in fertility for women that are not single or that had their first pregnancy. The subsamples of single women and women that had a previous pregnancy show a positive coefficient on the fertility rate outcome.

TABLE 2.4: Difference in differences estimates, fertility outcomes

	(log) Number of births	Fertility rate
Mean	6.539	34.312
Baseline	-0.020	-1.050
	(0.037)	(0.707)
	340	340
Had previous pregnancy	-0.001	3.000
	(0.042)	(5.616)
	340	340
First pregnancy	-0.056	-3.997
	(0.040)	(4.524)
	340	340
Not single	-0.178	-5.725
	(0.090)	(3.411)
	340	340
Single	0.169	3.733
	(0.159)	(4.490)
	340	340

Notes: This table shows separate estimations of equation 2.2 using ordinary least square regressions, for a baseline sample and for four subsamples. Each estimation is presented in a group of three rows, where the first number shows the estimated coefficient for the interaction term between a treatment and post dummies, standard errors are clustered by state and calendar time are shown below between parenthesis, and the number of observations.

Taken together, the previous results show that fertility was negatively impacted around the time of the implementation of the IS program, particularly for women that are not single or that had their first pregnancy. The IS program provides women with a safe alternative for (emergency) reproductive planning that can impact them differently depending on whether they have support or not to continue with an unplanned pregnancy. In the absence of a safe abortion method, women without support might see pregnancy termination as the only alternative regardless of health risks, in contrast, women that have support might be more likely to continue their pregnancies to avoid health risks. Women who have a partner may be more likely than single women to have support to continue with an unwanted pregnancy. Hence, if the pregnancy is undesired at the time, women who have a partner might respond more to the availability of a safe abortion method and terminate a pregnancy. In line with this hypothesis, initial results from the implementation in Rivera (Fiol et al. (2012)) show that unwanted pregnancies occurred mostly because of interference with their life plans (68%), economic problems (27%), and the absence of a stable partner (27%).

2.4.2 Labor outcomes

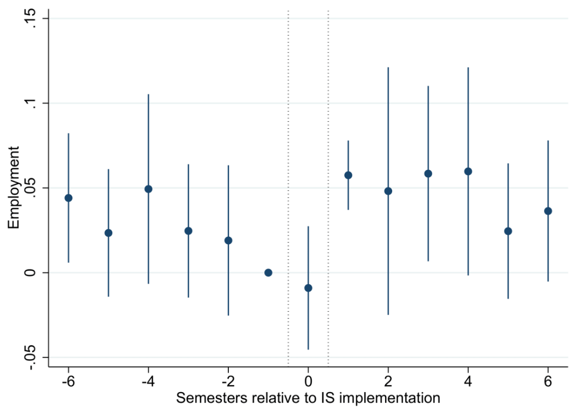
The IS program provides women access to safe abortion methods, and we expect that this would allow women to stay in the labor force or to commit more hours to work in the absence of a baby. The decision whether to continue with their pregnancy could also impact the choice between part-time and full-time employment.

We begin with the estimation of the restricted dynamic difference-in-differences model (equation 2.1) for our three labor outcomes, as shown in figure 2.9. Figure 2.9a shows that the difference in the employment rate was decreasing before the implementation of the IS program, and that it increases afterwards. This jump is large in magnitude and statistically significant, although it looks like it jumped on the semester following the IS program implementation. This can reflect the fact that we would observe a difference only after a woman that would've given birth does not give birth, which would be after one semester of the implementation. Figures 2.9b and 2.9c show no clear changes in the trend of the hours worked and part-time work outcomes, respectively.

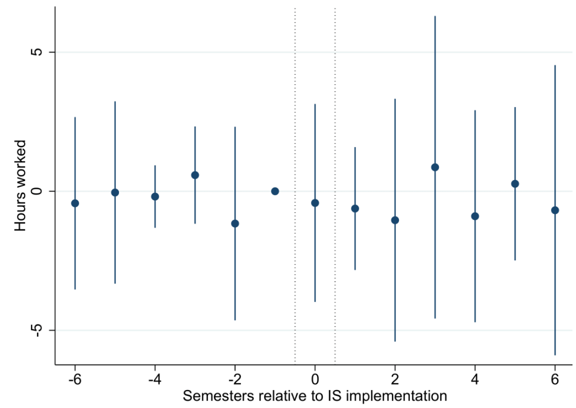
We make a closer examination of our labor outcome results by estimating the static difference-in-differences model (equation 2.2) for the baseline sample and a placebo sample of older women ages 45 to 60, which we use as a falsification test. The latter sample was

FIGURE 2.9: Difference in differences estimates, labor outcomes

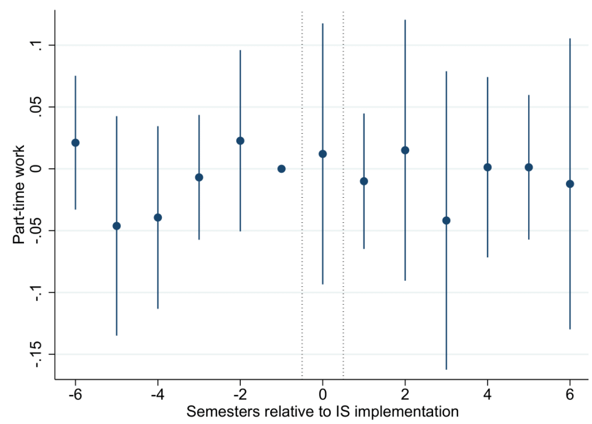
(a) Employment



(b) Hours of work



(c) Part-time work



Notes: The figures show the estimated coefficients for the lags and leads of the relative time interacted with a treatment dummy for the labor outcomes as in equation 2.1, using ordinary least square regressions. Standard errors are clustered by state and calendar time. Period 0 represents the treatment date and we normalize the coefficient for period -1 to zero. We use births records data from 1999 to 2014 in Uruguay.

not available for the birth outcomes analysis, since birth records data only includes a subset of women - those that give birth. Additionally, we disaggregate the baseline and placebo samples into subsamples for marital status and having children before the implementation of the program.

Table 2.5 shows the estimated coefficients of the interaction term from equation 2.2 for all the above mentioned subsamples. The coefficients from the baseline estimation have the expected sign, although only the employment coefficient is statistically significant, therefore we focus on this outcome. We observe that the probability of being employed increases about 3% in treated states after the IS program implementation. The subsample analysis for the employment outcome shows that the effect is driven by women that are not single, aligned with our results for the fertility outcomes. The magnitude of the coefficients are very similar for the subsamples of women with or without previous children. The estimated coefficients for the placebo sample of women age 46-60 and their subsamples are mostly much smaller in magnitude than the childbearing age group and not statistically different than zero .

TABLE 2.5: Difference in differences estimates, labor outcomes

	Employment		Hours worked		Part-time work	
	(1)	(2)	(3)	(4)	(5)	(6)
	Age: 16-45	Age: 46-60	Age: 16-45	Age: 46-60	Age: 16-45	Age: 46-60
Mean (baseline)	0.568	0.611	37.150	37.835	0.369	0.369
Baseline	0.032	0.003	0.404	0.259	0.005	-0.004
	(0.012)	(0.013)	(0.825)	(1.300)	(0.023)	(0.031)
	95,824	41,325	54,469	25,249	54,469	25,249
Kids under 14	0.036	0.042	0.563	1.464	0.004	0.007
	(0.012)	(0.027)	(1.080)	(1.693)	(0.027)	(0.044)
	62,363	13,928	34,133	8,859	34,133	8,859
No kids under 14	0.038	-0.004	0.081	0.461	0.008	-0.021
	(0.022)	(0.016)	(1.079)	(1.332)	(0.033)	(0.031)
	33,461	27,397	20,336	16,390	20,336	16,390
Not single	0.041	0.008	1.065	0.691	-0.010	-0.011
	(0.009)	(0.013)	(0.956)	(1.323)	(0.025)	(0.030)
	54,410	29,360	33,101	16,807	33,101	16,807
Single	0.022	-0.009	-0.413	-0.565	0.023	0.012
	(0.019)	(0.032)	(0.847)	(1.722)	(0.027)	(0.049)
	41,414	11,965	21,368	8,442	21,368	8,442

Notes: This table shows the estimated coefficient for the interaction term between a treatment and post dummies as in equation 2.2, using ordinary least square regressions, for a baseline sample and four subsamples. Standard errors are clustered by state and calendar time. Each group of three rows presents the estimated interaction term, the standard error, and the number of observations for a separate regression. The first three rows present the main estimates in odd-numbered columns, with placebo samples in even-numbered columns, and the rows below show the subsample estimates for two dummies.

Our dynamic and static analysis shows a positive impact on employment that is not observed on our falsification sample, and where the effect is driven by women that are not single, aligned with our results for the fertility outcomes. This suggests that the IS program could have an impact through women having an additional alternative for family planning that was safe that would have allowed them to start a family (or having another child) at the time that was most convenient for them.

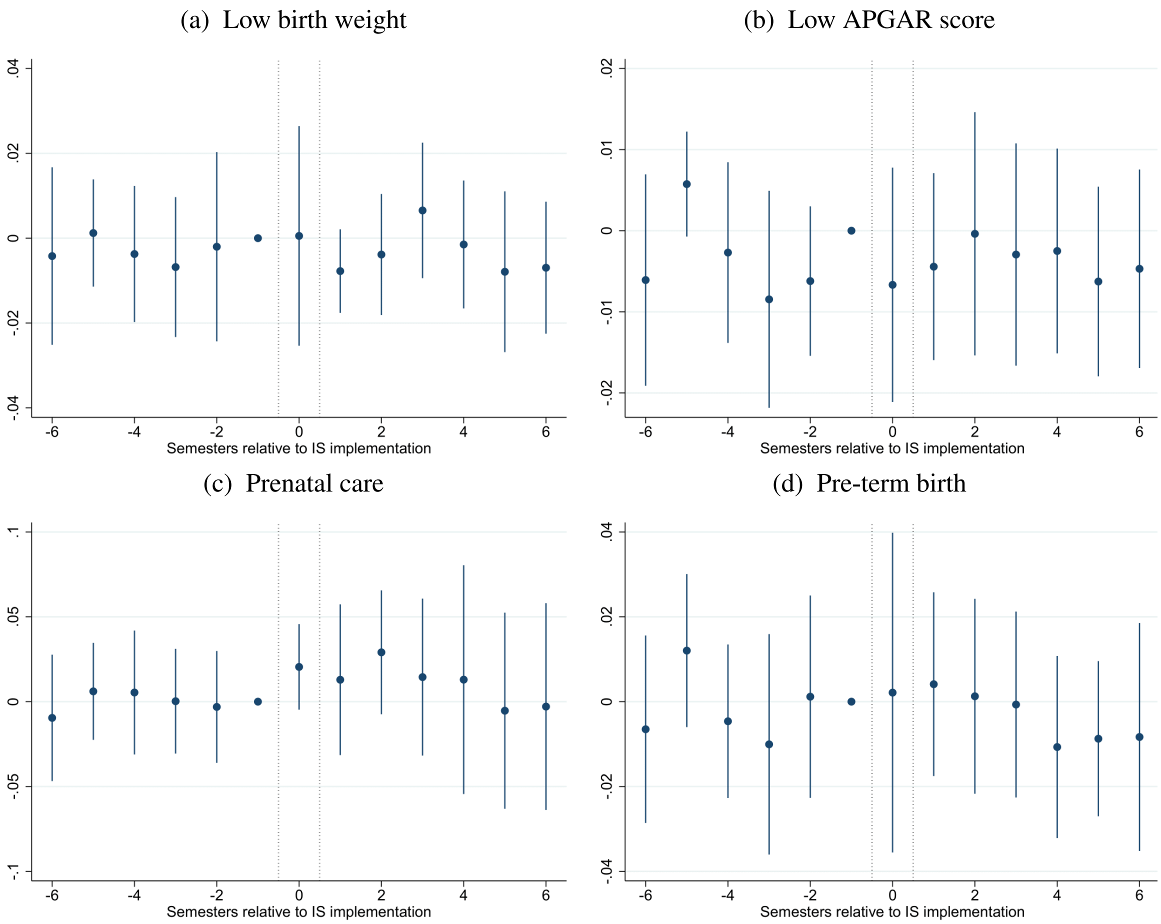
2.4.3 Birth outcomes

We have discussed how the IS program assists pregnant women and provides information about the options women have regarding a pregnancy, including both pregnancy termination and continuation. One possible impact of this provision of information and access to a safe abortion method is that it can lower the cost of pregnancy termination in terms of health risks and access to medical supervision. Impacts like these can affect which women end up giving birth after the IS program implementation, as we showed in the fertility outcomes results. Presumably, this would incite women who are in a better place for having children to do so which would translate into better birth outcomes.

We argue that the IS program could be improving birth outcomes by changing the composition of pregnant women through two channels. On one side, pregnant women would be more likely to have wanted pregnancies, and hence to have healthier pregnancies that lead to better birth outcomes. On the other side, for women that considered an abortion but decided to continue with their pregnancies, the IS program provides an early first contact with pregnancy-related health care services that might foster adequate prenatal care.

We begin by examining the path of the treatment effect relative to the implementation date for our four birth outcomes, from the estimation of the dynamic difference-in-differences model (equation 2.1). Figure 2.10 presents suggestive evidence that the implementation of the IS program had a positive impact on some birth outcomes. Figure 2.10a shows a decrease in the incidence of births with low birth weight right after the implementation, but it fluctuates afterwards. Figure 2.10c shows that recommended prenatal care had an increase around the implementation date. We do not observe a clear change in the trend for the low APGAR score indicator or for pre-term births, although, the latter shows a decrease afterwards.

FIGURE 2.10: Difference in differences estimates, birth outcomes



(a) Low birth weight (b) Low APGAR score (c) Prenatal care (d) Pre-term birth

Notes: We plot the estimated coefficients for the lags and leads of the relative time interacted with a treatment dummy for the birth outcomes as in equation 2.1, using ordinary least square regressions. Standard errors are clustered by state and calendar time. Period 0 represents the treatment date and we normalize the coefficient for period -1 to zero. We use births records from 1999 to 2014 in Uruguay.

