Redistributive Tax and Transfer Policies in Developing Countries

by

Juliana Londoño-Vélez

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Emmanuel Saez, Chair
Professor Edward A. Miguel
Professor Gabriel Zucman
Professor Hilary Hoynes

Spring 2019

Redistributive Tax and Transfer Policies in Developing Countries

Abstract

Redistributive Tax and Transfer Policies in Developing Countries

by

Juliana Londoño-Vélez

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Emmanuel Saez, Chair

This dissertation studies redistributive tax and transfer policies in developing countries. It examines how these policies can be shaped to level the playing field, promote equal opportunity, and reduce inequalities. To tackle these questions, it exploits large administrative datasets and original survey data, and focuses on empirical settings where plausibly exogenous variation can be leveraged using program evaluation methods and quasi-experimental techniques.

In recent years, wealth taxation has received renewed interest as a way to raise revenue and address inequality. The first chapter of my dissertation studies how wealthy individuals respond to personal wealth taxes and wealth tax enforcement. This chapter is co-authored with Javier Avila-Mahecha. We use Colombian tax return microdata from 1993 to 2016 linked with the leaked "Panama Papers." We exploit discrete jumps in wealth tax liability and reforms varying exemption cutoffs and tax rates. We find clear evidence that individuals lower their reported wealth to reduce their tax burden, especially by misreporting items subject to less third-party reporting. We complement this analysis by studying offshore sheltering in Colombia's most relevant tax havens. Offshore entities have become more popular since the reintroduction of wealth taxes in Colombia. These entities are predominantly used by the wealthiest taxpayers, at least in part to hide assets from the tax authority. Finally, better enforcement helps recover tax on offshore wealth. A voluntary disclosure scheme revealed wealth worth 1.7 percent of GDP, most of it having been hidden offshore. The wealthiest individuals disproportionately disclosed under the scheme and, as a result, pay more taxes. Halfway through the scheme, the Panama Papers news story broke, raising disclosures by more than 800 percent. The higher perceived detection probabilities, coupled with harsher noncompliance sanctions, improved wealth tax collection and enhanced tax progressivity at the top.

In addition to progressive taxation, governments may reduce inequality by leveling access to colleges through financial aid. The second chapter, co-authored with Fabio Sanchez and Catherine Rodriguez, exploits a large-scale financial aid program in Colombia for low-income high-achievers to study how aid affects postsecondary enrollment, college choice, and student

composition. Our regression discontinuity estimates show aid eligibility raised immediate enrollment by 56.5 to 86.5 percent, depending on the complier population. This rise, driven by matriculation at private high-quality colleges, closed the SES enrollment gap among high-achievers. Moreover, a difference-in-differences approach suggests enrollment of aid-ineligible students also improved because college supply expanded in response to heightened demand. With ability stratification largely replacing SES stratification, diversity increased 46 percent at private high-quality colleges.

This large expansion in socioeconomic diversity has the potential to deeply affect interactions between low- and high-income students at elite private colleges. In the third chapter, I shift attention to the high-income classmates of aid recipients and ask, does socioeconomic diversity shape their perceptions of inequality and their preferences for redistribution? I collect original survey data and conduct survey experiments to measure these outcomes, which are not available in the administrative records. I exploit the plausibly exogenous timing of the aforementioned financial aid program and test the effect of this shock in peer characteristics. Specifically, I test the effect of exposure to low-income classmates on high-income students' interactions with low-income peers, and their perceptions of inequality and poverty, beliefs of social justice, and redistributive preferences. I find significant positive effects of exposure to low-income peers on interactions among students with heterogeneous family backgrounds, and an increase in perceived inequality, poverty and upward social mobility, and meritocracy in college admissions. I also find a stronger support for redistribution. Lastly, I rule out any negative effects on academic performance in college.

To my parents, Juan Luis and María Zulema.

Contents

List of Figures

List of Tables

Acknowledgments

I am very fortunate to have been trained as an economist at Berkeley. The faculty is truly outstanding in their dedication to rigorous scholarship and excellent advising, always being incredibly generous with their constructive feedback. I have so much admiration and respect for them, and I hope to be able to follow their example. In particular, I am profoundly grateful to Emmanuel Saez for his tremendous guidance, thoughtful mentorship, and constant support throughout all these years. I am also deeply grateful to Ted Miguel for his encouragement at all stages of my PhD. Gabriel Zucman and Hilary Hoynes have provided invaluable feedback to improve my work. I have benefited immensely from insightful conversations with Alan Auerbach, Patrick Kline, David Card, Christopher Walters, Jesse Rothstein, Fred Finan, and Danny Yagan.

I am also grateful to my co-authors, Catherine Rodríguez, Fabio Sánchez, and Javier Ávila-Mahecha, who made this dissertation possible.

My experience these years in Berkeley has been wonderful thanks to many friends and colleagues, including Nicholas Y. Li, Alessandra Fenizia, Jonathan Schellenberg, Avner Shlain, Daniel Haanwinckel, Evan K. Rose, Itai Trilnick, Zarek Brot-Goldberg, Darío Tortarolo, Arlen Y. Guarín, Maxim Massenkoff, Carla Johnston, Jonathan Holmes, Francis Wong, Erin Kelley, Hadar Avivi, and Ceren Baysan. I have enjoyed the camaraderie of the members of Women in Economics at Berkeley.

My family has given me infinite love, inspiration, and support. My father, Juan Luis, is the greatest role model of all. My mother, María Zulema, has been my rock throughout all the adventures, and her enthusiasm and positivity have been paramount in my life. My sister, Daniela, and my brother, Juan Felipe, have celebrated my triumphs and supported me during my struggles. My aunts, uncles, and cousins are my backbone. Lucía and Zulema gave me admirable examples of life. My mother-in-law, Debbie, has been the greatest gift.

I am sincerely thankful to the fantastic Staff, Vicky Lee, Camille Fernandez, Heather Iwata, and Patrick Allen, for all their help. I also want to acknowledge generous financial support from the Center for Equitable Growth, the Robert D. Burch Center for Tax Policy and Public Finance, and the Center for Effective Global Action.

Lastly, I would like to thank Yotam, the love of my life and my best friend, for the endless love and support he selflessly gives me every single day. Life is wonderful next to him.

Chapter 1

Can Wealth Taxation Work in Developing Countries? Quasi-Experimental Evidence from Colombia

1.1 Introduction

Progressive wealth taxation has received renewed interest as a tool to raise revenue and curb inequality.[1] Following Piketty's (2014) call for a global wealth tax and, more recently, Senator Warren's proposal to tax wealth in the United States, much of the discussion has focused on whether wealth taxes are enforceable if taxpayers (legally) avoid or (illegally) evade them. For instance, taxpayers might hide assets in tax havens or underreport their wealth, jeopardizing the feasibility of taxing wealth in practice. Indeed, tax sheltering by wealthy individuals raises the resource cost of levying taxes, offsets wealth tax progressivity, and hinders its redistributive appeal—concerns that are exacerbated in developing countries with high inequality and low tax compliance. While governments may have tools to control enforcement, we know little about the effectiveness of these tools in practice.

A nascent empirical literature estimating behavioral responses to wealth taxes takes enforcement as given and largely ignores key sheltering mechanisms used by wealthy taxpayers, such as offshoring to tax havens.[2] Indeed, it is inherently challenging to reliably measure and identify offshore sheltering responses to taxation (inter alia, secrecy is precisely what makes offshore shell corporations attractive). Further, how individuals respond to taxes depends not only on the tax incentives but, critically, also on the enforcement environment (Slemrod, 2017). Yet there is surprisingly scant evidence on how wealthy individuals re-

[1] Wealth taxes are levied on the stock of (financial and non-financial) assets, net of liabilities.

[2] See Seim (2017); Jakobsen, Jakobsen, Kleven, and Zucman (2018); Brulhart, Gruber, Krapf, and Schmidheiny (2017); Zoutman (2015); Durán-Cabé, Esteller-Moré, and Mas-Montserrat (2017).

spond to stronger enforcement, partly because sources of exogenous variation on wealth tax enforcement have so far been elusive.

This paper breaks new grounds on these issues providing quasi-experimental evidence on behavioral responses to personal wealth taxes and wealth tax enforcement in Colombia. We use extensive tax microdata on wealth covering all income and wealth tax filers between 1993 and 2016, while leveraging Colombia's numerous tax and enforcement policy reforms. First, we estimate how elastic reported wealth is with respect to wealth taxation and the drivers of this elasticity, i.e., real versus reporting responses. We then delve into studying offshoring to Colombia's most relevant tax havens by linking our tax returns to the Panama Papers leak and tracking the clients of Mossack Fonseca, one of the world's five largest wholesalers of offshore secrecy. Lastly, we evaluate how offshore evasion responds to improvements in enforcement, exploiting exogenous shocks to detection probabilities and the punishment for cheating. This piece of evidence is critical to evaluate the feasibility of wealth taxes in a globalized world, given the fear that wealth taxation triggers capital flight to tax havens and the recent global crackdown on offshore tax evasion through voluntary disclosure schemes and mandatory information reporting requirements (e.g., the US Foreign Account Tax Compliance Act—FATCA). In conjunction, these analyses provide the necessary ingredients to assess whether wealth taxes can work, even when the stakes are particularly high.

In the first part of the paper, we exploit variation from discontinuities in Colombia's wealth tax schedule and reforms modifying who pays the wealth tax and what share of wealth is taxed. The Colombian wealth tax schedule features discrete jumps in tax liability at given thresholds of reported wealth, i.e., tax *notches*. For instance, in 2010, an income taxpayer reporting just below 1 billion pesos (USD 520,830) owed no wealth tax, while a taxpayer reporting an additional peso owed 1 percent of all taxable net wealth, i.e., a tax bill of 10 million pesos (USD 5,208.3). Multiple tax reforms modified both the location of the bracket cutoffs and the wealth tax rate, which ranged from 0 to 6 percent; a very large policy experiment relative to previous studies.[3] Critically, we observe wealth holdings before and after reforms for individuals above and below the cutoffs because taxpayers annually report end-of-year *wealth* (e.g., taxable and non-taxable bank deposits, equities, business assets, real estate, vehicles, debt) in their *income* tax statements regardless of whether they owe wealth taxes.

Leveraging the variation in wealth tax rates across brackets and time, we apply the bunching techniques to estimate the elasticity of substitution between truthfully reporting individual wealth holdings to the tax authority. If individuals do not respond to taxation, reported wealth will be distributed smoothly around the notch points. If, instead, individuals avoid the jumps in wealth tax liability by reporting below the notch points, they will bunch just below them—the quantity of bunching suggesting the responsiveness to the tax policy (see Saez, 2010; Chetty, Friedman, Olsen, and Pistaferri, 2011; Kleven and Waseem, 2013). Reassuringly, we can use the observed pre-reform distribution as our counterfactual density,

[3] Assuming the return to wealth is 6.6 percent, a wealth tax rate of 6 percent implies a 90.9 percent tax on the return to wealth. In contrast, the top wealth tax rate ranged between 1 and 2.5 percent in Europe.

alleviating concerns about implicit functional form or preference assumptions (Blomquist and Newey, 2017). We find large and immediate bunching in response to the wealth tax. In our main analysis, the marginal buncher reports 21 percent lower wealth because of the tax notch. A one percent increase in (one minus) the wealth tax rate raises reported wealth by 2 percent, slashing up to a fifth of the wealth tax revenue. Despite these behavioral responses, in the short run, Colombia remains below its revenue-maximizing wealth tax rate (in the long run, individuals might also potentially adjust their real behavior, like dissaving or making bequests).

To decompose our estimated behavioral response, we use a two-stage least squares approach that characterizes changes in taxpayer reporting behavior induced by wealth taxation. We compare the type of wealth individuals report before and after tax notches are introduced, characterizing taxpayers who bunch in response to the notches. This is akin to an analysis of the compliers characteristics (see Abadie, 2002; Diamond and Persson, 2016). Bunchers avoid the wealth tax by (artificially) inflating liabilities and underreporting business assets not subject to third-party reporting. Consistent with reporting responses driving the bunching elasticity, bunchers exploit the differential coverage of third-party reporting to underreport the type of wealth less likely to be detected by the tax authority.

In the second part of the paper, we focus on offshoring assets to tax havens. We exploit data from the Panama Papers, which lists the names of shareholders of entities incorporated by Mossack Fonseca in Panama and over 20 other jurisdictions between 1977 and 2016.[4] Thanks to the publication of the data by investigative reporters and cooperation with the Colombian tax authority, we match the data to our tax returns using personal names. Leveraging the date individuals incorporated these offshore entities, we examine the interaction between wealth taxation and the dynamics of offshore entity incorporation and tax sheltering at the top. The reintroduction of the wealth tax was followed by a tenfold increase in Colombian entities incorporated in tax havens. These offshore entities were predominantly used by very wealthy individuals: even if underreporting, the wealthiest 0.01 percent is 24 times as likely to appear named in the Panama Papers than the wealthiest 5 percent. Offshore entities were used at least in part to hide assets and minimize the tax burden: although Colombia's residence-based tax system requires reporting all foreign assets, an event study analysis shows the reported value of assets drops 10.9 percent as soon as they incorporate their entity offshore.

In the third part of the paper, we examine how wealthy individuals responds to improvements in enforcement. To crack down on offshore evasion and recover tax on hidden wealth, Colombia implemented a voluntary disclosure program similar to those recently implemented in the United States and Europe (OECD, 2010, 2015). Generous tax incentives were awarded to disclosures of unreported assets and/or nonexistent liabilities in 2015, 2016, or 2017, while evaders who did not come forward faced strict penalties if caught cheating. Wealth worth

[4] Geographic proximity and political stability contributed in making Panama the most popular destination for foreign assets owned by Colombians, after the US. Even before the leak, Colombians were almost 15 times as likely to report owning foreign assets in Panama than in the Virgin Islands or Switzerland.

1.73 percent of GDP was disclosed under the scheme. Admission of noncompliance rises sharply with wealth: the wealthiest 0.01 percent was 55 times more likely to disclose under the scheme than the wealthiest 5 percent. In all, two-fifths of taxpayers in the wealthiest 0.01 percent disclosed some hidden wealth.[5] At least a third of evaders' wealth had been hidden abroad, concomitant with the pervasiveness of offshore tax evasion at the top.[6] Critically, awarding tax incentives for disclosures had persistent effects on wealth and income tax compliance. A difference-in-differences analysis shows disclosers reported 49.2 percent more wealth as well as more capital income relative to nondisclosers three years after their first disclosure. As a result, disclosers also paid 39 percent more income taxes, further raising revenue collected from the wealthiest tax filers and enhancing the progressivity of the tax system as a whole.

With compliance responding not only to the tax incentives, we test how wealthy taxpayers react to exogenous increases in the perceived risk of detection and the punishment for cheating. Halfway through the disclosure program, the Panama Papers news story broke. The Colombian tax authority reacted by scrutinizing Mossack Fonseca and its clients, contacting taxpayers named in the leak and requesting documentation of their offshore activities and transactions. Three weeks after the leak, the governments of Colombia and Panama announced a tax information exchange agreement between the two countries—a move the tax haven had resisted for years. Exploiting the exogenous timing of the leak, we compare outcomes between wealth tax filers named (treated) and not named (control) in the leak before and after it occurred. The difference-in-differences coefficient shows that the Panama Papers and subsequent events induced a 27 percentage point increase in the likelihood of disclosing, i.e., a ninefold increase relative to the pre-event control mean. Disclosures of foreign assets in particular increased more than fifteenfold. Consequently, taxes paid by these individuals more than doubled after the leak.

Six months later, Colombia criminalized tax evasion for the first time. If convicted, tax evaders could face up to nine years in prison. Critically, we find that most disclosers admitted to noncompliance immediately after the criminalization of tax evasion and before the disclosure window expired, arguably at least in part due to the harsher sanction for noncompliance.

We contribute to three main literatures. First, there is the aforementioned empirical literature on wealth taxes in Europe (Seim, 2017; Jakobsen et al., 2018; Brulhart et al., 2017; Zoutman, 2015; Durán-Cabé et al., 2017). Compared to previous studies, we address offshore tax evasion and examine how enforcement policy improves tax compliance among wealthy taxpayers. Further, we examine responses from multiple wealth tax changes, some of which are much larger than in the existing literature. Finally, our estimated elasticities—the

[5] Top 0.01 percent individuals are almost three times as likely to disclose under these schemes in Colombia than in Norway and Sweden (Alstadsater, Johannesen, and Zucman, forthcoming).

[6] Accounting for hidden offshore wealth increases measured wealth inequality. A conservative correction shows that the top 1 percent share of total wealth rises from 40.6 to 43.2 percent. Because offshore wealth is extremely concentrated at the top, the rise is more pronounced for the wealthiest 0.1 percent.

first in a developing country context—can be interpreted as an upper bound on responses occurring in settings with a more sophisticated enforcement technology.

Second, a burgeoning literature focuses on globalization and offshore tax evasion (Alstadsater et al., forthcoming; Zucman, 2015), and the effectiveness of the global wave of crackdowns on tax havens through enhanced cross-country information exchange agreements (Johannesen, Langetieg, Reck, Risch, and Slemrod, 2018), reporting requirements (Johannesen and Larsen, 2016; Johannesen and Zucman, 2014), and the amnesties and voluntary disclosure programs that have accompanied these enforcement initiatives (Bayer, Oberhofer, and Winner, 2015; Langenmayr, 2017; Alstadsater, Johannesen, and Zucman, 2018b). We are the first to shed light on the interactions between wealth taxation, offshoring, and stronger enforcement. Moreover, we show how the events triggered by Panama Papers leak improved tax compliance among very wealthy individuals. To our knowledge, this is the first direct evidence of whistleblowing contributing to tax compliance.

Lastly, we contribute to the growing literature on tax design and compliance in developing countries. While most of this literature focuses on firm behavior (Naritomi, 2016; Brockmeyer and Hernandez, 2016; Pomeranz, 2015; Bachas and Soto, 2018), we examine the behavior of high net worth individuals, an understudied population that is critical for the design and evaluation of redistributive fiscal policy. It is precisely in these settings, where few individuals at the top end up holding the vast majority of wealth (Alvaredo, Chancel, Piketty, Saez, and Zucman, 2018), that progressive wealth taxation can potentially be a powerful redistributive tool.

The remainder of this paper is organized as follows. Section 1.2 describes the institutional context and our data. Section 1.3 presents bunching responses to personal wealth taxes. Section 1.4 exhibits offshore sheltering using evidence from the Panama Papers. Section 1.5 studies responses to wealth tax enforcement. Section 1.6 discusses the implications of wealth concealed offshore for the study of inequality. Finally, Section 1.7 concludes.

1.2 Background and Data

Institutional Context

Colombia is an upper middle income country, with GDP per capita of 14,154 US dollars at purchasing power parity in 2016 (World Bank International Comparison Program database). Total tax revenues represent 19.8 percent of GDP, while recurrent taxes on personal net wealth have constituted 0 to 0.27 percent of GDP between 2002 and 2017 (Figure A.1).[7] In addition to annual taxes on net worth, Colombia levies other taxes on capital (e.g., property tax, inheritance tax) and capital income (e.g., rental income, realized capital gains).

Direct taxes on income and wealth are collected by the central government tax authority, *Dirección de Impuestos y Aduanas Nacionales* (DIAN, for its Spanish acronym). Income

[7] The equivalent 2016 shares in Spain, Norway, and Switzerland were 0.18, 0.43, and 1.0 percent, respectively (OECD, 2018).

and wealth taxes in Colombia are individually based and have never allowed joint filing for married couples. The tax authority records *wealth* information in *income* tax statements every year, regardless of whether income taxpayers are subject to the wealth tax. This is done to calculate the minimum income tax base or "presumptive income," which is based on net worth.[8] Income taxpayers are required to annually self-report end-of-year financial assets (e.g., cash, bank deposits, stocks, bonds, unlisted securities, financial assets held abroad), non-financial assets (e.g., real estate, land, large durables, non-corporate business assets, non-financial assets held abroad), and debt (e.g., mortgages, inter-personal debts).[9] Self-reporting assets is crucial because it provides taxpayers with sheltering opportunities. Income tax filers reporting (taxable and nontaxable) net worth above a cutoff are eligible for the wealth tax and file a separate tax statement.[10] This exemption threshold has been very high and excludes more than 99 percent of adults from the wealth tax. For instance, in 2017, only 0.2 percent of adults paid the wealth tax—a smaller fraction than other wealth-taxing countries (OECD, 2018).

Colombia has a long tradition of taxing net wealth of individuals and firms, as detailed in Appendix A.2. Recurrent wealth taxes were first introduced in 1935 and kept in place until 1992. A decade after its abolition, President Uribe Vélez reintroduced wealth taxation in 2002 to finance *Seguridad Democrática*—the administration's security effort against drug trafficking, guerrilla, and paramilitary groups—earmarking its revenues for defense and security expenditures. The wealth tax was levied on taxable net worth, that is, net worth minus two allowances: the value of principal residence (up to a limit) and the net equity value of shares in domestic companies (this avoid double wealth taxation of firms and individuals).

The wealth tax was designed as a piecewise linear schedule, with each bracket associated with a fixed *average* tax rate. For instance, in 2010, individuals reporting net worth below 1 billion pesos (2010 USD 520,830) were exempt from the wealth tax, whereas those reporting an additional peso paid 1 percent of all their taxable wealth, i.e., a tax bill of USD 5,208. This wealth tax rate increases to 1.4 percent at 2 billion pesos in reported wealth, 3 percent at 3 billion pesos, and 6 percent at 5 billion pesos. The notched wealth tax schedule thus produces discontinuous jumps in tax liability at bracket cutoffs, as depicted in Figure 1.1, Panel (a). A series of reforms in the last two decades modified both the wealth tax rates and the bracket cutoffs, as illustrated by Figure 1.1, Panel (b).

While income is largely covered by third-party reporting in Colombia, there is only partial third-party reporting of wealth. Most financial wealth is subject to third-party reporting: end-of-year savings and checking account balances, loans, bonds, deposits, listed equities, voluntary pension contributions, and mortgage debt are reported by financial insti-

[8]There is a legal presumption that a taxpayer's taxable income is no less than a fixed share of her net worth reported the previous year, minus two allowances (e.g., 6 percent in 1999–2006, and 3 percent in 2007–2016).

[9]A pervasive informal sector, outdated cadastral values, and high filing thresholds imply that only a fraction of tax units (adults aged 20 and above) file income tax returns in Colombia (Alvaredo and Londoño-Vélez, 2014). For instance, in FY 2016, all but 6.6 percent of tax units were excluded from filing income taxes.

[10]Because wealth tax eligibility and rates are determined by taxable *and* non-taxable net worth in our setting, shifting from taxable to non-taxable assets will be a second-order strategy for Colombian taxpayers.

tutions. In contrast, non-financial assets such as real estate and vehicles are subject to less third-party reporting; while taxpayers should report the same values as in the property and vehicle taxes, this information is not systematically cross-verified by the tax authority. Finally, several wealth components have virtually no third-party reporting, such as cash, large durables, unlisted equities, non-corporate business assets (e.g., inventories), inter-personal debts, and—until recent developments in tax information exchange agreements with other countries—assets held abroad.

Despite technological improvements in third-party reporting since 2006 (reviewed in Appendix A.2), enforcement capacity is still limited. The few staff handling third-party reports and the tax technology available are not enough to systematic cross-check items reported in the wealth tax return using available third-party reported information. Unlike some OECD countries, there is no dedicated unit for managing the tax affairs of high net worth taxpayers (OECD, 2017). Moreover, taxpayers reporting total assets in aggregate form in a single box in the income tax return makes it more difficult to detect the sources of year-to-year changes. The tax authority carries out randomized audits and requests documentation for all reductions in wealth holdings that are not evidently compatible with changes in other positions of the tax return. However, we are not aware of the exact number of wealth tax audits performed nor the number of verification or audit activity for high net worth taxpayers, as the tax authority discloses relatively limited information regarding its verification and audit actions.[11] This offers scope for tax evasion through underreporting or altogether failing to report assets, or fabricating liabilities.[12] To our knowledge, no rigorous estimates of the extent of wealth tax evasion exist for Colombia.[13]

Data

Our data come from four main sources. Our first dataset is individual-level administrative tax microdata covering the universe of income tax filers between FY 1993 and FY 2016. These comprise 20.5 million observations (taxpayer-years) in a longitudinal panel of taxpayers. Our records contain information on the majority of items recorded in individual income tax declarations, and include total assets and debt owned by December 31 every year. Before

[11] This opacity may be optimal from a policy perspective if taxpayers overestimate audit probabilities (Bérgolo, Ceni, Cruces, Giaccobasso, and Perez-Truglia, 2018). Overall, the number of completed audits per 100 active taxpayers is extremely low in Colombia compared to OECD countries: 0.38 for the personal income tax, 1.68 for the corporate income tax, and 0.54 for the value-added tax (OECD, 2017).

[12] One way of fabricating a liability is to report a debt contract with a friend or relative such that the debt is held by someone whose wealth placed them above the threshold while the asset was held by someone located below.

[13] The incentives for individuals to own wealth through an entity rather than directly strengthened in recent years because the wealth tax for firms was progressively phased out starting 2015, to be completely eliminated by 2018 (Law 1739/2014). However, this will affect individuals mostly after 2018, i.e., after our period of study.

2004, assets are decomposed into six broad categories.[14] The income tax return was modified in 2004 and, since then, this level of wealth disaggregation is only required for taxpayers keeping records.[15] This subset of taxpayers are business owners involved in retail and other commercial ventures, and represent 10–15 percent of income taxpayers every year.

The second dataset is composed of individual-level wealth tax returns for all filers in wealth tax years from 2002 to 2017. Individuals with reported net worth above a cutoff in the income tax statement are required to submit a wealth tax return. For these individuals, and a handful of voluntary filers, this dataset includes the decomposition of taxable and non-taxable net wealth, and wealth tax liability. Between 2015 and 2017, Colombia offered tax benefits to tax evaders voluntarily disclosing hidden assets and inexistent liabilities. Information on these disclosures is also available in our data.

The third dataset is the individual-level information return on foreign assets. Since 2015, all income taxpayers owning any foreign asset must file a separate information tax reform. Tax filers report the type of asset held abroad, the location of the asset, and the value of the foreign asset (the level of disaggregation of this information depends on the value of assets owned). These records have also been made available to us.

The last dataset comes from three massive leaks published by the International Consortium of Investigative Journalists (henceforth ICIJ). The largest one comes from Panamanian law firm Mossack Fonseca (i.e., the "Panama Papers"). Around one-third of the offshore entities were incorporated through Portcullis TrustNet (now Portcullis) and Commonwealth Trust Limited (i.e., the "Offshore Leaks"). The remainder come from a trouve of data from the official corporate registry of the Bahamas (i.e., the "Bahamas Leak"). These microdata cover nearly 40 years—from 1977 through to early 2016—and link to individuals and companies in more than 200 countries and territories. The information includes, inter alia, the names of the real owners of offshore entities, the entity contact postal address, the entity incorporation and inactivation dates.

Regrettably, the ICIJ dataset has a number of limitations. First, it is restricted to offshore entities created by only the handful of laws firms and offshore service providers mentioned above. Second, not every officer of a company that appears in the three leaks shows up in the public database. This is either because information about ownership cannot easily be extracted in a systematic manner, or because the law firm or offshore service provider failed to collect the necessary information about the real owners of companies.[16] Third, information on the amount of wealth stored in the offshore entity or taxes evaded is not included. There are legitimate reasons to create a company in an offshore jurisdiction, and many law-abiding

[14] This decomposition is as follows: (i) cash, deposits in savings or checkings bank accounts, certificates of deposit, and other investments (e.g., bonds, life insurance, voluntary retirement fund); (ii) accounts receivable; (iii) stocks and contributions; (iv) inventories; (v) fixed assets (e.g., real estate, land ownership, vehicles, boats); and (vi) other assets (e.g., jewelery, art, industrial and intellectual property rights).

[15] Figure A.2 plots the decomposition of wealth by asset type for FY 2016.

[16] Indeed, anonimity and the strategic veil of secrecy is precisely what makes offshore corporations so attractive to some. For this reason, a sizeable amount of offshore entities are assigned to the Bahamas, the British Virgin Islands, etc. (Alstadsater et al., forthcoming).

individuals declare them to their tax authorities when it is required. For instance, during the more violent 1990s and early 2000s, wealthy Colombians may have preferred safekeeping their wealth abroad. The investigations conducted by the authorities are still ongoing.

Despite these limitations, the Panama Papers provide valuable information (and arguably a lower bound) on the extent to which Colombian citizens—law-abiding or otherwise—use offshore entities. We merge these data with administrative income and wealth tax records using individual names. The Panama Papers included information from 1,752 shareholders of offshore entities with a personal or entity contact address in Colombia, and we are able to match 1,208 individuals to their tax records using personal names, i.e., a match rate of 70 percent. This is partly thanks to the naming custom involving two surnames—a paternal surname, followed by a maternal surname—often practiced in Colombia. The pool of unmatched individuals represents cases where the full name does not uniquely identify an individual in the tax records, or where the individual—whether required to or otherwise—did not file an income tax record in Colombia between 1993 and 2015.

1.3 Bunching Responses to Wealth Taxes

In this section, we leverage quasi-experimental variation in wealth taxes introduced by the notched tax schedule and the tax reforms described in Section 1.2 to estimate the elasticity of net wealth with respect to the net-of-wealth tax rate. To fix ideas, consider the 2010 wealth tax reform, which lowered the exemption threshold from 3 to 1 billion pesos (2010 USD 1,562,400 to 520,830) and introduced two notches. Taxpayers previously exempt from the wealth tax suddenly faced average wealth tax rates of 1 percent if reporting 1 to 2 billion pesos or 1.4 percent if reporting 2 to 3 billion pesos. Figure 1.2 plots the density of tax filers by bins of reported wealth in 2009 (before the reform) and 2010 (after the reform). The distribution of taxpayers is smooth in the absence of wealth taxes (gray curve). In contrast, the introduction of the two tax notches is followed by the immediate emergence of excess and missing masses just below and above the notch points (blue curve). This bunching identifies a direct behavioral response to wealth taxes.

A priori, unlike earnings responses to income taxes, which potentially conflate real and sheltering responses, immediate bunching in the distribution of wealth—a stock—predominantly reflects reporting. It is difficult for individuals to immediately bunch below the notch points using real responses (e.g., investment) because wealth partly depends on asset prices, which are highly uncertain and fluctuate throughout the year (Jakobsen et al., 2018). This motivates our model of wealth underreporting, which we describe below.

The Elasticity of Reported Wealth: Theory and Evidence

Building on Kleven and Waseem (2013) and Almunia and Lopez-Rodriguez (2018), we propose a stylized model to examine the problem of utility-maximizing individuals that can underreport their wealth to shelter their fortune from taxation and incur a resource cost.

We use this framework to estimate how individuals respond to a discontinuous increase in wealth tax liability—a tax notch—at an arbitrary reported wealth threshold.

Conceptual Framework

Consider an economy with a continuum of individuals of measure one. Individuals have (latent) true wealth W, while the government levies a proportional tax τ on the wealth individuals report W_r. The wealth tax liability implies a proportional (average and marginal) tax rate on reported wealth, such that $T(W_r) = \tau W_r$. Since the tax authority does not perfectly observe true wealth, individuals may attempt to misreport it to reduce their tax burden, such that $0 \leq W_r \leq W$. Misreporting wealth implies a (direct and indirect) resource cost, which is captured by the convex cost function $C(1-W_r/W)\cdot W$. This function captures the intuition that, first, the cost of misreporting is rising in the share of unreported wealth $1 - W_r/W$ (e.g., it is more costly to underreport all of your wealth than just half of it). Second, holding this fraction constant, the misreporting cost is rising in true wealth W (e.g., it is more costly to misreport 10 percent of wealth for an individual owning 5 billion dollars than for an individual owning 500,000 dollars).

A positive wealth tax rate $\tau > 0$ depresses W_r below W, with the strength of the effect determined by elasticity e, the parameter of interest. Intuitively, if $e \to 0$, then individuals report their true wealth ($W_r \to W$), while if $e \to \infty$, individuals report no wealth at all ($W_r \to 0$). We initially assume elasticity e and resource cost function $C(\cdot)$ are the same for all individuals, but relax this assumption later. Under this homogeneity assumption, all the variation in W_r is due to differences in W. There is a smooth distribution of W in the population captured by a distribution function $F(W)$ and a density function $f(W)$. We denote $H_0(W_r)$ and $h_0(W_r)$ the distribution and density functions for reported wealth associated with the baseline linear tax system. Given a smooth tax system (i.e., no notches and no kinks), the smooth wealth distribution converts into a smooth reported wealth distribution.

Suppose that a proportional tax notch is introduced at reported wealth cutoff W_r^* so that $T(W_r) = \tau W_r + \Delta\tau \cdot W_r \cdot \mathbb{1}(W_r > W_r^*)$ where $\Delta\tau$ is the proportional tax notch and $\mathbb{1}(\cdot)$ is an indicator for being above the cutoff. Figure A.3, Panel (a), illustrates the implications of a proportional tax notch in a budget set diagram with two individuals, L and H, who have "low" and "high" wealth, respectively. Individual L chooses reported wealth W_r^* under both tax regimes, while individual H is the marginal buncher: she chooses reported wealth $W_r^* + \Delta W_r^*$ before the tax change and is exactly indifferent between reducing her reported wealth to bunch at W_r^* (which reduces the expected tax burden but implies a resource cost), or remaining at the interior point W_r^I and facing higher taxes. As a result, L (H) has the lowest (highest) pre-notch reported wealth among those who locate at the notch point. Every individual between L and H locates at the notch point: all individuals who had reported wealth in the interval $(W_r^*, W_r^* + \Delta W_r^*]$ before the introduction of the notch will bunch. There is a hole in the post-notch density distribution as no individual is willing to locate between W_r^* and W_r^I, as depicted in Figure A.3, Panel (b).[17]

Since there is a direct mapping between the true wealth distribution $f(W)$ and the pre-notch reported wealth distribution $h_0(W_r)$, we can define the number of bunching individuals at the notch point as

$$B = \int_{W_r^*}^{W_r^* + \Delta W_r^*} h_0(W_r) dW_r \approx h_0(W_r^*) dW_r^* \tag{1.1}$$

where the approximation assumes that the counterfactual density $h_0(W_r)$ is roughly constant on the bunching segment $(W_r^*, W_r^* + \Delta W_r^*)$. The number of bunching individuals depends positively on the increase in taxes at the notch and negatively on the resource cost of sheltering wealth from taxation.

We now relax the homogeneity assumptions and allow both the reported wealth elasticity and the cost of sheltering wealth to vary across individuals.[18] Individuals may face different resource cost functions $C(\cdot)$ through various channels. For instance, assets owned by individuals may be more or less covered by third-party reporting and therefore more or less manipulable. They might also depend on individuals' access to the offshore wealth management industry, preferences (e.g., risk aversion, honesty), misperception, adjustment costs, or inattention. As a result, individuals with the same underlying wealth W will face different incentives to bunch. If an individual has a prohibitively high resource cost, she might not react to the tax notch because the (perceived or real) costs of misreporting are higher than the expected tax savings of bunching.

We leverage the strong incentives created by tax notches to quantify the response that would be observed if individuals overcame these prohibitively high costs. Unlike kinks, notches can create a region of strictly dominated choice $(W_r^*, W_r^* + \Delta W_r^D]$ in which it is possible to increase individual utility by moving to notch point W_r^*, making these choices dominated under any parametric form for individual preferences.[19] Therefore, the presence of any individual located in the dominated range is directly attributable to the presence of high optimization frictions (Kleven and Waseem, 2013). Frictions also imply some individuals that respond do not bunch *exactly* at the notch point, thus creating a diffuse excess mass rather than a point mass at W_r^*, as illustrated in Figure 1.3, Panel (a). For instance, because wealth is a stock and not a flow, it is arguably more costly to respond to the notch and more

[17] To simplify the exposition, Figures 1.3 and A.3 assume that the notch is associated with a small change in the *marginal* wealth tax rate above the cutoff, so that intensive responses by those who stay above the notch can be ignored. This implies that pre- and post-notch densities coincide above $W_r^* + \Delta W_r^*$.

[18] Appendix A.3 considers another extension of the baseline model, namely, heterogeneity in elasticities but not in the resource cost of sheltering wealth from taxation.

[19] The width of the dominated range ΔW_r^D is defined such that reported wealth level $W_r^* + \Delta W_r^D$ ensures the same level of net-of-tax wealth $W - T(W_r)$ as the notch point W_r^*, that is, $(1 - \tau - \Delta\tau)(W_r^* + \Delta W_r^D) = (1 - \tau)W_r^*$. Therefore $\Delta W_r^D = \Delta\tau \cdot W_r^*/(1 - \tau - \Delta\tau)$. Appendix A.3 shows that even individuals with $e = 0$ should be bunching in the dominated range. Therefore, any remaining mass in that segment must be the result of prohibitively high resource costs.

difficult to locate precisely below the notch (inter alia, asset prices are not controlled by the taxpayer).[20]

Denote $a(W_r, e)$ the share of individuals at reported wealth level W_r and elasticity e with sufficiently high resource costs that they are unresponsive to the notch. We then have excess bunching

$$B = \int_e \int_{W_r^*}^{W_r^* + \Delta W_{r,e}^*} (1 - a(W_r, e))\, \tilde{h}_0(W_r, e) dW_r de \approx h_0(W_r^*)\,(1 - a^*)\, E\left[\Delta W_{r,e}^*\right] \qquad (1.2)$$

where the approximation assumes a locally constant counterfactual density and, in addition, a locally constant share of individuals with "large" resource costs, $a(W_r, e) = a^*$ for $W_r \in (W_r^*, W_r^* + \Delta W_{r,e}^*)$ and all e. Then, $E\left[\Delta W_{r,e}^*\right]$ is the average "structural" response not affected by frictions while $(1 - a^*)\, E\left[\Delta W_{r,e}^*\right]$ is the average observed response attentuated by resource costs.

We estimate the share of individuals with prohibitively large resource costs a^* from the strictly dominated range $(W_r^*, W_r^D]$: $a^* \equiv \int_{W_r^*}^{W_r^D} h(W_r) dW_r / \int_{W_r^*}^{W_r^D} h_0(W_r) dW_r$. The reported wealth response that would materialize if individuals overcame resource costs is then proportional to $B/(1 - a^*)$. This means that the larger the number of bunching individuals B and the smaller the hole in the dominated region (i.e., the higher the share of non-bunchers a^*), the larger the response to tax notches. Further, note that because the utility gain of bunching at the notch point, for a given elasticity e, is monotonically *decreasing* in $W_r > W_r^*$ and converges to zero at $W_r^* + \Delta W_{r,e}^*$, while a^* is monotonically *increasing* on bunching segment $(W_r^*, W_r^* + \Delta W_{r,e}^*)$. This implies that estimating a^* from the dominated range *understates* average resource costs, *overstates* the share of bunchers 1-a^*, and therefore provides a *lower bound* of the response ΔW_r^* that would materialize in the absence of resource costs. This is what Kleven and Waseem (2013) denote the "bunching-hole method."

If, instead, none of the missing mass can be explained by low elasticities (i.e., elasticities are homogeneous at $e = \bar{e}$) and it is all driven by frictions, then the "structural" response ΔW_r^* corresponds to the reporting response of the marginal buncher H. This ΔW_r^* can be determined as the point of convergence between observed and counterfactual distributions: $W_r^* + \Delta W_r^*$ is estimated as the point where excess mass is exactly equal to missing mass, as we describe below. This represents an *upper bound* on the average "structural" response ΔW_r^*.

Identification

We can uncover the elasticity e given knowledge of the notch parameters, τ and $\Delta\tau$, and the reporting response ΔW_r^*. This elasticity can be non-parametrically identified without relying on a specific utility functional form, or it can be identified assuming some additional

[20] An alternative interpretation is that taxpayers explicitly avoid bunching *exactly* at the notch if they believe doing so increases the likelihood of auditing. Under this interpretation, bunching farther away from the notch is reflective of risk aversion or an explicit strategy to pass undetected.

parametric structure. We present the former approach below, and develop the structural approach in Appendix A.3. Specifically, we apply the reduced-form approach described in Kleven and Waseem (2013) and Kleven's online technical note to relate the wealth reporting response ΔW_r^* to the change in the *implicit* marginal tax rate between W_r^* and $W_r^* + \Delta W_r^*$ created by the notch.[21] Treating ΔW_r^* as if generated by a hypothetical kink $1 - t^*$ between W_r^* and the tangency point of individual H's indifference curve $I(W_r)$ assuming $W_r^I \approx W_r^* + \Delta W_r^*$ (i.e., the interior incentives are small), it can be shown that

$$e_R \equiv \frac{\Delta W_r^*}{W_r^*} \cdot \frac{1-t^*}{\Delta t^*} \approx \left(\frac{\Delta W_r^*}{W_r^*}\right)^2 \cdot \left(\frac{1-\tau}{\Delta \tau}\right) \cdot \frac{1}{2} \tag{1.3}$$

Estimation

Figure 1.3, Panel (b) illustrates the bunching estimation to obtain the reporting response ΔW_r^* and corresponding elasticity. We follow previous bunching studies and slice the data into bins of reported net worth and count the number of taxpayers located in each bin to generate an empirical density $h(W_r)$. The counterfactual distribution $h_0(W_r)$ is obtained from a regression of the following form

$$c^j = \sum_{i=0}^{p} \beta_i \cdot (W_r^j)^i + \sum_{i=W_r^l}^{W_r^u} \gamma_i \cdot 1[W_r^j = i] + \eta^j \tag{1.4}$$

where c^j is the number of individual taxpayers in bin j, W_r^j is the reported net worth level in bin j, and p is the order of the polynomial. The excluded range $[W_r^l, W_r^u]$ corresponds to the area that is affected by the notch point either because of excess or missing mass. The counterfactual distribution is estimated as the predicted values from specification (1.4) omitting the contribution of the dummies in the excluded range, that is, $c^j = \sum_{i=0}^{p} \hat{\beta}_i \cdot (W_r^j)^i$. Excess bunching and missing mass are estimated as the difference between the observed and counterfactual bin counts in the relevant reported net worth ranges, $\hat{B} = \sum_{j \in [W_r^l, W_r^*)} (c^j - \hat{c}^j)$ and $\hat{M} = \sum_{j \in [W_r^*, W_r^u]} (\hat{c}^j - c^j)$.

The lower limit W_r^l is determined both visually and exploiting pre-reform data. The upper limit $W_r^u = W_r^* + \Delta W_r^*$ is estimated by imposing the restriction that the excess bunching equals the missing mass, $\hat{B} = \hat{M}$. This is equivalent to assuming that all responses to the tax notch are on the intensive margin. Starting from a low initial value of the upper bound $W_r^u \approx W_r^*$ and an initial estimate of the counterfactual $\hat{c}^{j0}$, the upper bound is increased in small increments and the counterfactual reestimated every time until $\hat{M}^k = \hat{B}^k$.[22] The estimated upper bound W_r^u is the counterfactual reported wealth of the marginal

[21] See derivations in Kleven's online technical note here.

[22] In the empirical application there is a finite number of bins, so we impose the condition that the absolute difference between $\hat{B}$ and $\hat{M}$ be "close" to zero, i.e. $|\hat{B} - \hat{M}| < 0.03$.

taxpayer that responds to the tax change. Total excess bunching $\hat{b}$ is then $\hat{B}$ relative to the counterfactual.[23]

Finally, as in other bunching papers, extensive responses and multiple notches are two main concerns we would potentially have to deal with.[24] Extensive responses could occur if taxpayers reduced their reported wealth below the income tax filing requirement. However, the filing cutoff is significantly below the wealth tax notches and also depends not only wealth but also on income and card expenditures. Therefore, extensive margin responses would occur only if individuals jointly reduced income, wealth, and card expenditures below the filing cutoffs. This is very unlikely in the short term.[25] Lastly, the presence of multiple notches may be an issue if bunchers jump more than one notch at a time, as $\hat{B}$ and $\hat{M}$ would no longer match. Reassuringly, the panel structure of our data allows us to test how big of a concern this is. We find that only a handful of bunchers jump more than one notch at a time.

Evidence from the 2010 Wealth Tax Reform

We first leverage the variation from the 2010 tax reform to estimate how individual reported wealth responds to wealth taxation. This reform introduced a one-time wealth tax levied on taxpayers reporting to own 1 billion pesos (2010 USD 520,830) or more in (taxable and non-taxable) wealth by December 31, 2010. This exemption threshold was significantly lower than previous years', thus adding more taxpayers into the wealth tax base. Individuals who paid no wealth tax in 2009 now faced two wealth tax notches (see Figure 1.1).

For our wealth bunching estimation, Panels (a) and (b) in Figure 1.4 plot the distribution of individuals around the first and second notches, respectively. The red vertical line marks the notch point in each panel. The gray line is the counterfactual distribution, estimated as a fifth-order polynomial, as specified in (1.4). The estimated parameters are displayed separately. The following patterns emerge from these panels. First, the notches are associated with large and sharp bunching just below the cutoff and missing mass above the cutoff. For the first notch, the excess mass $\hat{b}$ is 4.9 times the counterfactual, meaning there is 5.9 times the expected density in the absence of the notch. The estimated standard error of $\hat{b}$ is 0.17

[23] Standard errors are calculated by a bootstrap of the entire estimation procedure. We draw 1,000 random samples with replacement, and define the standard error of the estimated excess bunching $\hat{b}$, share $\hat{a}^*$, reported wealth response $\Delta\hat{W}_r^*$ (lower and upper bounds), reduced-form elasticity estimate $\hat{e_R}$, and structural elasticity estimate $\hat{e_S}$ (developed in Appendix A.3) as the standard deviation in the distribution of estimates of each variable.

[24] Note that an additional practical problem that is common in the bunching literature is the higher frequency in the reporting of "round numbers." However, as Figures 1.4 and 1.7 demonstrate, our data does not display "round-number" problems often present in other applications.

[25] We test for extensive margin responses to tax notches by checking whether the probability of disappearing from our sample the year of the tax reform is smooth around the tax notch using bins of pre-reform net worth (bin size is current 10 million pesos). Specifically, we use rdrobust command from Cattaneo, Calonico, and Titiunik (2014) and confirm that the probability of dropping from the sample the year the tax reform is introduced is smooth around notches for all years (not reported).

with an implied t-statistic of 28.82, so the null hypothesis of no bunching at the notch is strongly rejected by the data.[26] For the second notch, $\hat{b}$ is 4.3 (the standard error is 0.37). Second, behavioral responses are significantly attenuated by large optimization frictions: 43 percent and 57 percent of individuals in dominated regions do not bunch in response to the first and second notch, respectively. This implies that the degree of bunching absent these frictions $B/(1-a^*)$ is 1.75 and 2.33 times larger than observed bunching, respectively.

We now turn to the estimation of reported wealth responses ΔW_r^*, combining the non-parametric evidence with the conceptual framework from Section 1.3. We bound these responses and elasticities as previously detailed: a lower bound is obtained from the bunching-hole method based on $B/(1-a^*)$, while an upper bound is obtained from the response of the marginal buncher (i.e., the point of convergence between counterfactual and observed distributions). Table 1.1 presents the estimated parameters. For each reform year, the table shows the notch point (column 2), whether this notch is also the wealth tax exemption threshold (column 3), the average tax rate jump (column 4), the size of the dominated range (column 5), the share of taxpayers in dominated ranges that are unresponsive to the tax notch (column 6), the lower and upper bounds on the reporting responses (columns 7 and 8, respectively), and the bounds on the elasticities based on the reduced-form formula (1.3) (columns 9 and 10).

Table 1.1 shows that, if individuals in the dominated region overcame optimization frictions, the reported wealth response would be 11 percent higher than the first notch point (=110/1,000 million pesos) and 5.5 percent higher than the second notch point (=110/2,000 million pesos). The marginal buncher responding to the first notch would have reported 20 percent more wealth under a smooth tax system (=200/1,000 million pesos), reducing wealth tax tax revenue by 2010 USD 20.6 million or 3.5 percent of the total personal wealth tax revenues collected that year. For the second notch, the marginal buncher would have reported 9 percent more wealth (=180/2,000 million pesos), reducing wealth tax revenue by 2010 USD 2 million or 0.3 percent of personal wealth tax revenues collected that year. These reported wealth responses are all highly statistically significant. The difference between the marginal buncher's responses to the first and second notch is not significant.

The implied elasticities of reported net worth with respect to the net-of-tax rate can be obtained by applying the reduced-form approximation, equation (1.3). These elasticities are 0.6 using the bunching-hole method and 2.0 using the convergence method, and both are statistically significantly different from zero at the 1 percent level. This latter elasticity implies that, for the marginal buncher, a 1 percent increase in the net-of-tax rate raises reported wealth by 2 percent. In contrast, because $\Delta\tau$ is smaller at the second notch, the estimated elasticities are smaller (between 0.37 and 1.0), but only precisely estimated using the bunching-hole method.

[26] Figure A.4 shows heterogeneity in bunching across samples with different opportunities to shelter wealth from taxation (capital rentiers, wage-earners, and all other taxpayers). The figure shows that bunching is most pronounced among capital rentiers and least pronounced among wage-earners, who might have less sheltering opportunities than other taxpayers (Jakobsen et al., 2018).

An advantage of our setting is that we observe wealth before and after tax policy changes for those affected and not affected by them. This enables us to test key assumptions for the bunching procedure to identify behavioral responses. Panels (c) and (d) in Figure 1.4 compare the counterfactual densities from cross-sectional data and equation (1.4) (black line) with the counterfactual density using pre-reform data from 2009 when there are no notches or kinks (gray line). The figures confirm that, first, the distribution of reported net worth is smooth in the absence of wealth tax notches. Second, our estimated parameters are robust to using this counterfactual distribution of reported net worth. For the first notch, estimates using the 2009 distribution as counterfactual density are somewhat larger (e.g., $\hat{W}_r^u$ is 1250 versus 1200), although these differences are mostly not statistically significant. Specifically, upper-bound ΔW_r and corresponding elasticities are not statistically different, while lower-bound ΔW_r and corresponding elasticities are different at the 5 percent level. For the second notch, none of the differences are statistically significant.

Interestingly, taxpayers remain bunched below the exemption cutoff several years after the 2010 reform, even though wealth would not be taxed again until 2014. Figure 1.5 plots the distribution of reported wealth in years 2009 through 2012. The figure shows that wealth taxation triggers bunching that persists over time and remains salient even in years when wealth is no longer taxed (2011 and 2012). Further, taxpayers remain piled up below the 1,000 million peso cutoff in *current* prices. This implies reporting less wealth in *real* terms.

Lastly, we compare responses over time leveraging the panel microdata. Figure 1.6 compares taxpayer density across reported wealth bins under a smooth tax schedule in Panel (a) versus a notched tax schedule in Panel (b). Darker bins represent higher relative taxpayer density. Two key insights emerge from the figure. First, in the absence of wealth notches, taxpayers report owning similar amounts of wealth from year to year, although there is year-to-year variation that is not necessarily related to changes in the wealth tax schedule. In contrast, the introduction of wealth taxation reduces year-to-year variation, making the 45 degree line significantly more salient. Second, the emergence of two horizontal darker areas just below the 1,000 million and 2,000 million notch points in 2010 showcases the "bunchers," that is, taxpayers with wealth above the bracket cutoffs before the reform who report just below them immediately after. We will use this variation to characterize the changes in wealth reporting induced by wealth taxation in Section 1.3.

Evidence from Wealth Tax Reforms in 2003, 2006, 2010, and 2014

We now exploit variation introduced by tax reforms in 2003, 2006, 2010, and 2014 to compare bunching responses and elasticities across wealth levels, notch sizes, and notch saliency. The results, displayed in Table 1.1 and Figure 1.7, can be summarized as follows. First, between 35 percent and 74 percent of individuals in the dominated range do not respond to the wealth tax notch. If they overcame their high optimization frictions, excess bunching would be 1.54 to 3.85 times larger. Second, reported wealth responses ΔW_r are large at all notches, almost always precisely estimated, and often not statistically significantly different from each other. The largest responses, in *percentage* terms, come from the bottom notches, that is, the less wealthy taxpayers. These responses are triggered by relatively small tax notches (e.g., 0.05 percent in 2014), which mechanically translates into large elasticities for this group of individuals. This implies that elasticities are generally decreasing in reported wealth. For example, in 2010, the upper-bound reduced-form elasticity estimated from the first notch is statistically significantly different from the equivalent elasticities estimated from the third and fourth notches.

Decreasing bunching elasticities contrast with the conventional wisdom that wealthy taxpayers, who have access to more aggresive avoidance opportunities, are very responsive to taxation. There are several ways to rationalize this result.[27] First, taxpayers may be more elastic to more salient notches; the first notch is the exemption cutoff and arguably the most salient notch. Second, bunching estimates local elasticities driven by the sample of compliers around the tax notch. As the next section shows, individuals in the top brackets have a larger share of their wealth in the form of listed equities and portfolio securities covered by third-party reporting, which are harder to underreport.[28] Third, bunching captures an immediate response to wealth taxation in our context. However, some sophisticated tax sheltering schemes, such as setting up an offshore shell company, have larger fixed costs and thus may not instantaneously respond to wealth tax notches. Further away from the notch at the top, the infra-marginal taxpayers always have incentives to create offshore accounts.

In Appendix A.4, we complement the bunching analysis by examining responses over a longer time horizon, leveraging the panel microdata as well as variation introduced by the four wealth tax reforms. We compare outcomes across time between taxpayers reporting above or below the bracket cutoffs in a difference-in-differences design. We find that most of the overall effect on reported wealth is concentrated in the year of the reform, consistent with a one-time avoidance adjustment driving the main reduced-form effects. Interestingly, the gap between the treated and control groups—which conflates mechanical and behavioral responses—persists several years after a reform *even when wealth is no longer taxed*. For instance, a one-off wealth tax in 2010 generates a persistent gap the following years. This is partly due, as discussed above, to the persistence of bunching triggered by wealth taxation.

Misreporting Wealth Subject to Less Third-Party Reporting

Intuitively, while earnings responses to income taxes potentially conflate real and sheltering responses, bunching in the distribution of reported wealth—a stock—predominantly reflects sheltering. It is difficult for individuals to immediately bunch below notch points using real responses (e.g., changes in investment) because wealth partly depends on asset prices, which are uncertain and fluctuate throughout the year. In this section, we exploit our panel

[27] Decreasing bunching elasticities have also been documented in other settings studying earnings (Kleven and Waseem, 2013; Saez, 2010) and firm revenue (Bachas and Soto, 2018) responses to tax kinks and notches.

[28] Table A.1 confirms that taxpayers around the last notch have twice the share of assets in stocks than taxpayers in the first notch. Moreover, taxpayers who respond to the last notch by bunching below the cutoff have a different wealth composition than those remaining above it; specifically, they have a smaller share in stocks.

microdata as well as a reform raising in the incentives to bunch to identify the type of wealth taxpayers manipulate to place themselves exactly below the notch point.

Ex ante, it is unclear which type of wealth (e.g., financial, non-financial) should be more responsive to taxation. Financial assets are presumably more liquid and easier to adjust to tax changes; however, they are also generally subject to more third-party reporting and thus harder to misreport. In contrast, non-financial assets such inventories and real estate have values that are harder to "mark to market" for tax authorities (Brulhart et al., 2017). In Colombia, inventories in particular are subject to no third-party reporting, which implies taxpayers may be more likely to manipulate them (Chetty et al., 2011; Chetty, Friedman, and Saez, 2013). In all, the question of what type of wealth individuals use to avoid wealth taxation is an empirical one.

We identify and characterize taxpayers who bunch in response to the introduction of wealth tax notches leveraging the panel structure of our data. In our setting, these bunchers denote the subpopulation of individuals who respond to wealth taxes by bunching below the cutoffs, i.e., the compliers in a potential outcomes framework. Always-takers represent individuals who locate below the notch even in the absence of wealth tax notches, while never-takers represent individuals who do not bunch in spite of wealth tax notches. Tax filer i is located in the bunching region B_{it} if she reports wealth between W_r^l and the tax notch W_r^* in year t. Because being located in the bunching region can occur in the absence of wealth taxes, individuals with $B_{it} = 1$ are a mix of compliers and always-takers, as illustrated in Figure A.5. Tax filers located above W_r^* are the never-takers. We pool individuals filing income tax returns before and after the wealth tax reform (2008 through 2010) and exploit the variation over time in the likelihood of being located in the bunching region due to the introduction of wealth taxation in 2010. We characterize bunchers using the following IV specification:

$$Y_{it} = \alpha_1 + \gamma_1 t + \beta_1 B_{it} + \epsilon_{it} \tag{1.5}$$

where Y_{it} is the amount of debt or asset type (e.g., bank deposits, real estate, inventories) expressed as a share of total assets reported that year, t is a time trend that accounts for changes in wealth composition, B_{it} is an indicator for being located in the bunching region, and ϵ_{it} is the error term. Because the incentives to bunch are exogenously shocked by the wealth tax reform, B_{it} is instrumented with a post reform dummy $Z_{it} = \mathbb{1}(t = 2010)$. Standard errors are clustered at the taxpayer level to account for serial correlation.

A limitation of our setting is that, while debt is reported by all taxpayers, the decomposition of asset types is available only for the subsample of taxpayers keeping records, as explained in Section 1.2. These taxpayers, who represent 17.5 percent of our estimation sample, report six categories of assets: fixed assets (e.g., real estate, land), stocks and contributions, inventories, bank deposits, accounts receivable, and other assets.[29]

[29] Figure A.6 plots the density of record-keeping taxpayers by reported wealth in 2010. Although these taxpayers arguably face stronger resource costs of sheltering due to the adjustment costs imposed by record keeping, the estimated parameters are not significantly different from other taxpayers. However, the excess mass is visibly less diffuse for these individuals, i.e., the bunching segment $[W_r^l, W_r^*)$ is smaller. This suggests

Table 1.2 presents the results separately by each wealth category, focusing on taxfilers around the first notch. Column (1) plots the average outcome of bunchers in the absence of wealth taxes. Most of bunchers' assets are made up of real estate, land, and vehicles (52.7 percent). The rest is composed of, in decreasing order, stocks and contributions (17 percent), inventories (10.9 percent), bank deposits (8.6 percent), accounts receivable (6.6 percent), and other assets (3.8 percent). Liabilities are worth 9.4 percent of total assets. Compared to these bunchers, Columns (3) and (4) suggest always-takers and never-takers are similar in their observed wealth composition.

Column (2) plots the 2SLS coefficient β_1 from specification (1.5) representing the percentage point change in wealth composition induced by wealth taxation. First, bunchers do not seem to reduce their bank deposits—the most liquid type of assets–in response to tax notches. They also do not reduce their fixed assets if, for instance, they were to sell or give away their car or residence to avoid the wealth tax. Instead, bunchers do reduce their reported inventories by 22 percent ($= -0.024/0.109$). Because inventories are not subject to third-party reporting, misreporting them is likely to pass undetected by the tax authority. Further, bunchers inflate their liabilities by 35 percent ($= 0.033/0.094$) in response to the introduction of the tax notch. Intuitively, if the debt finances assets (e.g., mortgage to obtain more housing), an increase in debt would be offset by a similar increase in the value of assets, leaving net worth unchanged. If, instead, the debt does not finance assets, taking on more debt reduces wealth. Such would be the case if, for instance, bunchers provided untruthful accounts of nonexistent liabilities. While our data does not allow decomposing debt to test this mechanism, anecdotal evidence suggests taxpayers fabricated interpersonal debt, that is, debt owed to friends or family members rather than financial institutions. In fact, this wealth tax evasion mechanism well-known to the tax authority partly motivated the voluntary disclosure program analyzed in Section 1.5. These pieces of evidence suggest bunchers avoid wealth taxation by manipulating the reported values of assets and liabilities subject to less third-party reporting to artificially place themselves just below the exemption cutoff. Specifically, when enforcement capacity is weak, wealth reported to the tax authority instantly falls upon taxation—even if tax rates are relatively small.

Interpretating Wealth Elasticities and Short-Term Revenue Loss

To put our estimated elasticities into perspective, we first compare them to those of previous studies. As an illustration, our bunching estimation using the 2010 eligibility cutoff shows the elasticity of reported wealth with respect to the net-of-tax rate is 2. This estimate is an order of magnitude larger than estimates exploiting kinks in the wealth tax schedule (Seim, 2017; Jakobsen et al., 2018), which is consistent with tax notches being more salient than kinks. In fact, the estimates from studies exploiting kinks are similar to those we derive for wealthier taxpayers. In contrast, our elasticity of 2 is closer to that found using a difference-

that, conditional on bunching, business owners keeping records might be better able to target the cutoff than taxpayers not keeping records, possibly because they have somewhat more control over their reported wealth.

in-differences approach in Catalonia (Durán-Cabé et al., 2017).[30] Lastly, our elasticities are an order of magnitude smaller than the elasticity of 34.7 estimated using cross-canton comparisons in Switzerland by Brulhart et al. (2017).[31]

Importantly, a positive tax on wealth affects the return to wealth. For instance, a 1 percent tax on wealth corresponds to a 20 percent tax on capital income, assuming a rate of return on wealth of 5 percent. Wealth taxes can therefore be interpreted as isomorphic to taxes on capital income. In particular, the elasticity of wealth is equivalent to the elasticity of capital income multiplied by a factor of $(1-\tau_W)/r(1-\tau_K)$:

$$\epsilon_K = \epsilon_W \cdot \frac{r(1-\tau_K)}{1-\tau_W} \tag{1.6}$$

where ϵ_K is the elasticity of capital income, ϵ_W the elasticity of wealth, r the rate of return on wealth, τ_K the capital income tax rate, and τ_W the wealth tax rate (see derivation in Appendix A.3). In our example, $\tau_W = 0.01, r = 0.05, \tau_K = 0.2$, and $\epsilon_W = 2$ from our main analysis imply $\epsilon_K = 0.08$. Thus, our relatively large elasticities with respect to $1-\tau_W$ translate into fairly small conventional elasticities with respect to $1-\tau_K$.[32]

In addition, our estimated wealth elasticities enable us to recover the revenue loss due to short-term behavioral responses to wealth taxes, as well as the short-run revenue-maximizing wealth tax rate. We apply the framework laid out in Saez, Slemrod, and Giertz (2012) to our tax notches in order to evaluate the efficiency cost of a wealth tax reform. Starting from an *average* tax rate τ for taxpayers with reported wealth above W_r^*, increasing the top tax rate by $\Delta\tau$ mechanically raises revenue by $dM = NW_r\Delta\tau$ where N is the number of taxpayers in that bracket and W_r the average wealth in that bracket.

Taxpayers react to wealth tax notches by bunching below the cutoffs. These behavioral responses reduce reported wealth by $\Delta W_r = W_r \cdot \sqrt{2 \cdot e \cdot \frac{\Delta\tau}{(1-\tau)}}$, where e is the elasticity of reported wealth with respect to the net-of-tax rate. Using equation (1.3) for ΔW_r, taxpayers' behavioral responses reduce tax revenue by $dB = -N\Delta\tau\Delta W_r = -N\Delta\tau W_r\sqrt{2 \cdot e \cdot \frac{\Delta\tau}{(1-\tau)}}$.

[30] Elasticities estimated using difference-in-differences designs are often an order of magnitude larger than bunching estimates, as discussed in Kleven and Schultz (2014); He, Peng, and Wang (2018); Aronsson, Jenderny, and Lanot (2017).

[31] Brulhart et al. (2017) report a semi-elasticity of taxable wealth with respect to the wealth tax rate of 34.5. With the average wealth tax rate being 0.476 percent, this implies an elasticity of taxable wealth with respect to the net-of-tax rate of of 34.7 ($= 34.5/(1-0.00476)$).

[32] Brulhart et al. (2017) provide a verbal description linking wealth elasticities and capital income elasticities.

Hence, the total effect on tax revenue is

$$\begin{aligned} dT &= dM + dB \\ &= N\Delta\tau W_r - N\Delta\tau W_r\sqrt{2 \cdot e \cdot \frac{\Delta\tau}{(1-\tau)}} \\ &= NW_r\Delta\tau\left(1 - \sqrt{2 \cdot e \cdot \frac{\Delta\tau}{(1-\tau)}}\right) \end{aligned}$$

For instance, in FY 2010, $\tau = 0, \Delta\tau = 0.01$, and the upper bound of e is 2, as described above. This means that, based on our elasticity estimate, at most 20 percent of the projected tax revenue increase is lost through behavioral responses. Further, we estimate that the largest loss of tax revenues due to behavioral responses occurred in FY 2014 (21 percent, using the highest elasticity estimated from the first notch). We thus conclude that at most one-fifth of the projected tax revenue increase is lost due to short-term responses to wealth taxation.

Finally, what do our results imply for the revenue-maximizing (marginal) wealth tax rate? Again using our highest elasticity estimate from Table 1.1, the highest revenue-maximizing marginal wealth tax rate is $1/(1 + a \cdot e) = 1/(1 + 1.676 \times 4.41) = 11.9$ percent, where we have estimated $a = \frac{E[W_r|W_r>W_r^*]}{E[W_r|W_r>W_r^*]-W_r^*} = 1.676$ among individuals with reported wealth above notch point W_r^* of 1,000 million pesos in FY 2014.[33] This suggests Colombia is below its short-run revenue-maximizing wealth tax rate. Note, however, that this revenue-maximizing wealth tax rate regards short-run responses only. In the short run, past savings decisions are fixed and estimated behavioral responses capture mostly sheltering responses to wealth taxes. However, in the long run, individuals overcome frictions and adjust their saving, investment, migration, and bequest decisions. As a result, real responses may generate larger distortions in the long run. With the observed elasticity increasing in time, the long-run revenue-maximizing wealth tax rate is likely well below 11.9 percent.

1.4 Offshoring Assets to Tax Havens: Evidence from the Panama Papers

Previous evidence in the literature suggests wealthy individuals often have access to sophisticated tax sheltering strategies and often hide assets offshore in tax havens. In this section, we focus on this particular tax sheltering mechanism. We identify taxpayers having incorporated offshore entities using data from the Panama Papers, and track their reported wealth across time.

We exploit the leaked Panama Papers microdata, which includes information the ultimate beneficial owners of offshore entities incorporated with the help of Mossack Fonseca.

[33] Note that this formula applies a tax system with marginal tax rates. If the government were to keep its notched tax schedule, the formula would need to be adjusted because the bunching response from notches has a first-order, strictly negative effect on tax revenue (Lockwood, 2018).

While these entities were incorporated in more than twenty different jurisdictions, Panama represents a particularly desirable destination for wealthy Colombians to hold assets offshore, offering geographic proximity, the convenience of Spanish as the official language, and political stability. Unsurprisingly, Panama constitutes Colombians' most preferred destination for holding reported foreign assets, after the United States (see Figure A.7). Panama is also Colombia's most relevant tax haven. Even before the leak, Colombians were almost fifteen times as likely to report owning assets in Panama than in the Virgin Islands or Switzerland. Thus, the information offered by the Panama Papers is particularly relevant to study offshoring by Colombians.[34]

The Dynamics of Offshore Entity Incorporation and Wealth Taxation

Figure 1.8, Panel (a), plots the evolution of the top wealth tax rate in Colombia between 1995 and 2015 (dashed blue line) and the number of offshore entities incorporated by Colombians through Mossack Fonseca each year (solid black line). Offshore entity incorporation rose tenfold between the reintroduction of wealth taxes in 2002 and 2015. Further, the two series closely trail each other: more offshore entities are incorporated when the top statutory wealth tax rate increases. This rise is particularly salient in 2010, when the top wealth tax rate skyrocketed to 6 percent, and in 2014, following the reintroduction of progressive wealth taxes that year.[35] The flow of new incorporations remained particularly high since 2010, reaching a peak of 270 new incorporations in 2015 and a cumulative total of 1,784 entities since 1973.

Panel (b) compares the flow of offshore entities incorporated by Americans, Brazilians, Venezuelans, and Mexicans. As far as we can observe in the Panama Papers and the Offshore Leaks, Colombians incorporate more offshore entities every year, even relative to more populous countries like Brazil, Mexico, or the United States.[36] Figure A.8 extends this comparison with more countries and confirms that Colombia's stark increase in the relative flow of offshore entities cannot not replicated elsewhere.[37]

Who Are The Shareholders Of Offshore Entities Created By Mossack Fonseca?

To identify the shareholders of offshore entities created by Mossack Fonseca, we merge the microdata from the Panama Papers to our individual income tax returns using personal

[34] Trusts are less prevalently used in Latin America than in Anglo-Saxon countries. For those reluctant to loose control over their assets (for instance, due to mistrust in others or in the rule of law), foundations represent an attractive alternative. Catering to the Latin American elite, Panama has made the foundation form one of its specialties within the offshore industry (Harrington, 2016).

[35] An alternative (but not necessarily mutually exclusive) explanation for the more frequent incorporation of offshore entities by Colombians through Mossack Fonseca could be supply-driven if, for instance, Mossack Fonseca decided to more aggressively lure its Colombians clients into creating offshore entities in recent years.

[36] Given Colombians may use other asset management firms not included in these leaks or located in other offshore financial centers, Figure 1.8 represents a lower bound on total offshoring.

[37] The only other Latin American country with a similarly large flow of offshore entity incorporation is Uruguay, included in the European Union's 2017 "Grey List" due to concerning tax practices.

names, as detailed in Section 1.2. Table 1.3 presents descriptive statistics separately for individuals filing at any moment between 1993 and 2016 (Column 1) who do *not* appear named in the Panama Papers leak, and those that do appear named in the leak (Column 2). We observe more than 3.3 million taxpayers, that is, around 9.54 percent of the total number of tax units.[38] On average, individuals not appearing in the Panama Papers filed income taxes in Colombia for only 6.2 years during our period of study, relative to 15.8 years for the 1,208 taxpayers identified in the Panama Papers. Tax filers in the Panama Papers are more likely to be male and born after 1985 (these differences are statistically significant). They are also twice as likely to be "capital rentiers," meaning most of their income comes from owning capital. Tax filers in the Panama Papers are also more likely to be wage-earners, and less likely to report another activity codes (e.g., business-owners).

The most striking differences between taxpayers named and not named in the Panama Papers leak are in their average wealth and capital gains. Even if underreporting to the tax authority, individuals named in the leak are more than seven times as wealthy as others. More than two-thirds of them are among the wealthiest 1 percent of adults in Colombia, and over one-fourth are among the top 0.1 percent. To further illustrate this point, Figure 1.9, Panel (a), plots the likelihood of appearing named in the Panama Papers by the average wealth reported to the tax authority across all filing years. Appearing in the Panama Papers rises steeply with wealth. This probability is 0.02 percent for P95–P99 individuals with average reported net worth between 0.14 and 0.41 billion pesos (USD 47,261–136,869), 0.1 percent for the next wealthiest group, and 1.7 percent for individuals in the top 0.01 percent of net worth (these differences are statistically significant). That is, one in sixty individuals in the wealthiest 0.01 percent are identified in the Panama Papers, and this likelihood is 24 times higher than for the top 5 percent as a whole. Therefore, Colombians named in the Panama Papers are among the wealthiest individuals in the country.

Individuals Hid Assets in These Offshore Entities

Admittedly, many of the activities carried out through offshore entities are perfectly legal, and there are security-related and other legitimate reasons for a Colombian to store wealth offshore. Indeed, Colombia allows owning foreign assets, but requires taxpayers report them to the tax authority (i.e., a residence-based tax system). To test whether individuals change their reporting behavior once they create an offshore entity, we leverage information on the date of entity incorporation and the full time series of assets reported by individuals to the tax authority.

Specifically, we use an event study design that compares reported assets across individuals before and after incorporating an offshore entity. Let W_{it} be the value of assets reported to the Colombian tax authority by taxpayer i in FY t, and define the year in which individual i incorporated an offshore entity through Mossack Fonseca as e_i. Define $D_{it}^k = \mathbb{1}(t = e_i + k)$

[38] As noted in Section 1.2, only a fraction of tax units are required to file income taxes in Colombia due to high exemption thresholds, outdated cadasters, and a pervasive informal sector.

as an indicator variable that equals 1 if event e_i took place k years ago and 0 otherwise.[39] In our main specification, $k \in [-5, 1]$, to avoid dropping the significant number of offshore entities incorporated in the later years of our sample, as depicted in Figure 1.8.[40] The event study specification then takes the following form:

$$W_{it} = \alpha_i + \gamma_t + \sum_{k=-5}^{1} \beta_k D_{it}^k + u_{it} \tag{1.7}$$

where α_i and γ_t are individual and year fixed effects, respectively, and u_{it} is the error term. We cluster standard errors at the individual level because individual-specific errors are likely to be serially correlated. Any decrease in reported assets after entity incorporation ($\beta_0, \beta_1 < 0$) is consistent with these entities being used to hide assets from the tax authority.

Figure 1.10, Panel (a), plots the β_k in equation (1.7) over time and provides estimates of the mean assets in "event time" after having taken out individual- and year-specific effects. We scale the estimates with respect to mean reported wealth in 2001, the year preceding the introduction of the wealth tax in Colombia. The figure shows there are no pre-event trends: reported assets are not affected by offshore entity incorporation before it actually occurs, which lends some support against the strict exogeneity assumption of the timing of the event.[41] Moreover, the figure suggests that the value of total assets reported to the tax authority drops by 10.9 percent the year the taxpayer incorporates an offshore entity (the t-statistic is -2.38).[42] A joint F test on post-event years 0 and 1 strongly rejects the null hypothesis that the coefficients are equal to zero ($p = 0.0088$).

To explore heterogeneity, Panel (b) splits the sample by above- and below-median assets reported in event time $k = -5$. The figure shows that the immediate drop depicted in Panel (a) is driven by the wealthiest 50 percent of individuals in the Panama Papers. For these individuals, reported assets fall by 13.7 percent the year of offshore entity incorporation (the t-statistic is -2.02).[43] In contrast, for the bottom 50 percent, reported assets are *increasing* leading up to the event, with the trend breaking upon entity incorporation and the event having no statistically significant impact on reported assets.[44]

[39] The handful of tax filers incorporating offshore entities before 2000 are given an entity incorporation date of 2000 to preserve their anonimity.

[40] As a robustness check, Figure A.9 plots β_k coefficients for different event time windows.

[41] A joint F test on pre-event years -5 to -2 cannot reject the null hypothesis that the coefficients are equal to zero ($p = 0.8389$.)

[42] We test whether opening an offshore is associated with a drop in either reported income or income taxes owed by substituting W_{it} for these variables in specification (1.7). Large standard errors do not allow us to reject the null hypothesis of no statistically significant differences across event time in either case (not reported).

[43] A joint F test on post-event years 0 and 1 for above-median taxpayers strongly rejects the null hypothesis that the coefficients are equal to zero ($p = 0.0327$).

[44] A joint F test on pre-event years 2 through 5 cannot reject the null hypothesis that the coefficients are jointly equal to zero ($p = 0.3852$), and neither can a joint F test on post-event years 0 and 1 ($p = 0.2542$).

1.5 How Responsive is Tax Evasion to Better Enforcement?

In this section, we examine how better enforcement helps recover tax on hidden offshore wealth. Announced in December 2014, Colombia's voluntary wealth disclosure program allowed tax evaders to come clean with the tax authority.[45] Tax filers could voluntarily disclose any hidden wealth (e.g., underreported or unreported foreign or domestic assets) and/or inexistent debt in the wealth tax return filed in 2015, 2016, or 2017.[46] Unlike voluntary disclosure programs put in place across OECD countries, Colombia allowed disclosers to waive unpaid income and wealth taxes from past years. Disclosers paid a one-time explicit penalty of 10, 11.5 or 13 percent of the value of the disclosed assets and liabilities if disclosed in 2015, 2016, or 2017, respectively.[47] Disclosers also paid wealth taxes upon disclosure, and could face higher presumptive income taxes the following year. If caught misreporting after 2018, individuals would face monetary penalties worth 200 percent of owed taxes.[48]

How Tax Evasion Varies With Wealth

The program encouraged 11,927 individuals to disclose 15.76 trillion pesos (USD 5.28 billion) in hidden assets and inexistent liabilities between 2015 and 2017. This is equivalent to 1.73 percent of GDP being disclosed under the scheme. Specifically, 11,050 individuals disclosed assets worth 14.76 trillion pesos (USD 4.95 billion), 87 percent of which had been located abroad (1.4 percent of GDP). Liabilities worth 999.86 billion pesos (USD 335.08 million) were declared inexistent by 1,380 individuals; 97.6 percent of which was (mis-)reported to be in Colombia. The government collected 1.93 trillion pesos (USD 647.23 million) in penalty revenues, equivalent to 0.21 percent of GDP.

Table 1.4 presents descriptive statistics for disclosers and non-disclosers who filed an income tax return before the scheme (FY 2013) and filed a wealth tax return in either 2015,

[45] The 2015–2017 wealth disclosure program represents Colombia's first comprehensive effort to encourage evaders to regulare their tax affairs by disclosing hidden assets and fake debts (Article 35, Law 1739/2014). In 2003, taxpayers voluntarily disclosing hidden assets and/or fake debt were subject to harsher tax treatment (Article 6, Law 863/2003, ruled constitutional by *Sentencia C-910/04*). Subsequent attempts to encourage such disclosures in 2012 and 2013 were blocked by the Constitutional Court, arguing evaders were being given overly generous tax benefits (Article 163, Law 1607/2012, ruled unconstitutional by *Sentencia C-833/13*). Because disclosers were charged a penalty, the 2015–2017 voluntary disclosure program was ruled constitutional on August 26, 2015 (*Sentencia C-551/15*).

[46] Disclosers reported their hidden assets and inexistent liabilities in separate boxes in the wealth tax return (form 440, see Figure F.12), which we use to identify program participants.

[47] See Figure A.10 for a timeline of the events taking place around the voluntary disclosure program.

[48] The government's communication strategy explicitly did not encourage repatriation of capital invested abroad, by expressly stating that the aim of the program was *not* to have disclosers repatriate their offshore assets but rather to have them declare such assets to the tax authority for income and wealth tax purposes.