Then, we estimate equation 2.2 for the birth outcomes for the baseline sample and by subsamples. The sign of the coefficients in the baseline estimation shows that the IS program had a positive impact on birth outcomes, although they are not statistically significative. We observe a decline in births with low birth weight, low APGAR score, or that are pre-term, while prenatal care shows a positive coefficient. The subsample analysis shows that the effect is driven, for most outcomes, by women that had a previous pregnancy. For the marital status categorization, prenatal care improvement is driven by women that are not single, while the other outcomes show larger impacts for single women. The latter is striking since single women showed an increase in fertility measures, and hence there are more pregnancies with better birth outcomes.

TABLE 2.6: Difference in differences estimates, birth outcomes

	Low birth weight	Low APGAR score	Prenatal care	Pre-term birth
	(1)	(2)	(3)	(4)
Mean (baseline)	0.079	0.045	0.794	0.089
Baseline	-0.003	-0.007	0.014	-0.005
	(0.003)	(0.004)	(0.017)	(0.004)
	272,469	271,025	270,752	269,085
Had previous pregnancy	-0.003	-0.005	0.020	-0.004
	(0.004)	(0.004)	(0.017)	(0.004)
	175,641	174,506	174,761	173,454
First pregnancy	-0.001	-0.009	0.006	-0.006
	(0.005)	(0.006)	(0.020)	(0.007)
	95,329	95,059	95,090	94,138
Not single	-0.001	-0.005	0.033	-0.004
	(0.005)	(0.005)	(0.018)	(0.005)
	181,506	180,639	180,503	179,240
Single	-0.012	-0.013	0.002	-0.012
	(0.005)	(0.003)	(0.016)	(0.006)
	81,288	80,782	80,642	80,190

Notes: This table shows the estimated coefficient for the interaction term between a treatment and post dummies as in equation 2.2, using ordinary least square regressions, for a baseline sample and four subsamples. Standard errors are clustered by state and calendar time. Each group of three rows presents the estimated interaction term, the standard error, and the number of observations for a separate regression. The first three rows present the main estimates and the rows below show the subsample estimates for two dummies.

Overall, our results show that after the IS program implementation took place fertility decreased while birth outcomes slightly improved. We argue that the latter might occur due to a change in the composition of women that give birth, presumably fostering the pregnancies of women that actually wanted them at that time. For the implementation in Rivera, Fiol et al. (2012) show 7 out the first 87 treated women decided to continue their pregnancies after receiving the abortion counseling, therefore it is plausible that the IS program might have been an early first contact with health care services for those women that built up to better birth outcomes.

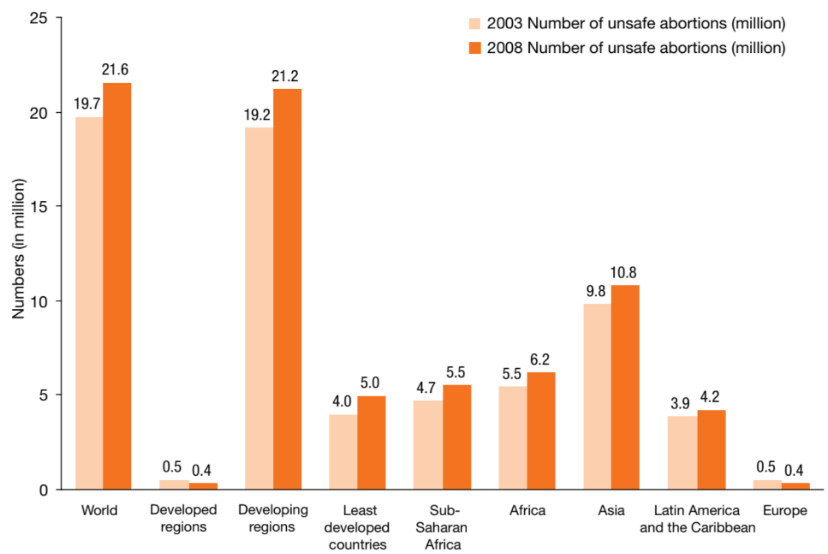
2.5 Discussion

Currently, a dichotomy exists between safe and legal abortion versus unsafe and illegal abortion, but this does not need to be case necessarily. In this paper we study the IS program in Uruguay that aimed to reduce abortion risks using the best medical practices available and without the need of changing the ruling laws. Our results show a decrease in fertility and a positive impact on women's birth outcomes and employment, particularly for not single women. These results suggest that even without the legalization of abortion the access to safe abortion methods can impact women's life choices. A comprehensive assessment of this policy would require to study other aspects of women's lives, such as their educational attainment.

2.6 Appendix

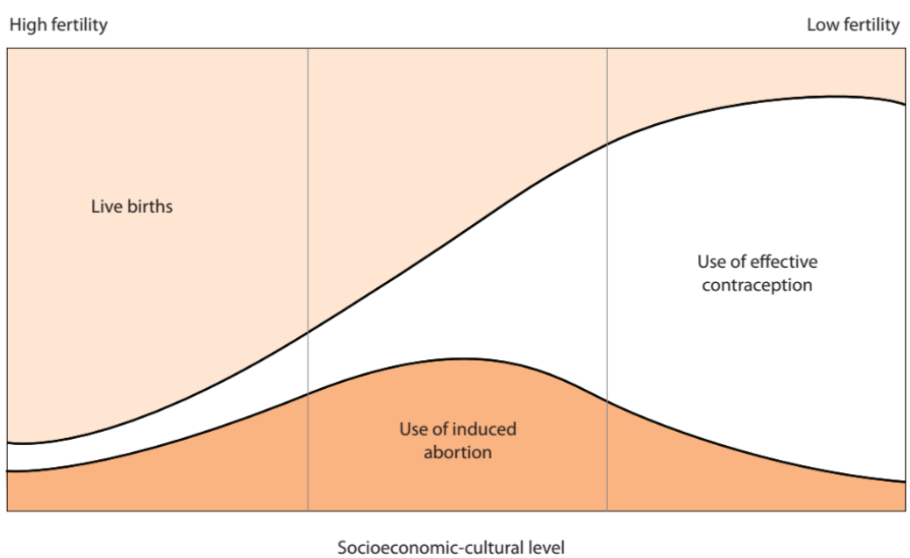
FIGURE 2.11: Appendix figures

(a) Estimated number of unsafe abortions, globally and major regions



Source: Figure 1 in Ahman & Shah (2011).

(b) Transition trajectory from high to low fertility



Source: Figure 3 in Ahman & Shah (2011) (adaptation from Requena (1966)). The figure hows the relative levels of induced abortion, effective contraception and live births.

FIGURE 2.11: Appendix figures (continued)

(c) Map of Uruguay

Source: Google images.

CHAPTER 3

DRIVERS OF PORTFOLIO CHOICE: EVIDENCE FROM RETIREMENT SAVINGS ACCOUNTS

3.1 Introduction

Individuals in a defined contribution (DC) pension system face complex financial choices which they can circumvent in two broad ways: outsourcing their investment decisions or independently by using heuristics. In the former group, return chasing is a common driver among independent investors that arises from the mistaken belief that past performance will persist in the future. Among the latter group, investment decisions are often driven by financial service providers, which are becoming widely popular despite the fact that it is hard for individuals to distinguish good from bad advice. Many drivers and heuristics behind these choices have been identified in the literature along with their wealth accumulation implications. Nevertheless, they have often been studied in settings where individuals face a wide set of portfolio choices, which hampers the measurement of the portfolio returns of the choice set, and thus the characterization of these drivers and its impacts.

In this paper, I characterize followers and return chasers, and analyze the impact of active choices on wealth accumulation. Among followers of financial advice, I find that high income men are more likely to follow financial advice and to keep following future recommendations. I show that the returns of the past three months before the switch drive the return chasing behavior, when returns increase about 0.0069 prior to reallocation. Men and high income individuals tend to switch after larger return differences. Analyzing the cumulative returns after each switch, I find that followers earn similar cumulative returns than the ones they would get by staying in the default option. For return chasers there are not significant differences two months after the rebalance, but after a year, low income return chasers perform significantly worse than higher income individuals.

To study following financial advice and return chasing as drivers of investment choices and its impact, I turn to the Chilean defined contribution pension system which has two main advantages. First, it is a simple setting where portfolio returns of the choice set are easily tractable. This is due to several features: mandatory participation for formal workers, savings must be allocated in at most two out of five funds otherwise they are assigned to a default investment plan, and savings cannot be withdrawn until retirement. Also, each of the funds is offered by all the five firms in the market and the returns across firms are highly correlated, hence the most relevant decision is about the portfolio choice and not the firm choice. Second, there is a main financial advisor that is widely known. "Felices y Forrados" (FyF) were by far the most popular financial advisor in the period studied and had many subscribers to their services. Their recommendations consist of timely advice to switch between funds, which were easily tractable in this context. All their subscribers receive the same recommendation at the same time.

The main dataset I use comes from a representative panel of monthly retirement savings accounts over seven years, that includes individuals' monthly balances for each fund and the firm in which they are contributing. This allows me to identify when individuals make active choices, and when they stay in the default investment option. Even though the set of demographic variables in this dataset is limited, it includes gender, income, and age, which I use to explore heterogeneity in the drivers. I complement this dataset with return data at the daily level for each fund and each firm. Additionally, I collected information about the content of the recommendations from the financial advisor's website. This information is publicly available after the date in which the recommendation was e-mailed to all the subscribers of the service.

I begin by showing that I can identify switches related to FyF's recommendations, and then I characterize the followers of financial advice. First, I construct an algorithm to identify the followers based on the match of the direction and timing of the individual switches and FyF's recommendations. Second, I construct two statistics to portray the followers: (i) the share of individuals that ever follow a piece of advice, and (ii) the share of recommendations they follow in the next months after they were labeled as followers. I estimate these statistics on demographic groups of age, gender, high or low income (above or below the median), and their interactions. I find that high income males are more likely to be followers, and that after being identified as followers they follow 14.1% more recommendations than other demographic groups. Also, they keep following advice more consistently over time. Women follow half the share of future advice with respect to men. There are some differences in the age groups, with older males that are followers following fewer recommendations in the future.

Next, I characterize the return chasers using a statistic motivated by the individuals' actual behavior. I use an event study at the time of the switch to show that individuals follow past return trends. Specifically, they reallocate funds when their current portfolio performs worse in the last three months prior to the switch. Then, I measure this behavior in a more systematic way using a statistic. I construct a statistic for the average return difference between destination and origin portfolio in the last three months before the reallocation. I estimate the statistic using two different methods to address the fact that past returns are correlated within months, and also between months. First, I aggregate the data to compute the net flow between the most conservative (fund E) and most risky (fund A) funds for each date. Second, I use subsamples of the panel data such that the switches do not overlap in the period prior to the switch. Using both methods, I regress the statistic on the demographic groups and their interactions to explore heterogeneity among these dimensions. The estimation of the statistic shows that individuals in this market exhibit return chasing behavior. In the aggregate case, the statistic shows that switches from fund E to A are preceded by an increase in the return difference of 0.018. The analysis of the subsamples of the panel data shows that for all fund reallocations this jump in the return difference prior to the switch is of 0.0069. This behavior plays a higher role among high income individuals. In contrast to previous studies, there are no significant differences by gender.

To explore whether these active decisions hurt individuals, I compare the cumulative returns after the rebalance to the returns of a benchmark, the default investment option. I compare cumulative returns for the short and long run, defined as two and twelve months after the switch, for both the followers and the return chasers. Since I do not know the exact day in which the switches occur, I compare cumulative returns assuming individuals switch in either the first or last day of the month, and I show that the results hold in both cases. For the followers of financial advice, the frequency with which they follow recommendations and how fast they react to the recommendations can have different impacts on their wealth accumulation. I compare the cumulative returns after they follow a recommendation for the first time in the sample, and additionally, I use an event study at the daily level to study how the advisors' strategy affects cumulative returns. For the return chasers, the frequency with which they rebalance their portfolio or pay attention to the past returns, can affect their retirement wealth. I compare their cumulative returns after the first time I observe them switch funds.

Overall, the followers and return chasers' active choices lead to lower cumulative returns on average than their default investment option. For the followers, I do not find significant differences over the long run by income groups. I find that a two-day delay in the implementation of the advisors' recommended strategy decreases significantly the cumulative returns of the advisor's recommended portfolio. For the return chasers, I find that they do as good as the default strategy in the short run, but they do worse in the long run. Twelve months after each switch, high income individuals' losses are almost three times smaller than those of low income individuals. This pattern is not observed among followers.

This paper contributes to several different literatures. First, I contribute to the literature on individual decision making in an investment and retirement savings context. Several studies show that individuals often rely on heuristics to make complex financial choices, and that the framing of the alternatives has large impacts in the investment decisions (Benartzi & Thaler (2007), Benartzi & Thaler (2001), Egan et al. (2016), Barber & Odean (2008)). This is in close relation to the documented fact that most individuals stay in the default option when available (Madrian & Shea (2001), Choi et al. (2002)). Engström & Westerberg (2003) study the implementation of the DC pension system in Sweden and find that a large share of individuals choose actively their portfolios (most of them are females), in contrast to the vast literature that focuses on 401(k) plans in the US. I find that in a mature DC system high income individuals and males are leading the active choices. In the context of individuals investors, Calvet et al. (2009) find that wealthy and educated investors tend to do more rebalancing, and that they follow past trends. A largely discussed heuristic for portfolio choice is return chasing behavior. For example, Clark-Murphy et al. (2009) show that individuals are influenced by recent historical returns using several trailing periods. I argue that the relevant past period is the last three months before the fund reallocation. Since I can identify the portfolio returns, I complement the analysis of determinants of portfolio choice with a study of the effect on returns of the active decisions, and the heterogeneity of the effects among demographic groups.

Second, I contribute to the literature on the nature of independent services that arise to support individuals making investment decisions. Financial literacy, as discussed in Hastings et al. (2013), plays a key role in individuals investment choices and motivate individuals to seek financial advice, but its overall effect is not clear. Financial advice is a costly and imperfect substitute for financial literacy. Chalmers & Reuter (2010) find that younger, less highly educated, and lower income employees are more likely to follow financial advice. Mullainathan et al. (2012) and Egan et al. (2016) show that financial advisors can act in their personal interest or engage in misconducts, which may hurt individuals. Da et al. (2016) show that financial advisors can lead to large coordinated trading that affect stock prices. This paper adds to this literature by showing that although there are some differences in cumulative returns among income groups, followers are concentrated among high income individuals.