2016, or 2017.[49] This subsample includes 55,098 individuals who did not disclose under the scheme (Column 1), and 11,210 who did (Column 2). Columns (3)–(5) presents the summary statistics separately for disclosers by year of first disclosure. Despite facing higher penalties, most disclosers revealed their wealth in 2017, i.e., immediately before the disclosure window expired. Otherwise, the table suggests that while disclosers and non-disclosers are similar in demographic characteristics, disclosers are almost 20 percent more likely to self-report their economic activity code as either capital rentiers or wage-earners than any other category. Crucially, disclosers are eight times more likely to appear in the Panama Papers than non-disclosers. Indeed, 37.5 percent (453 of 1,208) of taxpayers identified in the Panama Papers admitted past noncompliance and participated in the voluntary disclosure program, a point we return to in Section 1.5.[50]

Further, Table 1.4 suggests that, even before disclosing hidden wealth, disclosers are wealthier than non-disclosers: the latter have on average 2.84 billion pesos (2017 USD 950,909) in assets, while the former have 3.29 billion (2017 USD 1,102,095) in assets, i.e., 16 percent more wealth. The same is true for net worth. To further illustrate this point, Panel (b) in Figure 1.9 plots scheme partipation by net worth (FY 2013 net worth including disclosures). The probability of disclosing hidden wealth under the scheme rises with wealth: individuals in the top 0.01 percent are 55 times more likely to disclose under the scheme than the top 5 percent. In all, two-fifths of individuals in the wealthiest 0.01 percent admitted noncompliance and disclosed under the same (gray dashed line).[51] Once again, individuals identified in the Panama Papers are significantly more likely to disclose wealth under the scheme across all wealth levels, with 71.4 percent of those in the wealthiest group disclosing under the scheme (black solid line).

The magnitudes of these disclosures are also rising in wealth, as depicted in Figure 1.11. Panel (a) shows that disclosures represent 0.01 percent of wealth for individuals in P95–P99 and 15.3 percent of wealth for the top 0.01 percent. In other words, 15.3 percent of wealth of the top 0.01 percent had been concealed from the tax authority, with foreign assets representing the lion's share of concealed wealth. Panel (b) reproduces this estimate for disclosers only and shows that, on average, disclosures represent around 30 percent of

[49] 93.6 percent of program participants filed personal income taxes immediately before the scheme (i.e., FY 2013).

[50] Not all taxpayers identified in the Panama Papers participated in the voluntary disclosure program. First, being a client of Mossack Fonseca does not imply tax evasion. In Colombia and most other countries, it is legal to own offshore accounts, as long as they are duly declared on tax returns. Thus, tax-compliant client may have already been reporting their offshore entity to the Colombian tax authority. Second, the Panama Papers included Colombians having incorporated their offshore entity as far back as 1973; thus, some taxpayers could have deactivated their offshore entity by the time the disclosure scheme was introduced (and allegedly no longer hold assets offshore). Finally, risk-loving taxpayers may have chosen not to participate in the disclosure scheme and continue evading.

[51] Figure A.11 plots scheme partipation by pre- and post-disclosure net worth (that is, net worth in FY 2013 excluding and including disclosures). Ranking taxpayers by pre-disclosure wealth, only 0.05 percent of adults in P95–P99 disclosed under the scheme, while 10.3 percent of P99.9–P99.95 did so—and this difference is statistically significant. Disclosure probability continues to rise by wealth group, with one in four of the wealthiest 0.01 percent admitting prior noncompliance and disclosing under the scheme.

wealth for all but the top group, for whom disclosures amount to 37.5 percent of wealth. For this group, foreign assets represent 85 percent of disclosures and 31.9 percent of their wealth, reflecting the pervasiveness of offshore tax evasion at the very top of the distribution. These figures thus confirm that hiding wealth offshore had been a key mechanism to evade taxes among the Colombian elite.[52]

Improvements in Wealth and Income Tax Compliance

We have shown that taxpayers participating in the scheme disclosed foreign and domestic assets. Of direct tax policy interest is the effect these disclosures had on income reported and subjected to tax. We study wealth and capital income reporting behavior by linking the disclosure information with individuals' income tax returns, and comparing outcomes between disclosers and non-disclosers across time. To maximize the number of post-disclosure years for which we can observe reporting behavior, we focus on individuals who first disclosed in 2015.[53] Specifically, we compare outcomes using a balanced sample of 44,958 taxpayers (1,777 of whom first disclosed in 2015 and 43,181 of whom never disclosed) using the following difference-in-differences regression specification:

$$\log(y_{it}) = \alpha_i + \sum_{k=-4}^{3} \beta_k \cdot D_{it}^k + \nu_{it} \tag{1.8}$$

where $\log(y_{it})$ is a log-approximation (the inverse hyperbolic sine transformation) of outcome y reported by individual i in year t, α_i are individual fixed effects, D_{it}^k is an interaction term between year t and the discloser dummy, and ν_{it} is the error term.[54] We cluster standard errors at the individual level because individual-specific errors are likely to be positively serially correlated. The β_k coefficients are our main parameters of interest and identify the

[52]Table A.2 compares the location and type of assets reported by wealth tax filers who reported foreign assets in 2017. The table shows wealth tax filers who disclosed under the scheme in 2015–2017 are 63.6 percent more likely to report owning foreign assets in tax havens (e.g., Barbados, Bermuda, Cayman Islands, Monaco, Panama, Switzerland, the Virgin Islands) than those who did not disclose under the scheme. Disclosers are significantly more likely to report owning foreign financial assets than non-disclosers, with this difference being particularly large for portfolio securities and trusts. In contrast, disclosers are significantly less likely to report owning foreign non-financial assets (real estate and vehicles).

[53]Note that, compared to individuals who first disclosed in 2016 or 2017, our main estimation compares individuals who disclosed hidden wealth to the tax authority in the absence of (1) the Panama Papers leak, and (2) harsher punishment for evaders. That is, it is a selected sample of tax evaders responding to the tax benefits awarded by the disclosure scheme. Table A.4 and Figure A.12 present results equivalent to Table 1.5 and Figure 1.12 but for taxpayers first disclosing hidden wealth in 2016.

[54]To accommodate zeros in the dependent variable, we follow Johannesen et al. (2018) and use the inverse hyperbolic sine transformation, which is preferred when using administrative data and focusing on high-income groups. For positive ranges of y_{it}, the β_k coefficients in specification (1.8) can be interpreted exactly as if we were using a log specification, i.e., as the difference between reported log wealth reported at time t and reported log wealth had disclosure not occurred. As a robustness exercise, Table A.3 presents results using different functional form specifications.

percentage change in reported outcomes of disclosers relative to non-disclosers and the year immediately before the disclosure scheme. For instance, we expect β_k to be positive and statistically significant if foreign and domestic assets that generate taxable income and had not previously been reported are now being reported for tax purposes.

Figure 1.12, Panel (a), plots the event study coefficients and associated 95 percent confidence intervals from specification (1.8) using reported gross wealth (total assets) and net wealth (total assets minus debt) as the outcome variable. In both series, the difference between disclosers and non-disclosers is close to zero and not statistically significant before disclosure, thus validating the parallel trend assumption of our difference-in-differences specification. Moreover, in both series, the outcome jumps sharply in 2015 for disclosers relative to non-disclosers. The larger rise for net worth is consistent with disclosures of fake liabilities under the scheme. Importantly, there is no sign that either outcome decreases after the initial surge, and both remain significantly higher three years after policy rollout. Table 1.5 presents estimates of the difference in outcomes between disclosers and non-disclosers before and after policy rollout, i.e., the difference-in-differences coefficient, collapsing the event time in specification (1.8) to a simple post reform dummy. Columns (1) and (2) present the estimates for gross and net wealth, respectively. Disclosers report around 33.38 percent ($= 100 \times (e^{0.288} - 1)$) more assets and 49.2 percent ($= 100 \times (e^{0.4} - 1)$) more wealth relative to non-disclosers after the scheme is introduced.

As discussed above, if individuals are disclosing hidden assets, they should also be reporting the capital income (e.g., domestic interest income, capital gains) received by asset ownership. This, coupled with presumptive income taxes on net wealth, imply tax compliant disclosers should also pay more income taxes. Importantly, 2015 disclosers were required to truthfully report their income in the tax—not calendar—year of their disclosure. We thus expect wealth disclosures in 2015 to affect reported capital income starting 2016. Panels (b) through (d) from Figure 1.12 plot the event study coefficients for regular and irregular income categories. Panel (b) plots interest income, total gross income, taxable income, and income taxes owed. Panel (c) plots dividends and foreign income, which are reported as separate income categories starting in 2014. Panel (d) plots irregular income, a category that combines long-term realized capital gains, inheritances, and *inter vivos* gifts. Again, the figures suggest there is no difference in the trend of income categories prior to disclosure, followed by a large increase in reported capital income starting 2016.[55]

Table 1.5 presents the associated difference-in-differences coefficients and associated standard errors in Columns (3)–(12). The table shows that disclosers report 174.6 percent ($= 100 \times (e^{1.01} - 1)$) more foreign income and 51.6 percent ($= 100 \times (e^{0.416} - 1)$) more interest income than non-disclosers once the disclosure scheme is introduced.[56] In all, the total amount of income reported increases by 14.1 percent ($= 100 \times (e^{0.132} - 1)$), raising taxable income by the same amount. This translates into 39 percent ($= 100 \times (e^{0.329} - 1)$)

[55] There is no impact on dividend income, presumably because dividends—which are subject to third-party reporting—should have already been fully reported.

[56] Interest income should be expected to increase if disclosers repatriate their offshore assets upon disclosure.

more regular income taxes owed. Reported irregular income net of costs also increases by 38 percent ($= 100 \times (e^{0.322} - 1)$), such that disclosers are paying more irregular income taxes than non-disclosers after policy rollout.[57]

Our results therefore indicate that wealth disclosures were associated with large increases in capital income reporting and income taxes owed. Prior to the disclosure scheme, non-compliant taxpayers underreported capital income by not reporting returns to foreign assets. Upon disclosure, disclosers report all (or at least more of) their true capital income and also pay more income taxes. Our results thus rule out substitution towards legal income tax avoidance; on the contrary, wealth disclosers are disclosing more capital income and paying more taxes on that income relative to non-disclosers once the disclosure scheme is introduced.

As a result of the higher taxes paid by disclosers, the overall progressivity of the tax system improved. Figure 1.13 plots the effective tax rate (total income and wealth taxes as a share of net wealth) by wealth group. First, despite having a progressive income tax schedule, the progressivity of the tax system is low in the absence of wealth taxes (see Alvaredo and Londoño-Vélez, 2014). The effective tax rate hovers just over 1 percent regardless of wealth: on average, the top 0.01 percent paid the same share in taxes as the top 1 percent of the distribution. Second, the reintroduction of progressive wealth taxes raised the effective tax rate for the wealthiest 0.05 percent of individuals, i.e., those above the exemption threshold. Wealth taxes can therefore complement progressive personal income taxes in reinforcing tax progressivity at the top. Third, the penalties associated with program participation that year further raised the effective tax rate from 1.4 percent to 2.6 percent for the wealthiest 0.01 percent. The disclosure scheme thus appears to have been an effective way to reduce evasion and generate more revenue from the wealthiest tax filers.

Tax Evasion Responds to Other Shocks in Enforcement

Tax evasion is likely to not only depend on the tax incentives, studied in the previous section, but also critically, on the information and enforcement environment. In this section, we analyze how shocks in the perceived detection probability and the punishment for noncompliance affected tax compliance at the top of the distribution.

On April 3, 2016—that is, one month before the second deadline to disclose under the scheme—the Panama Papers news story broke and the names of Mossack Fonseca's clients were thrust into the public spotlight.[58] The Colombian tax authority reacted by scrutinizing

[57] These spillover results are robust across different functional form specifications, as shown in Table A.3. With the inverse hyperbolic sine function, however, some of the aforementioned effect may be due to changes in reporting from zero to positive amount. Because this extensive margin of capital income reporting is of interest, Table A.3 substitutes the outcome for a binary variable indicating whether the taxpayer reported any positive value of the dependent variable. We find a 4.4 percentage point increase in the probability of reporting any positive foreign income, a 1.8 percentage point increase in reporting any positive interest income, and a 1.2 percentage point increase in owing any income tax (see Figure A.13). We also detect some extensive margin impacts on reported irregular income.

[58] Colombians mentioned in the Panama Papers scandal involved businessmen, politicians, members of congress, lawyers, and journalists. The leak drew media coverage and public interest in the matter. As an illustration,

Mossack Fonseca and its clients, contacting taxpayers named in the leak and requesting documentation of their offshore activities and transactions.[59] Taxpayers named in the Panama Papers were allowed, and in fact encouraged, to participate in the disclosure program. Three weeks after the leak, the governments of Colombia and Panama announced a tax information exchange agreement between the two countries, a move that the tax haven had resisted for years.[60] The agreement involved on-demand and automatic exchanges of information starting June 2017 and 2018, respectively. These events arguably raised the perceived threat of detection.

Exploiting the exogenous timing of the Panama Papers, we identify the causal effect of the leak and subsequent events on tax compliance among very wealthy individuals. We use a difference-in-differences approach that compares outcomes between wealth tax filers named (treated) and not named (control) in the leak before and after it occurred. To fix ideas, Figure 1.14 plots the probability of first disclosing under the scheme by wealth group before (2015) and after (2016) the leak for taxpayers named and not named in the leak. While the likelihood of disclosing under the scheme was similar prior to the leak, the leak substantially raised disclosures for taxpayers named in the Panama Papers. Interestingly, however, the leak does not appear to have raised disclosures among taxpayers unnamed in the leak. Hence, we do not find empirical support to the claim that the leak itself deters evasion among wealthy individuals.

Equation (1.9) presents the OLS specification used to estimate the causal effects of the Panama Papers leak on tax compliance:

$$y_{it} = \alpha + \gamma \mathbb{1}(\text{In Panama Papers})_i + \lambda \mathbb{1}(\text{After Leak})_t + \beta \cdot \mathbb{1}(\text{DID})_{it} + \mu_{it} \tag{1.9}$$

where y_{it} is the outcome of interest, β the parameter of interest, and μ the error term. Because the leak was unanticipated by both taxpayers and the Colombian tax authority, we can interpret β as identifying the causal response to the Panama Papers leak and subsequent events.

The results from specification (1.9) are plotted in Table 1.6. Column (1) shows the Panama Papers leak raised willingness to disclose hidden wealth by 27.4 percentage points. On a base of 3.3 percent for wealth tax filers not appearing in the Panama Papers before the leak, this translates into a more than 830 percent increase in wealth disclosures induced by

Figure A.14 plots the number of times individuals in Colombia searched for the term "Panama Papers" on Google.

[59]The first formal charges of illicit enrichment, fraud, money laundering, among others, would not be filed by Colombia's district attorney's office against individuals related to the Panama Papers until October 4, 2017. Later that month, a number of individuals would be sent to house arrest. See Figure A.10 for a timeline of events taking place between 2014 and 2017.

[60]This announcement came after lengthy negotiations dating from 2014. On October 7, 2014, the deadline for Panama to sign an information exchange treaty with Colombia expired, leading Colombia to declare Panama a tax haven (Decree 1966/2014). Financial transactions with Panama would thus be taxed at withholding at 33 percent in lieu of 10 percent. A diplomatic crisis between the two countries ensued, leading Colombia to remove Panama from its list of tax havens on October 21, 2014 (Decree 2095/2014).

the Panama Papers. Consistent with a drop in offshore tax evasion in particular, Column (2) shows that disclosures of foreign assets increased more than fifteenfold (= .296/.0192). As a result, Column (3) shows wealth taxes increased by 29.8 percent ($= 100 \times (e^{0.261} - 1)$). Finally, including disclosure penalties raises taxes owed to 134 percent ($= 100 \times (e^{0.85} - 1)$), as shown in Column (4).

Six months after the Panama Papers leak, in December 2016, Colombia criminalized tax evasion for the first time. If convicted, tax evaders could face sentences of up to nine years in prison.[61] Most disclosers admitted to noncompliance after tax evasion was criminalized in 2017 (see Figure A.15), arguably at least in part due to the harsher sanction for noncompliance. However, larger disclosures in 2017 also coincides with the expiration of the disclosure window. Unfortunately, if taxpayers disclose this year for other reasons, such as procrastination, we cannot attribute the rise in disclosures solely to the harsher punishment of tax evasion.

1.6 Implications for the Study of Wealth Inequality

We end with a brief discussion of our findings for the study of wealth inequality. We combine tax records and household survey data to analyze how measured wealth inequality is affected by accounting for unreported offshore wealth.

In Colombia, measuring top wealth shares, such as the fraction of total wealth held by the top 1 percent, faces challenges due to severe data limitations. Unlike in many developed countries, there is no aggregate wealth measure to construct the total amount of wealth in Colombia (i.e., the denominator). We cannot compute aggregate wealth as the total wealth reported in tax records because, as discussed, less than 6 percent of tax units file taxes. This implies that we must refer to household survey data to capture wealth for non-filers, which is a second-best alternative and has limitations of its own (see Saez and Zucman, 2016). Further, even in spite of the improvements in enforcement, offshore wealth may remain underreported in tax records for the purposes of reducing the tax burden. To the extent that wealthier individuals are disproportionately likely to hold foreign assets, our measures of top wealth shares will underestimate inequality if we do not account for unobserved offshore wealth. Appendix A.5 describes in detail these issues and how we deal with each of them. We argue that our measures will likely be biased downward, meaning that, in addition to interpreting our estimates as a first, albeit imperfect attempt at measuring wealth inequality in Colombia, they should also be considered potentially conservative.

We estimate that the top 1 percent today has 40.6 percent of total wealth in Colombia (Figure 1.15, Panel (a)). Given the richest 1 percent has roughly 20 percent of total income (Alvaredo and Londoño-Vélez, 2014), wealth in Colombia is twice as unequally distributed as

[61] Evaders caught underreporting their assets or overreporting their liabilities over 7,250 monthly minimum wages (2017 5.66 billion pesos or USD 1.9 million) could face between 48 and 108 months in prison, and pay a penalty worth 20 percent of the value of the hidden or misreported assets or liabilities (Article 338 of Law 1819/2016).

income. In fact, this baseline estimate places Colombia as one of the most unequal countries for which similar wealth inequality measures exist, second only to the United States, from whom the top 1 percent owns 41.8 percent of total wealth.[62] Interestingly, the similarity between Colombia and the US in terms of income inequality is reproduced again in terms of wealth inequality.

Further up the wealth hierarchy, we find that the top 0.1 percent has 15.85 percent of total wealth (Figure 1.15, Panel (b)). This is less than the equivalent American share of 22 percent (Saez and Zucman, 2016). Even further up, the top 0.01 has 6.01 percent of total wealth in Colombia, which is in fact roughly half of the United States' equivalent share of 11.2 percent. Therefore, while the two countries share similar top 1 percent shares, wealth is less concentrated at the extreme top of the distribution in Colombia than in the United States.[63]

As we have discussed, foreign assets may continue to be underreported in tax records in spite of the recent enforcement initiatives. Indeed, while the 2015–2017 voluntary disclosure program incentivized some taxpayers to disclose (at least part of) their assets hidden in tax havens, it is likely that other individuals continue choosing to hide their fortunes from the tax authority. We would therefore like to put bounds on the total amount of offshore wealth that could potentially remain hidden abroad, and analyze its implications for the measurement of wealth inequality.

We begin from the macro estimate for total offshore wealth by Colombians from Alstadsater, Johannesen, and Zucman (2018a). Using fiduciary deposits data from the Central Bank of Switzerland in 2003–2004 as well as cross-border bank deposits data from offshore financial centers in 2007, Alstadsater et al. (2018a) estimate that total offshore wealth by Colombians is 9.0 percent of GDP. This places Colombia just below the world average of 9.8 percent of GDP kept offshore. Adding up total offshore wealth reported by individuals in the tax return for foreign assets amounts (tax form #160), offshore assets represented 2.8 percent of GDP in FY 2017. That is, less than one-third of the baseline measure of offshore wealth is reported to the tax authorities.[64] This means that 6.2 percent of GDP remains concealed offshore.

Who owns this unreported offshore wealth? Our estimates, using data from the Panama Papers, the 2015–2017 voluntary disclosure scheme, and foreign asset information returns, all show that offshore wealth is extremely concentrated at the top. Specifically, the disclosure scheme shows that 99 percent of disclosed offshore wealth is owned by the wealthiest 1

[62] Applying the capitalization method to income tax microdata, Saez and Zucman (2016) estimate that the top 1 percent had 41.8 percent of total wealth in 2012. Using household survey data, the wealthiest 1 percent of households are estimated to own 42.5 percent of total wealth (Balestra and Tonkin, 2018). Further, comparing our estimates with measures of household wealth inequality from the OECD Wealth Distribution Database, Colombia appears to be twice as unequally distributed as France, Luxembourg, and Norway.

[63] Furthermore, we estimate that the wealth-to-income ratio $\beta = W/Y$ is 213 percent in Colombia. This is half the equivalent estimate for the United States, where household β is 430 percent (Saez and Zucman, 2016), but similar to Mexico, where β is 237 percent. In fact, Colombia's wealth-to-income ratio today is similar to that in Germany, Canada, and Italy in the 1970s.

[64] Half of this amount (1.4 percent of GDP) was disclosed thanks to the voluntary disclosure scheme.

percent, with the wealthiest 0.01 percent owning 58 percent of disclosed offshore wealth (see Figure E.8). We assume that the distribution of *unreported* offshore wealth is similar to the distribution of offshore wealth disclosures made during the 2015–2017 disclosure program by each net wealth group, and recompute top wealth shares accordingly.

Note, however, that the estimate from Alstadsater et al. (2018a) is based mostly on data from 2007, that is, a period preceding hefty wealth taxation in Colombia. As Figure 1.8 illustrated, some individuals may have responded to the high wealth tax rates by obscuring their wealth offshore. This implies that basing our estimate from Alstadsater et al. (2018a) represents a lower bound on unreported offshore wealth. To construct an upper bound, we assume that the increase in unreported offshore wealth is one-half the rise in the stock of offshore entities created by Mossack Fonseca between 2007 and 2015. This inflates our measure of unreported offshore wealth to 15 percent of GDP, placing offshore wealth owned by Colombians—both reported and unreported, expressed as share of GDP—above the equivalent shares owned by Americans, Frenchmen, and Germans.

Figure 1.15 shows that conservatively accounting for unreported offshore wealth increases our estimates for top 1 percent share by three percentage points, from 40.64 to 43.17 percent.[65] The upper bound raises this share to 46.4 percent, as reported in the last bar of Panel (a). The resulting increase in measured top 0.1 percent shares is even more dramatic, from 15.85 percent to 19.06–23.17 percent, as reported in Panel (b). Therefore, accounting for unreported offshore wealth places Colombia's top share estimates closer to those of the United States at the very top.

1.7 Conclusion

Progressive wealth taxation has received renewed interest as a tool to raise revenue and address wealth inequality. However, the feasibility of levying this tax depends on how wealthy individuals respond to it, an empirical question few papers have examined due to the major challenges regarding measurement and identification. While a handful of researchers have recently estimated behavioral responses to wealth taxes in Europe, it is unclear what policy implications, if any, their findings have for developing countries, where enforcement capacity is severely limited and few individuals own a very large fraction of total wealth. Further, offshoring assets to tax havens represents might potentially be a significant threat to wealth taxation at the very top of the distribution, yet there is no evidence on the extent to which offshore evasion responds to wealth taxation and enforcement. Thus, understanding how wealthy individuals response to wealth taxes and wealth tax enforcement is key to assessing whether wealth taxes can work in these settings.

We contribute to this discussion by providing quasi-experimental evidence of behavioral responses to wealth taxes and enforcement policy in a developing country. We exploit ex-

[65] Intuitively, given that the wealth-to-income ratio $\beta = W/Y$ for Colombia is 213 percent, and that unreported wealth is 6.2 percent of GDP and is owned solely by the top 1 percent, then the top 1 percent share of total wealth corrected for hidden wealth is almost 3 percentage points higher.

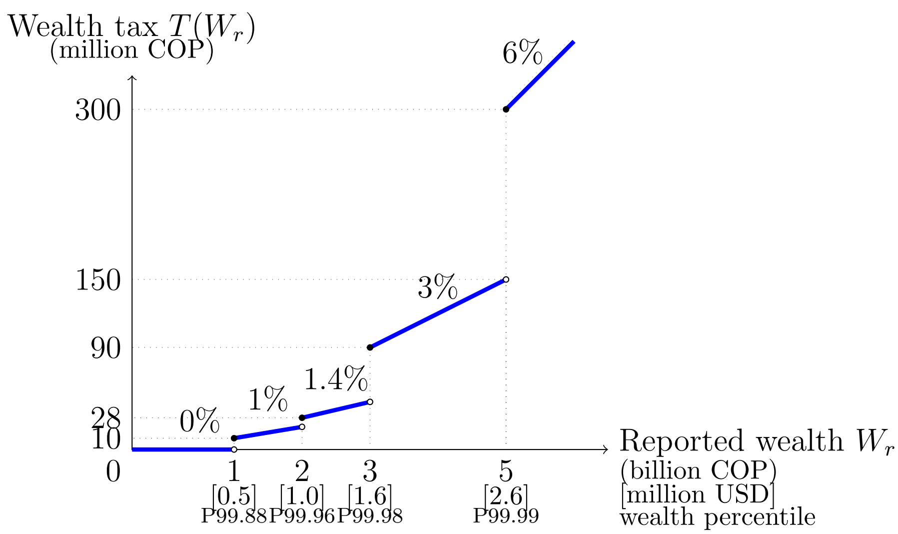
tensive administrative tax microdata on the assets and debts of wealthy individuals linked with microdata from the Panama Papers leak to shed light on offshore sheltering. We find clear evidence that individuals respond to both wealth taxes as well as better enforcement. In spite of significant sheltering responses to wealth taxation, we find that enforcement policy can enhance tax compliance among wealthy individuals and thus improve wealth tax collection.

Our results point to the critical importance of investing in tax enforcement capacity. First, requiring taxpayers report their assets separately to the tax authority facilitates enforcement. Second, enforcement capacity may also be improved thanks to a wider coverage of third-party reporting, if coupled with systematic cross-validation of reported information and increased scrutiny of high net worth taxpayers. Third, improving compliance at the very top requires cracking down on offshore tax evasion. Policies promoting financial transparency and foreign asset reporting, as well as voluntary disclosure schemes, may aid the tax authority in collecting new information about offshore assets and income and generate more revenue from wealthy taxpayers. Yet, for such programs to be effective in improving compliance in the shorter and longer term, stricter enforcement needs tough noncompliance sanctions and a credible threat of detection, for example, by exploiting an automatic exchange of tax information and whistleblower data. Further, getting tax havens to cooperate and provide information about foreign owned financial assets in their jurisdictions is key—hence the paramount importance of understanding the effectiveness of recent initiatives like the United States's FATCA. Through stronger enforcement, wealth taxes can complement progressive income taxes to reinforce progressivity at the top.

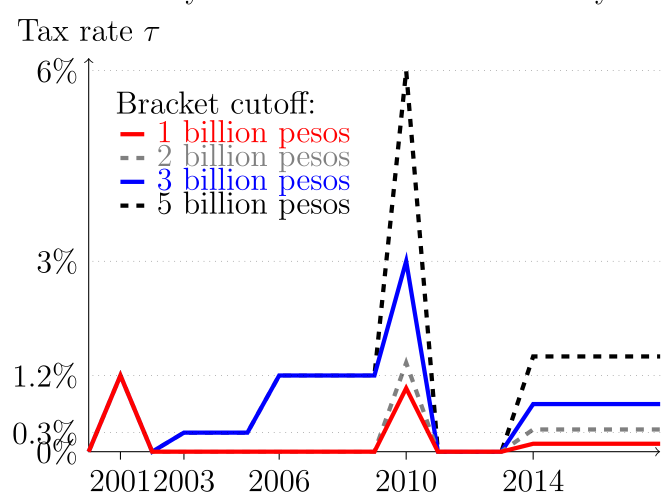
Figures and Tables

Figure 1.1: The Personal Wealth Tax Schedule in Colombia

(a) Wealth Tax Liability as a Function of Reported Net Wealth (FY 2010)

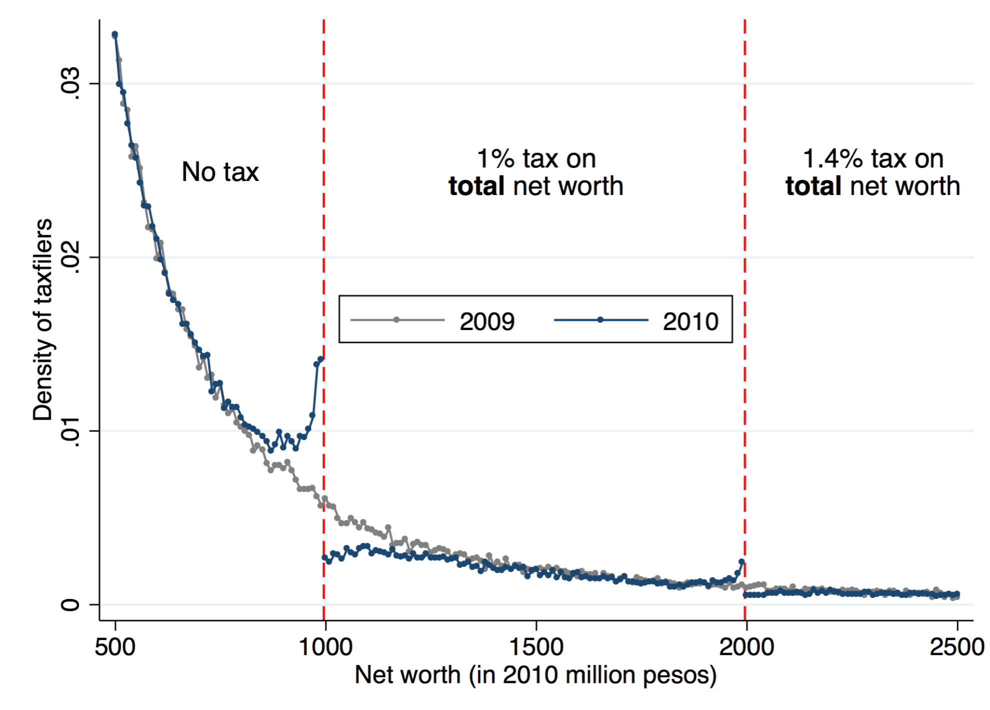


(b) Evolution of Statutory Annual Wealth Tax Rates by Bracket Cutoff



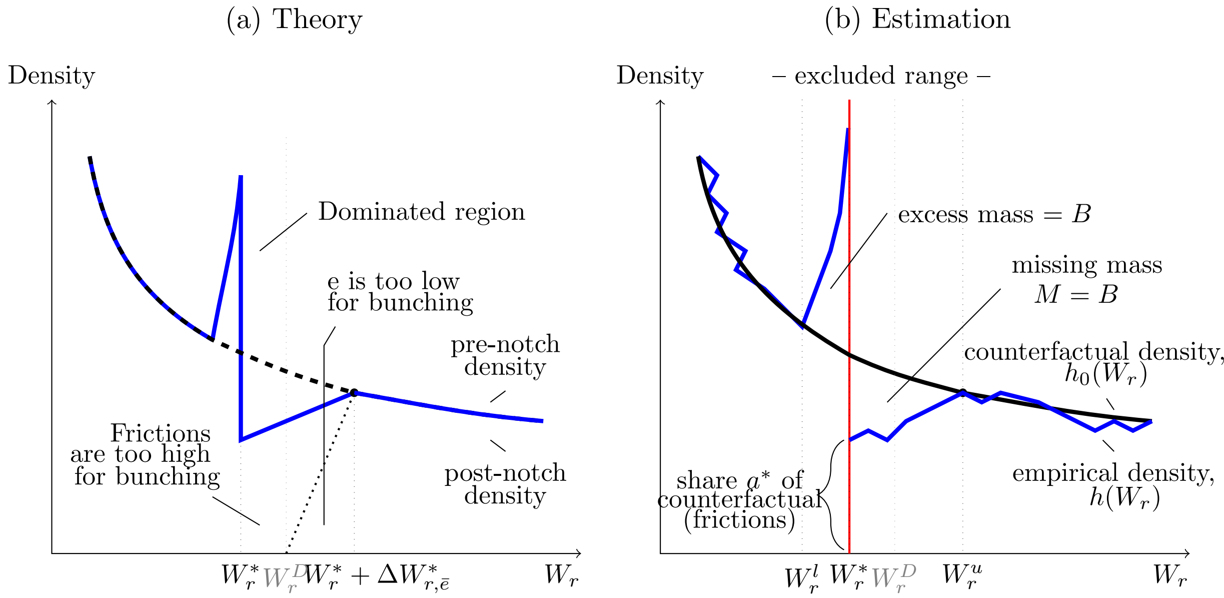
Notes: These figures depict the personal wealth tax schedule for Colombia. Panel (a) plots wealth tax liability by reported wealth W_r in FY 2010. Each bracket of W_r is associated with a fixed average tax rate on taxable net wealth. As a result, wealth tax liability $T(W_r)$ jumps discretely at the notch points. That year, the wealth tax brackets affected the top 0.12%, top 0.04%, top 0.02%, and top 0.01%, respectively. Panel (b) plots the statutory wealth tax rate FY 2000–2018. Wealth tax eligibility is determined using (taxable and non-taxable) net worth in all years but 2001, when it is determined using gross wealth. For 2007–2009, eligibility is established in 2006. In 2015–2018, eligibility is established in 2014. Tax brackets are expressed in current values for all years except 2004 and 2005 (2003 pesos). The tax schedule refers to average tax rates for all brackets in FY 2001–2010. In FY 2014–2018, only the first bracket is an average tax rate; the rest are marginal rates.

Figure 1.2: Distribution of Reported Net Worth in 2009 (Before Reform) and 2010 (After Reform)



Notes: This figure overlays the distribution of tax filers by reported wealth before and after a wealth tax reform. The reform introduced the wealth tax for all taxpayers reporting 1 billion pesos (USD 520,830) or more in wealth. Two notches were introduced at 1 and 2 billion pesos (red vertical lines), generating jumps in the wealth tax liability at these thresholds. While the distribution of tax filers is smooth in the absence of these tax notches (2009), there is an immediate emergence of excess and missing masses just below and above them, respectively (2010). This observed bunching of taxpayers below the notch points is a direct behavioral response to wealth taxation. Bin width is 2010 10,000,000 pesos (2010 USD 5,208.30 in 12/31/2010). *Source:* Authors' calculations using administrative data from DIAN.

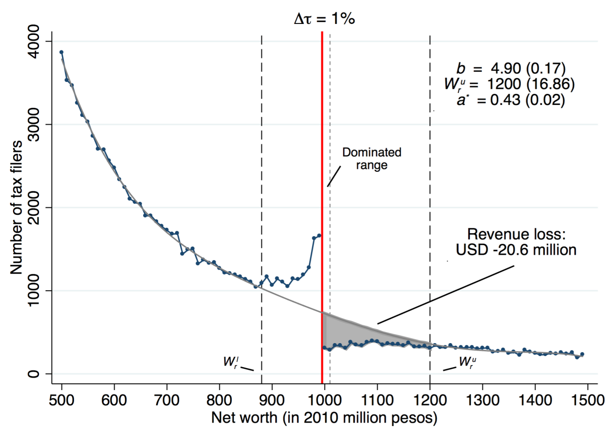
Figure 1.3: Bunching Theory and Estimation: Density Distribution of Reported Wealth



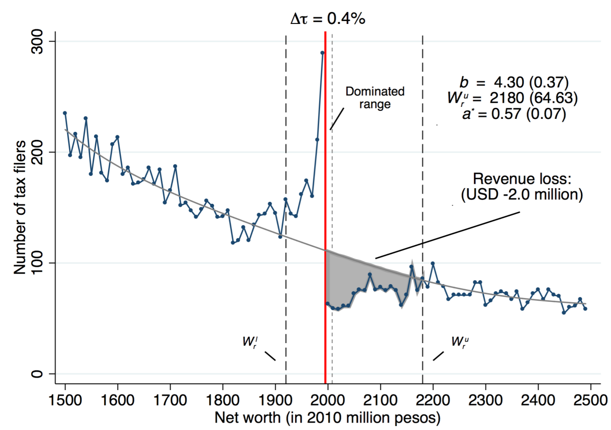
Notes: These figures illustrate the bunching approach to a proportional tax notch ($\Delta\tau > 0$) that discontinuously raises tax liability for those reporting wealth above W_r^*. For simplification, this notch is associated with a small change, such that intensive responses by those who stay above the notch can be ignored. Panel (a) depicts the theoretical effect of this notch on the density distribution of reported net wealth in the presence of heterogeneous elasticities and optimization frictions. Before the notch, the distribution of reported wealth is smoothly decreasing around the cutoff (dashed black line). A group of individuals responds to the notch by underreporting wealth below W_r^*. The notch thus generates excess mass at W_r^* and corresponding missing mass between the interval $(W_r^*, W_r^* + \Delta W_r^*]$ (solid blue line). Some individuals in the range $(W_r^*, W_r^D]$ cannot bunch below the notch point due to high optimization frictions (to compare this result with the baseline case of no frictions and homogeneous elasticities, see Figure A.3). Panel (b) illustrates the bunching estimation. The blue line represents the (hypothetical) empirical density of taxpayers. The black solid line represents the counterfactual density, which is estimated by either fitting a flexible polynomial to the empirical density and excluding observations in a range $[W_r^l, W_r^u]$, or by using observed pre-reform density. We denote W_r^u as the net worth of the marginal buncher, obtained with the point of convergence method such that the excess mass (B) below W_r^* and the missing mass (M) above W_r^* are equal ($B = M$). A lower bound on the estimated elasticity is obtained by scaling excess mass B by $1 - a^*$, where a^* is the share of individuals in the range of the dominated region $(W_r^*, W_r^D]$ who are unresponsive to the notch.