Third, I contribute to the literature on the costs and benefits of defined contribution retirement plans, in which workers themselves manage the allocation of their savings. Ahmed et al. (2016) uses simulations to show that choice in private accounts lead to decreases in income, with respect to no choice. Barber et al. (2009) find that individual investors lose about 3.8% in returns in a year, 66% of which comes from fees and taxes, while the rest comes from portfolio choices. I find that on average individuals underperform the default option by -1.71% in a year. This loss does not come from any fees or taxes, which are independent of the balance in the savings accounts, and it increases to -2.6% for low income individuals. Nevertheless, note that these are unadjusted returns.

From a public policy perspective, understanding what drives the choices of these individuals and how these choices affect their wealth accumulation it is important and can guide the design of DC pension systems. Moreover, understanding if these effects are heterogeneous would help to direct the efforts to the groups of individuals that are hurt by their actions. As shown above, the choices individuals make over their working life impact their wealth accumulation, and these decisions are often associated with financial literacy and thus with education and income. Therefore, the drivers of portfolio choice are also particularly relevant for income inequality at the retirement age.

The remainder of the paper is organized as follows. Section 3.2 provides details on the institutional background and the data used in this paper. Sections 3.3 and 3.4 presents the characterization and evidence of the two portfolio choice drivers: followers of financial advice and return chasers, respectively. Section 3.5 documents the effects of the drivers on wealth accumulation. Section 3.6 concludes.

3.2 Institutional background and Data

The Chilean pension system has two main advantages to study return chasers and followers of financial advice. First, it is a simple setting that facilitates the computation of portfolio returns for the choice set of individuals. Second, one widely known financial advisor operates in this market, and gives timely recommendations for a small yearly fee. In order to identify followers and return chasers and their portfolio returns, I use a representative panel of individual monthly balances for each of the five funds, and the share prices by fund and firm. The subsections below provide more details.

3.2.1 Chilean pension system

The Chilean Pension system is a defined benefit contribution system in which individuals can choose to invest in at most two out of five available funds. If individuals do not choose actively, they are assigned to a default investment plan, which is very similar to target-date funds in 401(k) plans. Formal workers in this system contribute 10% of their taxable income to an individual savings account, which are managed by one of the Pension Fund Administrators (AFP) in the market. Individuals enrolled in this system have to make two choices: (1) choose one AFP to manage his savings account[1], and (2) choose a portfolio allocation. Since returns are highly correlated among firms for each fund, the relevant choice is the second one.

The choice set in this setting is limited and tractable. Each AFP offers five different funds: A, B, C, D, and E[2]. Individuals are allowed to allocate their savings in at most two of the five funds offered by their AFP. Individuals can reallocate their savings as many times as they want since it is not capped. The composition of these funds is regulated by the laws governing the pension system and the funds differ mainly in the shares allocated to bonds and equity. According to the regulation[3], Fund A has a larger proportion of its

[1]Since 2010, new enrollees in this system are assigned to the AFP with the lowest fee, according to a bidding process. They can choose to move their savings to another firm after 18 months.

[2]This occurs since 2002, therefore it is a constant feature for the time period of this study.

investments in equities, with a minimum of 40% and a maximum of 80%. This percentage progressively decreases for funds B, C, and D, reaching a range of 0-5% for fund E. In practice, the share that the AFPs invest in equities is close to the maximum levels, for all the firms. Also, foreign investment is limited to 20-30% of each AFP's total investment.

If individuals do not make an active choice to allocate their savings, they are assigned to a default allocation according to their age and gender. The default allocation transition is shown in table 3.1. Under the default option, individuals' money is moved to more conservative funds as they age. In practice, this transition occurs gradually over 4 years[4]. Note that Funds A and E are always actively chosen, and that the riskiest fund is unavailable for older individuals. This default allocation is very similar to the target-date funds available in the US through 401(k) plans. Nevertheless, in the Chilean system the investment alternatives are the same no matter where individuals are employed. This feature of the system makes it easy to track the portfolio options and choices, and therefore the portfolio returns. Pension systems similar to the Chilean, such as the Mexican or the Peruvian, offer fewer investment options and do not allow to choose more than one fund (Impavido et al. (2010)).

TABLE 3.1: Default allocation transition, Chilean pension system

	Men	age≤35	35<age≤55	55<age
	Women	age≤35	35<age≤50	50<age
Risk ↑	Fund A			Not allowed
	Fund B	Default		
	Fund C		Default	
	Fund D			Default
	Fund E			

For most of the time since 2004, six AFPs have been functioning in this market. Prices vary by firm, which may affect individuals decisions about where to invest their contributions. Each fund manager charges the same monthly fee to every enrollee, and the median variable fee is 1.5% of the individuals' taxable income[5]. Note that this fee is not deducted

[3]Source: www.spensiones.cl

[4]20% of the balance is moved to the next fund in the default plan each year, until the transition is completed after four years.

from the individual account balance, and it is only charged when people make contributions. This is the only fee in this system in practice, since no fees are charged when reallocating funds, even though AFPs are legally entitled to charge them. Illanes (2017) presents evidence about the switching costs, and argues that individuals *perceive* large costs from switching firms for this context that are much larger than the actual fees charged.

Returns also vary by firm, although the monthly returns in the last decade are very correlated among firms, as shown in figure 3.9a in appendix 3.7.2. For example, the average return for fund C across firms is 0.57%, and the correlation among firms is 0.964. The average for all funds is 0.96, but without one outlier observation[6], the mean correlation goes up to 0.974. Therefore, when focusing on returns to make an investment decision, is not as important which firm you choose, but what fund you choose. This should not be very surprising since the dispersion of returns in this system is also regulated; the monthly returns of each AFP should not be lower than the annualized real return of the last 36 months of all funds of the same type minus a buffer of 2% for conservative funds or 4% for riskier funds.

Each AFP sends its enrollees account statements four times a year, and their contents are strictly regulated. These statements include information about the account balance and returns, and are sent roughly during the same week by all AFPs. If individuals want to know their balances or fund returns at any other time, they can access them at any time through their online account in the AFP or the Superintendence of Pensions website.

3.2.2 Financial advisor: "Felices y Forrados"

To my knowledge, the most relevant financial advisor in this market is a firm called "Felices y Forrados" (FyF, literally "Happy and Loaded"). They entered the market in July 2011,

[5]There was also a monthly fixed fee until September 2008 that was deducted from the account balance. From January 2007 onwards the median fixed fee was about 50 cents (CLP$320), and three firms charged a zero fee.

[6]The outlier observation is the return of a firm in the month prior to the fusion with another AFP (the last month in which it was functioning).

and charge a fee of US$20 a year for timely advice provided to every subscriber, offering the same advice to everyone regardless of his age or gender. Currently, they offer two plans: a “Premium Plan” that gives same day advice, and a “Basic Plan” that offers a two day delay advice. Keep in mind that the AFPs take about four business days to process switching requests. It can be longer if more than 5% of the money in a given fund is requested to be reallocated on the same day. The number of subscribers of FyF increased rapidly after they entered the market, reaching about 80,000 in 2013. Afterwards it has been decreasing, reaching about 60,000 subscribers in 2017[7]. In the period studied, as well as until today, they recommend switches between the most conservative and most risky funds, as shown in table 3.2.

TABLE 3.2: Timeline of financial advice from FyF

Recommendation date	Direction: Origin	Direction: Destination
July 27, 2011	A	E
October 12, 2011	E	A
November 22, 2011	A	E
January 11, 2012	E	A
March 29, 2012	A	E
June 19, 2012	E	A
June 28, 2012	A	E
July 19, 2012	E	A
August 29, 2012	A	E
January 2, 2013	E	A
April 2, 2013	A	E
July 17, 2013	E	A
August 16, 2013	A	E
September 6, 2013	E	A

* Source: www.felicesyforrados.cl

3.2.3 Data sources

The main data source comes from the Superintendence of Pensions, Chile. I use a monthly panel of a representative sample of about 26,740 individuals over the years 2007 to 2013.

[7]Sources: www.elmostrador.cl (2013/04/25), and www.eldinamo.cl (2017/03/24)

The data includes the full savings history of enrollees, for each of the savings accounts they might have in the AFPs[8]. For 93% of the individuals I observe the full 84 months of information, and the mean number of months is 81. Additionally, I use the share prices by funds and firms at the daily level, and the content of FyF's recommendations.

The panel of individual data includes monthly balances separately for each of the five available funds, for each individual, which allows me to identify fund switches and individuals that stay in the default option. The latter is inferred from the fund balances and the age and gender of individuals, following the information presented in table 3.1. Then, I use the transitions under the default investment option to distinguish passive from active fund reallocations. Note that if an individual actively choses his corresponding default fund (eg. by switching back to it), he will not be automatically moved as he ages. About 30% of the individuals-month observations in the data are out of default, and 84% of them allocate their contributions into one fund, as shown in table 3.9 in appendix 3.7.1.

The data provides few demographic variables: date of birth, date of death, gender, and income. The latter is matched for 52% of the sample. To overcome this lack of data, and since the missing income data does not necessarily imply that the income was zero, I use the average income for each year, for each individual. This measures the *potential* income of individuals. Since higher income levels correlate with higher education, and thus with financial literacy, the measure of *potential* income would be appropriate as it is related to education and financial literacy. An alternative, would have been to use the monthly data and setting the missing data to zero income. In appendix 3.7.4 I discuss the choice of this measure of income in more detail, and shows that in the majority of the cases there is no difference in switching behavior as the availability of income information changes. To create the demographic categories that will be used during our analysis that are time dependent I use the information at the month of the switch. For instance, in the analysis of

[8]Besides the mandatory retirement savings accounts for formal workers, the focus of this paper, other savings accounts that can be held in the pension system are voluntary savings accounts, and voluntary retirement savings accounts for independent workers.

fund reallocations, the age group of an individual is defined at the time of the switch.

Another source of data is the share prices by funds and firms at the daily level. This allows us to compute returns by fund and firm at both the monthly and daily levels. Individuals' portfolios are easily identifiable from the funds' balances, and therefore the individuals' portfolio return is computed for each month. Observing portfolio returns is difficult, as discussed in Campbell (2016), and the Chilean pension system provides a good setting to overcome this due to the small set of investment options available. I also compute the return that each individual would have had if they pursued the default investment strategy. Note that, the returns I use are not adjusted by risk, since the focus of this analysis is on the wealth accumulation of the individuals' savings account.

The focus here is to identify active portfolio choices made by these individuals, where individuals can switch funds on a daily basis. The monthly data does not allow to identify the day in which the reallocation was made. I identify the switches from funds' balances changes in which a fund with zero balance in the previous month now has a positive balance, and the other way around. To be conservative, I assume that individuals switch funds the last day of the month in which I observe the changes in balances. This implies that the month in which the individual starts gaining the return of the destination portfolio for the full month, $t = 0$, is the month *after* the observed fund reallocation. The role of this assumption will be discussed in the Results section.

3.2.4 Participants of the pension system

The data used in this paper excludes any other source of savings individuals might have in the financial system. Nevertheless, it constitutes the most relevant savings source for pensions[9]. In the sample used just 1.7% of the individual-month observations show the existence of a voluntary savings account (VSA), corresponding to 7% of individuals. Also, since withdrawals are allowed for the VSAs, in many cases its balance is very low. The

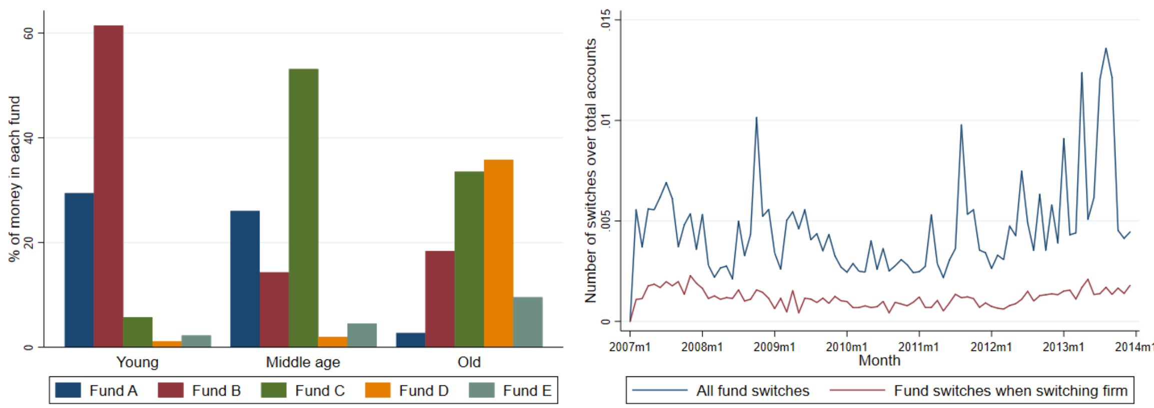
[9]For instance, the National Accounts Report 2016 of the Central Bank of Chile show that the increase in household financial assets comes largely from contributions to the pension funds.

total value of the pension system funds accounts for 70% of the Chilean GDP in 2016[10] , which shows that most of the pension wealth is in these pension accounts.

The data shows that most individuals are passive choosers, while the active choosers are concentrated among high income individuals. The majority of the money in this system is allocated into the default fund. Figure 3.1a shows this for the different age groups. The money out the default portfolio is largely allocated to riskier funds, which corresponds to funds to the left of the default in the figure.

FIGURE 3.1: Aggregate allocations and reallocations

(a) Shares of money allocation by fund and age (b) Reallocations over time



Notes: The figure on the left shows the share of money allocated into each fund, for the different demographic groups. The default fund for young people is fund B, fund C is for middle aged individuals, and fund D is for old people. The figure on the right the share of fund reallocations over all accounts in each month, for all fund switches and for the fund switches that occur simultaneously with a firm switch.

Similarly to what occurs in other defined contribution systems, most individuals are passive choosers. Figure 3.1b shows the switches over time as a share of the total accounts in each month. The first thing to notice is that the scale in the vertical axis is quite small, since a minority of individuals choose funds actively. Second, the switches have several spikes, some of them related to the financial crisis, and others related to financial advice, as I will show later. Only a small share of these switches occurs at the same time of firm switches.

[10] According to the Superintendence of Pensions, Chile, this amount is US$ 165 billions.

Finally, to start characterizing the active choosers, table 3.3 shows that the switchers are concentrated among higher income individuals, but there are not many differences by gender or age. The next sections explore if these groups of individuals behave differently in terms of their portfolio choices.

TABLE 3.3: Demographic characteristics of active choosers

	None	**At least once**
Gender		
Female	81%	19%
Male	80%	20%
Age group		
Young	79%	21%
Middle age	81%	19%
Old	81%	19%
Income percentile		
Below 50th	88%	12%
Above 50th	75%	25%

Notes: This table shows the share of individuals that are switchers and those who are not, by demographic groups.

3.3 Followers of financial advice

In this section I analyze following financial advice as a portfolio choice driver. The complex financial choices that individuals face in DC pension systems, provides a context where most individuals choose passively and there is evidence of low financial literacy. Therefore, financial advisors can potentially have a large influence on individuals. I focus I show that it is possible to identify the switches that follow FyF's recommendations by setting up an algorithm that matches the timing and direction of fund rebalances between individuals and FyF's financial advice. Then, I define two statistics to characterize the followers. First, the share of followers, defined as the share of individuals that ever follow a piece of advice. Second, the share of advice that the followers followed in the next months, for each recommendation month. Overall, high income males are more likely to be followers and to keep following advice, than other demographic groups.

3.3.1 Identifying followers

To analyze individuals' responses to financial advice, the algorithm uses the direction and timing of the recommendations (according to table 3.2) and of the individual switches to identify the switches of the followers. Thus, I can identify and match the switches in the date and direction of FyF's recommendations. It is important to note that under this approach we cannot distinguish if investors are independently responding to the same underlying shock that may motivate the financial advisors' recommendation, or if another piece of information not considered by FyF is driving the individual reallocations. Nevertheless, by comparing the switches in months with and without financial advice, I show that FyFs recommendations are associated with a large share of reallocations.

In the period studied, the financial advisor only recommended reallocations between the most risky (A) and most conservative (E) funds. In months prior to FyF's entry, the switches between funds A and E represent on average just a 14.3%, showing that reallocations were not often concentrated in the directions of FyF's recommendations. After FyF enters the market, 44.0% of switches are in the advised direction in months where there is a recommendation, and 28.1% in months without recommendations[11]. This share is increasing over time, as FyF's becomes more widely known.

Even though the majority of switches match FyF's recommended direction in months with a recommendation, this could be driven by different things. On one hand, the set of switchers may be fairly constant over time but all of them switch in a herd when FyF sends an advice. On the other hand, some individuals would always switch but others switch only in months with recommendations, and therefore FyF's advice could be driving more people to switch. Of all the switchers in the period where FyF is offering advice, 54.3% of them switch only in months with recommendations, 38.2% switch only in months with no

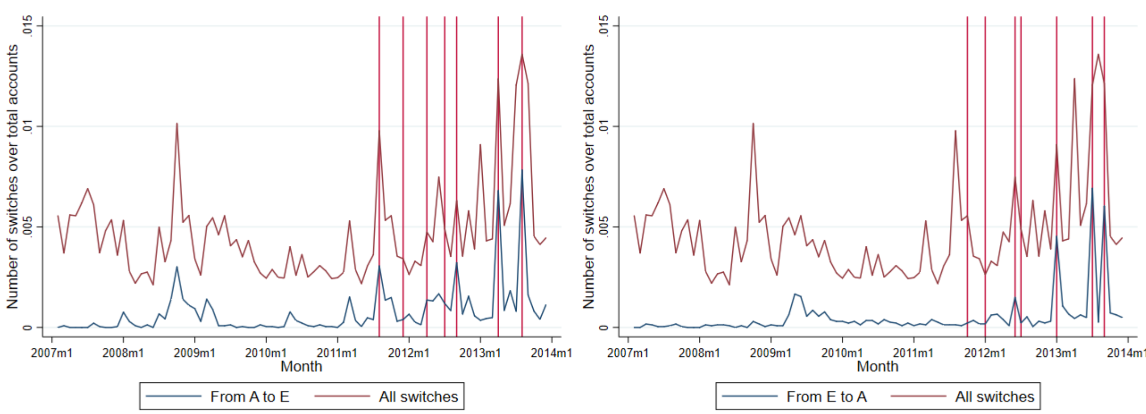
[11]This share computed for non-recommendation months is 17.9%, for the full period in the sample. When excluding the financial crisis period (December 2007 to June 2009), this share decreases to 17.2%.

recommendations, and 17.6% switch in both types of months.

Figure 3.2 explores these switches at a more detailed level by presenting the number of switches in the advised direction over the total number of accounts in each month[12]. The vertical solid lines represent the month in which a recommendation in the corresponding direction was sent by FyF. The big spikes in the switches share coincide with the months with recommendations. Moreover, in 2013 the spikes from FyF recommendations increase total switching by around 100%, which shows that people seem to be responding to FyF recommendations. Before FyF' entry, the large spikes in switches mostly occurred during the financial crisis.

FIGURE 3.2: Reallocations between funds A and E and recommendations

(a) From fund A to E (b) From fund E to A



Notes: These figures show the number of switches in FyF's advised direction over the total number of individuals in each month. Vertical solid lines indicate FyF recommendations in the corresponding direction, A to E on panel (a) and E to A on panel (b).

3.3.2 Characterization of followers

After identifying the switches related to FyF recommendations, I define two statistics to characterize the followers. The first one, the share of followers, is defined as the share of individuals that ever follow a piece of advice, or that follow advice at least once. The

[12]Similar results are obtained using the number of switches in each direction over the total number of indivuals who ever reallocate funds. Figures 3.9c and 3.9d in appendix 3.7.2 show the plots using the number of switchers as denominator.

second statistic is the share of future advice followed, defined as the share of advice that the followers followed in the next months, for each recommendation month. I compute this using data at the individual-recommendation date pair level and then calculate the share of recommendations that the individual followed afterwards, at each recommendation date. To implement this, I use all observations after the date of the first recommendation (July 2011) and estimate regressions for the two statistics at the individual level including all the demographic groups and their interactions.

Table 3.4 presents the main analysis to characterize the followers using the two statistics, by regressing them on the demographic groups and their interactions. Columns (1) to (3) show the analysis for the share of followers. High income individuals are the main group driving the results; they are 1.6% more of likely to be followers, and they follow 14.1% more future recommendations, with respect to other demographic groups. Columns (4) to (6) characterize the share of future advice followed. There are no significant differences by gender; old males and high income old individuals show significant negative coefficients for the share of recommendations followed. It should not be surprising that older followers follow fewer future recommendations, since 10 years before the retirement age they cannot choose the riskiest fund.

Next, I show the time pattern of the share of advice followed for each of the eight recommendations in the period studied. Figure 3.3 presents the results. Since followers are defined as ever following a piece of advice, 100% follow advice for the first time. Then, 40% of them follow advice a second time, and this share decreases as the number of followed recommendations increases. Note that this does not require to follow recommendations consecutively, also, the the share of followers is not for a single recommendation. As showed before, most of the followers are high income individuals. The share of women following each number of recommendations is roughly half the share of men. Finally, middle age individuals follow advice more consistently than other age groups.

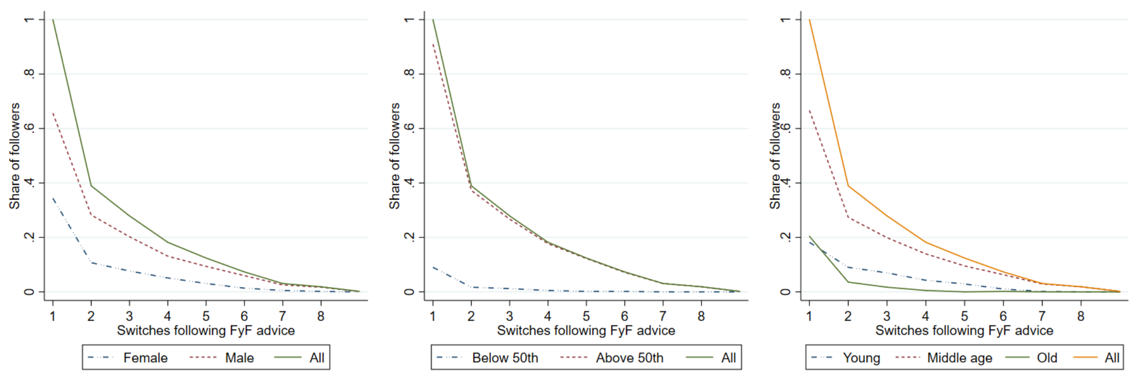
TABLE 3.4: Estimation of followers' statistics

	Followers			Following		
	Coeff. (1)	SE (2)	p-value (3)	Coeff. (4)	SE (5)	p-value (6)
Constant	0.0102	(0.0033)	0.0019	0.2058	(0.0460)	0.0000
Main effects						
Male	-0.0071	(0.0046)	0.1237	0.0266	(0.0432)	0.5387
Above 50th percentile	0.0244	(0.0049)	0.0000	0.0668	(0.0499)	0.1809
Middle age	-0.0047	(0.0038)	0.2199	-0.0545	(0.0500)	0.2757
Old	-0.0047	(0.0039)	0.2315	-0.0555	(0.0493)	0.2608
Interactions						
Male#Above 50th	0.0164	(0.0037)	0.0000	0.1405	(0.0316)	0.0000
Male#Middle age	0.0072	(0.0055)	0.1928	-0.0561	(0.0389)	0.1491
Male#Old	0.0056	(0.0063)	0.3781	-0.1600	(0.0372)	0.0000
Above 50th#Middle age	0.0007	(0.0050)	0.8876	-0.0136	(0.0528)	0.7972
Above 50th#Old	-0.0020	(0.0058)	0.7328	-0.1505	(0.0521)	0.0039

Notes: This table presents the regression analysis of the two statistics constructed: (1) the share of followers, defined as following ever a piece of advice; and (2) the share of future advice followed. The standard errors are clustered by individual and recommendation.

FIGURE 3.3: Followers switches, by demographics

(a) Gender (b) Income (c) Gender



Notes: These figures show the share of individuals that follow different, and possibly non-consecutive, recommendations by demographic groups.

3.4 Return chasers

In this section I present the second portfolio choice driver analyzed. Return chasing is a bias where individuals chase past returns in the mistaken belief that they will persist in the future. I use an event study to develop a statistic that measures the extent of return chasing behavior based on the return differential between the destination portfolio and the origin portfolio before the rebalance. Then, I show that higher income individuals change funds when the return differential is larger than what triggers a change for lower income individuals. There are some differences by age groups, but not by gender, in contrast to previous studies.

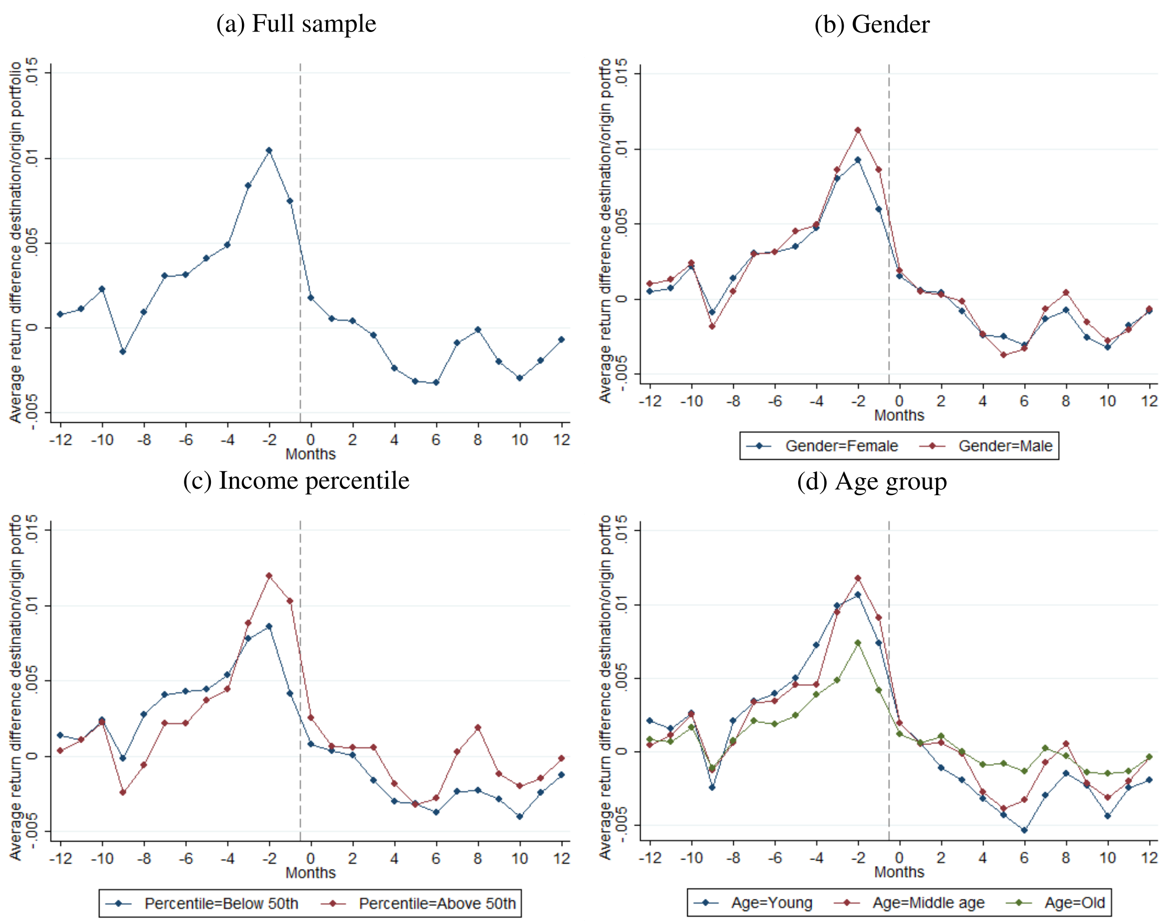
3.4.1 Measuring return chasing: an event study

If individuals are actually doing some market timing, we would observe that they switch to a given fund when its returns are increasing. A first approach to analyze whether individuals follow short-term trends in their investment reallocation decisions is to use an event study of the returns around the time of the switch. I define the event time such that the switch occurs on month $t = 0$, which is the first month in which the individuals receive the return of the destination portfolio for the full month. The underlying assumption is that individuals switch on the last day of the month in which I observe the reallocation, and thus the month in which individuals start obtaining the return of the destination portfolio is the month *after* the observed reallocation.