Figure 1.4: Wealth Bunching Estimation at First Two Notches in 2010, and Robustness using Pre-Reform Counterfactual

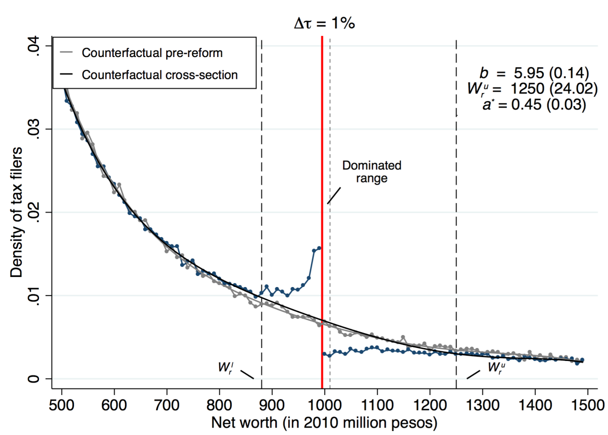
(a) First Notch



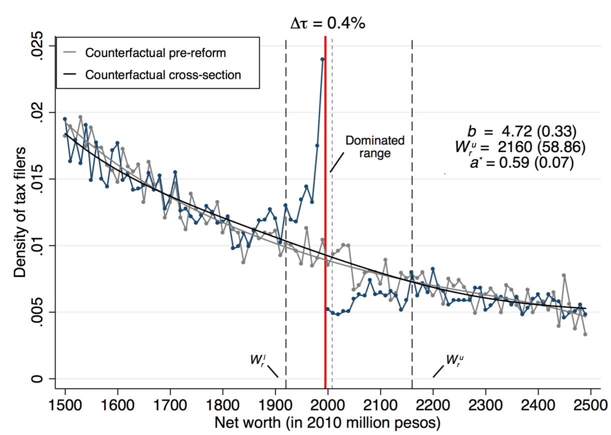
(b) Second Notch



(c) Robustness: First Notch



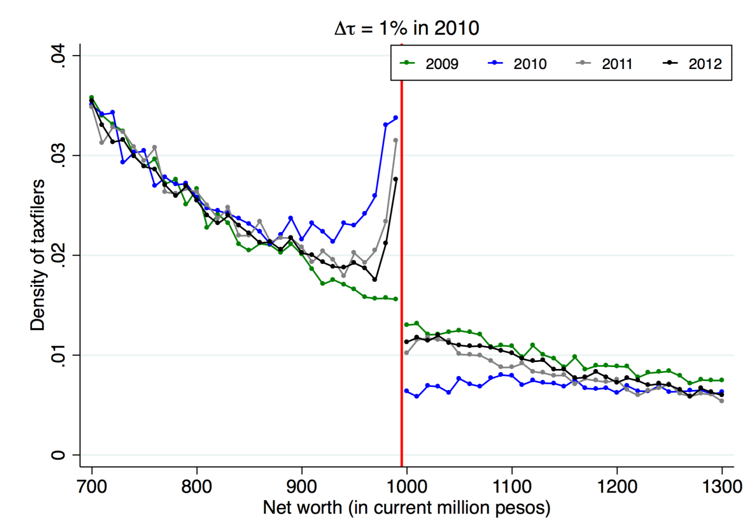
(d) Robustness: Second Notch



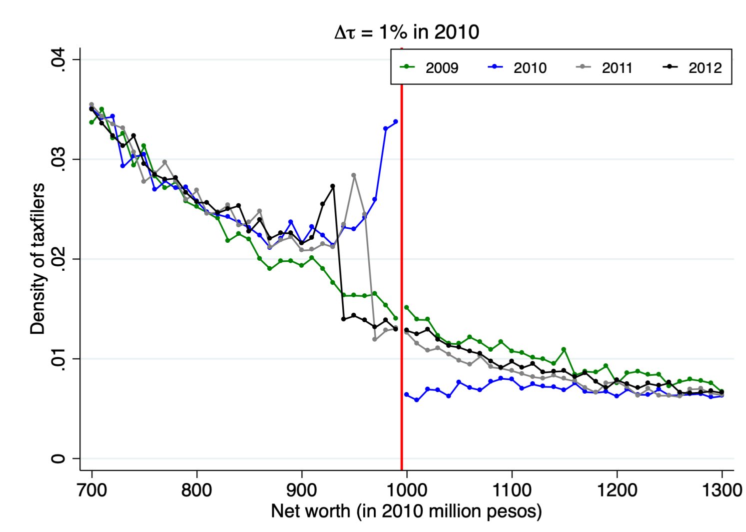
Notes: These figures display taxpayer density by 2010 net worth around the first notch in Panels (a) and (c), and the second notch in Panels (b) and (d). The counterfactual densities are obtained from the regression of a polynomial of degree 5 on all data points outside the $[W_r^l, W_r^u]$ interval in Panels (a) and (b), and using pre-reform data (in gray) in Panels (c) and (d). b is the excess mass as a share of the counterfactual and W_r^u is the net worth of the marginal buncher, obtained with the point of convergence method. The lower bound W_r^l is determined visually. The upper bound W_r^u is estimated from an iterative process: starting from $W_r^u = W_r^*$, we obtain the counterfactual and estimate the excess mass (B) below the threshold and missing mass (M) above the threshold. For low W_r^u, the excess mass is larger than the missing mass ($B >> M$). We iteratively increase W_r^u until the two masses are equal ($B = M$). a^* represents the share of individuals in the dominated range that do not bunch due to optimization frictions. The standard errors in parentheses are estimated from 1,000 bootstrap samples with replacement. Revenue losses are obtained by multiplying $\tau \times W_r$ by the shaded area between the counterfactual and observed densities. Bin width is 10,000,000 pesos (2010 USD 5,208.30 in 12/31/2010). The estimated parameters are summarized in Table 1.1. *Source:* Authors' calculations using administrative tax microdata from DIAN.

Figure 1.5: Bunching Persists Even When Wealth is Not Taxed

(a) Bunching in Current Pesos



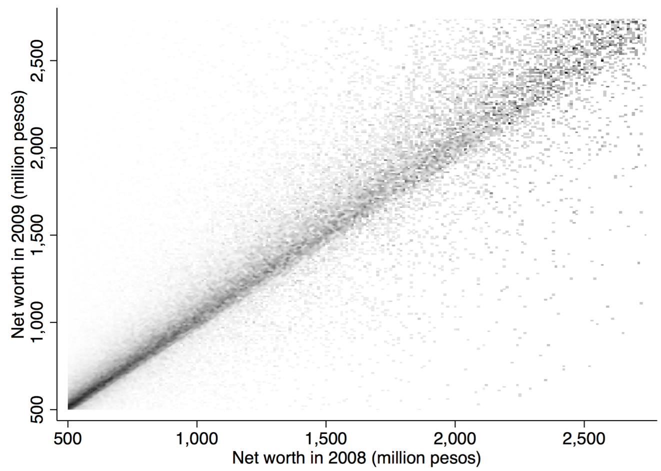
(b) Bunching in Constant Pesos



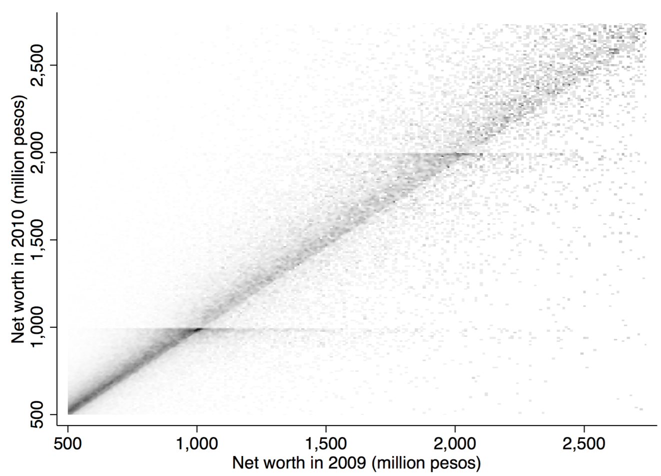
Notes: This figure shows the distribution of reported wealth before and after a reform that taxed wealth owned in 2010 for those reporting at least 1 billion pesos (USD 520,830) that year. Tax filer density by bins of 10,000,000 pesos is plotted in 2009 (no wealth tax for these taxpayers), 2010 (a 1 percent wealth tax), 2011 (no wealth tax), and 2012 (no wealth tax). Pesos are expressed in *current* terms in Panel (a) and in *constant* 2010 terms in Panel (b). Panel (a) shows the immediate emergence of excess mass below the exemption cutoff in 2010. Critically, bunching persists two years after wealth was taxed, even though there was no wealth tax those years. Panel (b) shows that, because taxpayers bunch below the cutoff in *current* pesos, they report less wealth in *real* terms. *Source:* Authors' calculations using administrative tax microdata from DIAN.

Figure 1.6: Tax Filer Density Before and After a Wealth Tax Reform

(a) Tax Filers in 2008 and 2009 (Not Subject to the Wealth Tax)



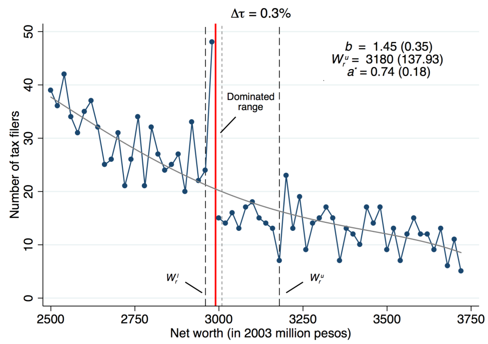
(b) Tax Filers in 2009 and 2010 (Subject to the Wealth Tax)



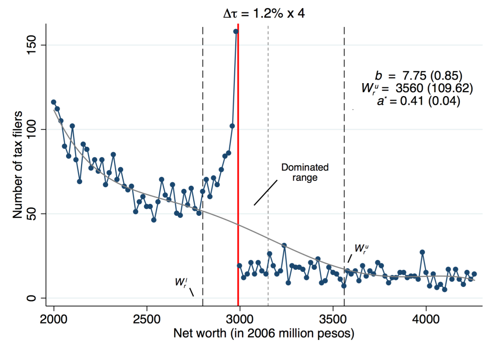
Note: These figures present density heatmaps for taxpayers reporting 500 million to 3,000 million pesos in bins of 2010 10 million × 10 million pesos (USD 5,208.30 in 12/31/2010) in 2008, 2009, and 2010. Bins are re-weighted by the total number of taxpayers in the column, so that darker bins represent higher relative taxpayer density. Panel (a) is restricted to individuals who filed in 2008 and 2009 (no notches or kinks). Taxpayers report similar values of wealth from year to year in the absence of wealth tax notches. The presence of diffuse mass in lieu of a precise 45 degree line suggests wealth rises and falls year-to-year, and that such variations are not necessarily related to changes in the wealth tax schedule. Panel (b) is restricted to individuals in 2009 and 2010 (a notched schedule was introduced in 2010). Taxpayers report more similar values of wealth relative to the previous year, rendering the 45 degree line more salient. Taxpayers also respond by bunching just below the two notches to avoid owing more wealth tax. *Source:* Authors' calculations using administrative data from DIAN.

Figure 1.7: Wealth Bunching Across Time and Notches: 2003, 2006, 2010, and 2014

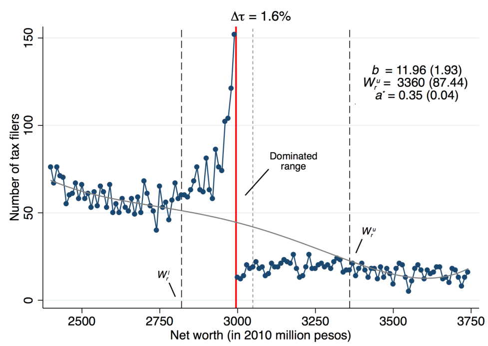
(a) 2003, First (and Only) Notch



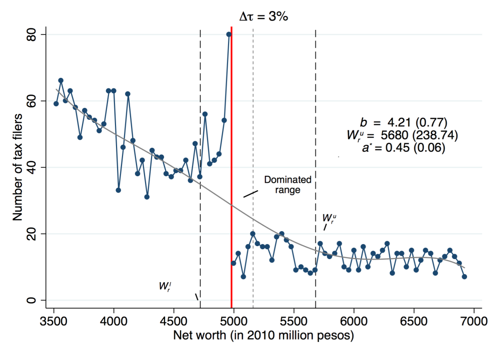
(b) 2006, First (and Only) Notch



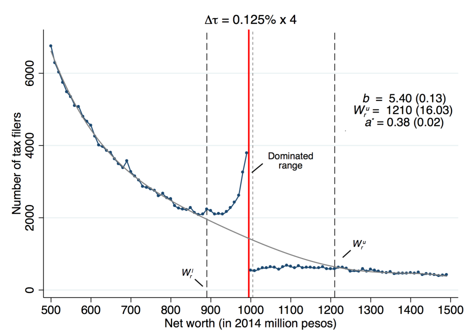
(c) 2010, Third Notch



(d) 2010, Fourth Notch



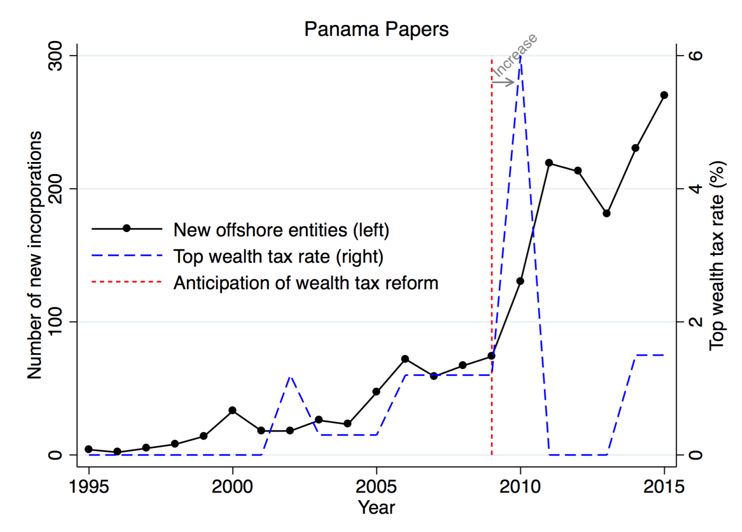
(e) 2014, First Notch



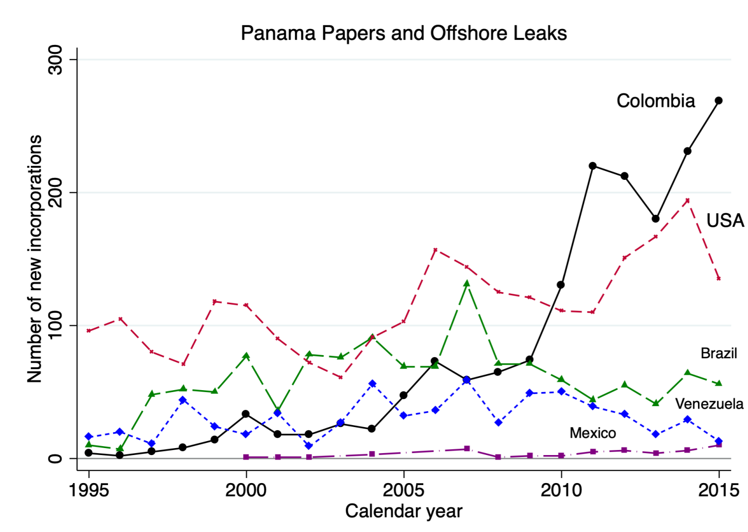
Notes: These figures display taxpayer density by net worth in 2003, 2006, 2010, and 2014, and fit the counterfactual distribution. b is the excess mass as a share of the counterfactual and W_r^u the net worth of the marginal buncher, obtained with the point of convergence method which requires the excess and missing masses around the threshold be equal. a^* is the share of individuals in the dominated range that do not bunch. The tax notch $\Delta\tau$ is $1.2\% \times 4$ in 2006 because individuals reporting wealth of 3 billion or more that year were subject to the wealth tax until 2009 (inclusive). The tax notch is $0.125\% \times 4$ in 2014 for similar reasons. Standard errors in parentheses are estimated from 1,000 bootstrap samples with replacement. Bin width is 20 million pesos in Panels (a) and (b), 10 million pesos in Panels (c) and (e), and 40 million pesos in Panel (d). The estimated parameters are summarized in Table 1.1. *Source:* Authors' calculations using administrative tax microdata from DIAN.

Figure 1.8: The Use of Offshore Entities

(a) Colombia



(b) Colombia versus Selected Countries

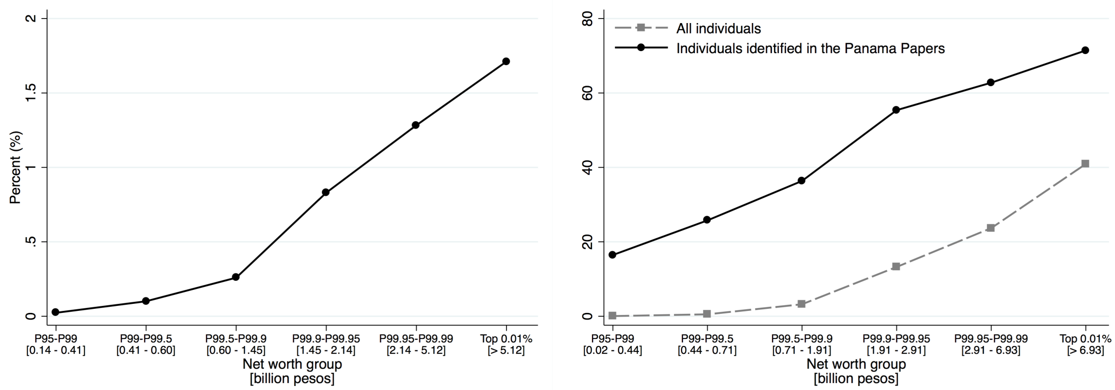


Note: These figures, based on the Panama Papers microdata, compare the flow of offshore entities created by individuals from Colombia and other countries through Panamanian law firm Mossack Fonseca. Panel (a) plots the number of new Colombian offshore entities that are incorporated every year (black solid line) and the top statutory annual wealth tax rate (blue dashed line). The figure suggests wealth tax changes are associated with a more frequent incorporation of offshore entities. Panel (b) compares Colombia's offshore entity incorporation flow with that of several other countries (the red vertical line marks the reintroduction of net wealth taxes in Colombia). The figure suggests Colombians stand out in their use of offshore entities, even relative to larger and wealthier countries. Both panels include active and inactive offshore entities.
Source: ICIJ. Accessed June 12, 2017.

Figure 1.9: Panama Papers and Disclosures of Hidden Wealth Under Voluntary Disclosure Scheme, by Wealth Group

(a) Probability of Appearing Named in the Panama Papers

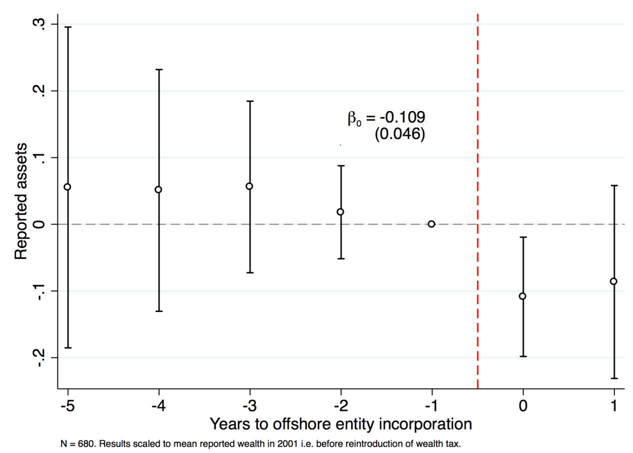
(b) Probability of Disclosing Hidden Wealth Under Scheme



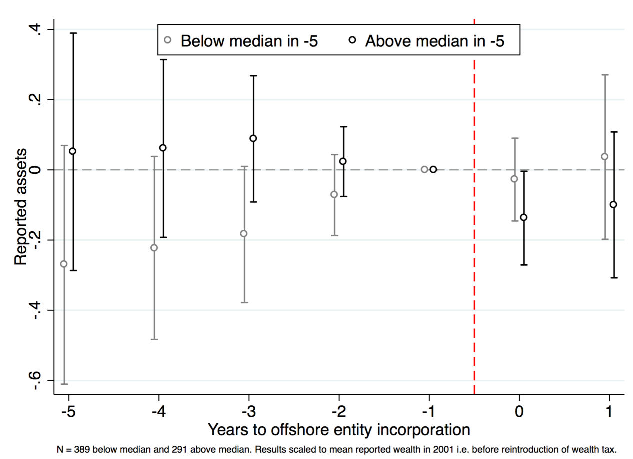
Notes: Panel (a) plots the fraction of tax units in Colombia identified in the Panama Papers by bins of net worth. Individuals are ranked by their average net worth reported for any year in which they have filed an income tax return. The sample includes 1,208 shareholders of offshore entities created by Mossack Fonseca who could be exactly matched to their individual income tax return filed at any point between 1993 and 2015. The figure shows that the likelihood of appearing in the Panama Papers is increasing in reported wealth (the differences across wealth groups are always statistically significant). One in sixty individuals in the wealthiest 0.01 percent (1.7 percent) are identified in the Panama Papers. Panel (b) plots the probability of participating in voluntary disclosure program by wealth bins for all tax filers (gray dash line), and tax filers identified in the Panama Papers (solid black line). Individuals are ranked by FY 2013 net worth plus any disclosures made under the scheme. The figure shows that two-fifths percent of individuals in the wealthiest 0.01 percent disclosed hidden wealth under the scheme. This share is 71 percent for individuals identified in the leak. The sample in Panel (b) is restricted to 1,633,383 individuals filing the income tax return in FY 2013, and includes 11,210 disclosers and 1,085 individuals in the Panama Papers (of which 434 disclosed hidden wealth). The exchange rate for 1 billion pesos is USD 335,130 in 12/31/2017. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Figure 1.10: The Effect of Opening an Offshore Entity on Reported Assets

(a) All Tax Filers



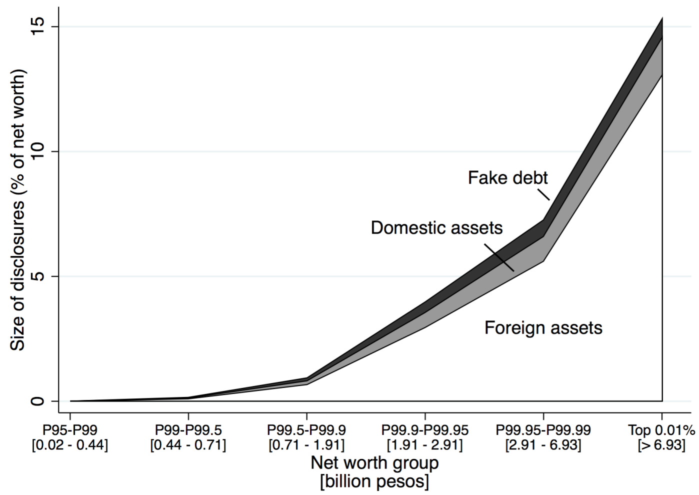
(b) Heterogeneity: Above- vs Below-Median Reported Assets in Event Time -5



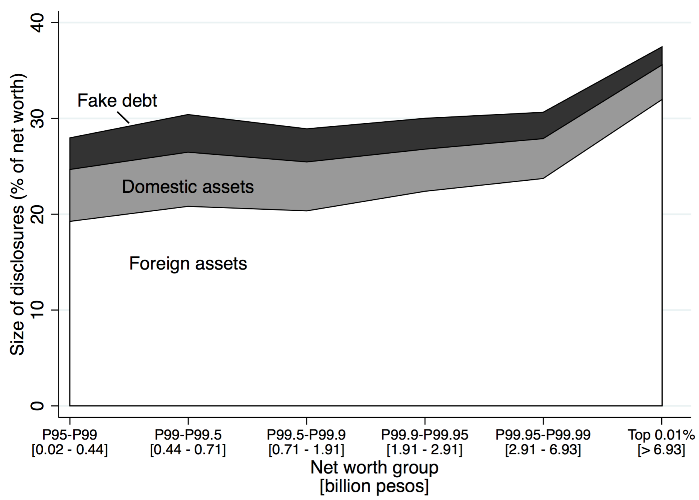
Notes: These figures compare total assets reported to the Colombian tax authority before and after incorporating an offshore entity in Panama and other tax havens through Mossack Fonseca. Panel (a) presents the β_k coefficients from event study specification (1.7). The outcome variable is total reported assets scaled with respect to mean reported wealth in 2001 (i.e., the year preceding the reintroduction of wealth taxation in Colombia). An "event" is defined as the year an individual incorporates an offshore entity for taxpayers with only one offshore entity. The sample is balanced in event time and excludes taxpayers not appearing in the Panama Papers. The figure suggests reported assets drop by 10.9 percent the year a taxpayer incorporates an offshore entity through Mossack Fonseca. This drop is consistent with wealth obfuscation for the purpose of reducing the tax burden. Panel (b) presents results separately for individuals with above- and below-median reported assets in event time -5. The figure suggests that the result is driven by the wealthiest individuals: for this group, reported assets significantly drop by 13.7 percent the year of offshore entity incorporation. As a robustness check, Figure A.9 plots β_k coefficients for different event time windows. *Sources:* Authors' calculations using administrative tax microdata from DIAN and ICIJ.

Figure 1.11: Size of Disclosed Hidden Wealth in 2015–2017 Relative to Post-Disclosure Net Worth, by Post-Disclosure Top Wealth Group

(a) All



(b) Disclosers Only

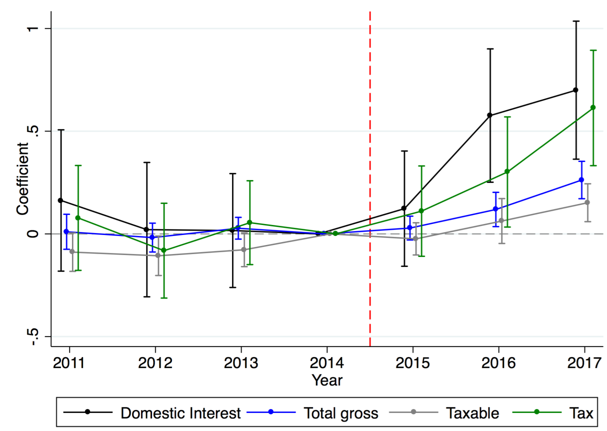


Notes: These figures show the magnitude of hidden foreign and domestic assets and fake liabilities disclosed during the 2015–2017 voluntary disclosure program. The sample is restricted to 1,633,383 individuals filing income taxes in FY 2013 (they may or not file a wealth tax return in 2015–2017), of which 11,210 participated in the disclosure program ("disclosers"). Tax filers are ranked by pre-program (FY 2013) net worth *plus* disclosures. Panel (a) shows the size of disclosed hidden assets and fake liabilities between 2015–2017 as a share of FY 2013 net worth *plus* disclosures. The figure suggests 15.3 percent of wealth had been hidden for among the top 0.01 percent. Panel (b) plots these shares for disclosers only. The figure suggests that hidden wealth represented 37.5 percent of disclosers' wealth for those in the top 0.01 percent. The exchange rate for 1 billion pesos is USD 335,130 in 12/31/2017. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

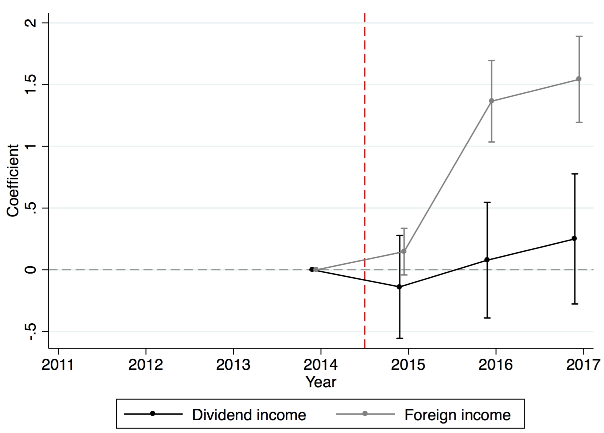
Figure 1.12: The Impact of a Voluntary Disclosure Program on Reported Wealth and Income

(a) Wealth: Gross and Net

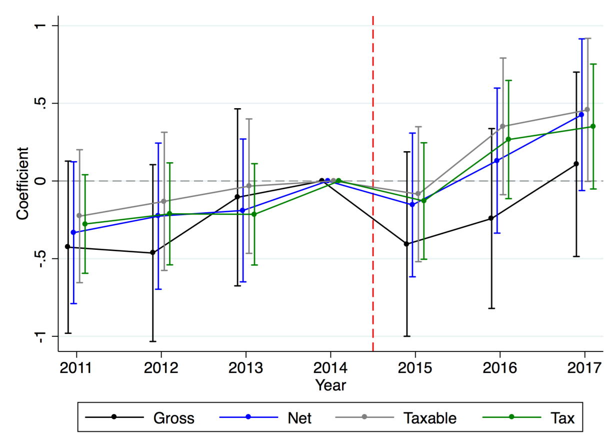
(b) Regular Income: Interest, Total Gross, Taxable, and Tax



(c) Regular Income: Dividend and Foreign

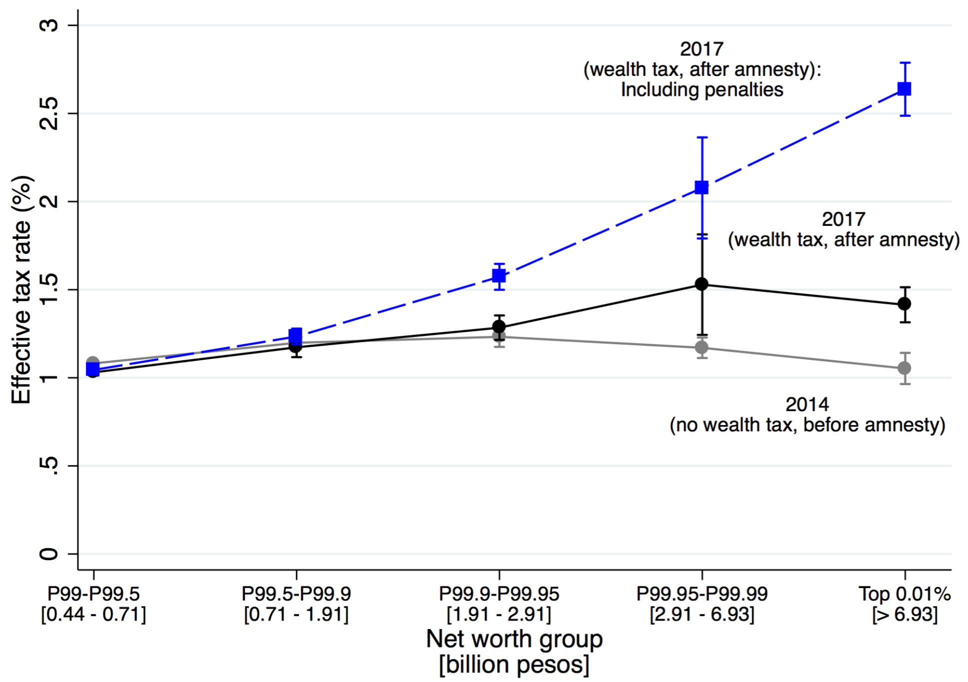


(d) Irregular Income (Includes Long-Rerm Realized Capital Gains): Gross, Net, Taxable, and Tax



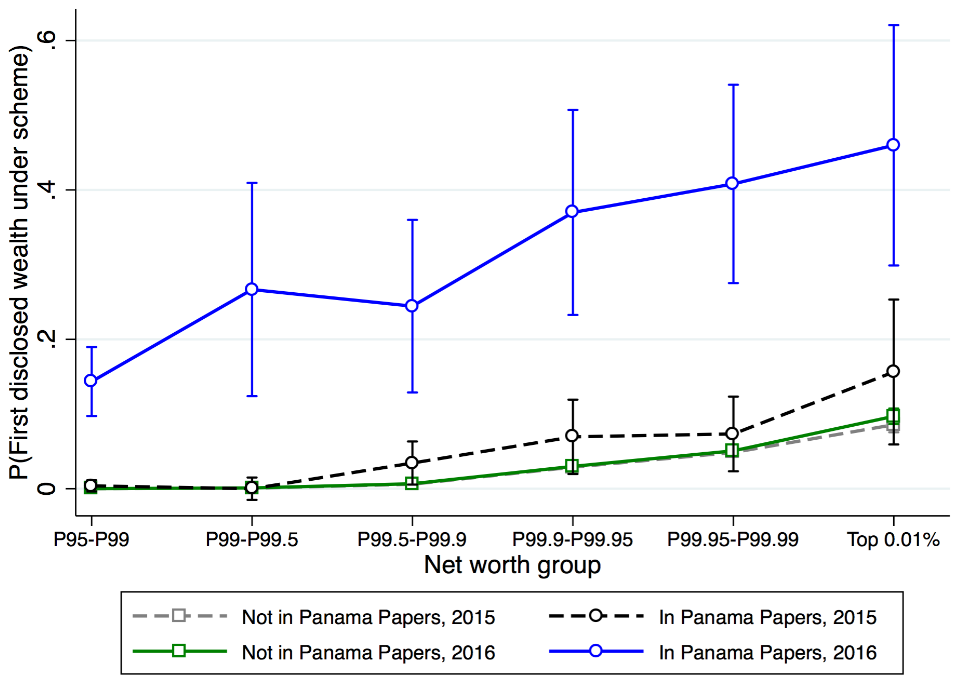
Notes: These figures present the effect of the voluntary wealth disclosure scheme on tax compliance. The outcome variable is reported wealth in Panel (a) and reported income categories in Panels (b)–(d). The figures compare outcomes between 1,777 individuals that voluntarily disclosed hidden wealth in 2015 and 43,181 that never disclosed under the scheme between 2015 and 2017. The outcome, an inverse hyperbolic sine transformation of a given income category, is regressed on individual fixed effects and a voluntary discloser dummy interacted with year fixed effects (2014 is the omitted category). Standard errors are clustered at the taxpayer level. The lines plot the coefficients on the interaction terms and 95 percent confidence intervals. The red vertical line marks the period individuals disclosed hidden wealth. The figure shows that the scheme raised wealth reported to the tax authority three years after initial disclosure, as well as reported capital income (interest income, foreign income, realized capital gains) derived from asset ownership. As a result, income tax liability increased. The sample is a balanced panel of 44,958 individuals that both filed income taxes annually between 2011 and 2017, and filed the wealth tax in 2015, 2016, or 2017. Tax filers that first disclosed assets and liabilities *after* 2015 (i.e., in 2016 or 2017) are excluded from the estimation sample. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Figure 1.13: Rise in the Effective Tax Rate for Wealthiest Tax Filers



Notes: This figure illustrates how wealth taxation and a voluntary disclosure scheme raised tax progressivity at the top of the wealth distribution. The figure plots average taxes paid on income and wealth in 2014 and 2017, expressed as a share of net wealth for subgroups of individuals in the wealthiest 1 percent of the distribution. Individuals are ranked by their net wealth reported before the voluntary disclosure scheme (FY 2013) including any disclosures made under the scheme. The gray curve plots income taxes in calendar year 2014 (FY 2013), before the wealth tax was reintroduced. The black curve plots income *and* wealth taxes in 2017 (FY 2016), while the dashed blue curve adds the penalties associated with the disclosure program that year. The figure shows that wealth taxation increases taxes paid by the wealthiest individuals, and that the voluntary disclosure scheme more than doubled the average effective tax rate for the wealthiest group of individuals. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Figure 1.14: The Panama Papers Leak Raised Disclosures of Hidden Wealth

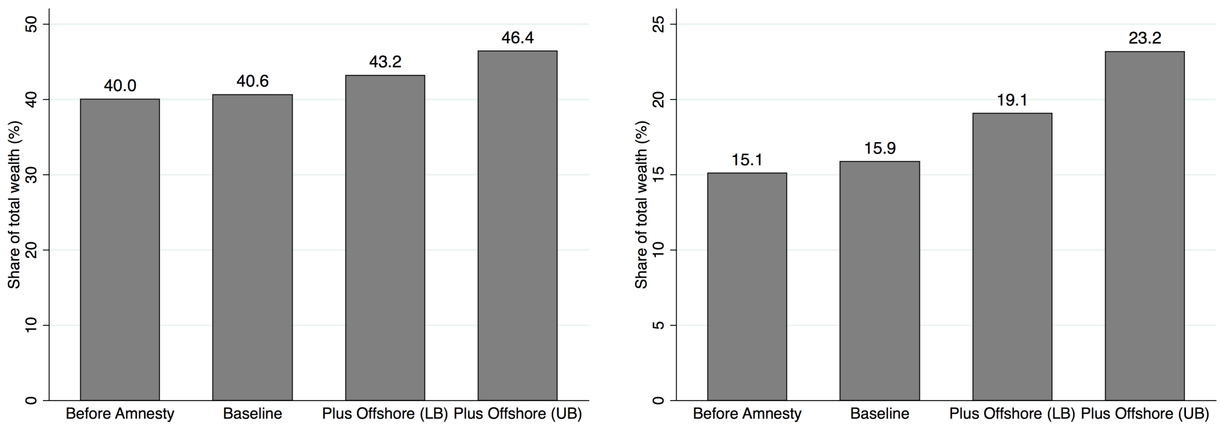


Notes: This figure illustrates the impact of the Panama Papers leak on tax compliance. The figure plots the probability of first disclosing hidden assets and/or fake liabilities in 2015 (before the leak) and 2016 (after the leak) for taxpayers in the Panama Papers (round marker) and taxpayers not in the Panama Papers (square marker) by wealth group. The vertical lines represent the 95 percent confidence intervals. The figure suggests that the Panama Papers leak in 2016 raised evaders' willingness to disclose hidden wealth for those named in the leak. The sample is an unbalanced panel of 2,421,936 individuals that either filed the income tax in 2014–2016 or filed a wealth tax return 2015–2017, and includes 11,927 disclosers and 1,167 individuals in the Panama Papers (of which 453 ever disclosed assets). Wealth groups are generated every year including disclosures. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Figure 1.15: Wealth Inequality in Colombia Including Hidden Offshore Wealth

(a) Top 1 Percent

(b) Top 0.1 Percent



Notes: These figures presents estimates of top wealth shares in Colombia for 2017. In the baseline estimates, the top 1 percent owns 40.64 percent of total wealth in Panel (a), and the top 0.1 percent owns 15.85 percent of total wealth in Panel (b). The figures compare wealth inequality excluding offshore assets disclosed during the amnesty (first bar). In addition, the figures present estimates for top shares corrected by including unreported offshore wealth. Using data from Alstadsater et al. (2018a) and the Panama Papers leak, the lower bound assumes unreported offshore wealth today represents 6.2 percent of GDP, while the upper bound assumes it represents 15 percent. *Sources:* Table A.5.