For the event study analysis, I keep individuals' events that are at least three months apart, to clearly distinguish the trends leading to a fund reallocation. I consider each of these events separately, and for each of them I construct an event window from one year before to one year after the switch. To measure return chasing, I analyze the return difference between the destination portfolio return and the origin portfolio return. This is done for the different demographic groups.

If the investment behavior of individuals is influenced by past trends, one may wonder how far individuals look back. Figure 3.9g in appendix 3.7.2 explores this by presenting an event study of the return of origin portfolio when switching away of it. On average, the return of the origin portfolio is decreasing in the three months prior to switch. Both the returns of the origin and destination portfolio matter for switches, so this is just half of the story, and therefore the difference in returns between these portfolios accounts for the full story. The event study for the return difference is shown in figure 3.4. The return differential is positive and is increasing in the three months before the switch at month $t = 0$, but it becomes negative a few months after the switch[13]. The other figures break down the event study into the available demographic groups. Higher income individuals exhibit greater return chasing. These results suggest that the extent of return-chasing varies in some demographic groups.

FIGURE 3.4: Event study around the time of the switch, destination vs origin portfolio



Notes: These figures present event studies around the time of the switch, for the full sample and for the different demographic groups. The plotted variable is the return difference between the return of the destination portfolio and the origin portfolio. Time-dependent demographic groups are computed using the value at the time of the switch. Obs=161,261.

[13]The average return differential for the first month after the switch is slightly positive (0.05%), and at the fourth month it is negative (-0.24%).

3.4.2 Characterization of return chasers

Motivated by the results in the previous section and to explore the return-chasing in a more systematic way I construct a statistic to measure the extent of the return chasing behavior. Recall that in both events studies in figure 3.4 it is possible to note that the trends present a jump right before the switch. By simply computing the average return of the past three months I would not be capturing this fact, and also, this metric would be consistent with a constant return difference over the full year prior to the switch. I define the statistic $\Delta\bar{r}$ as the difference between the destination and origin portfolio returns in the last three months compared to the return in the rest of the previous year, as shown in equation 3.1,

$$\Delta\bar{r} = \bar{r}_{[-1,-3]} - \bar{r}_{[-4,-12]} \tag{3.1}$$

where $\bar{r}_\tau$ is the average return differential for the interval of months $\tau = [m_{min}, m_{max}]$. $\Delta\bar{r}$ captures the deviation of the return difference with respect to last year's trend, just before the fund switch.

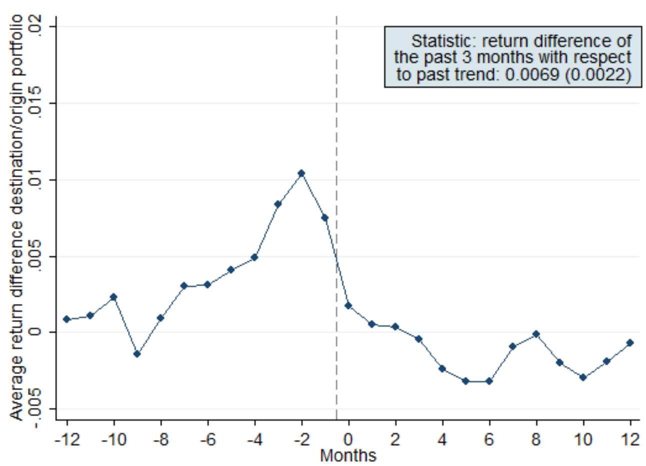
A concern when computing the statistic is that past returns are correlated within months, and therefore more switches may occur in particular months. A possible approach to estimate the statistic is to cluster the standard errors by the date of the switch (eg. January 2010). This is an imperfect approach since clusters of switches could be correlated because the statistic is computed by averaging the return differences in the past three months. In the case of individuals switching in two consecutive months, the statistic considers overlapping months between these individuals which is not accounted for by the clustering.

To overcome this issue I estimate the statistic using two different approaches. First, I aggregate the panel data into time series data, using the net flow of reallocations between the most risky and most conservative funds, fund A and fund E, respectively. These direction of reallocations are among the most common, and are also related to FyF's recommendations. I compute the statistic on the time series data using a regression with Newey-West standard errors.

For the second approach, I divide the panel data into three subsamples such that each of them contains only switches that occurs three months apart. I compute the statistic for the different demographic groups for each subsample, clustering by the date of the switch. This last approach, even though imperfect, allows me to study the extent return-chasing at the individual level.

This analysis computes the statistic $\Delta\bar{r}$ using an event study around the month of the switch to analyze whether individuals move, on average, into higher return portfolios. Recall that return chasing is measured as the return difference between the return of the destination and origin portfolios, with respect to the previous trend. Whenever the destination return is higher than the origin return, this difference will be positive. Bearing in mind the concerns about the correlation of returns within months for the estiamtion of the statistic, I show in figure 3.5 the estimation of the statistic for the full sample. On average, individuals switch after returns increase about 0.0069 prior to reallocation.

FIGURE 3.5: Event study around the time of the switch, full sample



Notes: This figure shows the event study around the time of the switch for the full individuals' panel data. The plotted variable is the return difference between the return of the destination portfolio and the origin portfolio. Standard errors of the statistic are clustered by date of the switch. Obs=161,261.

As mentioned before, I estimate the statistic using two approaches. First, I use the data aggregated to time series and show that individuals follow past trends of returns. Second, I use subsamples of the panel data and show that low income individuals show half the return chasing as the higher income individuals.

For the first approach, I use the aggregated time series data for the net flow of reallocations between funds A and E. For each date, I construct the statistic $\Delta\bar{r}$, and an indicator variable that takes a value of one if the number of switches towards fund A is greater than that towards fund E. I use Newey-West standard errors with a three month lag to correct for the correlation in returns from consecutive months[14]. Table 3.5 shows the results. In months where the net flow of switches was leaning towards fund A, the difference in returns before the switch is 0.030 larger than in months in which the net flow of switches was in the other direction. Therefore, the estimated value of the statistic for return chasing from fund E to A is 0.018.

TABLE 3.5: Estimation of return chasing statistic, aggregated flow

Statistic from E to A	Coeff. (1)	Newey-West Std.Error (2)	p-value (3)
Flow from E to A	0.0299	(0.0138)	0.0338
Constant	-0.0117	(0.0072)	0.1079

Notes: This table presents the estimation of the statistic defined in equation 3.1 using the time series data. The statistic is computed using the return difference between fund A and fund E, and the variable "Flow from E to A" is an indicator variable for the months in which the number of switches from E to A was larger than those from A to E. Column (1) shows the coefficients, and column (2) shows the Newey-West standard errors using a three month lag.

Figure 3.6a shows the event study for the switches from fund E to fund A, using the panel data and clustering the standard errors by the date of the switch. Note that the statistic from table 3.5 which was computed using the time series data takes the same value as the one in this figure. The magnitude of statistic for switches between funds A and E is half a percentage point larger than that of all the followers' reallocations, as shown in figure 3.6b[15]. These differences can be explained by the fact that switches between funds E and

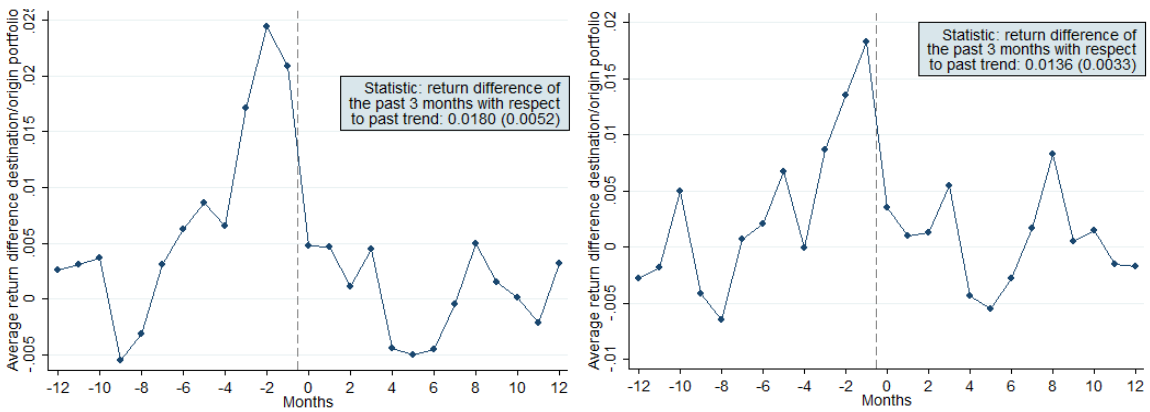
[14]Similar results are obtained using a twelve month lag.

A should have the largest return difference since fund A is the riskiest fund and fund E is the most conservative fund.

FIGURE 3.6: Event study around the time of the switch, funds E to A and followers

(a) Switches from E to A

(b) Switches of followers



Notes: These figures show the event study around the time of the switch for the switches from fund E to fund A, and for the followers. The plotted variable is the return difference between the return of the destination portfolio and the origin portfolio. In the left graph this is the difference between funds A and E. The graph on the right shows every switch made by followers, not only those labeled as following advice. Standard errors of the statistic are computed clustering by the month of the switch.

From the previous section, figure 3.4 suggests that there different demographic groups have different extents of return chasing behavior. To study this, I turn to the second approach that uses subsamples of the panel data. Each subsample contains only switches that occurs three months apart such that the months in which the jump in returns occur do not overlap between switches. For each subsample, I estimate conditional means of the different demographic groups clustering standard errors by the date of the fund switch. Table 3.6 shows the results[16]. Two of the three samples show an unconditional mean significantly different than zero, and all of them have similar magnitudes. Overall, some clear patterns are common to the three subsamples. Significant differences arise among income percentiles, where low income individuals show half the return chasing of higher income individuals. Also, there are some differences between young and old age groups. In contrast to other

[15]Figure 3.5 shows that the statistic for all the individuals switches is even lower.

studies, there are no significant differences by gender.

TABLE 3.6: Estimation of return chasing statistic, by subsamples

Subsamples	Jan-Apr-Jul-Oct			Feb-May-Aug-Nov			Mar-Jun-Sep-Dec		
Demographic groups	Coeff. (1)	SE (2)	Diff. p-value (3)	Coeff. (4)	SE (5)	Diff. p-value (6)	Coeff. (7)	SE (8)	Diff. p-value (9)
Gender									
Female	0.0060	(0.0030)		0.0049	(0.0040)		0.0071	(0.0031)	
Male	0.0072	(0.0032)	0.1696	0.0073	(0.0050)	0.0172	0.0085	(0.0037)	0.1640
Age group									
Young	0.0069	(0.0041)		0.0065	(0.0057)		0.0072	(0.0036)	
Middle age	0.0077	(0.0033)	0.5067	0.0074	(0.0049)	0.5605	0.0097	(0.0039)	0.0685
Old	0.0043	(0.0025)	0.0529	0.0034	(0.0030)	0.0213	0.0046	(0.0024)	0.0427
Income perc.									
Below 50th	0.0047	(0.0028)		0.0037	(0.0044)		0.0045	(0.0027)	
Above 50th	0.0085	(0.0035)	0.0000	0.0082	(0.0048)	0.0000	0.0108	(0.0042)	0.0000
Mean	0.0067	(0.0031)		0.0063	(0.0046)		0.0079	(0.0034)	

Notes: This table presents conditional means of the statistic defined in equation 3.1, using three subsamples. Each subsample contains only switches that occurs three months apart, to avoid the overlap in the construction of the statistic that arises in consecutive months. Columns (1), (4), and (6) show the statistic for different demographic groups. Columns (2), (5), and (7) show the standard errors clustered by the date of the switch. Columns (3), (6), and (9) show the p-values.

Finally, even though previous results show that individuals switch funds when the destination fund is performing better than the origin one, one might wonder if they were indeed moving into the best option available. In the context of this study, return correlation across funds is high, so switching between the most conservative and the most risky funds is likely to be the strategy to switch to the best option available, but individuals are not always choosing those funds. I compute the statistic $\Delta\bar{r}$ for the difference between each of the five funds and the origin portfolio[17] and find that 1.9% of the individuals actually switch to an option that is at least as good as the best option available.

[16]Table 3.10 shows the results for the full sample.

[17]Which is a subset of the possibilities individuals have, since they could also allocate their savings into at most two funds. Nevertheless, a large majority of individuals allocate their savings into a single fund.

3.5 Wealth accumulation

Portfolio choice drivers such as following financial advice or return chasing are meant to aid individuals to circumvent the complexity of financial decisions, nevertheless the active choices driven by them may hurt individuals. Short gains may not be persistent over time. To understand if the active decisions are beneficial or not for the individuals, a direct way to do this is by analyzing the value of these switches. I compare them against a benchmark available in this setting: the default portfolio. Also, I show that the financial advisor follows a very sensitive market timing strategy, and hence it has to be tightly followed for best results.

3.5.1 Measuring the value of switching

To measure the value of rebalancing, I need a counterfactual of the active choices individuals make. I compare the rebalancing choices against a benchmark available in this setting: the default portfolio. Alternatively, I could use the returns of the origin portfolio as a benchmark, nevertheless, this choice hardly changes the results. With this information I estimate conditional means for the difference between the actual cumulative returns and the cumulative returns of the benchmark, for each demographic group.