Table 1.1: Summary of Notches, Responses, and Elasticities

						Response ΔW_r^* (mill. pesos)		Reduced-Form Elasticity e_R	
Year of Reform (1)	Notch Point (mill. pesos) (2)	Exemption Cutoff (3)	ATR Jump $\Delta\tau$ (%) (4)	Dominated Range ΔW_r^D (mill. pesos) (5)	Frictions a^* using ΔW_r^D (6)	Bunching-Hole Method (7)	Convergence Method (8)	Bunching-Hole Method (9)	Convergence Method (10)
2003	3,000	✓	0.3	9	0.74	120	180	0.27	0.60
					(0.18)	(80.09)	(137.93)	(0.62)	(1.76)
2006	3,000	✓	1.2×4	151	0.41	340	560	0.13	0.36
					(0.04)	(52.04)	(109.62)	(0.04)	(0.16)
2010	1,000	✓	1.0	10	0.43	110	200	0.60	2.00
					(0.02)	(7.91)	(16.86)	(0.09)	(0.35)
2010	2,000		0.4	8	0.57	110	180	0.37	1.00
					(0.07)	(24.10)	(64.63)	(0.19)	(0.99)
2010	3,000		1.6	49	0.35	220	360	0.17	0.44
					(0.04)	(40.82)	(87.44)	(0.09)	(0.20)
2010	5,000		3.0	160	0.45	360	680	0.08	0.30
					(0.06)	(105.16)	(238.74)	(0.06)	(0.23)
2014	1,000	✓	0.0125×4	5	0.38	110	210	1.21	4.41
					(0.02)	(6.31)	(16.03)	(0.15)	(0.71)

Notes: This table presents elasticity estimates at different wealth levels exploiting four wealth tax reforms taking place in 2003, 2006, 2010, and 2014. Column (1) presents the year of the wealth tax reform. Column (2) indicates the bracket cutoff, expressed in current million pesos. Column (3) indicates whether this cutoff also marks the eligibility threshold, below which taxpayers are exempt from the wealth tax. Column (4) presents the size of the wealth tax notch. Column (5) presents the dominated range in current million pesos, defined as $\Delta\tau \cdot W_r^*/(1-\tau-\Delta\tau)$. Column (6) presents the estimate of frictions (the fraction of individuals in dominated ranges who are unresponsive). Columns (7)–(8) present the reporting responses in current million pesos using bunching-hole and convergence methods, respectively. Columns (9)–(12) present elasticities based on the reduced-form formula (1.3) in columns (9)–(10). Table A.6 includes the elasticities based on a parametric equation for a comparison. *Source:* Authors' calculations using administrative tax microdata from DIAN.

Table 1.2: Compliers (i.e., bunchers) analysis for the first notch in 2010

	$E[Y(0)\|\text{Compliers}]$ (1)	$E[Y(1)-Y(0)\|\text{Compliers}]$ (2)	$E[Y\|\text{Always-takers}]$ (3)	$E[Y\|\text{Never-takers}]$ (4)
Fixed assets (real estate, land, vehicle)	0.527	-0.002	0.63	0.541
	(0.027)	(0.023)	(0.042)	(0.01)
Stocks and shares	0.17	0.015	0.091	0.161
	(0.02)	(0.019)	(0.027)	(0.007)
Inventories	0.109	-0.024	0.109	0.107
	(0.016)	(0.012)	(0.024)	(0.006)
Bank deposits, bonds, other investments	0.086	0.008	0.08	0.094
	(0.016)	(0.017)	(0.022)	(0.005)
Accounts receivable	0.066	0.025	0.074	0.08
	(0.016)	(0.013)	(0.02)	(0.005)
Other assets	0.038	-0.021	0.019	0.018
	(0.011)	(0.012)	(0.016)	(0.002)
Debt	0.094	0.033	0.079	0.121
	(0.005)	(0.004)	(0.009)	(0.002)

Notes: This table presents the results of a compliers analysis using the set-up illustrated by Figure A.5. In this setting, a complier refers to a taxpayer bunching below the exemption cutoff in response to the wealth tax. The sample is a balanced panel of 8,016 income tax filers reporting net wealth between W_r^l and W_r^u in 2008, 2009, and 2010. The endogenous variable is $B_{it} = 1$ if the individual has net wealth (in 2010 pesos) between W_r^l and W^*, i.e., the bunching region. Complier means in Column (1) are calculated as the coefficient on $1 - B_{it}$ in a 2SLS regression of $1 - B_{it}$ multiplied by Y_i and using 2010 as the instrument (Z_{it}). Column (2) presents the 2SLS coefficient β_1 from specification (1.5). Always-taker and never-taker means are calculated in analogous 2SLS regressions of $B_{it}(1 - Z_{it})Y_{it}$ on $B_{it}(1 - Z_{it})$ and $(1 - B_{it})Z_{it}Y_{it}$ on $(1 - B_{it})Z_{it}$, respectively, again using 2010 as Z_{it}. The first stage coefficient is 0.313 (t-stat 35.3) for debt, and 0.275 (t-stat 13.13) for all others, as only business owners keeping records report asset types separately. Standard errors are clustered at the taxpayer level. The table suggests bunchers inflate their debt and underreport inventories, which are not covered by third-party reports, to artificially place themselves below the wealth tax exemption cutoff. *Source:* Authors' calculations using administrative tax microdata from DIAN.

Table 1.3: Who are the Shareholders of Mossack Fonseca's Offshore Entities?

	Not in Panama Papers (1)	In Panama Papers (2)
Number of individuals	3,300,718	1,208
Number of years filed tax return	6.21	15.83
	[5.87]	[6.91]
Demographics		
Male (percent)	56.2	63.4
Born after 1985 (percent)	7.2	5.1
Rentier (percent)	13.21	27.40
Wage-earner (percent)	37.16	59.69
Other (percent)	49.63	12.91
Income and wealth (2017 millions)		
Gross wealth	263.11	1,880.42
	[1,034.38]	[3,511.60]
Net worth	209.72	1,533.36
	[878.76]	[3,245.76]
Irregular capital income	8.11	65.04
	[140.32]	[178.49]
P99 (percent)	9.54	68.96
P99.9 (percent)	0.95	28.75
P99.99 (percent)	0.09	4.47

Notes: This table presents descriptive statistics (means and standard deviations in brackets) for the 3.3 million income tax filers we observe between tax years 1993 and 2016 (Column 1) and for tax filers that appear named in the Panama Papers (Column 2). Rentier, Wage-earner and Other refer to economic activity codes, as self-reported by taxpayers to the tax authority. Rentier also includes individuals without an economic activity as well as dependents. Income and wealth values (in 2017 million pesos) and top percentile groups use individual net worth means across tax years. The exchange rate for 1 million pesos is USD 335.13 in 12/31/2017. *Sources:* Authors' calculations using administrative tax microdata from DIAN and ICIJ.

Table 1.4: Who Disclosed Hidden Assets or Inexistent Liabilities?

	Non-disclosers	Disclosers	Disclosers by year of disclosure		
			2015	2016	2017
	(1)	(2)	(3)	(4)	(5)
Number of individuals	55,098	11,210	2,179	2,708	6,323
Number of years filed tax return	19.69	19.19	19.08	19.38	19.15
	[5.00]	[5.62]	[5.58]	[5.55]	[5.67]
Demographics					
Male (percent)	60.58	57.76	61.03	56.99	56.95
Born after 1985 (percent)	0.0	0.0	0.0	0.0	0.0
Rentier (percent)	29.54	34.55	33.59	35.23	34.59
Wage-earner (percent)	24.29	28.72	24.55	31.09	29.13
Other (percent)	46.17	36.74	41.85	33.68	36.28
In the Panama Papers (percent)	0.43	3.87	1.88	10.60	1.68
Buncher (percent)	18.27	13.26	13.68	12.15	13.59
Pre-Disclosure Wealth					
Gross wealth	2,837.43	3,288.56	4,445.34	2,976.44	3,023.58
	[7,638.74]	[20,666.54]	[37,727.68]	[5,092.94]	[15,975.48]
Net worth	2,367.62	2,719.82	3,538.67	2,487.30	2,537.22
	[5,557.32]	[18,570.51]	[36,321.41]	[4,295.27]	[12,196.98]
Irregular capital income	98.29	95.20	110.60	72.94	99.42
	[678.19]	[740.71]	[526.18]	[311.14]	[914.13]
P99 (percent)	100.0	100.0	100.0	100.0	100.0
P99.9 (percent)	46.08	39.71	43.05	42.87	37.21
P99.99 (percent)	4.04	6.75	7.85	7.24	6.17

Notes: This table presents descriptive statistics (means and standard deviations in brackets) for 66,308 income tax filers in FY 2013 (i.e., immediately before the disclosure program) that filed a wealth tax declaration in either 2015, 2016, or 2017. Column 1 provides summary statistics for those that did not disclose hidden wealth or inexistent liabilities in 2015, 2016, nor 2017, while Column 2 provides summary statistics for those that ever disclosed between 2015 and 2017. Columns 3–5 provide summary statistics separately by year of first disclosure. Rentier, Wage-earner and Other refer to economic activity codes, as self-reported by taxpayers to the tax authority. Rentier also includes individuals without an economic activity as well as dependents. An individual is coded as a buncher if she is located within the bunching region $[W_r^l, W_r^*)$ in FY 2003, 2006 or 2010, as defined in Section 1.3. Wealth values (expressed in 2017 million pesos) and top percentile groups (using net worth) are based on values reported in the income tax statement in FY 2013 (calendar year 2014), i.e., immediately before the disclosure program. The exchange rate for 1 million pesos is USD 335.13 in 12/31/2017. *Sources:* Authors' calculations using administrative tax microdata from DIAN and ICIJ.

Table 1.5: The Impact of a Voluntary Disclosure Program on Wealth and Income Reported to the Tax Authority

	Wealth		Income						Capital gains and other irregular income			
	Gross (1)	Net (2)	Foreign (3)	Dividend (4)	Interest (5)	Total gross (6)	Taxable (7)	Tax (8)	Gross (9)	Net (10)	Taxable (11)	Tax (12)
DID	0.288***	0.400***	1.010***	0.061	0.416***	0.132***	0.132***	0.329***	0.069	0.322**	0.339***	0.339***
	(0.015)	(0.030)	(0.105)	(0.170)	(0.117)	(0.030)	(0.034)	(0.090)	(0.157)	(0.130)	(0.125)	(0.100)
N	314,706	314,706	138,004	138,004	314,706	314,706	314,706	314,706	314,706	314,706	314,706	314,706
R^2	0.66	0.572	0.614	0.753	0.629	0.686	0.547	0.641	0.264	0.246	0.246	0.242

Notes: This table presents the effects of the 2015 voluntary disclosure program on income and wealth reported to the Colombian tax authority. The dependent variables in columns (1) and (2) are taken from the wealth tax form 440, while those in columns (3)–(12) are taken from the individual income tax forms 110 and 210. Outcomes are expressed in log-approximation form using the inverse hyperbolic sine function. The table compares outcomes between 1,777 taxpayers that voluntarily disclosed hidden assets and inexistent liabilities in 2015 and 43,181 that did not disclose between 2015 and 2017. Each outcome is regressed on individual fixed effects, year fixed effects, and an interaction of the voluntary discloser dummy and post-reform years (2014 is the omitted category): $\log(y_{it}) = \alpha_i + \gamma_t + \beta \cdot \mathbb{1}(\text{Post} \times \text{Discloser}) + \nu_{it}$. The standard errors in parentheses are clustered at the taxpayer level. The sample is a balanced panel of 44,958 individuals that both filed income taxes annually between 2011 and 2017, and filed the wealth tax in 2015. Tax filers that first disclosed assets and liabilities strictly after 2015 (i.e., in 2016 or 2017) are excluded from the estimation sample. The number of observations with foreign income and dividend income is smaller than the rest because taxpayers report these two varibles as separate variables starting 2014. Wealth tax liability is not reported as an outcome because there is no wealth tax during most of the pre-program period. $*p < 0.1, **p < 0.05, ***p < 0.01$. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Table 1.6: The Effect of the Panama Papers Leak on Wealth Disclosures and Taxes Owed

	Dependent variable			
	$\mathbb{1}$(Disclosed any)	$\mathbb{1}$(Disclosed foreign asset)	log(Wealth tax)	log(Wealth tax plus penalties)
	(1)	(2)	(3)	(4)
β	0.274***	0.296***	0.261***	0.850***
	(0.027)	(0.025)	(0.043)	(0.088)
Control Mean	0.0328	0.0192	15.221	15.315
N	118,966	118,966	118,966	118,966
R^2	0.015	0.023	0.001	0.004

Notes: This table presents the effect of the Panama Papers leak on willingness to disclose hidden wealth and tax liability. Column (1) presents the likelihood of disclosing any hidden asset or fake liability for the first time. Column (2) plots the likelihood of disclosing foreign assets in particular. Columns (3) and (4) display wealth tax liability, including or excluding penalties associated with disclosing hidden wealth. These outcomes are expressed in log-approximation form using the inverse hyperbolic sine function. The table compares outcomes using a balanced panel of 59,483 individuals that filed the wealth tax return in 2015 (before the Panama Papers leak) and 2016 (after the Panama Papers leak), 504 of which appear in the Panama Papers and 58,979 of which do not. The difference-in-differences coefficient represents β from specification (1.9). The standard errors in parentheses are clustered at the individual level. Column (1) results suggest that taxpayers in the Panama Papers are 27.4 percentage points more likely to first disclose hidden assets and/or fake liabilities after the Panama Papers leak relative to taxpayers that do not appear in the Panama Papers before the leak. On a basis of 0.0328, this implies a more than ninefold increase in the likelihood of disclosing hidden wealth. $*p < 0.1, **p < 0.05, ***p < 0.01$. *Sources:* Authors' calculations using administrative tax microdata from DIAN and ICIJ.

Chapter 2

Upstream and Downstream Impacts of College Merit-Based Financial Aid for Low-Income Students: *Ser Pilo Paga* in Colombia

2.1 Introduction

The rising higher education premium observed in many countries suggests that college is increasingly important to financial well-being (Autor, 2014). However, there is clear evidence that students from disadvantaged backgrounds are severely underrepresented in the post-secondary pipeline worldwide, particularly at selective higher education institutions (HEIs; Chetty, Friedman, Saez, Turner, and Yagan, 2017; Hoxby and Avery, 2013; Ferreyra, Avitabile, Botero, Haimovich, and Urzúa, 2017). This often persists despite extensive financial aid available to high-achieving, low-income students offered by governments and HEIs in developed countries.[1] In developing countries, credit markets are generally less well-developed; financial aid and student loan programs are more recent developments. Even so, research on the efficacy of such programs is more nascent (Solis, 2017). Nonetheless, it is precisely in these contexts that financial aid policies may have the largest enrollment potential.

[1]See Dynarski (2003); Kane (2003); Angrist, Autor, Hudson, and Pallais (2014); Bettinger, Gurantz, Kawano, and Sacerdote (2019); Fack and Grenet (2015); Solis (2017); Melguizo, Sanchez, and Velasco (2016). In fact, the enrollment gap between high- and low-income students has persisted despite a plethora of alternative initiatives, such as class-based affirmative action (Kahlenberg, 2014; Alon and Malamud, 2014), information provision (Hoxby and Avery, 2013; Hastings, Neilson, and Zimmerman, 2015; Hoxby and Turner, 2014), intensive college counseling (Castleman and Goodman, 2017), personal assistance in completing FAFSA filing (Bettinger, Long, Oreopoulos, and Sanbonmatsu, 2012), reduction in application costs (Avery, Hoxby, Jackson, Burek, Pope, and Raman, 2006; Pallais, 2015), and making college entry exams mandatory (Goodman, 2016).

We study the impact of transitioning from a setting with very little financial aid to one with full scholarship-loans for high achieving, low-income students. We exploit a large-scale program introduced in Colombia between 2014 and 2018 called *Ser Pilo Paga* (SPP). SPP has four key characteristics. First, beneficiaries are extremely high achieving, scoring in the top decile of the standardized high school exit exam (taken by all seniors, regardless of their postsecondary intentions). Beneficiaries also belong to the poorest 50 percent of households, as measured by the main proxy-means testing system for social assistance. Second, the program is highly visible and has simple, easy-to-understand eligibility rules and application procedures. Students are readily able to ascertain their own aid eligibility from their test score and household wealth index. Third, SPP is generous and large in scale. The program covers the full tuition cost of attending *any* four- or five-year undergraduate program in *any* government-certified "high-quality" university in Colombia. SPP annually benefits roughly one-tenth of all first-year postsecondary enrollees and one-third of those entering "high-quality" universities immediately after graduating high school. Lastly, program eligibility strongly predicts overall access to financial aid because students in Colombia, as in many other developing countries (Solis, 2017), have little alternative sources of aid.

We exploit SPP's sharp merit and need requirements to identify the program's causal impacts on enrollment and college choice in a regression discontinuity (RD) design. We focus mainly on the first cohort, who cannot manipulate nor influence their scores around the eligibility cutoffs. SPP was announced after they took the national standardized exam (merit) and there was no time to request a re-evaluation of their household wealth index (need). We find very large immediate enrollment impacts of financial aid. For need-eligible students, having a test score just above the cutoff raises immediate enrollment by 32 percentage points or 86.5 percent. For merit-eligible students, crossing the household wealth index cutoff raises immediate enrollment by 27.4 percentage points or 56.5 percent. These massive enrollment impacts virtually eliminated the socioeconomic status (SES) enrollment gradient among top decile test-takers.

Moreover, the program substantially altered students' college choices, shifting students from low- to high-quality HEIs. For instance, among need-eligible students, crossing the test score cutoff lowers enrollment at low-quality HEIs by 15.4 percentage points (57.7 percent) and raises enrollment at high-quality HEIs by 46.5 percentage points (426.6 percent). This implies aid substantially improved the quality of HEIs students are exposed to (e.g., better peers, more resources, higher graduation rates). Importantly, students disproportionately chose to enroll in private over public high-quality HEIs. This student sorting into private education is not explained by differential quality accreditation status. Instead, and consistent with previous literature (Riehl, Saavedra, and Urquiola, 2016), students appear to perceive private schools as more prestigious and producing greater value added—broadly defined—for students.

The large expansion of financial aid may have no impact on overall postsecondary enrollment if seats are fixed and aid recipients simply displaced other prospective students from colleges. Alternatively, aid may induce both demand and supply side effects that alter and perhaps improve outcomes for non-recipients too (i.e., a net social gain). On the demand side, financial aid could increase the option value of applying to colleges, with students sending their applications before determining whether they are eligible for aid. The advertising push associated with SPP could have also increased the perceived benefits of attending college. Peers of eligible students may also feel encouraged to attend college. On the supply side, colleges (especially private high-quality HEIs) may enlarge their cohorts in response to any rightward shift in demand. In addition, low-quality HEIs might fill the empty seats left by financial aid recipients with the next-best applicants. Such demand- and/or supply-side driven changes may affect enrollment aid-ineligible students and impact the postsecondary education system as a whole.

In the second part of the paper, we focus on these broader effects of financial aid on students eligible and ineligible for aid. We exploit rich administrative data linking the universe of high school test-takers to all postsecondary attendees, as well as college admission records, in the years before and after the program was introduced. Using a difference-in-differences approach that compares cohorts more or less exposed to financial aid expansion across time, we find evidence consistent with both demand- and supply-side mechanisms being at play. The demand for private high-quality undergraduate education significantly expanded following the expansion of financial aid, and these institutions responded by admitting and enrolling more students. Moreover, as aid recipients sorted out of low-quality HEIs, the empty seats left were filled by low-income, lower-performing applicants (test score deciles 9 and under). This suggests financial aid raised college attendance among both students eligible and ineligible for aid, thus producing not only equity gains but also net social gains.

As financial aid pushed low-income high-achievers into private high-quality HEIs, the student body composition at these colleges changed dramatically. Specifically, both student quality and socioeconomic diversity significantly increased at private, high-quality HEIs, with the share of low-income entering students increasing by a staggering 46 percent at these institutions.[2] That is, by relaxing their credit constraints, financial aid enabled low-income high-achievers to access expensive HEIs historically reserved for those who could afford them. In contrast, average student quality dropped at low-quality HEIs. Ability stratification thus largely replaced SES stratification in postsecondary schooling as a result of financial aid.

We end with two suggestive findings on the effect of financial aid on the behavior of low-quality HEIs and younger high school students. The exodus of high-ability aid bene-

[2] Using survey experiments and administrative data from a private, high-quality university, Londoño-Vélez (2016) analyzes how the socioeconomic diversity brought about by SPP in private, high-quality institutions affected high-income students' perception of inequality and preferences for redistribution.

ficiaries from low-quality HEIs put pressure on them to become more efficient and obtain High Quality Accreditation to attract high-achievers. Indeed, the number of institutional requests for High Quality Accreditation increased discontinuously after SPP was announced. However, the number of HEIs awarded this accreditation has only gradually increased over the last three years, and whether or not this reflects an actual quality improvement remains to be seen. Finally, relative pre-collegiate achievement improved among very low-income high school students following the expansion of aid. Comparing relative test performance among 2.7 million test-takers between 2012 and 2016 by socioeconomic background, very low-income students are 32 percent more likely to score in the top decile and 175 percent more likely to score in the top percentile two generations after policy rollout. These findings complement recent work by Laajaj, Moya, and Sanchez (2018), whose RD design causally shows that 2015 test-takers whose household wealth index renders them barely eligible for SPP performed better than those just barely ineligible by the wealth cutoff.[3] We interpret these findings as evidence against the notion that the overall pool of high-achieving, low-income students is inelastic (Hoxby and Avery, 2013; Kremer, Miguel, and Thornton, 2009; Angrist and Lavy, 2009; Angrist, Bettinger, Bloom, King, and Kremer, 2002; Barrera-Osorio and Filmer, 2016; Pallais, 2009). Given test scores are positively correlated with college degree completion and labor market outcomes in Colombia—even after controlling for baseline individual and college characteristics (MacLeod, Riehl, Saavedra, and Urquiola, 2017)—our results suggest SPP may bring future gains, promoting social mobility and reducing intergenerational inequality.

Our work contributes to the burgeoning literature that evaluates the impact of need- and merit-based financial aid on college attendance in developed and developing countries.[4] Our work is most closely related to Bettinger et al. (2019), Cohodes and Goodman (2014), Scott-Clayton and Zafar (2016), and Solis (2017), but builds on their work in three key aspects. First, we shed new light on the full effect of aid for low-income high-achievers, expanding the analysis to the impact of financial aid on both eligible and *ineligible* students. Indeed, a large expansion of financial aid may induce demand and supply responses that indirectly alter outcomes for non-recipients too. However, this has been difficult to evaluate in the United States, as financial aid has been gradually phased in (Dynarski and Scott-Clayton, 2013). Instead, we document the effects of drastically expanding financial aid available for low-income high-achievers. In our setting with *ex ante* little existing financial aid, we find

[3] Our descriptive result, from comparing performance by socioeconomic strata among *all* test-takers three years before and two years after SPP rollout, is consistent with the well-identified LATE estimate from Laajaj et al. (2018). While our approach is less well identified, we qualitatively reproduce their main findings while comparing the performance of low- and high-income students across time and two years after SPP was rolled out. We find that low-income students have crowded out high-income students from the top of the test score distribution. The magnitude of these results increases over time, as students, parents, and teachers have more time to reoptimize in response to the policy change.

[4] See Angrist et al. (2014); Angrist, Autor, Hudson, and Pallais (2016); Bettinger et al. (2019); Castleman and Long (2015); Dynarski (2003); Marx and Turner (2015); Fack and Grenet (2015); Melguizo et al. (2016).

that such a policy has extensive impacts on the population of eligible and ineligible students.

Second, our potential eligible population ranks higher on the relative ability distribution, since the financial aid programs in previous studies affect students with above-median test scores (Solis, 2017), above-median high school GPA (Bettinger et al., 2019; Scott-Clayton and Zafar, 2016), or top-quartile high school test scores in each school district (Cohodes and Goodman, 2014). In contrast, financial aid recipients in Colombia score in the 91st percentile and may enroll at top-ranked universities. Studying college choice as an outcome is particularly relevant in our context, since one of the main reasons preventing eligible students from enrolling in these top institutions is access to financial aid.[5]

Third, we estimate the effect of targeted financial aid on the entire population of high school seniors, i.e., students that *ex ante* may or may not be interested in attending postsecondary education. Our population thus differs from the self-selected subsample of high school students that express interest in postsecondary attendance, for instance, by filing a student loan application form (Bettinger et al., 2019; Angrist et al., 2014) or taking a college admission test (Solis, 2017). Because financial aid may induce high schoolers to be interested in postsecondary education, estimating the enrollment effects of aid for the entire population of high school students—including students who would be inframarginal in other studies—is of direct interest to policymakers concerned with questions of access.

Lastly, our work is related to the literature on school vouchers in secondary schooling. In particular, it contributes to research analyzing the effect voucher programs have on participants, student sorting across schools, as well as the impact that non-random migration of students from public to private schools may have on aggregate educational performance (Epple, Romano, and Urquiola, 2017; Urquiola, 2016). We contribute to this literature by offering a case of a voucher for postsecondary education that increases access to quality schooling for high-achieving students at the lower-end of the socioeconomic ladder. As such models predict, we find that school choice generates equity gains and reduces the stratified education provision wherein quality rises with socioeconomic status. But as high-achieving students sort out of no-college education and low-quality schools and enter private elite colleges, stratification by ability increases. Finally, we observe that financial aid restricting college choice to high-quality schools pressures low-quality schools to become more efficient at attracting high-ability students.

While the magnitude of our enrollment results is similar to what previous studies have found in Latin America (Solis, 2017), the response to financial aid observed in Colombia is much stronger than that documented in more developed countries. There are four main fac-

[5]Further, the fact that pre-collegiate achievement overlaps between financial aid recipients and non-recipients at high-quality colleges partly explains why we do not find evidence of mismatch (Goodman, Hurwitz, and Smith, 2017; Hoxby and Avery, 2013; Dillon and Smith, 2017), as detailed in Appendix B.3.

tors driving this difference. First, baseline levels of postsecondary attainment—even among high-achievers—are lower in Colombia relative to OECD countries (OECD and The World Bank, 2012) and the SES enrollment gradient is relatively steep.[6] Second, the program is visible, transparent, and simple; three key determinants of financial aid take-up and postsecondary enrollment (Hoxby and Turner, 2014; Bettinger et al., 2012; Dynarski and Scott-Clayton, 2013). Third, in addition to their high ability, eligible students are able to access top-ranked universities because admissions are mostly based on standardized test scores and rarely require qualitative inputs like letters of recommendation or essays. Lastly, and most importantly, Colombian students are *ex ante* severely constrained in their ability to finance their collegiate studies through credit markets. Few private HEIs offer resources to low-income high achievers, and only 11 percent of first-year undergraduate students had access to student loans before SPP. With binding credit constraints, low-income students were often squeezed out of collegiate opportunities. Together, these four factors account for the large enrollment impacts we document relative to previous studies of financial aid.

We thus interpret our findings as the result of relaxing credit constraints in a context of wide SES enrollment gaps with costly top-ranked private institutions (relative to per capita income) and a shortage of access to credit for low-income individuals. Our results suggest that, in the long run, SPP is likely to have important impacts on the education and labor markets in Colombia far beyond the ones here analyzed. Studying impacts will undoubtedly come in time when the required data becomes available.

The remainder of this paper is organized as follows. Section 3.2 provides some institutional background. Section 2.3 describes the data. Section 2.4 presents the impacts of financial aid on immediate postsecondary enrollment and college choice. Section 2.5 describes the upstream and downstream effects of financial aid on overall enrollment, student body composition, and pre-collegiate test performance. Section 2.6 concludes.

2.2 Background

Higher Education in Colombia

The postsecondary admissions process in Colombia starts with SABER 11, the national standardized high school exit exam. SABER 11 is generally analogous to the SAT in the United States, but differs in two important ways. First, SABER 11 is taken by more than 90 percent of high school seniors regardless of whether they intend on applying to a HEI. High school seniors take the exam in either the Spring or Fall semester according to their graduation date, as there are two graduating cohorts per year. Most private high school

[6] Comparisons by SEDLAC (CEDLAS and The World Bank) using household survey data show that the ratio in years of education between the top and bottom income quintiles was higher in Colombia than in any other Latin American country in 2015.

seniors take the exam in the Spring semester, while most public high school students take SABER 11 in the Fall semester. In all, 15 percent and 85 percent of students take SABER 11 in the Spring and Fall semesters, respectively.

Second, SABER 11 plays a larger role in admissions in Colombia than the SAT does in the United States. College applications and admissions are decentralized, and there is no central agency to process applications. While colleges decide and apply their own admission criteria and processes, nearly four-fifths of them use SABER 11 as an admission criterion (OECD and The World Bank, 2012). In fact, many schools award admission offers solely based on performance in this exam. There is no common date by which all applications must be submitted, or by which colleges send acceptance letters. Students may make multiple applications and prospective students apply to a college-major pair. Because there are two graduating cohorts, colleges admit new students every semester, and have separate admission processes and distinct SABER 11 cutoffs for each cohort.

All HEIs are required to obtain the Ministry of Education's *Qualified Registry* of minimum quality standards to provide their education services. Institutions can also voluntarily apply for a certificate of *High Quality Accreditation*, awarded by the National Accreditation Council (CNA, a Spanish derived acronym). The CNA, composed of members representing the academic and scientific community, sets the quality criteria and carries out the peer review evaluation process. This process is designed to encourage continuous self-evaluation, self-regulation, and institutional/program improvement (OECD, 2016).[7] Importantly, High-Quality Accreditation proxies quality of education provision, as measured by the college exit tests, and graduates' wage profiles (Camacho, Messina, and Uribe, 2016). One in ten HEIs had received High Quality Accreditation by October 1, 2014, the day SPP was announced.[8] We henceforth refer to these institutions as "high-quality" HEIs. Among these 32 high-quality HEIs, 12 were public and 20 were private. In terms of the proportion of students, roughly one-third of students enrolled in higher education attend a high-quality HEI.

Private universities in Colombia are expensive even by international standards (OECD and The World Bank, 2012). Very few private HEIs offer resources to low-income high achievers. Private HEIs are very expensive relative to public HEIs: their tuition fees are more than eightfold those of public HEIs (MEN, 2016), with the latter being able to charge low and means-tested tuition fees thanks to heavy government subsidization. The low fee of attending public, high-quality HEIs for low-income students helps explain why these HEIs have historically been oversubscribed, making their admission processes highly competitive.

[7]The High Quality Accreditation can be program-specific (e.g., the Economics undergraduate program at University of Los Andes) and/or institutional (e.g., University of Los Andes); the latter encompassing the former. Institutional accreditation is closely related to the official university ranking made by Colombia's Ministry of Education: 19 of the top-ranked 20 HEIs have High Quality Accreditation (MEN, 2016).

[8]This number would increase to nearly 50 two years after SPP was announced (see Section 2.5).

Despite progress in the past decades, educational credit markets and financial aid mechanisms remain substantially less developed in Colombia than in the United States and other developed countries. Before SPP, only 11 percent of first-year undergraduate students had a loan from ICETEX, the public institution providing student loans in Colombia (Ferreyra et al., 2017). In contrast, 70 percent of American undergraduate students receive some form of financial aid, and 30 percent borrowed federal loans in 2017 (CollegeBoard, 2017). High direct college costs, coupled with limited mechanisms for financing the full cost of college, squeeze many low-income Colombian students out of collegiate opportunities.[9]

As access to private, high-quality HEIs largely depends on students' own financial resources, sorting into private colleges in Colombia is strongly defined by the tuition rates they charge (Riehl et al., 2016). Conditional on enrolling, low-income students sort into low-quality (public and private) HEIs, and only a minority of extremely high-performing students access the highly competitive high-quality public HEIs. In contrast, high-income students are most likely to attend private HEIs—mostly high-quality, but sometimes also low-quality HEIs—regardless of their ability (Ferreyra et al., 2017). As students sort across HEI types, postsecondary education in Colombia becomes de facto severely segregated.

Ser Pilo Paga Financial Aid Program

In light of this situation, the Santos administration created *Ser Pilo Paga* (roughly, "hard work pays off" in Spanish), a merit-based financial aid program for low-income students. Announced October 1st, 2014 and extended for a quadrennium, SPP is a publicly funded program that covers the full tuition cost of attending a four- or five-year, degree-awarding undergraduate program at any university with High Quality Accreditation in Colombia. The loans are forgivable upon graduation and can be used to study any major. The academic cost of each SPP beneficiary is transferred by the central government to the university he or she attends. In addition, SPP recipients directly receive a biannual stipend of one to four times the minimum wage, depending on whether they migrated to a different metropolitan area to attend college.[10] Between 2014 and 2018, SPP benefited some 40,000 students, or

[9] For this reason, the Gini of access to postsecondary education is higher in Colombia than in many other Latin American countries, including Argentina and Chile (Ferreyra et al., 2017). Further, while gross enrollment rates have increased from 31.6 percent in 2007 to 49.4 percent in 2015, most of this increase has been due to the expansion of low-quality programs in low-quality HEIs, with questionable impacts on academic performance and future job prospects (Camacho et al., 2016).

[10] As in the rest of Latin America and the Caribbean, students tend to attend HEIs close to home. Almost 75 percent of students in Colombia attend a HEI located in their same department (or state) of residence at the time of finishing high school (Ferreyra et al., 2017). Students also tend to live at home while enrolled in postsecondary education. In addition to the SPP stipend, SPP recipients may receive a biannual subsidy of COP \$800,000 pesos (2015 USD 1 = COP 3,174) awarded by the National Planning Department upon completion of the academic semester, and an additional COP \$200,000 per semester if their college GPA is 3.5/5.0 or above. Furthermore, SPP recipients often benefited from additional in-kind subsidies offered by receiving institutions (e.g., free photocopies, reduced lunch fees).

roughly 10,000 individuals per year. Its large scale and an immense government advertisement campaign contributed to making SPP one of the most popular social programs in the country.

To become eligible, applicants must satisfy three conditions. First, they must score at least 310/500 in SABER 11. In Fall 2014, this meant scoring in the top 9 percent of test scores (Figure 2.1, Panel A). Due to rising program take-up and binding budget constraints, the government subsequently increased this test score cutoff to 318/500 in 2015 (i.e., top 8 percent), 342/500 in 2016 (i.e., top 4 percent), and 348/500 in 2017 (i.e., top 3 percent). Second, students must come from a disadvantaged household, as measured by the main proxy-means testing instrument used by the government to target social welfare program recipients, SISBEN. The student's SISBEN wealth index must be below a cutoff that varies with geographic location: 57.21 in the 14 main metropolitan areas; 56.32 in other urban areas; and 40.75 in rural areas (Figure 2.1, Panel B).[11] Third, applicants must have been admitted by a university with High Quality Accreditation.

Two program characteristics are worth noting. First, SPP was introduced by surprise almost two months *after* some 575,000 individuals had taken SABER 11 and only a couple of weeks before most colleges' application deadlines. Students could not manipulate their test scores to become eligible for SPP, lending credence to our identifying assumption of test scores being quasi-randomly assigned close to the SPP eligibility cutoff (validated in Section 2.4). Moreover, the surprise introduction explains why program take-up among eligible students was incomplete (59.4 percent), and why the first cohort of beneficiaries was less prepared for college relative to later cohorts (DNP, CNC, and de Los-Andes, 2016). Second, the program was remarkably large in scale. A year prior to policy rollout, 102,000 Fall 2013 test-takers enrolled in a HEI immediately after taking SABER 11. Of these, roughly 30,000 enrolled in a high-quality HEI (of which 16,600 did so in a private HEI and 13,300 in a public HEI). Thus, by providing scholarship-loans to roughly 10,000 students, the program aimed to benefit one in every three students attending a high-quality HEI in Colombia.

Finally, while SPP loans are forgivable upon graduation, students who drop out must repay the loan. A student is considered to have dropped out if he or she does not attend a high-quality HEI for three or more consecutive semesters. Data from ICETEX—the institution in charge of repayment schemes for SPP—indicate that, as of June 2018, 743

[11] SISBEN uses data from a proxy-means survey to assign households a single and continuous score from 0 to 100 (poorest to richest) based on housing quality, possession of durables, public utility services, and human capital endowments, among others. SPP's SISBEN cutoffs coincide with eligibility cutoffs of other social programs, such as the conditional cash transfer program "Familias en Acción" and humanitarian aid for victims of Colombia's armed conflict. Figure A.1 plots the distribution of test-takers in Fall 2014 by SISBEN-eligibility status. The figure shows that while the subsample of test-takers coming from disadvantaged households have a lower performance relative to the overall population of test-takers, a significant fraction of them score above the SABER 11 eligibility cutoff.

beneficiaries (1.9 percent of all beneficiaries) from the first three cohorts had dropped out from the program. This is less than one-tenth the average dropout rate among comparable college students (see Appendix B.3). The dropout rate is the highest among the first cohort (467 of 8,971 recipients enrolling in Spring 2015, i.e., 5.2 percent), partly because they began postsecondary schooling a longer time ago and partly because this cohort was the least prepared for college. On average, monthly loan repayments were around USD 80, for a total of USD 2 million (5.9 billion Colombian pesos) owed, i.e., USD 2,730 per dropout. 75 percent of this total amount was owed by the first cohort of SPP beneficiaries. By June 2018, 80 of the 743 dropouts had already paid back the loan, and the rest were in the process of loan repayment.

2.3 Data

We use administrative data from six main sources as well as survey data specially collected for the impact evaluation of SPP. First, we use data from the *Instituto Colombiano para el Fomento de la Educación Superior* (ICFES), the institution in charge of standardized testing in Colombia. It contains test scores and sociodemographic information (e.g., socioeconomic stratum, parental education, municipality of residence) for all SABER 11 test takers in the Spring and Fall semesters of 2011 through 2016.[12] ICFES data is then merged with data from the Department of National Planning (henceforth, DNP), which contains SISBEN scores. Together, these two sources allow the identification of the eligible population, i.e., students with a test score above the cutoff and wealth index below the geographic thresholds.

Third, we use the Ministry of Education's *Sistema para la Prevención de la Deserción en la Educación Superior* (SPADIES), which tracks students along the postsecondary education system. We use SPADIES data from Fall 2011 to Spring 2016, which provides a wealth of individual-by-semester level information on student observable characteristics, including enrollment status, HEI, major of study, share of courses passed, and graduation or dropout status. This data covers roughly 90 percent of all postsecondary enrollees; information from a handful of institutions is omitted due to poor or inconsistent reporting.

Fourth, we use data from ICETEX, the institution that manages national and international scholarships and grants for post-baccalaureate programs—including SPP—on behalf of public and private organizations. This data allows us to identify SPP beneficiaries and, in case of dropout from SPP, also observe their loan repayment behavior.

The above sources of data represent census information for the population of interest—SABER 11 test takers, SISBEN-eligible households, postsecondary students, and SPP bene-

[12] *Estratos* (strata) from the Colombian socioeconomic stratification system classify housing according to its physical characteristics and environment. Dwellings are classified into one of six strata, with strata 1 being the poorest and 6 the wealthiest.

ficiaries, respectively. Our main RD analysis focuses on Fall 2014 test takers—the first cohort of SPP recipients—for reasons detailed below. We thus use information from almost 575,000 students, out of which 9,166 are SPP beneficiaries (Appendix B.4 extends the analysis to include the second cohort of SPP). Moreover, to understand the drivers of institutional choice, we use survey data collected from a representative sample of 1,479 low-income, high-achieving high school seniors who took SABER 11 in Fall 2015. The data was collected considering a RD design using the second cohort of SPP recipients, surveying SISBEN-eligible individuals who scored above and below the SABER 11 eligibility threshold.

The analysis related to the upstream and downstream consequences of SPP uses the aforementioned census information of all high school exit test-takers and postsecondary enrollees in the years 2011–2016. This comprises almost 4 million individuals. Moreover, to explore the impact of SPP on college application and admissions processes, we use administrative data on application and admissions directly collected by us from a sample of High Quality Accredited universities in Colombia.

Fifth, we use earnings records for formal sector workers during 2008-2013 from the Ministry of Education's *Observatorio Laboral para la Educación* (OLE). OLE uses data from the Ministry of Social Protections PILA database on contributions to pension and health insurance funds, and includes data on monthly wages, employment status, and four-digit economic activity codes for nearly all college graduates between 2001-2013 in Colombia. Lastly, we use information from balance sheets and financial accounts provided by HEIs to the Ministry of Education. This data includes annual information regarding, for instance, spending per student and research spending per faculty member. We also use data on requests made by HEIs to receive High Quality Accreditation from the Ministry of Education, as well as Accreditations awarded by this Ministry across time.

2.4 Direct Impacts of Financial Aid

RD Design and Validity

To estimate the causal impact of SPP on postsecondary enrollment, we exploit the SABER 11 and SISBEN cutoffs using a RD design. Let $Z_i = 1(R_i > k)$ be an indicator for SPP eligibility, where k is the point of a discontinuous assignment rule (e.g., SABER 11 score, SISBEN). Note that being SISBEN- and SABER 11-eligible are necessary but not sufficient conditions to receive SPP financial aid. The third eligibility requirement—admission at a high-quality HEI, which we do not directly observe—requires the student apply and be granted admission by such an institution. Students must provide the government proof that all three conditions have been satisfied to receive SPP.

Denote D_i as an indicator for whether an individual is a beneficiary of SPP. Since receiving SPP depends on both the SISBEN wealth index and the SABER 11 test score, this multidimensional RD setting can separately identify two types of compliers: (1) need-eligible students around the test score cutoff, and (2) merit-eligible students around the need cutoff (see Figure 2.2 for an illustration). There are many strategies for dealing with multidimensional regression discontinuities, as discussed by Wong, Steiner, and Cook (2013). In the economics of education, recent examples include Cohodes and Goodman (2014) and Bettinger et al. (2019). We follow previous studies and report estimates separately, collapsing the discontinuity into a single dimension for each student by defining the distance of SABER 11 (SISBEN) scores from the eligibility cutoff, given SISBEN- (SABER 11-) eligibility status.[13] We argue that the resulting "frontier" specific effects from this univariate approach are the preferred causal estimand over the frontier average treatment effect, i.e., a weighted average of the two univariate RD effects. This is because the running variables are neither in the same metric nor in the same content area. More importantly, the discontinuities represent different populations, and the heterogeneity in estimated impacts across these frontiers is informative.

We focus our analysis on the first cohort of SPP, that is, students who graduated from high school late 2014 and began college early 2015. The first cohort guarantees the highest internal validity: students were informed about the financial aid program *after* they had taken the SABER 11 exam. This eliminates concerns regarding non-random sorting across the eligibility cutoff. In contrast, later cohorts may react to the program by, for instance, exerting more effort in the high school exit examination or requesting an evaluation from local authorities to be included in SISBEN. Indeed, anecdotal evidence suggests there has been a significant rise in the number of households requesting SISBEN evaluations since 2015. We further avoid pooling different SPP cohorts because public knowledge of the program has increased over time. As a result, the program's take-up rate—the share of eligible students who become SPP beneficiaries—is higher in the most recent cohorts (see DNP et al., 2016). This raises concern about endogeneity due to time-varying unobservables and complicates pooling different cohorts.

The three key assumptions for the validity of the RD design are the following: (i) there is no evidence of manipulation in assignment to treatment near the discontinuity; (ii) any observed differences in the neighborhood of the discontinuity occur only as a result of the differences in the running variables; and (iii) the predicted discontinuity creates a large change in assignment to treatment as a function of the running variable. We address each of these assumptions in turn.

[13] We use data-driven (that is, fully automatic) local-polynomial-based robust inference procedures through "rdrobust". This command implements the bias-corrected inference procedure proposed by Cattaneo et al. (2014), which is robust to "large" bandwidth choices. It also offers robust bias-corrected confidence intervals for average treatment effects at the cutoff (Imbens and Kalyanaraman, 2012; Cattaneo et al., 2014).

First, an assumption often employed in RD is that there is no selective sorting across the treatment threshold. The skeptical econometrician might fear students control their test scores and/or wealth index and behave strategically so as to ensure that they are just above or below the eligibility thresholds, thereby violating this assumption. We argue against this concern in the case of SPP, as there is little scope for manipulation of these variables. As previously mentioned, SPP was announced on October 1, 2014, almost two months *after* students took SABER 11. Thus, students did not know the eligibility cutoff—nor in fact the mere existence of this financial aid program—at the moment they sat for the exam. Once the program and eligibility conditions were announced, students could not go back in time and re-take SABER 11 to become eligible. Similar arguments can be made for the SISBEN score: household SISBEN scores were assigned well before the program was announced. Even though SISBEN reclassifications are possible, neither students nor their families had the time to ask for re-assessments before the SPP application deadline in November 2014. Moreover, SABER 11 exams are centrally scored and raw scores transformed into scaled scores via an algorithm unknown to students, their families, or teachers. Finally, ICFES reformed SABER 11 in Fall 2014 such that students taking the exam that semester were not familiar with the scoring mechanism *ex ante*. We thus conclude that our identifying assumption that the running variables are quasi-randomly assigned close to the SPP eligibility cutoffs is a plausible one.

To formally test for manipulation of the running variable, we use the recent local polynomial density estimator proposed by Cattaneo, Jansson, and Ma (2016b,a).[14] The resulting robust-corrected p-values are 0.218 with SABER 11 as R_i, and 0.436 with SISBEN as R_i (see Table A.1 in the Appendix). This confirms there is no statistical evidence of systematic manipulation of the running variable.

Second, we expect the behavior of individuals to be correlated with Z_i only because of its correlation with D_i (i.e., the exclusion restriction necessary for Z_i to be a valid instrument for D_i). We test this possibility by examining whether observable covariates are different on either side of the discontinuity point. The results in Appendix Table A.2 suggest there is balance in most of the 25 covariates. Specifically, using SABER 11 (SISBEN) as the running variable, there is balance in 23 (19) of the 25 baseline characteristics. In the few cases where an imbalance is detected, these differences are small in magnitude. Overall, we cannot reject the joint null hypothesis of balance in covariates around the discontinuity. As a robustness

[14] In addition, we perform McCrary tests on the sample of SABER 11-eligible test-takers in Fall 2013 (placebo) and Fall 2014 (non-placebo), using SISBEN as R_i, and perform a t-test on the differenced outcomes. The resulting t-statistic is 0.224, which suggests there is no manipulation of SISBEN as result of SPP. However, when we attempt a similar comparison among SISBEN-eligible students using SABER 11 as the running variable, the resulting t-statistic is -7.365, which would point to manipulation of test scores among the treated cohort. The histograms in Panel A of Figure A.2 show that this result is likely due to changes in score rounding rules in the most recent version of SABER 11—which incidentally began in Fall 2014. The histograms thus give us confidence in concluding there was no manipulation of SABER 11 in 2014.

check, we perform a dimension reduction exercise where we predict core outcomes using all 25 observable characteristics, and then run the RD on the predicted outcomes. The result of this exercise, displayed in Appendix Table A.3, confirms that imbalances in baseline covariates do not explain the large impacts we document on immediate postsecondary enrollment.

Figure 2.3 presents the take-up rate of SPP, that is, $E[D_i|R_i]$ against the running variable R_i, the SABER 11 score for those eligible by SISBEN (Panel A), and the SISBEN score for those eligible by SABER 11 (Panel B). The figure shows the sharp eligibility rules; since no student below the cutoff received SPP, $P(D_i = 1|Z_i = 0) = 0$ (i.e., there are no always-takers). The eligibility cutoffs increase SPP receipt by 55.4 percentage points when using SABER 11 as the running variable and by 62.3 percentage points when using SISBEN as the running variable. Thus, although there is a discontinuity in the probability of receiving SPP financial aid, the eligibility cutoffs do not deterministically predict SPP receipt because there is incomplete take-up. This one-sided non-compliance was due in large part to the short timespan between the announcement of SPP and the high-quality university application deadline (in many cases, a couple of weeks), as suggested by qualitative field evidence from DNP et al. (2016).

Having validated our RD design, we now document how the population directly affected by this policy compares to the typical high school exit test-taker in Colombia. Table 2.1 characterizes the sample population as well as the subpopulations of compliers—individuals who respond to financial aid eligibility by receiving SPP—and never-takers for each running variable (Imbens and Rubin, 1997; Abadie, 2002). The mean baseline characteristics are presented for all 574,269 test-takers in Fall 2014 (Column 1). Consistent with need-eligibility, relative to the average test-taker, SISBEN-eligible students come from larger, less educated, and poorer families (Column 2). They are also less likely to attend high school full time or attend private high schools than the average test-taker. In contrast, SISBEN-eligible students scoring close to the SABER 11 cutoff (i.e., high performers) are more likely to be younger, non-ethnic minority males from smaller, more educated, and wealthier families, and are more likely to attend a private high school full time (Column 3). Compliers are very similar in observable characteristics to other students within the bandwidth, but have somewhat more educated parents (Column 4). Relative to compliers, never-takers are more likely to be male and twice as likely to be employed at the time of taking SABER 11, i.e., presumably why they did not take up financial aid (Column 5).

A comparison of Columns (4) and (8) in Table 2.1 illustrates the differences in observable baseline covariates between the two types of compliers. Compliers using SABER 11 as the running variable are significantly poorer and have less educated parents than compliers using SISBEN as the running variable. They are also more likely to attend a public high school part time. These stark observable differences underline why studying effects separately for the two populations of compliers is important. Due to their more disadvantaged background, credit constraints are likely more severely binding for compliers using SABER 11 as the running variable. We thus expect any impact on postsecondary enrollment to be larger in magnitude when using SABER 11 rather than SISBEN as the running variable.

RD Results: Immediate Enrollment, School Quality, and Choice

Having validated our RD design, we now estimate the effect of financial aid on immediate postsecondary enrollment. Figure 2.4 plots the probability of immediate enrollment in any HEI using SABER 11 (Panel A) and SISBEN (Panel B) as the running variables. The dots are cell means, and the lines are fitted values from a regression of immediate postsecondary enrollment on the running variable estimated separately on either side of the eligibility cutoff. The figures suggest there is a discontinuous jump in the likelihood of immediate postsecondary enrollment at the eligibility cutoffs. Financial aid encouraged postsecondary attendance by subsidizing students who would have not otherwise gone to college immediately after high school.

These reduced-form estimates are large and precisely estimated, as suggested by Column (1) in Table 2.2. Using SABER 11 as the running variable (Panel A), financial aid eligibility raises enrollment by 32 percentage points; on a base of 37, this implies an 86.5 percent increase in immediate enrollment.[15] As expected, financial aid has a smaller impact at the household wealth cutoff for sufficiently high performers than at the test score cutoff for sufficiently poor students: the corresponding reduced-form estimate is 27.4 percentage points (56.5 percent). As explained above, SISBEN-eligible individuals around the test score cutoff are significantly poorer and *ex ante* less likely to attend college than SABER 11-eligible individuals around the household wealth index: the control means are 37 percent versus 48.5 percent, respectively. As a result, financial aid has the largest enrollment impacts among low-income students whose test scores render them barely eligible for financial aid.[16]

The estimates in Table 2.2 could be confounded by an effect of passing the SABER 11 or SISBEN threshold which is not purely due to access to financial aid. This would be the case, for instance, if students were more likely to be offered college admission if they scored above the SABER 11 threshold, or if the SISBEN cutoff were correlated with other factors affecting college entry (e.g., other transfer programs). A way to detect this kind of bias is to perform a placebo test by running the same regression in the pre-period sample. Intuitively, for the RD design to identify the causal impact of financial aid on the outcomes of interest, then the running variables R_i cannot affect the outcome of interest in the absence of SPP. We test for a discontinuity in immediate postsecondary enrollment around the equivalent

[15] Table A.4 presents equivalent results, including all 25 baseline characteristics, and compares estimates with and without these controls. The inclusion of baseline covariates does not significantly affect the magnitude nor the significance of the RD results.

[16] We focus on the reduced-form estimates, which capture the effect of financial aid eligibility on the outcomes of interest. Insofar as need- and merit-eligibility represent financial aid offers (contingent on aid being used to study at a high-quality HEI), this answers the relevant policy question.

SABER 11 and SISBEN eligibility thresholds among students that took SABER 11 in Fall 2013, a year before SPP was created (a placebo test).

Figure 2.5 overlays these probabilities. Using SABER 11 as the running variable, the regression coefficient is -0.009 (p-value is 0.327) for Fall 2013 test-takers; using SISBEN it is 0.004 (p-value is 0.667). Moreover, the difference in immediate postsecondary enrollment among barely-ineligible students before and after policy rollout is not statistically significant below the cutoffs. This, coupled with the absence of a statistically significant discontinuity at either SABER 11 or SISBEN cutoffs in Fall 2013, lend credence to the identifying assumption that the jump in immediate postsecondary enrollment is *caused* by the financial aid program.[17] Remarkably, SPP raised immediate enrollment by 25 percentage points even among extremely high-performing low-income test-takers, e.g., those scoring 2.5 standard deviations above the mean (the top 1.5 percent). This suggests that outstanding students were not attending college due to binding credit constraints, and that financial aid relaxed these constraints.

A crucial characteristic of the SPP program is that it restricts institutional choice to universities awarded High Quality Accreditation. Figure 2.6 plots enrollment in high- and low-quality HEIs by SABER 11 test score (Panels A and C, respectively) and SISBEN wealth score (Panels B and D, respectively). The corresponding RD estimates are presented in Columns (2) and (5) in Table 2.2. The results suggest that financial aid eligibility raised enrollment in high-quality HEIs by 46.5 percentage points at the test score cutoff. On a base of 10.9 percent, this implies an increase of 426.6 percent. The estimated coefficient is similarly large and significant at the SISBEN cutoff, even though the baseline is more than twice as large. Contrastingly, enrollment in low-quality HEIs dropped in similar magnitudes for both types of compliers: the reduced-form estimate is -15.4 to -12 percentage points (-57.7 to -53.3 percent). Therefore, financial aid pushed students out of no-college education and low-quality education and into high-quality education. This is one of the most important features of SPP.

The shift to colleges with High Quality Accreditation also gears students towards larger returns to schooling investment and is concomitant with access to more selective institutions, higher-quality peers, and more generous student resources. Appendix B.2 analyzes how our results compare when using six different metrics of college quality. To focus on the impact financial aid has on the quality of the institution a student attends, we restrict the sample to test-takers who enrolled in a university immediately after high school. Aid has significant

[17] Further, Figure 2.5 informs about the pre-treatment selection. Panel A suggests that, among sufficiently poor students, the likelihood of immediate college enrollment is a strictly increasing function of test scores. In fact, test-takers scoring one standard deviation above the mean in Fall 2013 are more than twice as likely to immediately enroll in any HEI than those scoring exactly at the mean. In contrast, this enrollment gradient does not exist for sufficiently high performers around the SISBEN wealth cutoff; if anything, immediate enrollment *falls* for extremely low-income students.

improvements along this intensive margin, raising peer quality (mean high school test scores), university quality (college exit test scores, graduation rate, share of faculty with a doctorate), and resources students are exposed to (spending per student, research spending per faculty member). We conclude that, since college quality causally affects earnings (Saavedra, 2009; Hoekstra, 2009; Zimmerman, 2014), financial aid has the potential to significantly promote intergenerational mobility.

As a result of the gains in postsecondary attendance among low-income, high-achieving students, the SES enrollment gradient shrank following a financial aid expansion. Figure 2.7 plots immediate enrollment probabilities for SABER 11-eligible students by socioeconomic stratum (where 1 represents the poorest households and 6 the wealthiest households), immediately before and after policy rollout. Financial aid yielded a large equity gain: enrollment for bottom-strata students increased by 46.4 percent at any HEI (Panel A) and by 182.1 percent at high-quality HEIs (Panel B). In doing so, SPP leveled access to postsecondary schooling and high-quality schools among high-performing students.

Even though SPP required students attend a high-quality university, there was no restriction on whether it be a public or private HEI. Table 2.2 further decomposes low- and high-quality enrollment by public versus private HEIs. We again analyze the results separately for each type of complier. For sufficiently low-income students around the test score cutoff, Figure 2.8 shows that the aforementioned high-quality enrollment effect operates completely through enrollment at *private* high-quality HEIs. In fact, on a base of just 3.3 percent, aid eligibility raised private high-quality enrollment more than fifteenfold. Critically, for this population of compliers, enrollment in *public* high-quality HEIs remains virtually unaffected by financial aid (see Table 2.2, Panel A, Column (4)). For high performers barely above the wealth cutoff—who are more than twice as likely to attend public, high-quality HEIs than the controls from the former comparison—enrollment in public, high-quality HEIs *decreased* by 7.9 percentage points or 40.7 percent (see Figure A.3). These results suggest SPP induced students to sort across institutions, particularly from low- to high-quality HEIs, and often from public to private HEIs.

To explore what drives students to choose private over public high-quality HEIs in Colombia, we turn to survey evidence from SISBEN-eligible students who took SABER 11 in Fall 2015 and scored slightly above or below the SABER 11 eligibility cutoff. Those attending a HEI by Spring 2016 (68 percent of surveyed students) reported the main factors driving their institutional choice among a list of alternatives. Appendix Table A.5 displays these summary statistics. The most important factor is prestige, second only to availability of preferred major.[18] Importantly, prestige, academic quality, and better job prospects are more prevalent among students attending private versus public HEIs. Indeed, graduates from

[18] This is consistent with models in which applicants have endogenous tastes for colleges with good reputation, i.e., those with high-ability peers, because firms set their wages by inferring skill levels from the reputation of the college attended (see MacLeod and Urquiola, 2015). Indeed, evidence from Colombia suggests college

top private schools enjoy a wage premium over top public schools, even when controlling for individual-level characteristics (e.g., SABER 11 score, household SES) and college-level characteristics (see Riehl et al., 2016, and Appendix Section B.8). In contrast, affordability is one the most attractive feature of public HEIs, confirming tuition fees are a key determinant of student sorting across schools. In sum, survey evidence suggests that the higher demand for private postsecondary education is a response to the perception that private HEIs are more reputable and produce greater value added—broadly defined—for students.

In addition to the answers provided by survey respondents, there are other reasons why students may prefer to attend private over public universities. First, for a student on the undecided between public or private university, she may select the school that offers the highest subsidy in price value, i.e., the private school. Second, public HEIs often require applicants to sit for their own competitive entrance exam. In 2017, UNAL, Colombia's flagship public university, charged applicants 98,000 Colombian pesos (34 USD) for this exam. For applicants on the margin of attending a public versus private HEI, the investment required to prepare for this entrance exam may be deemed too costly. Third, test performance among SPP beneficiaries may be high enough to be admitted at private but not public HEIs. However, the bottom 5 percent of UNAL's Fall 2013 entering cohort (i.e., pre-SPP) had SABER 11 scores *below* the equivalent SPP cutoff. *Ceteris paribus*, if admission at public colleges were based solely on SABER 11, SPP beneficiaries would have been admitted at UNAL—and arguably any other top-ranked public school—had they chosen to apply.[19]

Before we turn to the broader effects of financial aid on secondary and postsecondary education, we briefly summarize our findings on medium-term postsecondary enrollment and persistence 1.5 years after taking the high school exit exam, analyzed in Appendix B.3.[20] First, financial aid significantly improved any enrollment within 1.5 years. While this outcome remains the same for aid-ineligible students before and after SPP, it jumps dramatically for students above the aid eligibility cutoffs. However, as the probability of any enrollment of aid-ineligible students increases over time (although not more so than they would have in the absence of SPP), the magnitude of the enrollment gains diminishes relative to the immediate enrollment results. For instance, for need-eligible students, aid eligibility raises any enrollment within 1.5 years by 19.1 percentage points. On a base of 60.9 percent, this

reputation determines initial wages as well as subsequent earnings growth (see Figure H.22 and MacLeod et al., 2017).

[19] Another explanation is the application deadline of some public HEIs—such as UNAL—expired before SPP was announced. While this deadline may have been binding for some SPP applicants, it is by no means the driver of the general gap between public and private institutions. Had potential SPP applicants wanted to attend UNAL, the number of UNAL applicants would have increased in Spring 2016 relative to Spring 2015, shrinking the gap between public and private HEIs in the second cohort of SPP. In practice, we observe the exact opposite: the number of UNAL applicants stayed constant between 2014 and 2016, and the public-private gap widened for the second cohort of SPP students (see Appendix B.4).

[20] With the first SPP beneficiaries being scheduled to graduate starting 2019, longer-term outcomes such as overall dropout rate and completion will only become observable in our data around 2021–2022.

implies a 31 percent increase in any enrollment within 1.5 years, or roughly a third of the effect on immediate enrollment.

Second, changing the dependent variable in the RD design for being enrolled in any HEI in the Spring 2016 term —as well as the type of HEI a student attends that term—leads to a similar conclusion: the impact of financial aid is positive and significant, although diminishes relative to immediate enrollment because control students somewhat catch up over time. However, these RD results conflate persistence with the positive extensive and intensive margin results we documented above (e.g., compliers attend better-quality HEIs with lower dropout rates). We therefore complement this analysis with OLS and IV-2SLS specifications, where we restrict the sample to immediate enrollees and control for relevant individual and institutional characteristics. Our preferred specification suggests SPP increased medium-term persistence by 15.8 percent. Partly by shifting students towards high-quality colleges with better graduation rates and more resources, and partly by requiring beneficiaries pay back the loan if dropping out, financial aid improved college persistence for these low-income high-achievers.

2.5 Upstream and Downstream Impacts of Financial Aid

Financial Aid Raised Overall Immediate Enrollment

The significant enrollment gains from financial aid might have little impact on overall college enrollment if aid recipients simply displaced non-recipients from high-quality HEIs (i.e., a zero-sum admission game). Alternatively, postsecondary attendance might also increase among students *ineligible* for aid if college demand and supply are responding to the large aid expansion, thus producing net social gains. To study enrollment changes across time among *all* high school students—both eligible and ineligible for aid—we move away from the cross-sectional RD design and instead use a difference-in-differences design that compares outcomes across all high school exit test-takers between 2011 and 2015 (roughly 3 million students).

Recall from Section 2.2 that high school seniors take SABER 11 in either the Spring or the Fall semester according to their graduation date, as there are two graduating cohorts per year. Immediate postsecondary enrollment decisions for each graduating cohort occur at different moments in a year and relatively independently from one another. Colleges admit new students every semester and have separate admission processes and distinct SABER 11 cutoffs for each cohort. This, coupled with the fact that SPP eligibility was based on Fall—not Spring—test performance, indicates Spring test-takers may serve as an adequate control group in a difference-in-differences model.

Our baseline empirical strategy compares enrollment outcomes between Spring (control) and Fall (treatment) test-takers across time:

$$y_{itm} = \alpha_0 + \alpha_1 1(\text{Fall test-taker})_i + \delta_t + \sum_{k \neq 2013} \beta_k 1(\text{Fall test-taker})_i \times \delta_t + \gamma_m + \epsilon_{itm} \quad (2.1)$$

where y_{itm} is outcome y for test-taker i in year t and municipality m, $1(\text{Fall test-taker})_i$ is an indicator for taking the SABER 11 exam in the Fall semester, δ_t are calendar year fixed effects, γ_m are municipality fixed effects, and ϵ_{itm} is the individual-specific error term. We normalize the results with respect to 2013, the year immediately before financial aid expanded.[21] We plot the event study coefficients β_k in the figures and the standard difference-in-differences coefficients in Table A.6, which summarize the difference in outcomes between the treatment group relative to the control group, before and after SPP.

Naturally, Spring test-takers will be different in observable and unobservable characteristics from Fall test-takers. For instance, Spring test-takers are *ex ante* more likely to graduate from private high schools, have wealthier families, more educated parents, and attend at least *some* college. However, the identifying assumption underlying our design is that the *trends* in the two groups evolve parallelly before policy rollout. Because we observe three periods before SPP (2011–2013), we can assess the validity of this assumption by plotting the difference in outcomes between treatment and control groups across time.

Were Spring test-takers somehow affected by financial aid, this could potentially threaten our identification strategy. For instance, if high school students can decide their graduation semester and/or when to take the SABER 11 exam in response to SPP, then control students might self-select into treatment. Yet, we do not find evidence consistent with strategic changes in test-taking behavior driving our main results. First, SPP was announced *after* high school seniors took SABER 11 in 2014, thus ruling out the hypothesis that the large impacts we document taking place that year are driven by changes in the composition of test-takers; if these exist, they would affect 2015 test-takers only. Second, graduation semester is fixed within school. Spring 2015 students could transfer to a Fall graduation school or, what is more plausible, re-take the SABER 11 exam in Fall 2015 and apply with those new scores. However, we find that the number of Fall test-takers remained stable before and after 2015.[22]

[21] We restrict our estimation sample to test-takers aged 14 to 23, as they are more likely to be high school seniors at the time of taking SABER 11.

[22] A separate concern is that HEIs may re-allocate spaces from Spring to Fall following scholarship rollout, thus leading to an upward bias because treatment is negatively affecting the Spring graduation cohorts (we thank an anonymous referee for pointing out this potential concern). However, as detailed in Appendix B.6, we do not find evidence of this: the number of Spring test-takers accessing HEIs immediately after high school remained stable after SPP.

In what follows, we present results separately for low- and high-income test-takers. Because we do not have SISBEN scores for all cohorts, we use socioeconomic stratum—a variable reported by virtually all test-takers—as a proxy for SES. We henceforth refer to "low-income" students as those from strata 1–3 and to "high-income" students as those from strata 4–6. Figure 2.9 plots immediate enrollment probabilities separately for low- and high-income students by their relative test performance using specification (2.1). Panel A confirms the findings from the RD design, namely that, thanks to SPP, low-income students scoring in the top decile are significantly more likely to enroll in high-quality HEIs (blue curve). However, enrollment appears to have also increased among lower-performing, low-income students who, by virtue of scoring below the top decile, are *ineligible* for SPP (gray and black curves). While these results are an order of magnitude smaller than those documented for the top decile, they are positive and statistically significant. In contrast, Panel B shows that, for high-income students, enrollment remained stable across time.[23]

This positive spillover effect of financial aid on low-income, aid-*ineligible* students could be a result of demand- and/or supply-side responses to aid. On the demand side, the advertising push associated with SPP could spill over into aid ineligible populations. For instance, it could increase the option value of applying to colleges, with students sending their applications before determining whether they are eligible for aid. It could also have increased the perceived benefits of attending college. Peers of eligible students may also feel encouraged to attend college. On the supply side, colleges (especially high-quality private HEIs) may enlarge their cohorts in response to any higher demand. Furthermore, low-quality HEIs might fill the empty seats left by aid-eligible students with the next-best applicants.

To explore demand- and supply-side responses to financial aid, we complement our post-secondary enrollment data with college admission records. Our results, summarized in Appendix B.6, suggest that, first, the demand for high-quality education significantly increased following the expansion of financial aid. Specifically, the number of undergraduate applications received by *private* high-quality HEIs skyrocketed, even doubling at some of these institutions two years after policy rollout. In contrast, the demand for low-quality HEIs and *public* high-quality HEIs appears to be largely unaffected by financial aid. Second, in response to the heightened demand, private high-quality HEIs expanded their supply: cohort size increased at these institutions and, as a result, so did overall enrollment. The effect is smaller and not statistically significant for comparable public HEIs.

Figure 2.10 further decomposes the positive enrollment effects among low-income stu-

[23] In fact, the black curve in Figure 2.9 suggests some high-income students may have been temporarily displaced from HEIs the year financial aid was introduced. Figure A.4 confirms this displacement took place at private high-quality HEIs. Appendix B.5 delves into this temporary displacement effect using admission records from one of the country's top-ranked private HEIs. We identify and characterize displaced applicants and track them to the colleges where they end up enrolled around the country. We find that displaced high-income applicants enrolled in lower-ranked, high-quality, private HEIs.

dents from Figure 2.9 by type of HEI. Panel A shows that enrollment at low-quality HEIs decreased for top-performing students and increased for lower-performing students (deciles 1–9). Instead, the coefficients for high-income students are close to zero and not statistically significant. This suggests low-quality HEIs—both public and private—fill the empty seats left by financial aid recipients with lower-performing low-income students. Panel B in Figure 2.10 presents the equivalent results for high-quality HEIs. Interestingly, enrollment also increased for top-decile students at public, high-quality HEIs (blue line). This suggests public, high-quality HEIs remain highly oversubscribed despite the outflow of SPP beneficiaries and, therefore, fill their empty seats with aid-ineligible but equally high-achieving applicants.[24] These results are consistent with both supply- and demand-side responses to financial aid expansion generating positive overall enrollment gains among low-income, aid-*ineligible* students.

College Student Body Composition: Student Quality and SES

As financial aid encouraged postsecondary attendance among low-income high-achievers, the college student body composition—specifically, student quality and socioeconomic status—changed drastically. To be consistent with previous analyses, we measure "student quality" as the share of new enrollees scoring in the top decile of the test score distribution and "class diversity" as the share of entering students from strata 1–3.[25] To test the effect of aid expansion on student quality and the share of low-income students, we compare these two outcomes within colleges and between treatment and control groups across time:

$$\begin{aligned} y_{jt} &= \phi_0 + \phi_1 1(\text{Spring entering class})_t + \delta_t \\ &+ \sum_{k \neq 2014} \beta_k 1(\text{Spring entering class})_t \times \delta_t + \gamma_j + e_{jt} \end{aligned} \tag{2.2}$$

where y_{jt} is the outcome of interest of HEI j and time t, 1(Spring entering class) = 1 for enrollment in the Spring term (i.e., Fall test-takers), δ_t and γ_j are calendar year and HEI fixed effects, respectively, and e is the error term.

As Section 2.5 suggested, demand for private, high-quality education expanded after SPP and, as a result, so did its supply. Yet because cohort size increased less than one-to-one with respect to applications, the admission rate of these institutions dropped, making them significantly more selective. For instance, the admission rate at the University of Los Andes, Colombia's flagship private university, dropped by one-half just two years after policy rollout

[24] In fact, Panel B in Table A.6 and Figure A.4 present the equivalent results for students from strata 4–6, and suggest public, high-quality HEIs filled their empty seats with *high-income* high achievers too. There is no statistically significant difference in low-quality enrollment before and after financial aid expanded for high-income students.

[25] Some researchers use the distribution of skill among graduates or average quality admitted students as a measure of college "reputation" (see MacLeod and Urquiola, 2015; MacLeod et al., 2017).

(see Appendix B.6). Figure 2.11 plots student quality by HEI type, using specification (2.2). Panel A shows student quality dropped at low-quality HEIs after SPP. This is consistent with low-quality HEIs filling the empty seats left by financial aid recipients with the next best, lower-performing applicants. Contrarily, student quality was virtually unaffected at public and high-quality HEIs, as these (generally oversubscribed) HEIs fill any empty seats with aid ineligible students of similar ability (see Table A.6). Finally, student quality increased at private, high-quality HEIs thanks to the influx of SPP beneficiaries.[26]

To test the effect of financial aid on class diversity, Figure 2.12 plots the β_k from specification (2.2) using the share of low-income new enrollees (i.e., strata 1–3) as the outcome variable by type of HEI. The share of low-income enrollees increased by 13.7 percentage points two years after financial aid expanded at private, high-quality HEIs. On a base of 29.9 percent in the control group prior to SPP, this represents a 46 percent increase in SES diversity at these elite institutions. The equivalent increase at private, low-quality HEIs is an order of magnitude smaller and not statistically significant. Moreover, SES diversity remained completely unaffected at public HEIs, whose student population historically is made up of students from low- and middle-income backgrounds, as suggested by the control means from Panel C in Table A.6.[27]

Together, these results provide suggestive evidence of important compositional effects due to non-random sorting of students from no postsecondary education and across college types. As private, high-quality HEIs "cream skim" the most able students away from non-education and low-quality schools, increased sorting by ability raises stratification by ability and widens the quality gap in equilibrium. It also substantially promotes socioeconomic diversity at these elite HEIs (see Londoño-Vélez, 2016).

Demand for High Quality Accreditation by Low-Quality HEIs

As with other voucher programs, institutions that are left out of the program are pressured to become more efficient (Epple et al., 2017). In the case of SPP, the requirement that students enroll at high-quality HEIs puts pressure on low-quality HEIs to receive High Quality Accreditation (but whether or not this reflects an actual improvement in the quality of education provided remains to be seen). Indeed, HEIs immediately responded to SPP's announcement by requesting High Quality Accreditation (although receiving this accreditation

[26] Figure A.5 complements these results by plotting, for each high-quality HEI, the difference in mean SABER 11 percentile among first-semester students immediately before and after SPP. The figure suggests that the magnitude of the positive ability impact is inversely proportional to the institutional ranking, as measured by the average quality of admitted students before SPP (Panel A).

[27] To further illustrate this point, Appendix Figure A.6 compares the SES distribution of entering cohorts before and after financial aid expansion in two flagship, private and public HEIs. The figure shows that the share of students from the bottom two strata *decreased* by 3.0 percentage points at a top-ranked public HEI (Panel C), while it increased by only 6.3 percentage points at another flagship public HEI (Panel D).

status did not systematically follow, as shown in Figure A.7). This suggests that SPP, by restricting college choice to high-quality HEIs, pressured low-quality HEIs to become more efficient to attract high-ability students.[28]

Improvements in Relative Pre-Collegiate Test Performance

The newfound possibility of a tuition-free postsecondary education brought about by SPP raised low-income high school students' incentives to perform well in SABER 11 (Laajaj et al., 2018). In this section, we compare relative performance in standardized testing across time by socioeconomic strata and document improvements among very low-income students after SPP was announced.

Figure 2.13 plots the percentage change in the share of students that score in the top SABER 11 percentiles by socioeconomic stratum between Fall 2012 and 2016. Panel A suggests that the share of very low-income students (strata 1 and 2) scoring in the top decile increased by 32 and 14 percent, respectively, between Fall 2012 and 2016. Moreover, low-income students crowded out their higher-income peers from the top of the distribution: the shares of top decile performers from strata 3 through 6 *decreased* by 1–2 percent.[29] Panel B, which plots the percentage change between Fall 2012 and 2016 in the share of test takers in strata 1 and 6 in the top SABER 11 percentile, shows that the improvement in test performance occurs at the very top of the test score distribution: the share of students in stratum 1 and 2 scoring in the top percentile increased during this period by 175 and 28 percent, respectively (see also Appendix Figure A.8).[30] Moreover, the magnitude of these results increases over time, as students, parents, and teachers arguably have more time to reoptimize in response to the policy change.

We interpret these findings as suggestive evidence that student effort responds to the higher incentives for performance in standardized testing provided by merit-based financial aid. This is consistent with Laajaj et al. (2018), who show that test score improvements in

[28] The expansion of financial aid could also induce other supply responses such as colleges raising tuition fees, which we explore in Appendix B.7. Even though the demand for high-quality private education expanded, we do not detect an increase in tuition fees for this type of education. We argue tuition fees are regulated in Colombia, even among private HEIs. Moreover, tuition hikes are not implemented in practice because HEIs are restricted in the amount they can raise tuition fees in real terms from year to year.

[29] Between Fall 2012 and 2016, the share of stratum 1 test-takers scoring in the top decile increased by 1 percentage point. Given only 3.1 percent of stratum 1 test-takers scored in the top decile *ex ante*, this represents a 32 percent increase in performance for the lowest-income test-takers. In contrast, the share of top decile test-takers decreased by 1.4 percentage points among stratum 6. Given 57.4 percent of stratum 6 test-takers scored in the top decile *ex ante*, this represents a 2.4 percent decrease in performance for the highest-income test-takers.

[30] The absence of bunching around the top decile cutoff could suggest students are not simply increasing their scores to barely make themselves eligible for SPP; *inter alia*, this would not be a dominant strategy because SPP's SABER 11 eligibility cutoff has increased over time (see Section 3.2).

Fall 2015 are concentrated precisely among SISBEN-eligible high school seniors. It is also consistent with previous studies documenting the extent to which scholarships and vouchers improve individual study effort in other contexts (Kremer et al., 2009; Angrist and Lavy, 2009; Angrist et al., 2002; Barrera-Osorio and Filmer, 2016). Even though we remain agnostic as to which mechanisms are behind such test score gains (e.g., improved learning, teacher effort, parental investment, test cramming), these gains may have positive medium- and long-run effects, insofar as higher SABER 11 scores have been shown to be associated with improved performance in college and better labor-market outcomes—even after controlling for baseline individual and college characteristics (MacLeod et al., 2017).

2.6 Conclusions

The Colombian higher education system was characterized by a severe segregation due to costly tuition fees and a dearth of financial support for low-income students. In this context, a large-scale, need- and merit-based financial aid program significantly improved postsecondary enrollment among low-income high-achievers. In fact, the program virtually eliminated the SES enrollment gradient among top decile test-takers. Moreover, providing beneficiaries choice over college types induces a shift from low- to high-quality HEIs and, to a lesser extent, from public to private HEIs.

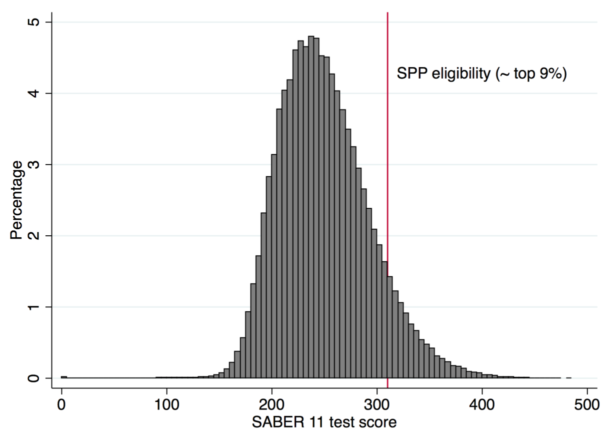
We also observe significant upstream and downstream effects of financial aid on secondary and postsecondary education which also affect students who are ineligible for aid. The program drastically changed the student body composition of colleges. It promoted class diversity at private, high-quality HEIs—institutions historically reserved for the lucky few able to afford them. As the demand for private, high-quality education expanded, financial aid raised entry competition and improved student quality at these HEIs. However, program beneficiaries did not fully crowd out high-SES students because the supply of private high-quality postsecondary schooling also expanded. Moreover, as low-quality HEIs filled their empty seats with the next-best applicants, immediate enrollment also rose for aid-ineligible students, and particularly low-income students just below the test score cutoff. Finally, the announcement of financial aid was followed by an increase in high school exit test performance by students from relatively poor backgrounds.

We posit that financial aid programs like SPP have the potential to shrink the SES enrollment gaps at selective colleges and positively impact secondary and postsecondary education systems. Ultimately, though, we care about how students' long-run outcomes are affected (e.g., college exit test scores, graduation rates, and earnings) and whether financial aid truly promotes social mobility. This also enables assessing whether such a program, which geared students towards elite colleges that are more costly to both the taxpayer and the student (if she were to drop out), is cost-beneficial. A key factor in answering these questions relates to the labor market returns for SPP participants and non-participants, given the numerous upstream and downstream impacts we documented and the potential for peer effects varying across HEI types. Given SPP participants and their college peers will only start graduating circa 2020, it remains too early to draw strong conclusions on this matter.[31] It is imperative that future research will explore these issues once data on longer-term outcomes become available.