I analyze the value of these switches by computing the cumulative returns of their choice and of the default investment plan. Since the data has monthly frequency, a key assumption is the day in which individuals switch. I assume that individuals switch on the last day of the month; on the other extreme, they could be switching on the first day of the month. I compute it for the short and long run after the rebalance, where the short run is defined as two months after the switch, and the long run as twelve months after. I only use events for which I observe the full twelve-month post switch period. I cluster the standard errors at the date of the switch. In this way, I assume independence across clusters (dates of switches) but correlation within clusters.

3.5.2 Effects on wealth accumulation

I study if the active choices affect individuals' wealth accumulation. I estimate conditional means for the difference between the actual cumulative returns and the benchmark, for different demographic groups. If individuals are switching to anticipate future trends, then high performance of the destination portfolio in the longer run would be consistent with this behavior. Overall, active choices of followers and return chasers lead to lower cumulative returns, and low income individuals are hurt significantly more.

Impact on cumulative returns for the followers

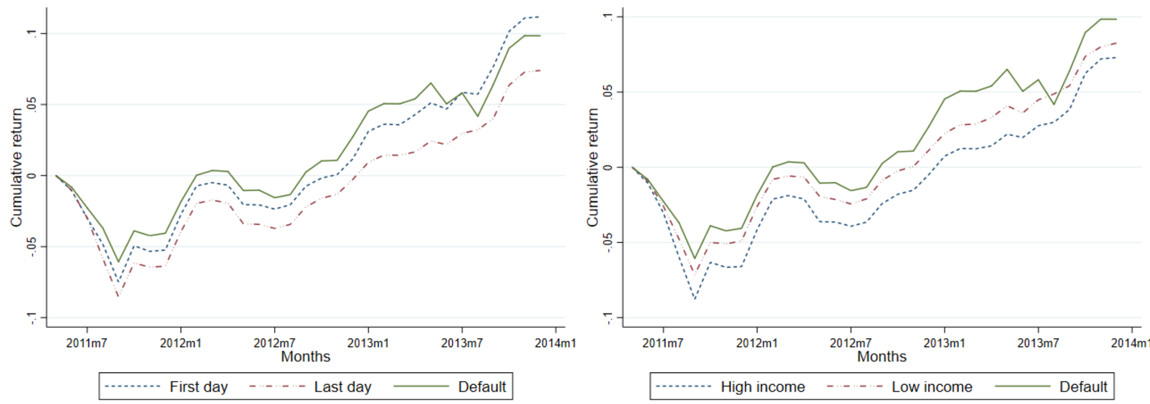
For the followers, who are presumably pursuing the tight market timing from FyF's advice, a key assumption is the day in which they switch. So far, I've assumed conservatively that individuals switch on the last day of the month. On the other extreme, they could be switching on the first day of the month. Figure 3.7a shows two curves that compare the cumulative returns of followers, if they switch either on the first or the last day of the month. The two curves show the upper and lower bounds for the cumulative returns. The t-test for equality of means between these two curves yields a p-value of 0.12, and therefore the null hypothesis of equality cannot be rejected. Note that the cumulative returns include all the switches made by the followers after the month in which they followed advice, not only the switches associated to FyF's advice.

In previous section, differences arose by income brackets. Figure 3.7b shows that followers of high and low income do not have significantly different cumulative returns. Testing for the equality of means between these two curves, yields a p-value of 0.59.

To explore these trends in a more detailed level, I compare the conditional means of the actual cumulative returns with the default portfolio after each switch in the short and long term, two and twelve months respectively. Results are presented in table 3.7. In the short run the coefficients are negative but not significantly different than zero for almost every demographic group. This implies that individuals are obtaining, on average, lower

FIGURE 3.7: Cumulative returns of followers, after first recommendation

(a) By day of month (b) By income



Notes: The figure on the left compares the cumulative returns of followers, if they switch on the first day of the month versus on the last day. The figure on the right compares the cumulative returns by the different income groups. Note that these includes all the switches they make, not only those labeled as following FyF's advice.

returns than the default option two months after the switch. Note that this analysis includes every followers' switch after FyF's entry, not only the reallocations following FyF's advice. Similar results arise when comparing the returns with respect to the origin portfolio (table 3.11 in appendix 3.7.1).

Another result from table 3.7 is that among followers, young individuals seem to be having larger losses than older ones in the short run. Notably, there are no significant differences among income groups. A plausible explanation for these results is that financial advice could compensate to some extent the financial literacy gap associated with high and low income levels. Nevertheless, these results need to be taken cautiously since they come from the sample of followers, and this group of individuals can be particular in other ways that could explain this fact.

Impact on cumulative returns for the return chasers

If individuals are switching to anticipate future trends, then high performance of the destination portfolio in the longer run would be consistent with this behavior. The previous sections showed that return chasers follow previous trends, and this section analyzes if they

TABLE 3.7: Cumulative returns of followers, destination vs default allocation

Followers	2 months			12 months		
	Coeff.	SE	Diff.p-value	Coeff.	SE	Diff.p-value
Subsamples	(1)	(2)	(3)	(4)	(5)	(6)
Gender						
Female	-0.83	(0.41)		0.02	(0.78)	
Male	-0.95	(0.41)	0.47	0.77	(1.07)	0.06
Age group						
Young	-1.55	(0.48)		1.06	(1.28)	
Middle age	-0.87	(0.48)	0.01	0.56	(1.04)	0.38
Old	-0.50	(0.24)	0.00	-0.13	(0.45)	0.05
Income percentile						
Below 50th	-0.70	(0.38)		0.30	(0.65)	
Above 50th	-0.98	(0.42)	0.16	0.57	(1.08)	0.54
Mean	-0.91			0.50		

Notes: This table shows the conditional means and the clustered standard errors of the cumulative returns for the destination portfolio with respect to the default portfolio, by demographic groups. Columns (1) and (2) show the estimation for two months after the switch Columns (4) and (5) present them twelve months after the switch. Columns (3) and (6) show the p-values that test for the difference between the corresponding coefficient and the baseline.

benefit from these active choices. I study how active choices affect wealth accumulation is to analyze how the cumulative returns of return chasers, defined as individuals that actively choose a portfolio at least once, perform in comparison to their corresponding default portfolio.

Figure 3.8a shows their actual cumulative returns after the first switch, and the counterfactual returns for staying in the default option. Even though in the first months after the first switch the two curves are quite similar, on average, they lose afterwards, and the gap increases over time.

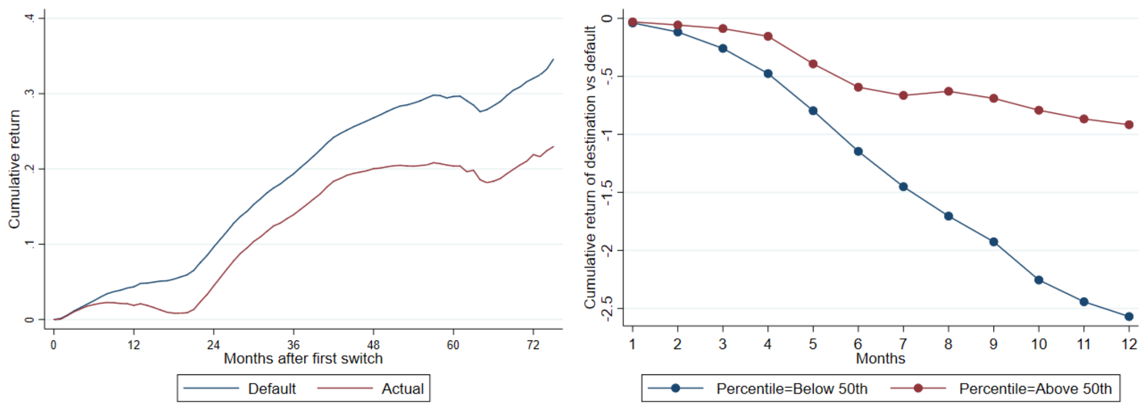
Figure 3.8b shows the cumulative return difference by income groups, for the year after each fund reallocation. The results highlight that each of the active choices can hurt low income individuals in particular, and thus can potentially worsen the income inequality among individuals at the retirement age. However, it is important to recall that high income individuals switch more often than low income individuals.

To explore the effect of each rebalance decision on cumulative returns, table 3.8 shows

FIGURE 3.8: Cumulative returns of return chasers

(a) After first switch
Actual and default portfolio

(b) After each switch, by income
Difference of destination and default portfolio



Notes: The figure on the left presents the cumulative returns of the return chasers after the first reallocation they make in the sample, for the default and actual portfolio choices. The figure on the right shows the average cumulative return difference between the destination and default portfolio, for the twelve months after each fund switch, by income group.

conditional means of the difference in returns for the short and long term patterns. As before, I focus on the difference between cumulative returns of the destination versus the default portfolio for two and twelve months after the switch. In the short run, every demographic group's destination portfolio performs similarly to the default portfolio. Nevertheless, the coefficients are not significantly different than zero, and the only significant difference among demographic groups is between old and young individuals. In the long run, every demographic group obtains a significant negative return difference. Interestingly, the difference among income groups is now significant. Individuals above the 50th percentile of income have cumulative returns -0.9 lower than the default portfolio, but they are still performing better than low income individuals, which have about three times that difference.

3.6 Conclusion

Using a representative panel of monthly retirement mandatory savings accounts over seven years from the Chilean pension system, this paper documents evidence on two drivers of

TABLE 3.8: Cumulative returns of return chasers, destination vs default allocation

Switchers	2 months			12 months		
	Coeff.	SE	Diff.p-value	Coeff.	SE	Diff.p-value
Subsamples	(1)	(2)	(3)	(4)	(5)	(6)
Gender						
Female	-0.10	(0.15)		-1.85	(0.46)	
Male	-0.07	(0.16)	0.68	-1.61	(0.45)	0.22
Age group						
Young	-0.24	(0.24)		-3.13	(0.65)	
Middle age	-0.08	(0.17)	0.10	-1.63	(0.48)	0.00
Old	0.02	(0.13)	0.01	-0.77	(0.38)	0.00
Income percentile						
Below 50th	-0.12	(0.18)		-2.57	(0.53)	
Above 50th	-0.06	(0.15)	0.40	-0.92	(0.38)	0.00
Mean	-0.09			-1.71		

Notes: This table shows the conditional means and the clustered standard errors of the cumulative returns for the destination portfolio with respect to the default portfolio, by demographic groups. Columns (1) and (2) show the estimation for two months after the switch Columns (4) and (5) present them twelve months after the switch. Columns (3) and (6) show the p-values that test for the difference between the corresponding coefficient and the baseline.

portfolio choices: following financial advice and return chasing. Understanding the effects on returns, and if these drivers affect individuals heterogeneously, can help directing public policy efforts to the groups of individuals that are hurt by their actions, and can also guide the public policy design in DC pension systems.

First, I find evidence that high income males are more likely to follow advice, to keep following any future advice, and to follow it more consistently over time. The followers' rebalancing of portfolios does not lead to higher cumulative returns a year after the reallocation. The financial advisor firm follows a tight market timing that does not clearly benefit individuals, because of the delay they have to incur in implementing the strategy.

Second, I show that switches are largely driven by a jump in the return difference between the destination and origin portfolio in the three months prior to the fund switch. The extent of return chasing is large for reallocations from the most conservative to the most risky fund. Active choices in this context lead to losses in cumulative returns one year after the fund switch, with losses significantly larger for low income individuals (-2.6% with respect to -0.9%).

The results highlight that each of the active choices can hurt low income individuals in particular, and therefore can potentially worsen the income inequality among individuals at the retirement age. To some extent, my findings are constrained to the monthly level aggregation. Information at the daily level would be useful to explore more clearly the drivers in this study, and also to study other drivers related to the projection bias such as the effect of weather or sports events in the investment decisions (eg. Hirshleifer & Shumway (2003), Edmans et al. (2007)). In this analysis I used the default portfolio as the main benchmark for the active decisions, which was useful to evaluate what the returns would have been under passive choice. Further research is required to analyze the impacts of these drivers of portfolio choice on welfare, as well as to study the relative importance of these drivers in wealth accumulation, and the overall effect on inequality.

3.7 Appendix

3.7.1 Tables

TABLE 3.9: Observations by default allocation and number of funds

	One fund	Two funds
Active chooser	26%	3%
Default option	58%	14%

Notes: This table shows the share of observations (at the individual-month level) by two categories: having their contributions allocated in the default option, and having them allocated in one or two funds.

TABLE 3.10: Estimation of return chasing statistic, full sample

All obs.	Coeff. (1)	SE (2)	Diff.p-value (3)
Gender			
Female	0.0059	(0.0020)	
Male	0.0076	(0.0024)	0.0025
Age group			
Young	0.0068	(0.0027)	
Middle age	0.0082	(0.0024)	0.0883
Old	0.0041	(0.0015)	0.0003
Income percentile			
Below 50th	0.0043	(0.0020)	
Above 50th	0.0091	(0.0025)	0.0000
Mean	0.0069	0.0022	

Notes: Columns (1) and (2) show the statistic by different demographic groups, of the full panel data. Column (3) shows the p-values that test for the difference between the corresponding coefficient and the baseline. Standard error are clustered by the date of the fund reallocation.

TABLE 3.11: Cumulative returns of followers, destination vs origin allocation

Followers	2 months			12 months		
	Coeff.	SE	Diff.p-value	Coeff.	SE	Diff.p-value
Subsamples	(1)	(2)	(3)	(4)	(5)	(6)
Gender						
Female	-1.57	(0.75)		-0.89	(0.82)	
Male	-1.76	(0.78)	0.52	-0.45	(1.02)	0.30
Age group						
Young	-2.07	(0.50)		-0.25	(1.03)	
Middle age	-1.78	(0.92)	0.49	-0.77	(1.00)	0.38
Old	-1.07	(0.65)	0.02	-0.34	(0.73)	0.89
Income percentile						
Below 50th	-1.31	(0.76)		-0.55	(0.80)	
Above 50th	-1.82	(0.77)	0.12	-0.63	(1.01)	0.87
Mean	-1.69			-0.61		

Notes: This table shows the conditional means and the clustered standard errors of the cumulative returns for the destination portfolio with respect to the origin portfolio, by demographic groups. Columns (1) and (2) show the estimation for two months after the switch Columns (4) and (5) present them twelve months after the switch. Columns (3) and (6) show the p-values that test for the difference between the corresponding coefficient and the baseline.