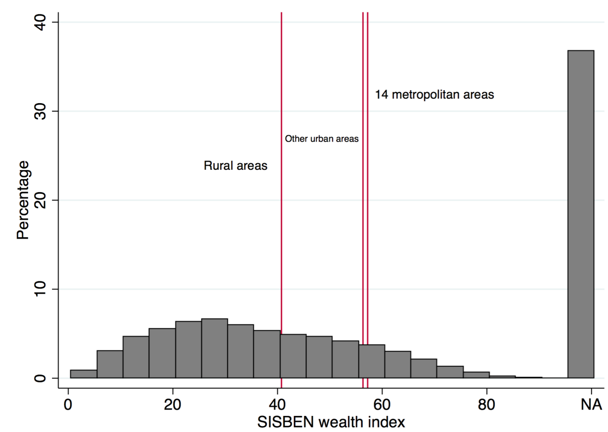
Figures and Tables

Figure 2.1: SPP Eligibility Conditions

(a) Merit: SABER 11 test score $\geq 310/500$



(b) Need: SISBEN wealth index $<$ threshold

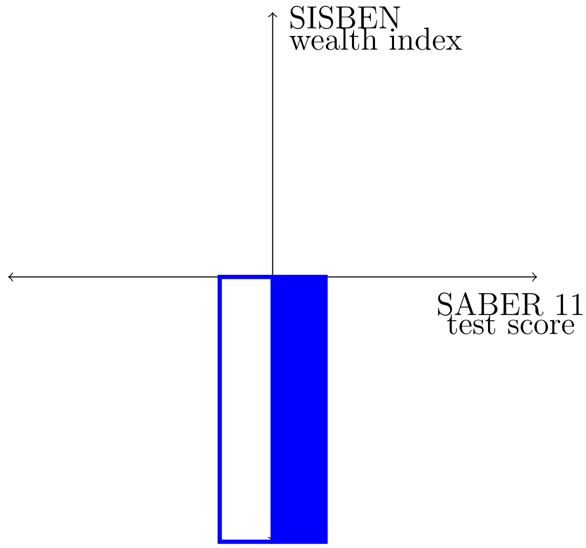


Notes: To be eligible for financial aid SPP, students must score above a cutoff in the national standardized high school exit exam, SABER 11. Their household wealth index, SISBEN, must also be below a cutoff. These figures show the distribution of SABER 11 test scores (Panel A) and SISBEN poverty index (Panel B) for Fall 2014 test-takers. The red vertical lines represent the SPP eligibility cutoffs. The figures suggest both variables are distributed smoothly around the eligibility cutoffs. Figure A.1 plots the distribution of test-takers in Fall 2014 by SISBEN-eligibility status. In Panel B, the SISBEN eligibility cutoff varies by applicant geographic location. Test-takers not included in SISBEN (e.g., individuals that do not receive welfare) do not have a SISBEN score and appear in Panel B as "N/A". *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

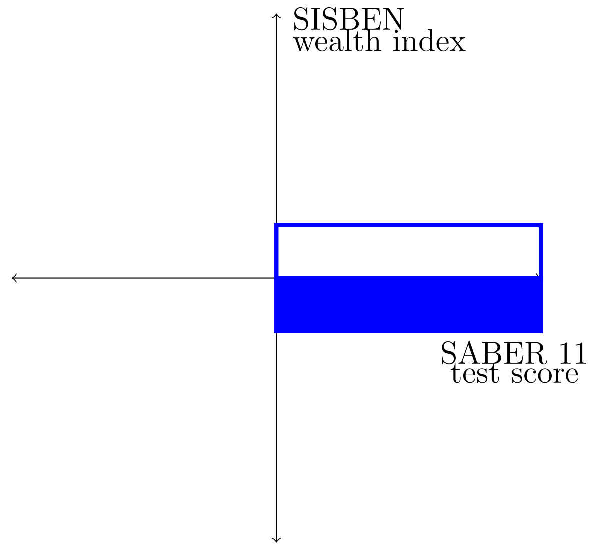
[31] Even though a first approximation of a cost-benefit analysis is presented in Appendix B.8 for the interested reader, the factors highlighted here, as well as others detailed in the Appendix, preclude us from a convincing assessment.

Figure 2.2: Illustration of the Two Types of Compliers

(a) SABER 11 as R_i



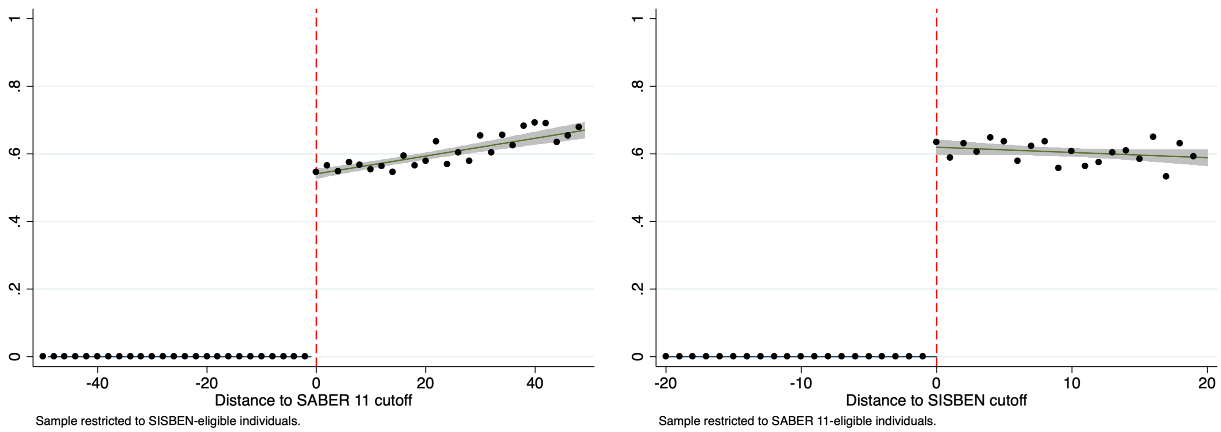
(b) SISBEN as R_i



Notes: This figure compares the two types of compliers of need- and merit-based financial aid program SPP. Panel A uses the SABER 11 test score as the running variable and compares students around the test cutoff who are SISBEN-eligible. Panel B uses the SISBEN wealth index as the running variable and compares students around the wealth cutoff who are SABER 11-eligible.

Figure 2.3: Discontinuity in the Probability of Receiving SPP Financial Aid

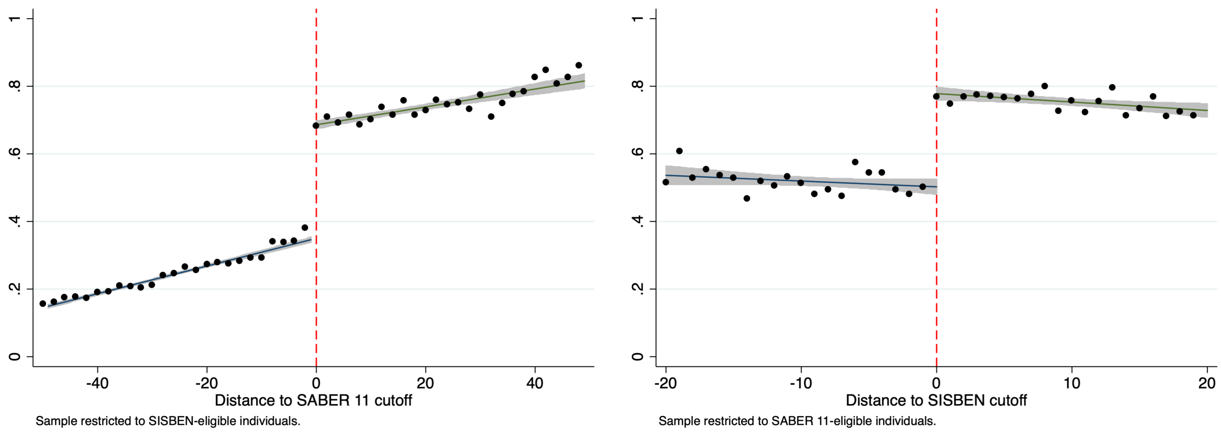
(a) R_i = SABER 11 test score (b) R_i = SISBEN wealth index



Note: The figures plot the take-up rate, that is, the probability of receiving SPP financial aid program as a function of the distance to the SABER 11 (Panel A) and SISBEN (Panel B) eligibility cutoffs, restricting the sample to SISBEN- and SABER 11-eligible students, respectively. The probability of being a SPP recipient increases from 0 percent to 55.4 percent using SABER 11 as the running variable (Panel A) and to 62.3 percent using SISBEN as the running variable (Panel B). Sample average within bin. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure 2.4: Immediate Postsecondary Enrollment

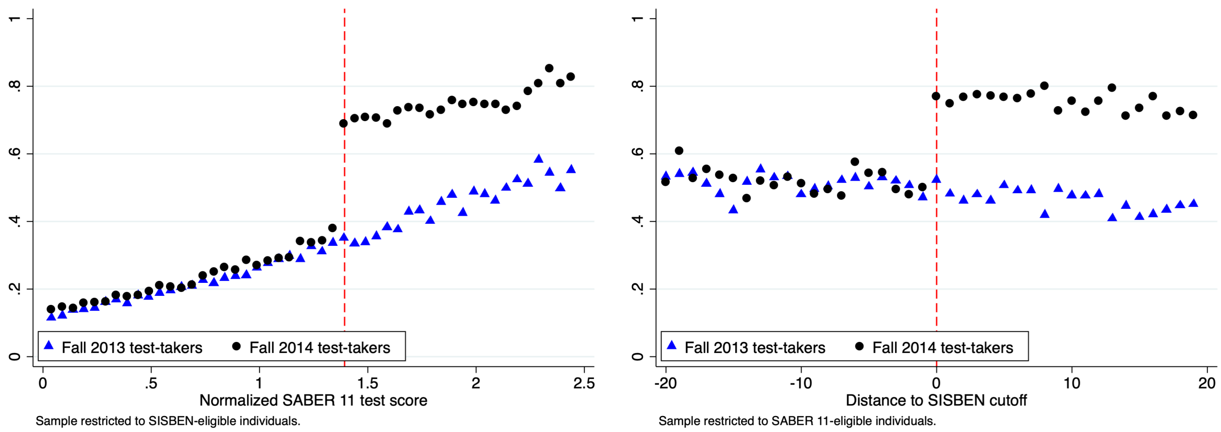
(a) R_i = SABER 11 test score (b) R_i = SISBEN wealth index



Note: The figures plot the probability of immediate enrollment in any postsecondary institution as a function of the distance to SABER 11 (Panel A) and SISBEN (Panel B) SPP eligibility cutoffs. The likelihood of immediately accessing any postsecondary institution increases by 32 percentage points (86.5 percent) using SABER 11 as the running variable (Panel A) and by 27.4 percentage points (56.5 percent) using SISBEN as the running variable (Panel B). See reduced-form estimates in Table 2.2. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure 2.5: Placebo Test using Pre-Treatment Period

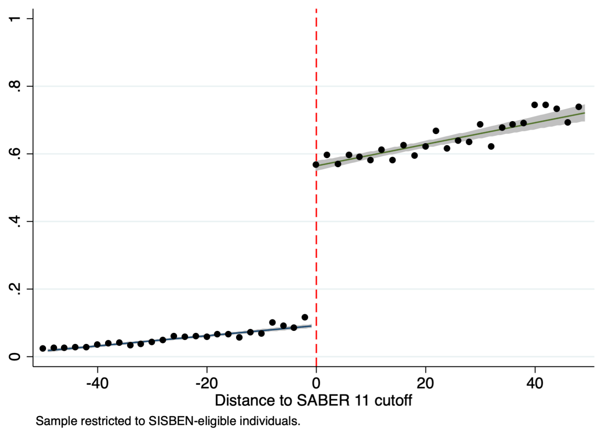
(a) R_i = SABER 11 test score (b) R_i = SISBEN wealth index



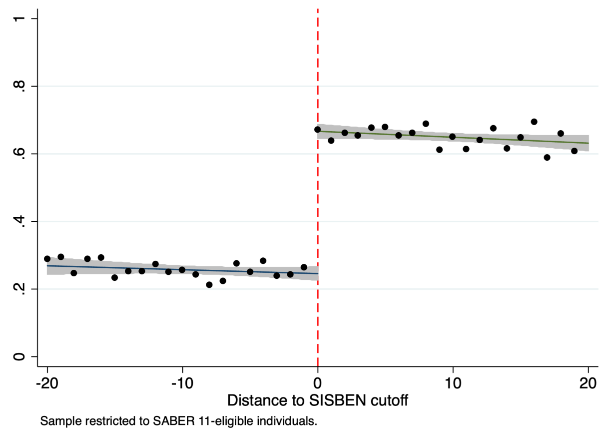
Note: These figures plot the probability of immediately accessing any postsecondary institution for test-takers in Fall 2013 (before SPP) and Fall 2014 (after SPP) as a function of the normalized SABER 11 score (Panel A) and the SISBEN (Panel B) eligibility cutoffs. The SABER 11 scoring mechanism changed in Fall 2014; for this reason, Panel A uses normalized SABER 11 score as the running variable. For Fall 2013 test-takers (placebo), the regression coefficients—estimated using package rdrobust (Cattaneo et al., 2014)—are -0.009 (robust p-value is 0.327) using SABER 11 as the running variable and 0.007 (robust p-value is 0.667) using SISBEN as the running variable. These results suggest there is no discontinuous change at the cutoffs in the likelihood of immediately attending postsecondary education in the year before SPP is implemented. Moreover, the differences in enrollment probabilities before and after policy rollout become statistically significant only above the cutoffs. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure 2.6: Immediate Postsecondary Enrollment: High- vs. Low-Quality Institutions

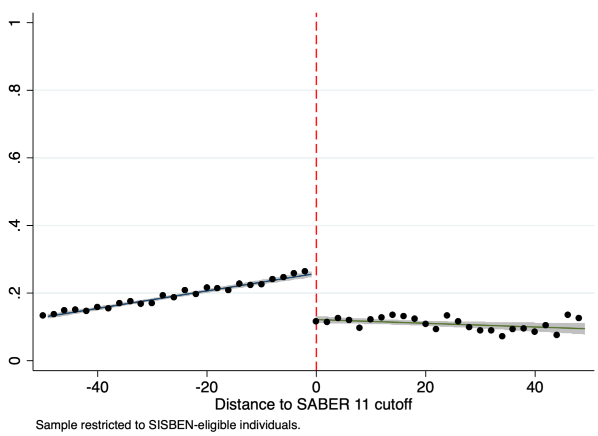
(a) R_i = SABER 11 test score: High Quality



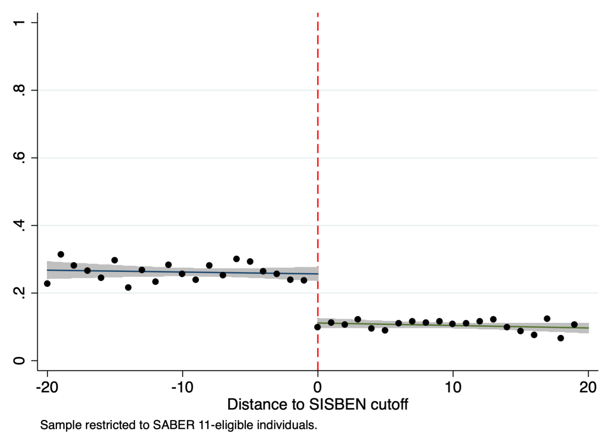
(b) R_i = SISBEN wealth index: High Quality



(c) R_i = SABER 11 test score: Low Quality



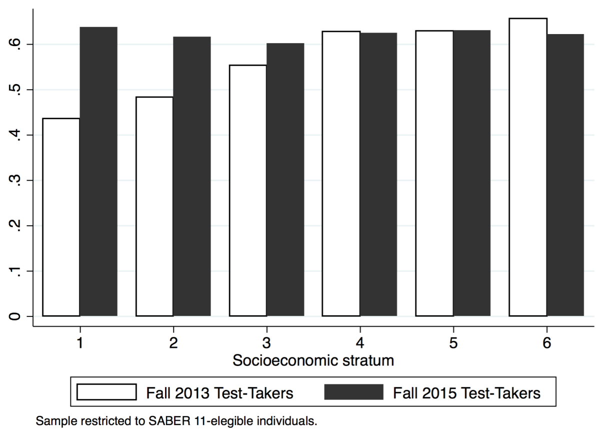
(d) R_i = SISBEN wealth index: Low Quality



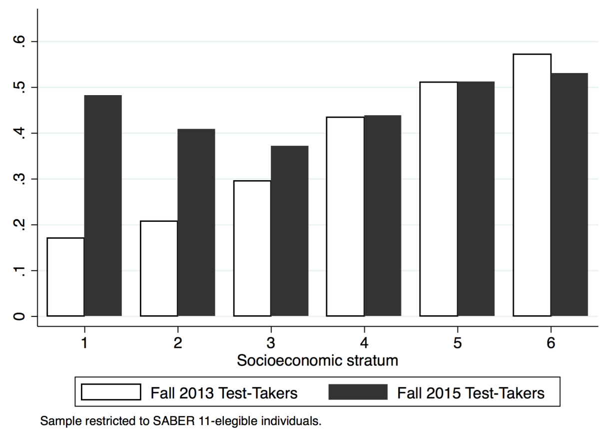
Note: The figures plot immediate enrollment probabilities by HEI quality for each running variable, SABER 11 test score and SISBEN poverty index. Panels A and B plot enrollment in high-quality HEIs as a function of the distance to the eligibility cutoffs. Panels C and D do the same for low-quality HEIs. The figures show that the likelihood of attending a high-quality HEI immediately after high school rose between 39.6 and 46.5 percentage points (152–427 percent), while the probability of attending a low-quality HEI decreased between 12 and 15.4 percentage points (53–58 percent). See reduced-form estimates in Table 2.2.
Sources: Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure 2.7: The Enrollment Gap Disappeared Among Top Students

(a) Any HEI



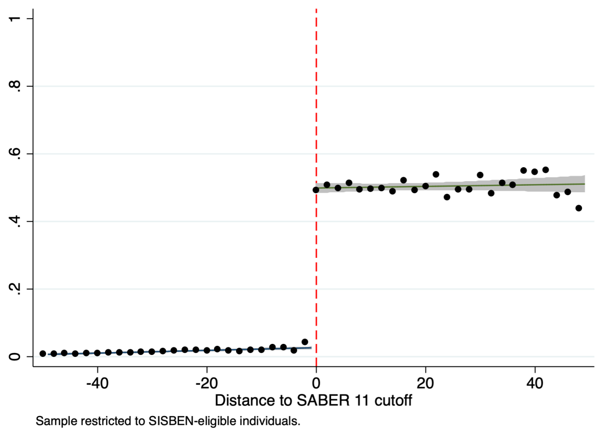
(b) High-Quality HEIs Only



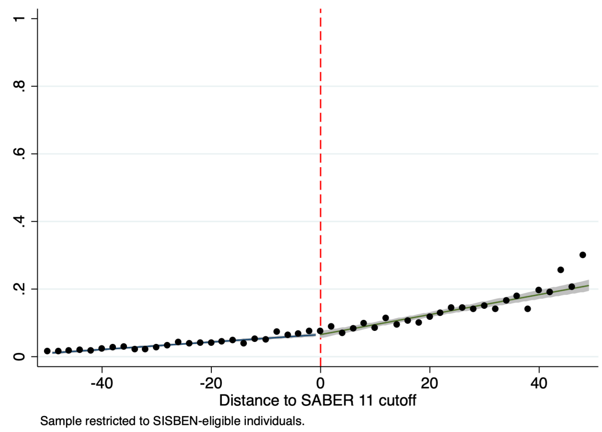
Note: The figures present the probability of immediately enrolling in any postsecondary institution (Panel A) and a high-quality institution (Panel B) among SABER 11-eligible test-takers in Fall 2013 and 2014 aged 14–23. The likelihood of immediately enrolling in postsecondary education increased from 43.5 percent to 63.7 percent for students in stratum 1, i.e., an increase of 46.4 percent (Panel A). The probability of attending a high-quality institution increased by 182.1 for students from stratum 1 (Panel B). *Sources:* Authors' calculations based on ICFES, DNP, MEN, and SPADIES (2016).

Figure 2.8: Immediate Postsecondary Enrollment: High- vs. Low-Quality, Private vs. Public Institutions (R_i = SABER 11 test score)

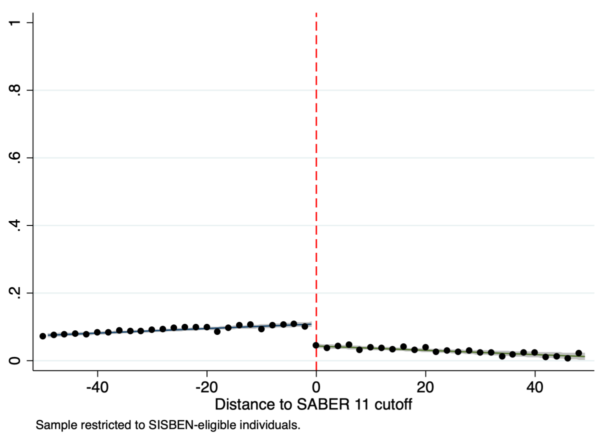
(a) Private High Quality



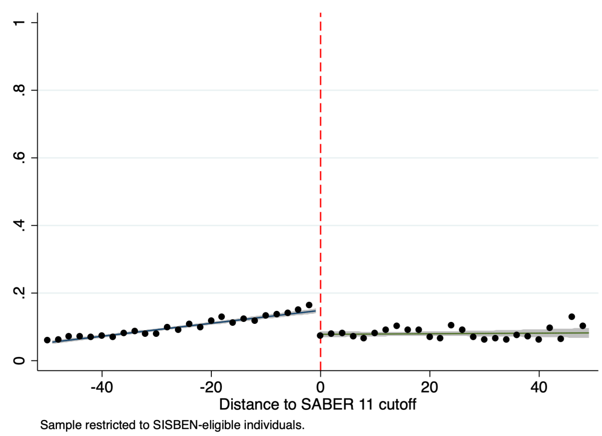
(b) Public High Quality



(c) Private Low Quality



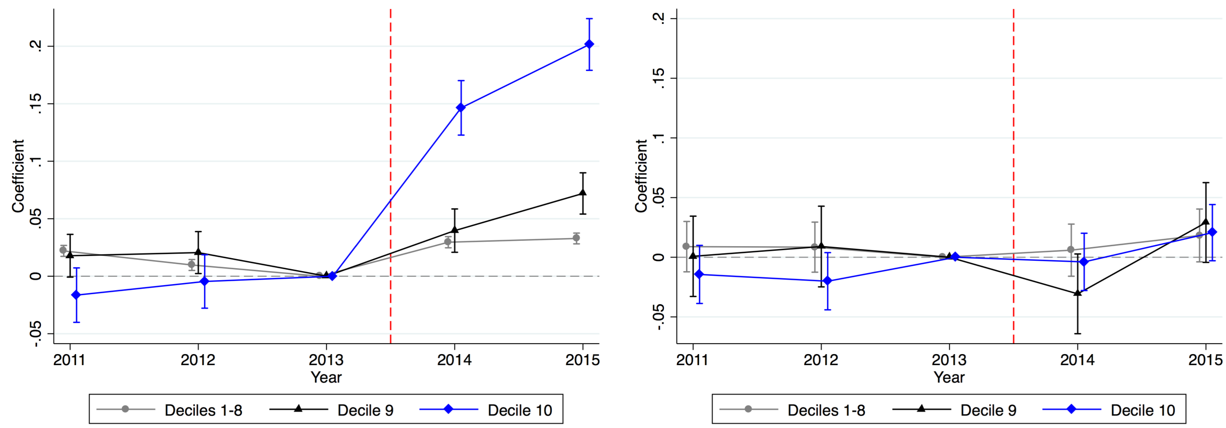
(d) Public Low Quality



Note: The figures plot the probability of immediate enrollment in a private or public, high- or low-quality postsecondary institution as a function of the distance to the SABER 11 test eligibility cutoff. The sample is restricted to SISBEN-eligible students. The likelihood of immediately attending a private, high-quality institution rose 46.6 percentage points (1412 percent), while the probability of attending a public, high-quality institution did not change. The likelihood of attending a private or public low-quality institution decreased by 6.3 and 8.7 percentage points (59 and 55 percent), respectively. See reduced-form estimates in Table 2.2. The equivalent figures using SISBEN as the running variable are displayed in Figure A.3. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure 2.9: Immediate Enrollment for Low- and High-Income Students by SABER 11 Decile

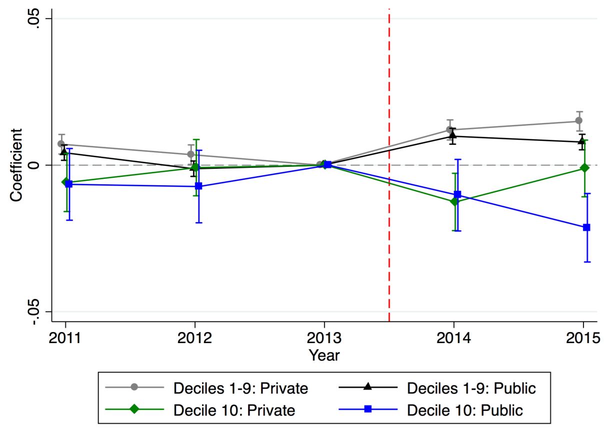
(a) Low-Income (Strata 1–3) (b) High-Income (Strata 4–6)



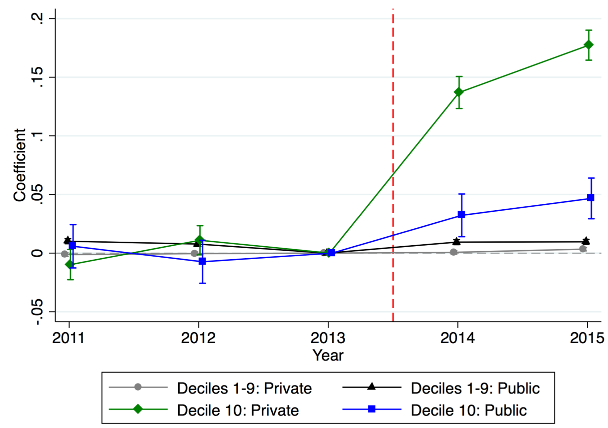
Notes: These figures plot immediate enrollment probabilities separately by socioeconomic stratum and test score performance among Fall test-takers (treatment) relative to Spring test-takers (control) before and after SPP financial aid is introduced (red vertical line) using specification (2.1). Panel A suggests that financial aid raised immediate enrollment for low-income students. This effect is strongest for top-performing students—i.e., decile 10, as these students are most likely to receive financial aid—but it is also positive and significant for lower-performing students. In contrast, Panel B shows that financial aid had little enrollment impact among high-income students, except a temporary displacement effect for decile 9. Figure A.4 confirms high-income students were temporarily displaced from private high-quality HEIs.
Sources: Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure 2.10: Low-Income Students Only: Enrollment by SABER 11 Decile and HEI Type

(a) Low Quality HEIs: Private vs Public



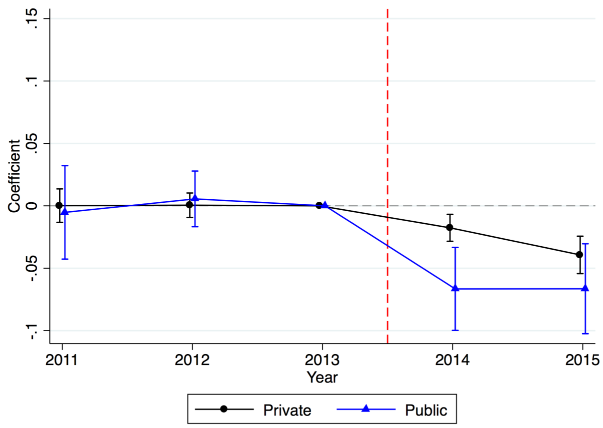
(b) High Quality HEIs: Private vs Public



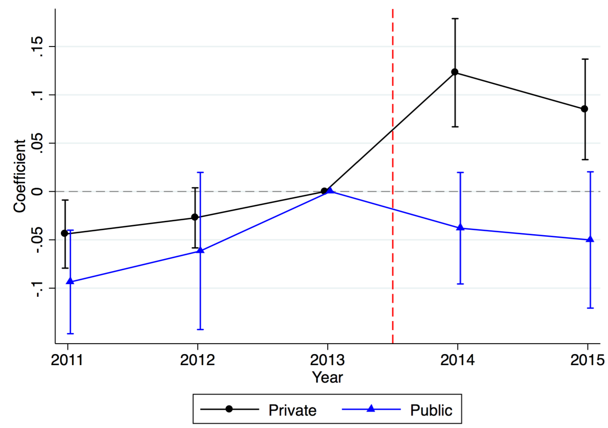
Notes: These figures plot, for low-income students (strata 1–3), the difference in immediate enrollment probabilities separately by test score decile and HEI type between Fall (treatment) and Spring (control) test-takers before and after SPP financial aid is introduced (red vertical line) using specification (2.1). Panel A suggests that while enrollment somewhat fell at low-quality HEIs for top-performers, it increased for lower-performing students (deciles 1–9). This is presumably because low-quality HEIs are filling the empty seats left by SPP beneficiaries with their next best, lower-performing applicants. Panel B confirms the RD results and shows that enrollment of low-income top-achievers at private, high-quality HEIs significantly increased two years after the expansion of financial aid. Interestingly, although SPP beneficiaries sorted out of public, high-quality HEIs, overall enrollment of top-decile test-takers increased at these HEIs, albeit by a significantly smaller amount. Figure A.4 plots the equivalent figures for high-income students. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure 2.11: Student Quality: Share of Entering Students Scoring in Top Decile by HEI Type

(a) Low Quality HEIs: Public vs Private

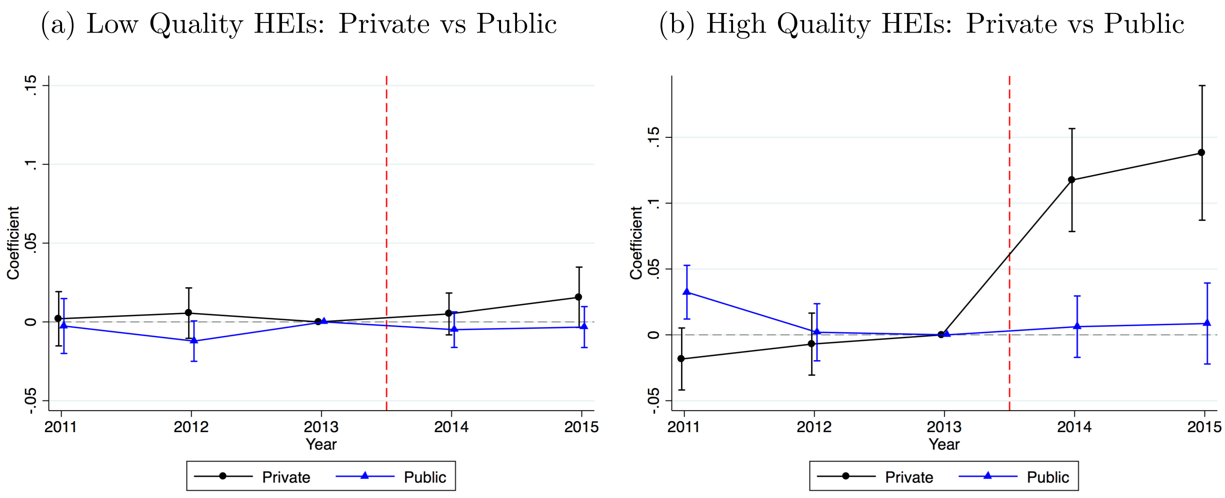


(b) High Quality HEIs: Public vs Private



Notes: These figures plot the share of new enrollees scoring in the top decile of SABER 11 test scores among Spring (treatment) relative to Fall (control) entering classes before and after SPP financial aid is introduced (red vertical line) using specification (2.2). Panel A suggests student quality dropped at low-quality HEIs, presumably because the empty seats left by financial aid recipients are being filled with lower-performing students. Panel B shows student quality was not significantly affected at public, high-quality HEIs, but increased at private, high-quality HEIs thanks to the influx of SPP beneficiaries. These results are qualitatively similar using average SABER 11 percentile as an alternative measure of student quality. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

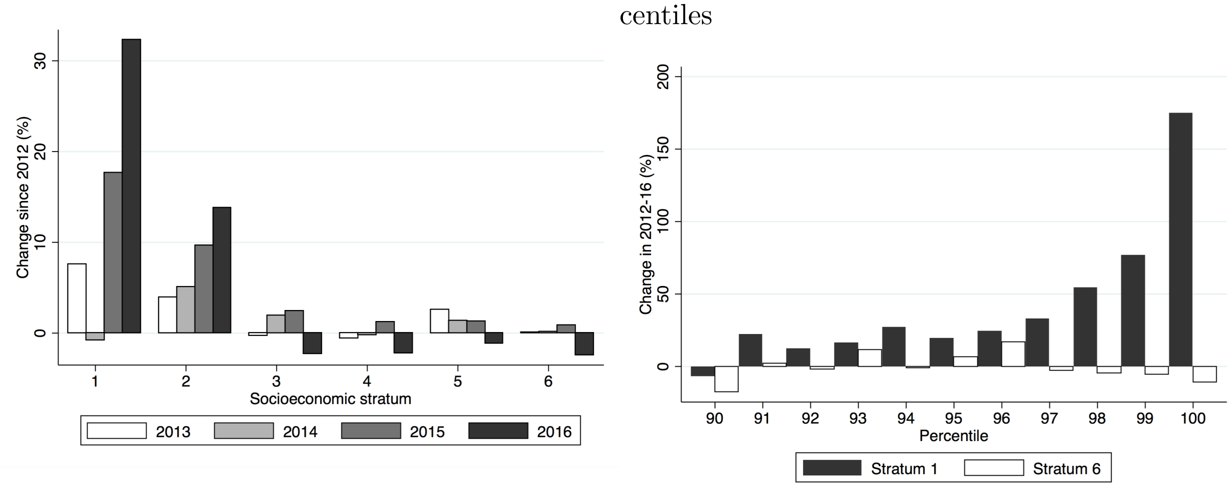
Figure 2.12: Class Diversity: Share of Entering Students from Strata 1–3 by HEI Type



Notes: These figures plot the share of entering students in the Spring (treatment) relative to Fall (control) entering classes who are low-income (i.e., strata 1–3, as opposed to strata 4–6) before and after SPP financial aid is introduced (red vertical line) using specification (2.2). While the share of low-income first-year enrollees did not change either in low-quality HEIs (Panel A) or in public, high-quality HEIs (Panel B, blue curve), this share increased by 13.7 percentage points at private, high-quality HEIs two years after financial aid expanded (Panel B, black curve). On a base of 29.9 percent, this represent a 46 percent increase in SES diversity at private, high-quality HEIs. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure 2.13: Gains in Test Performance for Low-Income Students

(a) Change in share of strata scoring in top decile (b) Change in share of strata 1 and 6 in top percentiles



Notes: These figures show improvements in SABER 11 performance among low-income students (strata 1 and 2) since the expansion of financial aid in Colombia. Panel A plots the percentage change in the share of students in each socioeconomic stratum that score in the top SABER 11 decile in Fall 2013 through 2016, using 2012 as baseline. The figure suggests the share of test-takers from the bottom stratum scoring in the top decile increased by 32 percent between 2012 and 2016. Panel B plots the percentage change between Fall 2012 and 2016 in the share of test takers in strata 1 and 6 scoring in the top SABER 11 decile, by percentile of the test score distribution. The figure shows there was an increase in the share of students in strata 1 scoring in the top decile, and particularly in the top percentile, where the increase was of 175 percent. The sample in all figures is restricted test-takers aged 14–23. *Sources:* Authors' calculations based on ICFES (2016).

Table 2.1: Characterization of Compliers and Never Takers

		SABER 11 as the Running Variable				*SISBEN as the Running Variable*			
	All	SISBEN eligible	B-width	Compliers	Never takers	SABER 11 eligible	B-width	Compliers	Never takers
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	0.547	0.575	0.485	0.463	0.448	0.448	0.434	0.459	0.432
Age	17.914	17.968	16.684	16.388	16.925	16.751	16.579	16.282	16.981
Ethnic minority	0.132	0.139	0.089	0.075	0.081	0.060	0.061	0.064	0.047
Employed	0.136	0.136	0.095	0.053	0.109	0.082	0.075	0.044	0.119
Family size	4.688	4.871	4.617	4.620	4.485	4.238	4.319	4.266	4.156
Mother ed: Primary	0.369	0.449	0.263	0.211	0.274	0.106	0.157	0.111	0.184
Mother ed: Secondary	0.431	0.436	0.489	0.482	0.509	0.336	0.445	0.414	0.398
Mother ed: T&T	0.090	0.067	0.133	0.155	0.106	0.172	0.177	0.202	0.160
Mother ed: Higher	0.110	0.048	0.114	0.150	0.110	0.386	0.220	0.274	0.259
Father ed: Primary	0.434	0.525	0.353	0.324	0.374	0.148	0.218	0.203	0.197
Father ed: Secondary	0.385	0.375	0.430	0.428	0.419	0.321	0.414	0.381	0.398
Father ed: T&T	0.072	0.051	0.101	0.116	0.089	0.140	0.151	0.186	0.143
Father ed: Higher	0.110	0.049	0.118	0.133	0.120	0.391	0.216	0.231	0.258
Hh SES: Stratum 1	0.424	0.559	0.365	0.336	0.337	0.122	0.150	0.103	0.129
Hh SES: Stratum 2	0.349	0.340	0.441	0.443	0.452	0.306	0.503	0.522	0.503
Hh SES: Stratum 3	0.177	0.093	0.180	0.206	0.191	0.349	0.306	0.344	0.330
Hh SES: Stratum 4	0.035	0.007	0.014	0.016	0.017	0.148	0.025	0.029	0.038
Hh SES: Stratum 5	0.010	0.001	0.003	0.003	0.005	0.052	0.005	0.005	-0.001
Hh SES: Stratum 6	0.004	0.000	0.000	0.000	0.002	0.023	0.000	0.000	0.001
School: Full day	0.196	0.139	0.194	0.206	0.182	0.454	0.318	0.326	0.323
School: Morning	0.513	0.548	0.606	0.607	0.621	0.428	0.510	0.482	0.526
School: Evening	0.070	0.076	0.015	0.007	0.017	0.006	0.007	0.006	0.010
School: Afternoon	0.156	0.164	0.173	0.173	0.168	0.108	0.159	0.177	0.135
School: Weekends	0.064	0.074	0.013	0.008	0.013	0.004	0.005	0.001	-0.002
Private school	0.254	0.156	0.165	0.170	0.173	0.525	0.330	0.383	0.360
N	574,269	299,475				53,632			

Notes: This table characterizes compliers from a regression discontinuity design. Column (1) presents mean baseline covariates from the universe of Fall 2014 SABER 11 test takers. Columns (2)–(5) compare characteristics using SABER 11 test score as the running variable, while Columns (6)–(9) do so using SISBEN poverty index as the running variable. Bias-corrected RD results estimated with package rdrobust (Cattaneo et al., 2014). The table does not present statistics for always takers because there was perfect left-hand side compliance, that is, there are no always takers. *Sources:* Authors' calculations based on ICFES, DNP, MEN, and SPADIES (2016).