TABLE 3.12: Cumulative returns of return chasers, destination vs origin allocation

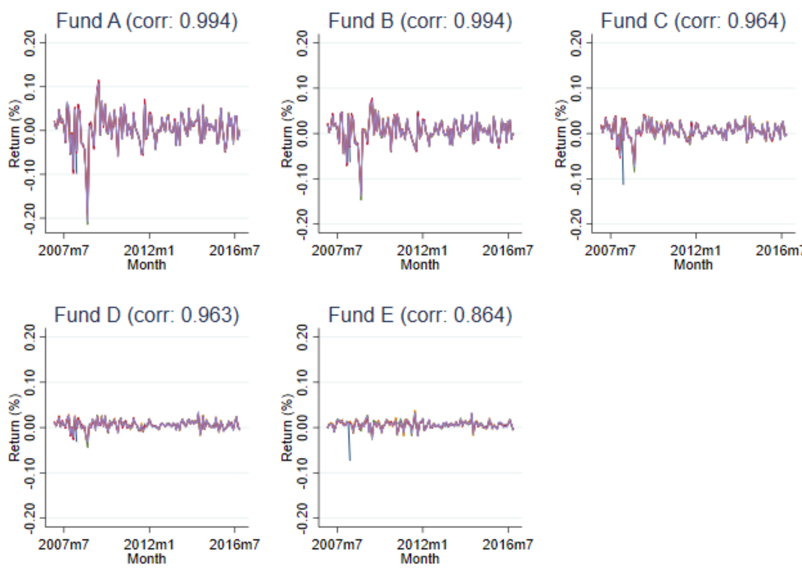
Switchers	2 months			12 months		
	Coeff.	SE	Diff.p-value	Coeff.	SE	Diff.p-value
Subsamples	(1)	(2)	(3)	(4)	(5)	(6)
Gender						
Female	0.06	(0.26)		-2.13	(0.71)	
Male	-0.00	(0.30)	0.55	-2.16	(0.81)	0.91
Age group						
Young	-0.21	(0.33)		-3.71	(0.92)	
Middle age	0.05	(0.31)	0.07	-2.20	(0.83)	0.00
Old	0.15	(0.27)	0.01	-0.79	(0.56)	0.00
Income percentile						
Below 50th	0.02	(0.28)		-2.95	(0.76)	
Above 50th	0.02	(0.30)	1.00	-1.40	(0.82)	0.00
Mean	0.02			-2.15		

Notes: This table shows the conditional means and the clustered standard errors of the cumulative returns for the destination portfolio with respect to the origin portfolio, by demographic groups. Columns (1) and (2) show the estimation for two months after the switch Columns (4) and (5) present them twelve months after the switch. Columns (3) and (6) show the p-values that test for the difference between the corresponding coefficient and the baseline.

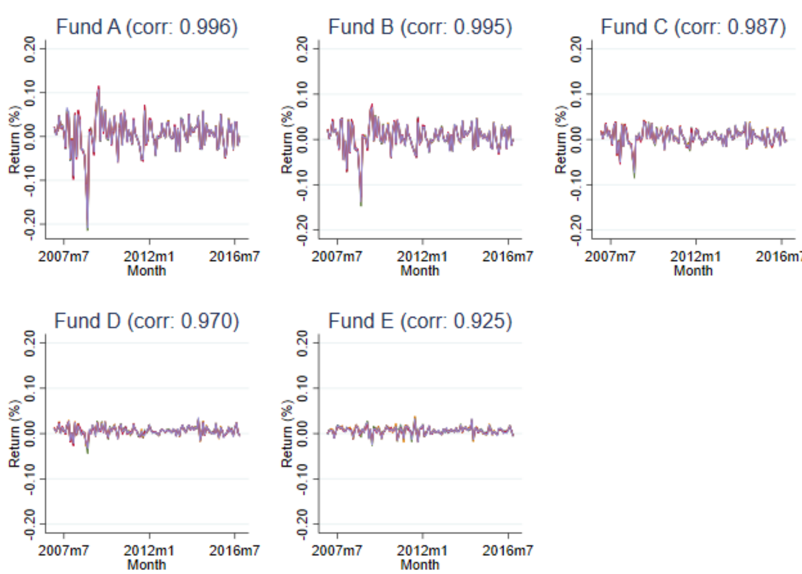
3.7.2 Figures

FIGURE 3.9: Appendix figures

(a) Monthly returns by AFP, all observations



(b) Monthly returns by AFP, without one outlier

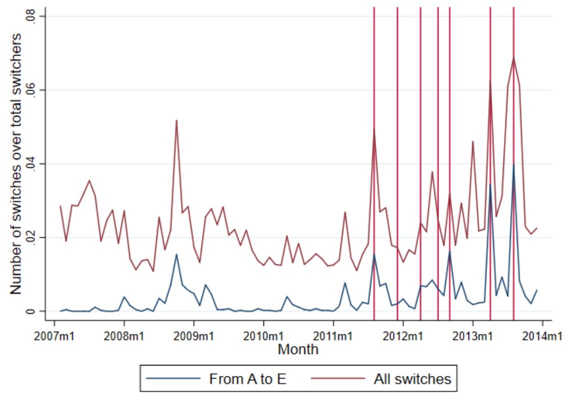


Notes: These figures present the monthly correlations of returns among firms, separately for each fund. In both panels it shows a high correlation of returns across firms. The panel above includes all the observations, and the bottom excludes one outlier observation. The last month in which one firm was functioning before the merger with another AFP.

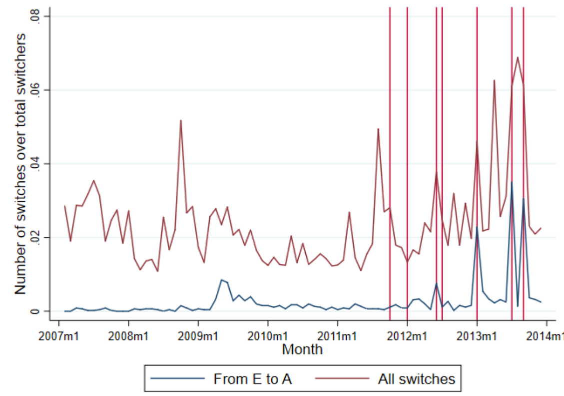
FIGURE 3.9: Appendix figures (continued)

Reallocations and recommendations

(c) From fund A to E



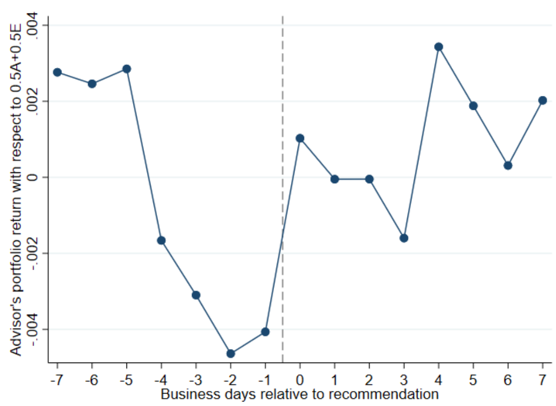
(d) From fund E to A



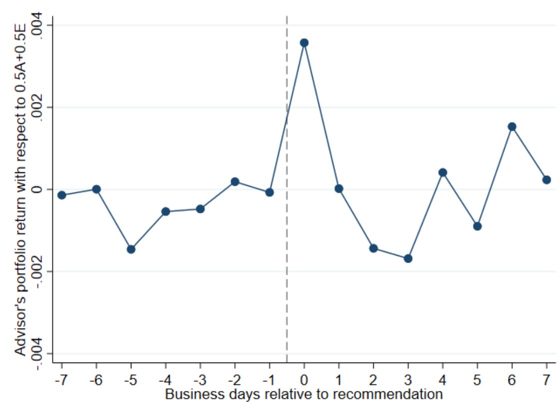
Notes: These figures show the number of switches in FyF's advised direction over the total number of switchers in each month. Vertical solid lines indicate FyF recommendations in the corresponding direction (A to E on the left panel and E to A on the right panel).

Event study, daily return

(e) Recommendations 1 to 5



(f) Recommendations 6 to 28

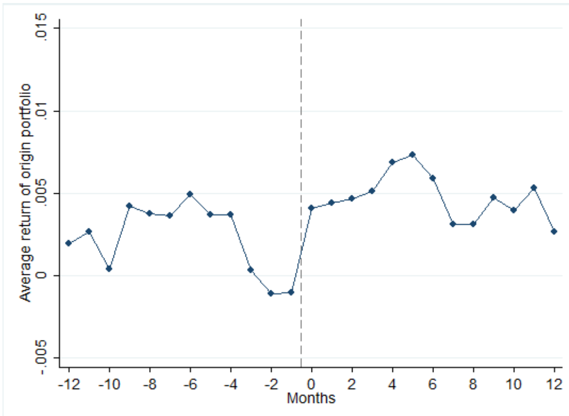


Notes: These figures show an event study for the advisor's portfolio return with respect to a benchmark portfolio (0.5A+0.5E), around the day in which a recommendation was sent. Each figure uses a different subset of recommendations, the left figure uses the first 5 recommendations, and the right figure uses all the other recommendations.

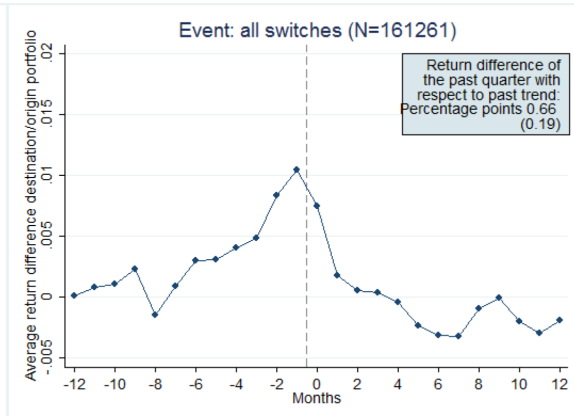
FIGURE 3.9: Appendix figures (continued)

Event study around the time of the switch

(g) Return of the origin portfolio



(h) Reallocation occurs on the first day of the month



Notes: These figures present event studies around the time of the switch. The figure on the left shows the return of the origin portfolio, when switching away of it. The figure on the right shows the difference in average returns between the destination and origin portfolios under the assumption that individuals switch on the first day of the month in which I observe the reallocation.

3.7.3 Income information

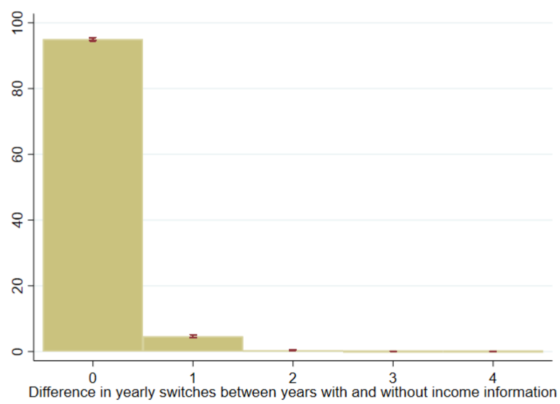
The data source used includes income information that is matched only for 52% of the individual-month observations. About 3.3% of the matched income data has a value of zero. This is indicative that the missing data is not clearly related to unemployment spells, and hence there is no reason to assign a zero value to the months with missing income. Instead of this and to use the most of the data, the approach I follow is to compute the income variables as means of the yearly available information. Under this approach, the computed income measures the *potential* income an individual can earn in a given month.

For the purposes of this paper, I'm interested in measuring income to the extent that it affects how individuals make investment choices. From the literature review, one important driver in individuals' choices is financial literacy, which is related to education, and thus to income level. Therefore, the measure of *potential* income is of interest in this context since it can be thought of as related to financial literacy.

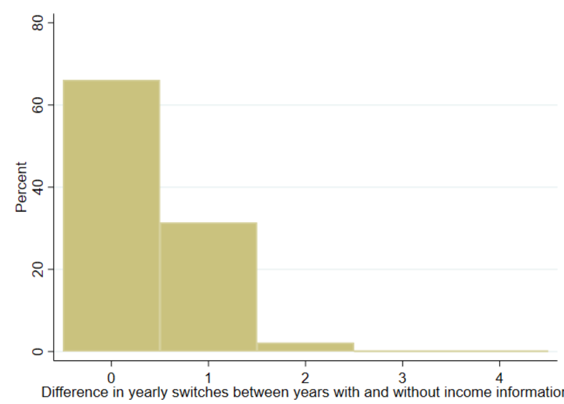
Nevertheless, one concern with this approach is that missing data might be related to individuals switching behavior. To explore this I compare the number of yearly switches that an individual makes in two consecutive years, in which there is recorded income information for at least one month in one of those years, and only missing income data for the other year. Figure 3.9i plots the share of individuals for each difference in switches between the two consecutive years. It shows that in most of the cases there was no difference in the number of switches. Only 99.6% of individuals and 97.5% of the switchers had at most one difference of switches.

FIGURE 3.9: Appendix figures (continued)

(i) Reallocations and income, Full sample



(j) Reallocations and income, Switchers



Notes: This figure presents the shares of individuals for each of the possible differences in the number of switches between years where an there is at least one month with income information and years with missing income information. The sample used includes individuals for which there is information in two consecutive years (one with income information and the other without), and the horizontal axis shows the difference in the number of yearly switches.

3.7.4 Financial advisor strategy

In this section, I take a short digression to show how the financial advisor's (FyF) strategy affects the followers[18]. Following FyF's recommended strategy impacts the individual's cumulative returns through two main channels: (1) the moment at which an individual started to follow the advice, and (2) the days of delay in implementing the strategy. The latter is quite relevant since it is not possible to reallocate funds in one day, this process takes about four business days.

To study how FyF's advising strategy can affect individuals, I compare FyF's portfolio returns of FyF at the daily level to capture more closely the market timing approach they follow. A good benchmark for their portfolio is to use a portfolio that is divided 50-50% between fund A and E, since they recommend switches between those funds. I use an event study for the timing of the advice at the daily level, for the period 2011-2016, where $d = 0$ is the day of the advice.

rom figure 3.9k we can learn two facts. First, reallocations occur when the advisors' portfolio is underperforming the benchmark. Second, the financial advisor seems to be following a very sensitive market timing approach. On the day they send the recommendation the return difference spikes, but then it goes back to the previous trend[19]. This is a key feature that affects individuals returns, since portfolio reallocations in practice take about four business days.

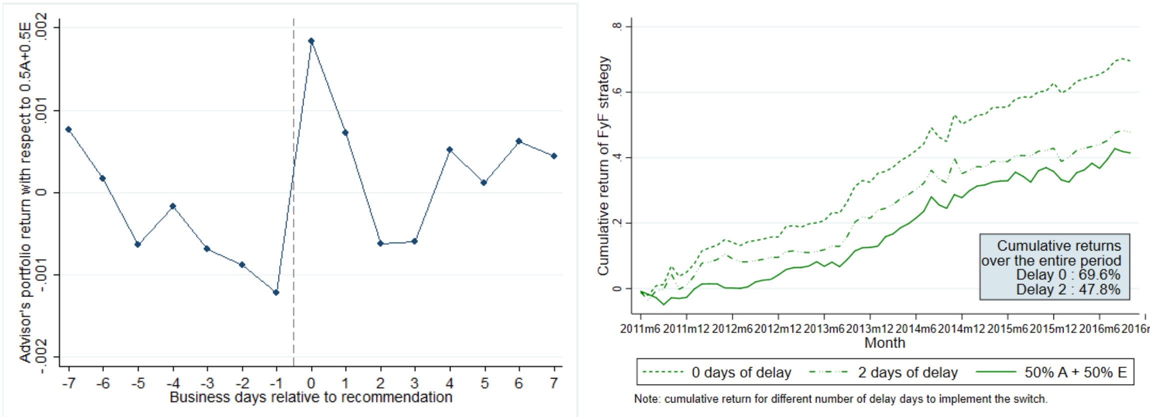
Figure 3.9l shows the unadjusted nominal cumulative return of FyF's portfolio considering zero and two days of delay in implementing the recommended strategy. As the number of days of delay increases, the cumulative return decreases and gets closer to the benchmark portfolio. This shows that actually there are no major gains in terms of re-

[18]For a detailed analysis of their strategy and how the fund reallocations impact stock prices, see Da et al. (2016).

[19]This daily event study can be split into two subsets of recommendations to show two different patterns in the advice. The first five recommendations show that the advisors' portfolio was doing much worse than the benchmark before sending the advice, and the next recommendations show a tight market timing of one day. This is shown in figures 3.9e and 3.9f in the appendix 3.7.2.

FIGURE 3.9: Appendix figures (continued)

(k) FyF's strategy: Event study, daily return (l) FyF's strategy: Implementation delay



Notes: The figure on the left shows an event study for the advisor's portfolio return with respect to a benchmark portfolio (0.5A+0.5E), around the day in which a recommendation was sent. The figure on the right shows the nominal cumulative returns of following FyF's strategy with zero and two days of delay in its implementation, along with the cumulative returns for the benchmark portfolio. Recall that the processing time for individuals' switches is about 4 days.

turns from following FyF's advice. Moreover, the four business days that take to process individuals' switches would decrease even more the cumulative returns over time.

BIBLIOGRAPHY

Ahman, E. & Shah, I. (2011). Unsafe abortion: global and regional estimates of the incidence of unsafe abortion and associated mortality in 2008.

Ahmed, J., Barber, B. M., & Odean, T. (2016). Made poorer by choice: worker outcomes in Social Security vs. private retirement accounts. *Journal of Banking & Finance*.

Alzúa, M. L. & Velázquez, C. (2018). Efectos educativos de la fecundidad adolescente: evidencia causal a partir de la legalización del aborto en uruguay.

Antón, J.-I., Ferre, Z., Triunfo, P., et al. (2016). *The impact of abortion legalisation on birth outcomes in Uruguay*. Technical report.

Baggaley, R. F., Irvine, M. A., Leber, W., Cambiano, V., Figueroa, J., McMullen, H., Anderson, J., Santos, A. C., Terris-Prestholt, F., Miners, A., et al. (2017). Cost-effectiveness of screening for hiv in primary care: a health economics modelling analysis. *The lancet HIV*, 4(10), e465–e474.

Barber, B. M., Lee, Y.-T., Liu, Y.-J., & Odean, T. (2009). Just how much do individual investors lose by trading? *Review of Financial studies*, 22(2), 609–632.

Barber, B. M. & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, 21(2), 785–818.

Benartzi, S. & Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American economic review*, (pp. 79–98).

Benartzi, S. & Thaler, R. H. (2007). Heuristics and biases in retirement savings behavior. *The journal of economic perspectives*, 21(3), 81–104.

Briozzo, L., Gómez Ponce de León, R., Tomasso, G., & Faúndes, A. (2016). Overall and abortion-related maternal mortality rates in uruguay over the past 25 years and their association with policies and actions aimed at protecting women's rights. *International Journal of Gynecology & Obstetrics*, 134(S1).

Briozzo, L., Vidiella, G., Rodríguez, F., Gorgoroso, M., Faúndes, A., & Pons, J. E. (2006). A risk reduction strategy to prevent maternal deaths associated with unsafe abortion. *International Journal of Gynecology & Obstetrics*, 95(2), 221–226.

Buchmueller, T. C. & Goldzahl, L. (2018). The effect of organized breast cancer screening on mammography use: Evidence from france. *Health economics*, 27(12), 1963–1980.

Bundorf, K., Polyakova, M., & Tai-Seale, M. (2019). *How do Humans Interact with Algorithms? Experimental Evidence from Health Insurance*. Technical report, National Bureau of Economic Research.

Calvet, L. E., Campbell, J. Y., & Sodini, P. (2009). Fight or flight? Portfolio rebalancing by individual investors. *The Quarterly journal of economics*, 124(1), 301–348.

Campbell, J. Y. (2016). Restoring Rational Choice: The Challenge of Consumer Financial Regulation. *American Economic Review*, 106(5), 1–30.

CDC (2009). The power of prevention: Chronic disease... the public health challenge of the 21st century. *Atlanta, GA: National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention*.

Chalmers, J. & Reuter, J. (2010). *What is the Impact of Financial Advisors on Retirement Portfolio Choices & Outcomes?* Technical report, National Bureau of Economic Research.

Choi, J. J., Laibson, D., Madrian, B. C., & Metrick, A. (2002). Defined contribution pensions: Plan rules, participant choices, and the path of least resistance. *Tax policy and the economy*, 16, 67–113.

Clark-Murphy, M., Gerrans, P., & Speelman, C. (2009). Return chasing as a driver in individual retirement savings investment choices: Evidence from Australia. *Journal of family and economic issues*, 30(1), 4–19.

Cram, P., Fendrick, A. M., Inadomi, J., Cowen, M. E., Carpenter, D., & Vijan, S. (2003). The impact of a celebrity promotional campaign on the use of colon cancer screening: the katie couric effect. *Archives of Internal Medicine*, 163(13), 1601–1605.

Da, Z., Larrain, B., Sialm, C., & Tessada, J. (2016). *Coordinated Noise Trading: Evidence from Pension Fund Reallocations*. Technical report, National Bureau of Economic Research.

Deri, C. (2005). Social networks and health service utilization. *Journal of Health Economics*, 24(6), 1076–1107.

Edmans, A., Garcia, D., & Norli, Ø. (2007). Sports sentiment and stock returns. *The Journal of Finance*, 62(4), 1967–1998.

Egan, M., Matvos, G., & Seru, A. (2016). *The market for financial adviser misconduct*. Technical report, National Bureau of Economic Research.

Einav, L., Finkelstein, A., Oostrom, T., Ostriker, A. J., & Williams, H. L. (2019). *Screening and Selection: The Case of Mammograms*. Technical report, National Bureau of Economic Research.

Engström, S. & Westerberg, A. (2003). Which individuals make active investment decisions in the new Swedish pension system? *Journal of Pension Economics & Finance*, 2(3), 225.

Fadlon, I. & Nielsen, T. H. (2017). *Family Health Behaviors*. Technical report, National Bureau of Economic Research.

Fernandez, M. M., Coeytaux, F., Gomez Ponce de León, R., & Harrison, D. L. (2009). Assessing the global availability of misoprostol. *International Journal of Gynecology & Obstetrics*, 105(2), 180–186.

Finseraas, H., Jakobsson, N., & Svensson, M. (2017). Do knowledge gains from public information campaigns persist over time? results from a survey experiment on the norwegian pension reform. *Journal of Pension Economics & Finance*, 16(1), 108–117.

Fiol, V., Briozzo, L., Labandera, A., Recchi, V., & Piñeyro, M. (2012). Improving care of women at risk of unsafe abortion: Implementing a risk-reduction model at the uruguayan-brazilian border. *International Journal of Gynecology & Obstetrics*, 118(S1).

Fischer, S., Royer, H., & White, C. (2017). *The Impacts of Reduced Access to Abortion and Family Planning Services: Evidence from Texas*. Technical report, National Bureau of Economic Research.

Goldin, C. & Katz, L. F. (2002). The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of political Economy*, 110(4), 730–770.

Hastings, J. S., Madrian, B. C., & Skimmyhorn, W. L. (2013). Financial literacy, financial education, and economic outcomes. *Annu. Rev. Econ.*, 5(1), 347–373.

Haushofer, J., John, A., & Orkin, K. (2019). *Can simple psychological interventions increase preventive health investment?* Technical report, National Bureau of Economic Research.

Hirshleifer, D. & Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *The Journal of Finance*, 58(3), 1009–1032.

Illanes, G. (2017). Switching Costs in Pension Plan Choice (Working Paper).

Impavido, G., Lasagabaster, E., Garc, M., et al. (2010). *New policies for mandatory defined contribution pensions: Industrial organization models and investment products*. World Bank Publications.

Maciosek, M. V., Coffield, A. B., Flottemesch, T. J., Edwards, N. M., & Solberg, L. I. (2010). Greater use of preventive services in us health care could save lives at little or no cost. *Health Affairs*, 29(9), 1656–1660.

Madrian, B. & Shea, D. (2001). The power of suggestion: inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics*, (pp. 18–116).

McMorrow, S., Kenney, G. M., & Goin, D. (2014). Determinants of receipt of recommended preventive services: implications for the affordable care act. *American journal of public health*, 104(12), 2392–2399.

Montoy, J. C. C., Dow, W. H., & Kaplan, B. C. (2016). Patient choice in opt-in, active choice, and opt-out hiv screening: randomized clinical trial. *bmj*, 352, h6895.

Montoy, J. C. C., Dow, W. H., & Kaplan, B. C. (2018). Cash incentives versus defaults for hiv testing: A randomized clinical trial. *PloS one*, 13(7), e0199833.

Mullainathan, S., Noeth, M., & Schoar, A. (2012). *The market for financial advice: An audit study*. Technical report, National Bureau of Economic Research.

Myerson, R. M., Lakdawalla, D., Colantonio, L. D., Safford, M., & Meltzer, D. (2018). *Effects of Expanding Health Screening on Treatment-What Should We Expect? What Can We Learn?* Technical report, National Bureau of Economic Research.

Nakagawa, F., Lodwick, R. K., Smith, C. J., Smith, R., Cambiano, V., Lundgren, J. D., Delpech, V., & Phillips, A. N. (2012). Projected life expectancy of people with hiv according to timing of diagnosis. *Aids*, 26(3), 335–343.

Nichols, B. E. & Meyer-Rath, G. (2017). Cost-effectiveness of early identification of hiv infection. *The Lancet HIV*, 4(10), e431–e432.

Oster, E. (2018a). *Behavioral Feedback: Do Individual Choices Influence Scientific Results?* Technical report, National Bureau of Economic Research.

Oster, E. (2018b). Diabetes and diet: purchasing behavior change in response to health information. *American Economic Journal: Applied Economics*, 10(4), 308–48.

Reitsema, M., Steffers, L., Visser, M., Heijne, J., van Hoek, A. J., Van Der Loeff, M. S., Van Sighem, A., Van Benthem, B., Wallinga, J., Xiridou, M., et al. (2019). Cost-effectiveness of increased hiv testing among msm in the netherlands. *Aids*, 33(12), 1807–1817.

Requena, M. (1966). Condiciones determinantes del aborto inducido.

Rotarou, E. S. & Sakellariou, D. (2018). Determinants of utilisation rates of preventive health services: evidence from chile. *BMC public health*, 18(1), 839.

Sedgh, G., Singh, S., Shah, I. H., Åhman, E., Henshaw, S. K., & Bankole, A. (2012). Induced abortion: incidence and trends worldwide from 1995 to 2008. *The Lancet*, 379(9816), 625–632.

Singh, S. & Maddow-Zimet, I. (2016). Facility-based treatment for medical complications resulting from unsafe pregnancy termination in the developing world, 2012: a review of evidence from 26 countries. *BJOG: An International Journal of Obstetrics & Gynaecology*, 123(9), 1489–1498.

Snyder, L. B. (2007). Health communication campaigns and their impact on behavior. *Journal of nutrition education and behavior*, 39(2), S32–S40.

Snyder, L. B., Hamilton, M. A., Mitchell, E. W., Kiwanuka-Tondo, J., Fleming-Milici, F., & Proctor, D. (2004). A meta-analysis of the effect of mediated health communication campaigns on behavior change in the united states. *Journal of health communication*, 9(S1), 71–96.

Wakefield, M. A., Loken, B., & Hornik, R. C. (2010). Use of mass media campaigns to change health behaviour. *The Lancet*, 376(9748), 1261–1271.

Wang, A., Clouston, S. A., Rubin, M. S., Colen, C. G., & Link, B. G. (2012). Fundamental causes of colorectal cancer mortality: the implications of informational diffusion. *The Milbank Quarterly*, 90(3), 592–618.

Zah, V. & Toumi, M. (2016). Economic and health implications from earlier detection of hiv infection in the united kingdom. *HIV/AIDS (Auckland, NZ)*, 8, 67.