Table 2.2: Immediate Enrollment in Postsecondary Education, by Type of Institution

	Any	High Quality			Low Quality		
		Any	Private	Public	Any	Private	Public
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: SABER 11 as the Running Variable						
RF	0.32	0.465	0.466	0	-0.154	-0.063	-0.087
	(0.012)	(0.012)	(0.011)	(0.007)	(0.011)	(0.007)	(0.009)
Mean Control	0.37	0.109	0.033	0.075	0.267	0.105	0.159
Observations	299,475	299,475	299,475	299,475	299,475	299,475	299,475
BW Loc. Poly.	29.679	24.804	23.861	32.065	23.562	30.554	25.059
Effect Obs Control	31,170	23,600	22,473	36,290	22,473	33,042	25,256
Effect Obs Treat	11,711	10,641	10,442	12,264	10,442	11,953	10,927
	Panel B: SISBEN as the Running Variable						
RF	0.274	0.396	0.477	-0.079	-0.12	-0.052	-0.076
	(0.027)	(0.024)	(0.02)	(0.018)	(0.022)	(0.015)	(0.016)
Mean Control	0.485	0.261	0.067	0.194	0.225	0.097	0.134
Observations	23,132	23,132	23,132	23,132	23,132	23,132	23,132
BW Loc. Poly.	9.028	10.954	12.224	11.604	8.05	9.176	10.329
Effect Obs Control	3,868	4,604	5,043	4,854	3,488	3,964	4,392
Effect Obs Treat	3,902	4,703	5,221	4,967	3,450	3,976	4,454

Note: This table presents the effect of financial aid eligibility on immediate postsecondary enrollment using a regression discontinuity design. The dependent variable is immediate enrollment by type of postsecondary institution (e.g., high-quality, low-quality, private, public). Panel A uses SABER 11 test score as the running variable restricts, restricting the sample to SISBEN-eligible students. Panel B uses SISBEN wealth index as the running variable, restricting the sample to SABER 11-eligible students. The reduced-form coefficient in Column (1) of Panel A suggests that, for individuals below a certain level of poverty, financial aid eligibility raises immediate postsecondary enrollment by 32.0 percentage points. On a basis of 37.0 percent, this implies a 86.5 percent increase in immediate enrollment. Bias-corrected RD results estimated with package rdrobust (Cattaneo et al., 2014). Robust standard errors in parentheses. These regressions exclude all baseline covariates. Table A.4 reproduces these reduced-form estimates and compares how the exclusion or inclusion of baseline covariates affects the reduced-form coefficient. *Sources:* Authors' calculations based on ICFES, DNP, MEN, and SPADIES (2016).

Chapter 3

Diversity and Redistributive Preferences: Evidence from a Quasi-Experiment in Colombia

3.1 Introduction

Higher education institutions are highly segregated across socioeconomic lines. This is particularly true in developing countries, where a dearth of need-based financial aid traditionally excludes low-income students from selective private institutions. Segregation in schools is undesirable because it widens the skill and wage gaps, blocks social mobility, and erodes social cohesion. In contrast, supporters of policies promoting racial and/or class diversity argue that these benefit not only disadvantaged students by leveling the playing field but also *non*-minorities. Indeed, engaging in a diverse student body has been said to enhance classroom dialogue, foster positive race relations, and generally enrich students' collegiate experience.[1] For these and other reasons, diversity, some would argue, is at the very core of the educational process, with the United States Supreme Court defending affirmative action as the primary tool to achieve diversity at selective higher education institutions.[2] Indeed, today a wide majority of Americans support affirmative action programs on college campuses, underscoring the value society places on diversity and integration in schools.[3]

[1]For instance, in Fisher v. University of Texas, the Supreme Court ruled in favor of considering race in college admissions, arguing it was "in the interest in the educational benefits that flow from a diverse student body". It reiterated that the attainment of a diverse student body "serves values beyond race alone, including enhanced classroom dialogue and the lessening of racial isolation and stereotypes." *Fisher v. University of Texas*, 2012, p. 6 (available here).

[2]Some developing countries have recently followed suit, including Brazil (whose Supreme Court imposed quotas for black, mixed-race and Amerindian students in all federal universities and technical schools in 2013) and India (the 2009 Right to Education Act reserves 25% seats to disadvantaged children in private schools).

[3]Pew Research Center (April 22, 2014) "Public strongly backs affirmative action programs on campus" (available here).

In spite of this, the identification of the causal effect of a more diverse student body on non-minorities' attitudes and academic outcomes has been an elusive feat for empiricists. First, diversity has been mainly promoted through affirmative action which, by definition, uses race (or, more recently, SES) as a factor in admissions. The advantage given to historically under-represented minorities makes it difficult for the econometrician to disentangle between the confounding effects of having 'diverse' peers versus peers of different ability. Moreover, since the extent to which race or SES is used as a factor in admissions varies considerably across universities, non-minorities can self-select into schools based on their affinity for having a more diverse group of peers.

I exploit a unique setting that allows overcoming these two identification challenges to estimate the causal effect of class diversity on perceptions of inequality, poverty and social mobility, beliefs of social justice, redistributive preferences, and academic outcomes. In an initial context of high inequality and severe de facto segregation in selective higher education institutions, a recent need-based financial aid program successfully promoted class diversity, generating a discontinuous and unprecedented jump in the presence of low-income students at elite universities in Colombia. At Bogotá's flagship selective institution, for instance, the fraction of poor students almost quintupled from 7% in 2014 to 33% in 2016. Importantly, the policy left the admissions process at these universities utterly unaffected, as low-income students were not given any preferential treatment in college admissions. This quasi-experiment provides an ideal setting to evaluate how perceptions and attitudes can be influenced by a more diverse student body.

I find that the policy was successful at promoting diversity at selective private institutions. Peer characteristics dramatically shifted, with almost one-third of freshmen being recipients of this financial aid program in Spring 2015. For rich students, one semester of exposure to poor students fostered interactions among students from heterogeneous social backgrounds. This reduced the biases in their perceptions of income distribution, increased perception of poverty rate and social mobility, raised belief in meritocracy in college admissions, and increased support for redistribution. Unlike with affirmative action, diversity promoted through need-based financial aid did not affect academic outcomes, with both dropout rates and grades remaining largely unaffected by the policy. Moreover, applications at this elite almost doubled as a result of the policy, significantly raising average ability of new cohorts and possibly inducing younger cohorts to exert more effort and devote more resources to preparing for the high school exit exam.

The results contribute to a rich literature on redistributive politics, peer effects, and diversity in schools. The literature in public finance and political economy has shown that preferences for redistribution respond to subjective perceptions about socioeconomic status (SES) (Meltzer and Richard, 1981; Cruces, Perez-Truglia, and Tetaz, 2013), social justice (Alesina and Angeletos, 2005; Alesina and Ferrara, 2005; Alesina, Glaeser, and Sacerdote, 2001), social mobility (Piketty, 1995; Benabou and Ok, 2001), the extent of inequality or poverty (Ariely and Norton, 2011; Kuziemko, Norton, Saez, and Stantcheva, 2015), and reference points (Charite, Fisman, and Kuziemko, 2015). I contribute to this literature by providing evidence on how a shift in peer characteristics significantly affects these outcomes six months after exposure to low-SES individuals. My paper is also related to the effect of integration and peer effects, and it is in line with Angrist and Lang (2004), who find no effect of Boston Metco school desegregation program on non-Metco students in host districts. Finally, my research is related to the literature studying whether interaction reduces inter-group prejudice. Rao (2013) studied affirmative action in private schools in Delhi and found that having poor classmates makes wealthy students more pro-social and generous, and less discriminatory against poor children. He also found that more diversity in the classroom negatively affected students' test scores in English (though not in Hindi or Math).

I seek to contribute to this literature by studying the impact of class diversity in a context of high inequality and severe de facto class segregation at selective private universities.[4] Moreover, I shed light on the question of whether policy can shape preferences among young adults who, unlike the population in Rao (2013), may exercise their democratic right to vote. Finally, the results derived from this policy speak more closely to the impact of need-based financial aid programs in other developing countries. This will become increasingly relevant as governments in other developing countries make more resources available to low-income students. For instance, Peru has recently implemented *Beca 18*, a need-based financial aid program that awards scholarships to 5,000 low-income high school graduates every year.[5] Similarly, in 2014 Ecuador launched its first pilot of *Política de Cuotas*, covering tuition fees for 500 low-income and under-represented minority students enrolled at one of the five participating private tertiary institutions.

The remainder of the paper is organized as follows. Section 3.2 provides some institutional background and describes the financial aid program. Section 3.3 introduces the data and describes the methodology used to identify the causal effect of diversity on perceptions and attitudes. Section 3.4 presents the results. Finally, Section 3.5 concludes.

[4]Colombia has long recorded one of the highest levels of income inequality in the world: its Gini coefficient is 0.54 and the top 1% of the population capture 21% of total income (Alvaredo and Londono-Velez, 2013). This acute income concentration translates into inequality in access to higher education, making income inequality self perpetuating (see Joumard and Londono-Velez (2013)).

[5]Unlike Colombia's SPP, Peru's *Beca 18* applicants follow a different admissions procedure than traditional applicants at selective higher education institutions. Specifically, participating institutions require *Beca 18* applicants to take an exam and write an essay to be considered for admission. Part of the reason for this is that Peru does not have a standardized high school exit exam. Moreover, once admitted, some institutions require *Beca 18* recipients to take one year worth of remedial classes.

3.2 Institutional Context: Higher Education in Colombia and *Ser Pilo Paga*

The college admission process in Colombia begins with SABER 11, the national standardized high school exit exam.[6] This exam is taken by virtually all high school seniors in the country in one of the two dates it is offered in a year. Similar in spirit to the SAT in the United States, the test is a mandatory requirement for those wishing to attend higher education and it has wide-spread use in admissions processes: 78% of Colombian tertiary institutions use SABER 11 scores as an admission criterion (OECD and The World Bank, 2012), though many institutions combine this information with other elements such as personal interviews, institution-specific exams or high school transcripts. College applications take place twice every year because there are two high school-graduating cohorts (Fall and Spring).[7] Applications to universities are decentralized and major-specific: students apply to a college-major pair.[8] In most institutions, course curriculum is relatively set within a major, leaving students considerably less freedom in choosing their courses than typical American universities. Finally, there is a mandatory field-specific college exit exam called *Saber Pro* that is required for graduation.[9]

Most top-ranked higher education institutions in the country are private and costly.[10] Indeed, the average private institution tuition fee is six-fold the size of the average public institution fee, and represents 55.4% of per capita gross national income. This is high compared to international standards (e.g., 4.9% in the United Kingdom, 32% in Chile, 42% in the US, see OECD and The World Bank (2012)). This, coupled with scarce resources available to the poor, excludes low-income students from selective private institutions in Colombia. As low-income high school graduates wishing to attend university resort to public institutions, these are over-subscribed and their admission rates are very low. In practice, this leaves the overwhelming majority of low-income students with the sole alternative of

[6] SABER 11 tests multiple subject areas including math, chemistry, physics, biology, social science, philosophy, and Spanish.

[7] In Colombia, there are two school calendars, called Calendar A and Calendar B. All official and public schools use Calendar A, as a result of which 95% of the annual 570,000 high school graduates start school in late January and end in November. Most private high schools use Calendar B, which leaves just 5% of high school graduates starting school in early September (OECD and The World Bank, 2012).

[8] It is difficult to ascertain application and admission rates accurately in Colombia because university applications are radically decentralized: not only does every institution decide and apply its own admission criteria and processes, but there is also no central oversight or collection of detailed data of the entry arrangements and criteria of each institution. Applications are sent by individual students to one or multiple higher education institutions. There is no common date by which all applications must be submitted, or by which all students will know whether they have been accepted (OECD and The World Bank, 2012).

[9] The *Saber Pro* exam was created in 2003 to measure the quality of higher education in Colombia. It was made mandatory for graduation in 2009.

[10] Specifically, 7 of the top 10 higher education institutions are private institutions, according to the ranking made by Colombia's Ministry of Education (available here).

attending medium- or low-quality tertiary institutions. The result is thus a severe de facto segregation of high-quality higher education institutions.

In light of this, the Santos administration announced the introduction of *Ser Pilo Paga* (henceforth SPP) in October 2014.[11] SPP is a need-based financial aid program that annually awards around 10,000 low-income but high-achieving students the means to attend a four-year undergraduate degree-awarding program at an accredited high-quality university in Colombia.[12] SPP scholarship-loan program covered tuition costs and provided recipients (henceforth referred to as 'Pilos') an average annual stipend of 3,523,488 pesos (US $1,475) in 2015,[13] costing the government around US $6,300 per student per year.[14] Importantly, the policy gave students free choice regarding their university and major. To become eligible, applicants must satisfy three criteria: (i) be high-achieving; (ii) be poor; and (iii) be admitted at a high-quality university. Specifically, applicants must have scored 310 or above in the SABER 11 test taken in August 2014, i.e., the top 9% of test takers (see Figure A.2). Second, they must have a national poverty SISBEN index scoring below a threshold, which varies with geographic location (see Figure A.3).[15] Lastly, they must must have received admission into one of the accredited high-quality universities in the country (in October 2014, there were 33 such institutions).

Granted, SPP is not the first need-based financial aid program to exist in Colombia. ACCES credit program, administered by the national student loan-granting institution ICETEX, has offered college loans to low-income students since 2003 (see Sanchez and Velasco (2014) for a review).[16] However, ACCES remains a small loan program, covering only about 10% of high school graduates from strata 1 and 2 (Melguizo et al., 2016). This explains why less than 10% of all students benefit from a loan in Colombia, compared to 35% and 43% in US and Canada, respectively (OECD and The World Bank, 2012). Moreover, while a handful of private institutions offer some resources to low-SES students, their capacity to satisfy the increasing demand for higher education remains largely limited. For instance, the

[11] "Pilo" is a Colombian colloquialism used to describe a person that is hard-working.

[12] The "high-quality" accreditation is awarded by the Ministry of Education.

[13] The size of the annual maintenance subsidy offered by the Ministry of Education varies between $400 and $1,600 depending on whether the recipient migrated to a different municipality to attend college. In addition to this subsidy, Pilos may receive a biannual subsidy of COP $800,000 pesos (US $320) awarded by the National Planning Department upon completion of the academic semester. Furthermore, recipients benefited by additional resources offered by receiving institutions (e.g., free photocopies, reduced lunch fees).

[14] Specifically, students receive a loan that becomes 100% forgivable upon graduation. That is, if the recipient drops out, he or she must pay back the total value of the loan (including interests).

[15] SISBEN is a proxy-means testing system to target social welfare programs in Colombia. Verified by local authorities, SISBEN based on dwelling characteristics, demographics, income, and employment at the individual and household level. The SISBEN threshold for SPP is 57.21 for applicants in the 14 main metropolitan areas; 56.32 for other urban areas; and 40.75 for rural areas.

[16] ICETEX offers student loans for students enrolled in technical, technological, university or postgraduate programs in national and/or international tertiary education institutions. The Institute also manages national and international scholarships and grants on behalf of various public and private organizations.

University of Los Andes' *Quiero Estudiar* scholarship-loan program, covers less than 5% of the student body.[17] Thus, prior to SPP, the ability of low-income students to finance their college education at selective private schools remained largely limited.

Five features of the policy change are particularly important for my analysis. First, insofar as SPP is not an *affirmative action program* but a *financial aid program*, Pilos receive no preferential treatment in college admissions processes. Thus, unlike in Rao (2013), rich students share their classroom with poor students that are ex-ante similar in cognitive ability (as measured by SABER 11); the key difference is their socioeconomic background. Second, colleges were not permitted to track students by SES, which guarantees Pilos are integrated in the same classrooms as non-Pilos. Third, the policy only applied to new admissions, which occur exclusively in the starting year (Freshmen). Thus, the policy did not significantly change the composition of cohorts that began schooling before Spring 2015. Finally, SPP was announced *before* most college admissions deadlines, but *after* students had taken the SABER 11 exam (see Figure 3.1). Specifically, students were informed of the creation of the financial aid program and its eligibility conditions almost two months *after* having sat for the SABER 11 exam; it was thus impossible for them to retake the exam to make themselves eligible for SPP. Finally, insofar as accreditation was awarded well in advance of the announcement of SPP, the ability of institutions to self-select into the program is severely limited.

I focus on an elite, private, and selective university in Bogotá (henceforth referred to as University X). Average annual tuition cost is 28 million Colombian pesos, or PPP $16,700, which far exceeds the country's per capita GDP of PPP $13,357.[18] There is no application fee; prospective students apply at no cost. Applicants apply to one major at most, and the SABER 11 score is the sole admission criterion.[19] Importantly, admission cutoffs are major-specific, unknown at the time of application, and reflect institutional capacity constraints (see Saavedra, 2009; Barrera-Osorio and Bayona-Rodriguez, 2015).[20] Finally, students at this university have relatively little freedom in choosing their courses, with the curriculum being relatively set within a major.

The success of SPP in expanding low-SES students' access to selective universities was unprecedented: by January 2015, only three months after SPP was publicly announced,

[17] *Quiero Estudiar* has been available at the University of Los Andes since 2006. It covers tuition costs for low-income students with the highest SABER 11 scores each semester. In practice, *Quiero Estudiar* serves only the extreme top right of the test score distribution of students in the country, making it a significantly more selective program than SPP.

[18] The tuition cost reaches 40 million pesos for Medicine majors.

[19] Exceptionally, in Spring 2016 applicants who were rejected from their first major choice could be considered for a second major. In practice, seats got filled up in the first round, and only a handful of those rejected by their first choice were admitted in their second choice.

[20] Figure A.1 in the provides evidence against bunching at the major-specific admission thresholds.

one-third of freshmen enrolled at University X were Pilos. The extent of this impact was largely unexpected both by governmental institutions, tertiary institutions, and other college applicants. In fact, it was only after ICETEX completed the student loan granting process and institutions published their enrollment statistics that the general public was made aware of the extent to which SPP had shifted entering students' characteristics at selective institutions.[21]

Specifically, the policy shifted peer characteristics at University X in two crucial ways. First, it raised the prevalence of poor students at this elite institution, which had historically been reserved for relatively wealthy students. Figure 3.2, which presents Spring freshmen students by their stratum (a measure of SES), shows that the share poor students (i.e., those in the bottom two strata) almost quadrupled from 7.1% to 27.3% between 2014 and 2015.[22][23]

Moreover, and as will become critical for my estimation strategy, the share of Pilos varied substantially by major, ranging from over 80% of freshmen Philosophy majors to 0% of Art History majors (see Figure 3.3). This variation reflects a combination of differences in the degree of competitiveness in the major-specific admission cutoffs as well as individual-specific tastes for majors.

The second main way in which SPP shifted peer characteristics at this university was by raising average ability of admitted students. Specifically, the increased probability of receiving tuition-free quality college education granted by SPP skyrocketed the number of applications sent by students from poor backgrounds to this university. Indeed, the number of applicants from relatively poor households doubled from 4,000 to almost 8,000 between Spring 2014 and 2015, and further to 10,000 the following year (see Figure 3.4). As expected, the number of applicants from relatively wealthy households was unaffected by the announcement of SPP and only slightly increased in 2016.[24] Notwithstanding this increase in number of applicants, class size remained relatively constant throughout this time period (see Figure 3.5). Consequently, the admission rate dropped from 57% in Spring 2014 to 44% in 2014 and finally 29% in 2016 (see Figure A.5 in the Appendix).[25] Parallelly, the admission thresholds significantly increased across majors. Figure 3.6 presents applicants' test scores,

[21]La Silla Vacía (January 13, 2015) "Las becas obligan a las universidades a actualizarse" (available here).

[22]"Stratum" is a measure of socio-economic status designed to target public service subsidies in Colombia. The system classifies dwellings into 6 strata (1 being the poorest) according to their physical characteristics and surroundings. While correlation with income is clearly imperfect, one advantage of using the strata system is straightforwardness: most Colombians are well aware of their stratum, making this information easy to collect.

[23]In fact, the share of poor students continued increasing, reaching 33.3% in Spring 2016 (see Figure A.4).

[24]The number of rich applicants did not increase between Fall 2014 and 2015 (not reported), suggesting that rich students are *not* avoiding Pilos by shifting their enrollment from the Spring to the Fall semester. Moreover, the share of rich admitted students that enrolled did not change in Spring 2015 (see Figure 3.7).

[25]As a reference, the 2015 freshmen admission rate at the University of California, Berkeley was 17% and 20% for California residents (available here).

standardized with respect to the universe of test-takers, for those seeking to enroll between Spring 2013 and Spring 2016. The figure shows that, while the distribution of applicants' scores did not change much prior to SPP (2013 and 2014), SPP raised the number of 2015 applicants just above the program's eligibility cutoff, thus shifting the admission cutoff towards the right. This phenomenon was reinforced in 2016, further making the cutoffs more selective and continuing to raise the average cognitive ability of admitted students.

Why did SPP have such a large effect on low-SES student enrollment at selective institutions? Figure 3.7 plots the share of admitted students that enroll at University X every Spring semester by socioeconomic stratum. While on average 30% of admitted students enroll, this fraction varies considerably between rich and poor students: rich admits are more than twice as likely to enroll as poor admits. This is explained in large part by the aforementioned financial constraints poor applicants face in Colombia. Indeed, the implementation of SPP relaxed credit constraints and served to shrink this gap by roughly one-half in Spring 2015 and by almost two-thirds one year later.

3.3 Data and Methodology

The data comes from the following sources:

1. **Survey Data:** Survey wave 1 was collected in August 2015, i.e., just over six month after SPP was implemented (see Figure 3.1). The survey sampled relatively wealthy undergraduate students (i.e., strata 4, 5, and 6) who began their studies at University X in Spring 2014, Fall 2015 and Spring 2015 (i.e., 3 cohorts). Survey wave 2 was collected in February 2016, i.e., one year after the first cohort of Pilos began their studies. This survey wave sampled from relatively wealthy undergraduate students (i.e., strata 4, 5, and 6) who began their studies between 2013 and 2016 (i.e., 7 cohorts). The survey questionnaires collected information on students' social and study networks, their perceptions and attitudes. Students were compensated in cash (wave 1) and in kind (wave 2), and were allowed to donate their compensation.[26]

2. **Administrative Data from University X:** Information on applicants' gender, SABER 11 test score, major, cohort, socioeconomic stratum, financial aid received, classes enrolled, academic performance (e.g., GPA), and enrollment status. Also, institutional information on class size, field-cohort admission cutoff, and transfers.

3. **Administrative Data from Ministry of Education:** Names of Pilos enrolled in a each university-program pair, as well some covariates (e.g., high school, municipality, SISBEN index). This database is merged with the names of survey participants' networks to construct a measure of intensity of interaction between Pilo and non-Pilo students.

[26] See survey response rates in Tables A.1 and A.2 in the Appendix.

4. **Administrative Data from SABER 11 High School Exit Exam:** This dataset is available for all students taking the standardized exam in 2003–2014 in Colombia. It contains information on students' test score, high school, and demographic information including students' parental education and socioeconomic characteristics.

My approach exploits the plausibly exogenous timing of SPP and the variation in the share of Pilos across majors to identify the average effect of having poor students in one's classroom. Restricting the sample to rich students (i.e., strata 4, 5, and 6), I estimate the following specification by OLS:

$$Y_{ism} = \alpha + \beta \text{Share Pilo Classmates}_{ism} + X'_{ism}\gamma + \delta_m + \epsilon_{ism} \tag{3.1}$$

where Y_{ism} is outcome Y for student i in semester s in major m, Share Pilo Classmates$_{ism}$ is the average share of classmates that are Pilo (the main treatment variable), X_{ism} is a vector of controls, δ_m are major fixed effects, and ϵ_{ism} is a student specific error term.[27] I cluster standard errors at the semester-by-major level, since this is the unit of treatment. The β coefficient is thus the average effect of having one percentage point increase in the share of classmates that are Pilo, and is the key parameter of interest. This approach identifies the average effect on rich college students of adding poor students to their classroom, which is a relevant estimate for policy.

This identification strategy faces the following potential challenges, each of which I briefly address below:

1. **Rich students may select into non-accredited universities based on their affinity for poor students.**[28] In practice, however, this mechanism is of limited concern. Aside from the potential long-term costs of attending a non-accredited university in lieu of an accredited high-quality university, SPP was implemented in a very short period of time, with the number of poor students enrolled remaining unknown until a few days before classes began. Moreover, it is difficult for students to be picky, since admission rates at top high-quality universities are rather low. Also, transfers across universities are rather uncommon. Finally, Figure 3.4 shows that the number of rich students applying to university X remained constant or, if anything, *increased* with SPP.[29]

[27] Controls include age, sex, SABER 11 z-score, migrant status, stratum, parental education, financial aid received.

[28] For instance, students who find that they particularly dislike poor classmates might transfer to a non-accredited institution in the early years of the program.

[29] Although unlikely, it is possible that students are avoiding Pilos by attending universities abroad. Survey wave 2 attempts to address this issue by asking students to list the universities they applied to. Less than 2% of the survey sample listed a foreign university. Among that 2%, Spring 2016 cohort is the *least* likely to have applied to a university abroad. Note, however, that my sample is biased by construction: I only observe students that enrolled at University X, even though 84% of those who applied to foreign universities were admitted there.

2. **The more selective admissions thresholds raise the average ability of admitted students.** The main concern with estimating effects on academic outcomes is that the policy caused admission cutoffs to rise, making universities more selective and raising average ability of admitted students (see Figure 3.6). I can control for this by including test scores in the regressions and restricting the sample to students that would have been admitted under the higher Spring 2015 thresholds.

3. **Students may endogenously select into classes.** Students may select into courses based on their affinity for poor students. In practice, this is of limited concern because course curriculum is relatively well defined, with students having little liberty to select their courses. Notwithstanding, it is possible to deal with this concern by fixing the predicted distribution Pilo classmates using the distribution of students among classes in the year prior to SPP. Section C.2 in the Appendix describes the methodology used, following Altonji and Card (1991).[30]

4. **Possible spillovers such as (unobserved) out-of-class interactions.** To the extent that these spillovers exist, they would bias against finding effects.

5. **Disproportional dropping out of poor students.** If Pilos drop out at a disproportionately higher rate (whether for financial, personal or psychological reasons), the intensity of the interaction between poor and rich students might be compromised. However, the data suggest that first-semester dropout rate was actually *lower* for Pilos than for non-Pilos.

3.4 Results

Table 3.1 shows that rich students are well-aware of the fraction of students in their classroom that are Pilo.[31] The perceived share of classmates that are Pilo among rich students in the 'treated' cohort (i.e., Spring 2015) is twice the share perceived by non-treated cohorts (i.e., Fall 2014 and Spring 2014). In fact, a one percentage point increase in the share of classmates that are Pilo leads to a 1.254 percentage point increase in the *perceived* share of Pilo classmates, after controlling for ability, SES, gender, parental education, and financial aid received (not reported). Moreover, SPP generated interactions between rich and poor students. Indeed, rich students in the 'treated' cohort have a higher chance of reporting to have Pilos in their social and study networks: 28.9% of Spring 2015 students in strata 4–6 report to have at least one Pilo among their closest five friends in college compared to 7.1% for Fall 2014 cohort and 1.5% for Spring 2014 cohort, while 33.2% of them report to have at least one Pilo among their five closest study partners, compared to 6.5% for Fall

[30] There was also no switching to majors with fewer Pilos (see Figure A.6 in the Appendix).

[31] Anecdotal evidence suggests that the Pilo status can be revealed not by skin color, but through speech: rich students state they know if a student is poor through their accent in Spanish.

2014 cohort and 3.62% for Spring 2014 cohort.[32] Finally, rich students from Spring 2015 cohort report having worked with a Pilo student on average three times, compared to only time for students in older cohorts. To the extent that peer group formation is inherently endogenous, Table 3.1 provides suggestive evidence that rich students exposed to Pilos have more heterogeneous peer groups.

Does this shift in peer characteristics affect rich students' perceptions of the income distribution? According to Cruces et al. (2013), assessment of an income distribution by an economic agent is a statistical inference problem: individuals observe the income levels of no more than a sub-sample of the population (a "reference group") and then apply Bayes' rule to infer the entire distribution from that information. Naïve agents – or those that fail to fully apply Bayes' rule and thus do not fully account for the selection process involved in the formation of the sample they observe – have systematically biased inferences. Moreover, this bias is significantly correlated with an individual's reference group. Insofar as selection into a reference group is a function of income, agents with "rich" reference groups (e.g., rich students in segregated elite universities) are more likely to observe higher-income individuals and vice-versa. In contrast, respondents with friends from heterogeneous social backgrounds are less prone to these biases.

Figure 3.8 illustrates the bias with a rich reference group. Taken from Cruces et al. (2013), the figure depicts the income distribution for the whole population and for a rich reference group. Naïve agents underestimate the actual cumulative income distribution for every income level.[33] In fact, naïve agents have biased estimates of many moments of the income distribution (e.g., mean, median, variance) and will also underestimate the proportion of individuals under the poverty line. If reference groups are more homogeneous in income than the total population (as is likely the case), perceptions about income inequality will be biased downward for all agents.

The extensive selection of relatively wealthy students into Colombian elite higher education institutions might cause a substantial bias in students' perception of the income distribution and poverty rate. However, insofar as SPP dramatically affects rich students' peer groups by making them more heterogeneous in socioeconomic characteristics, it may reduce said biases in the perception of the income distribution and the poverty rate.

Figure 3.9 summarizes the results from testing these hypotheses. The black bars repre-

[32] A back-of-the-envelope calculation suggests that if students selected their friends/study partners at random among their classmates, then the Spring 2015 Freshmen would have a 53.76% change of having at least one Pilo among their five closest friends/study partners. The Fall and Spring 2014 cohorts would have a 18.82% and close of 0% chance of having at least one Pilo among their five closest friends/study partners, respectively.

[33] Another way of saying this is that the distribution for the high-income reference group exhibits first-order stochastic dominance over the distribution for the whole population.

sent the distribution of socioeconomic stratum in Colombia.[34] The white bars represent the average responses when surveying relatively rich students (i.e., strata 4, 5, and 6), while the gray bars add the estimated coefficient from a regression using specification (3.1), collapsing the parameter of interest to an indicator for whether at least 5% of student i's classmates are Pilo.[35] The first finding is that rich students' perception of the income distribution is severely biased: they significantly underestimate the share of the population that is poor (i.e., strata 1 and 2) and overestimate the share of the population that is rich (i.e., strata 4, 5, and 6). This confirms that rich students at this elite university are naïve and do not fully apply Bayes' rule, making a biased inference on the distribution of income in a manner similar to that portrayed in Figure 3.9. Moreover, this is consistent with what Ariely and Norton (2011) and Cruces et al. (2013) find in the United States and Argentina, respectively.

The second finding is that exposure to poor students significantly reduces the magnitude of this bias, shifting rich students' perception of the income distribution closer to reality. Indeed, the coefficient from specification (3.1) on an indicator for whether at least 5% of your classmates are Pilo is positive and significant for the bottom stratum and negative and significant for top three strata (see Table A.4 in the Appendix). This suggests that exposure to peers with more diverse backgrounds makes reference groups more heterogeneous in income, bringing perceptions of the income distribution closer to the actual distribution in Colombia. Importantly, this is not a byproduct of being a more able cohort as a result of more selective admission thresholds: the regression controls for the increase in average ability documented in Figure 3.6 by dropping older students that would not have been admitted under Spring 2015 thresholds.[36][37]

If a shift towards a more diverse reference group lessens the bias in the perceived distribution of income, then it should also reduce rich students' underestimation of the poverty rate. Table 3.2 presents the results from a regression of the perceived share of Colombians that are poor on the share of Pilo classmates. Each row represents a separate regression: the first row uses the observed share of Pilo classmates, while the second row computes this

[34] The distribution of the population by socioeconomic stratum is taken from the stratum self-reported by SABER 11 test takers in Colombia in 2014. Insofar as high school graduates are a positively selected sample of the total population, this distribution underestimates the total amount of individuals in the bottom strata (high school dropouts).

[35] Results are presented in Table A.4.

[36] Another way of interpreting students' biased perception of the income distribution is that they incorrectly believe that the population is distributed uniformly among socioeconomic strata. Thus, the white bars in Figure 3.9 would reflect a misunderstanding of the strata system in Colombia rather than an actual biased perception of the income distribution caused by a failure to fully apply Bayes' rule. However, even if this were the case, the fact that the coefficients from specification (3.1) is positive and significant for the bottom stratum and negative and significant for top three strata suggests exposure to a more diverse peer group affects your perception of the distribution of strata in Colombia.

[37] These results appear to exist only in the short term (i.e., within six months of exposure to Pilos); the results do not hold in survey wave 2, which was collected six months later, i.e., after one year of exposure to Pilos.

share if there were no class sorting due to SPP.[38] Columns (1) and (4) pool together the entire survey sample. We might, however, be wary of pooling Spring and Fall cohorts together, given the vastly different observable and non-observable characteristics of these two cohorts (Figure A.7 in the Appendix, for instance, compares these cohorts' socioeconomic backgrounds). To make the sample more comparable, Columns (2), (3), (5), and (6) restrict the sample to Spring students only. Finally, Columns (3) and (6) control for higher ability of Spring 2015 cohort by restricting the sample of older students to those that would have been admitted under the more selective admission thresholds.

The results suggest that a one percentage point increase in the share of Pilo classmates raises the perceived share of Colombians below the poverty line by 0.349 percentage points, the coefficient being significant at the 1% level. Restricting the sample to Spring cohorts only raises this coefficient to 0.406, while restricting the sample to equally-able students lowers the coefficient back to 0.347. Predicting the distribution of Pilos reduces the size of the coefficient, as well as its significance. To explore potential nonlinear effects of exposure to Pilos on rich students' perceptions of the poverty rate, Table A.5 in the Appendix collapses the main independent variable of interest to an indicator variable that equals 1 if at least 5% of classmates are Pilo – the coefficient of interest is positive and highly significant across specifications.[39]

Does exposure to poor students raise perception of upward social mobility for the poor in Colombia? The results in Table 3.3 suggest that exposure to Pilos increases the perception of upward social mobility for strata 2 (which represents Pilos' average strata, as shown in Figure 3.2), but not strata 1: a one percentage point increase in the observed share of Pilo classmates increase the likelihood of perceiving upward social mobility for strata 2 by 0.7–1.1 percentage points. The results, however, are not robust to using the predicted share of Pilos in lieu of the observed share of Pilo classmates. If the treatment variable is collapsed to an indicator that turns 1 if at least 5% of classmates are Pilo, the coefficients are positive and significant across specifications (see Table A.6 in the Appendix).[40]

As discussed in Section 3.2, SPP caused a steep decline in admission rates, raising competition in college admissions and therefore average ability of admitted students. Moreover, by relaxing the credit constraints that contributed to segregation in private, selective institutions, SPP made the admissions process at private universities more meritocratic. Has this changed students' perception of meritocracy in college admissions?

[38] Section C.2 in the Appendix describes this procedure in detail.

[39] These results appear to exist only in the short term (i.e., six months of exposure to Pilos); the results do not hold in survey wave 2, which was collected six months later, i.e., after one year of exposure to Pilos.

[40] These results appear to exist only in the short term (i.e., six months of exposure to Pilos); the results do not hold in survey wave 2, which was collected six months later, i.e., after one year of exposure to Pilos.

Table 3.4 presents a linear probability model in which the dependent variable is a dummy that takes the value of 1 if the respondent agrees with the phrase *"The most talented students are admitted at the best universities in Colombia"* and 0 otherwise. The results suggest that a one percentage point increase in the observed share of Pilo classmates increases the probability of agreeing with this statement by exactly one percentage point. Neither the magnitude nor the significance of this coefficient is affected by the sample, and it is robust to dropping older students that would not have been admitted under the higher thresholds in Spring 2015. Controlling for possible endogenous course selection slightly reduces the magnitude of the coefficients but the results remain statistically significant.[41]

Does exposure to poor students affect preferences for redistribution? Table 3.5 suggests that a one percentage point increase in the share of Pilo classmates increases support for taxing the rich by 0.9 percentage points (Column (1)) and 1.2 percentage points when controlling for higher ability in the newer cohort (Column (2)). The results remain significant (although again the magnitude of the coefficient decreases somewhat) when using the *predicted* distribution of Pilo classmates. In addition, exposure to poor students raises support for subsidizing the poor, although the coefficient is at best only marginally significant under this specification. However, collapsing the treatment variable to a dummy for whether at least 5% of classmates are Pilo recovers strong significance across all columns: having at least 5% of Pilo classmates increases support for taxing the rich by 11.1–19.5 percentage points and for subsidizing the poor by 7.2–12.6 percentage points (see Table A.8 in the Appendix).[42]

Does this stronger support for redistribution caused by rich students' exposure to Pilos materialize into a higher willingness to donate a share of their own money to Pilos? Survey wave 1 allowed participants to donate a share of their compensation of 10,000 pesos (US $3.4, or the price of a cheap lunch in Bogotá) to fund "low-income, high-achieving students at high-quality universities in Colombia". Interestingly, the results are negative and non-significant across specifications (see Table 3.6).[43]

To delve deeper into these seemingly contradicting findings, survey wave 2 (collected six

[41] Table A.7 in the Appendix suggests that having at least 5% of observed (predicted) Pilo classmates increases belief in meritocracy in college admissions by 11.9–14 (7.5–13.1) percentage points, the coefficients being statistically significant. Moreover, these results remain statistically significant six months later, as documented by survey wave 2 using a larger sample of students: a one percentage point increase in the share of Pilo classmates increases the perception of meritocracy in college admissions by 0.7–1.1 percentage points (not reported).

[42] The positive effect of exposure to Pilo students on rich students' support for subsidizing the poor remains even after one year of continuous exposure: survey wave 2 suggests that a one percentage point increase in the share of Pilo classmates increases support for subsidizing the poor by 0.72–0.93 percentage points, and the coefficient is statistically significant.

[43] Results do not vary if the treatment variable is collapsed to an indicator for whether at least 5% of classmates are Pilo (not reported).

month later) allowed survey respondents to donate the full amount of their compensation of 12,600 pesos (US $4.3, or the cost of a burger combo at a popular burger chain in Bogotá) to an organization of their choosing among the following alternatives: (i) SPP scholarships or the university's own need-based financial aid program; (ii) *Fundación Ayuda por Colombia*, which contributes to the educational and emotional development of poor children and youth; and (iii) GiveDirectly, an organization that directly sends money to the extreme poor. Results confirmed that exposure to Pilos had no statistically significant effect on rich students' donations to SPP or other need-based financial aid programs at their university (not reported).[44] However, it *did* increase the likelihood of donating to GiveDirectly, a non-profit organization that sends money directly tot the extreme poor: being exposed to at least 5% of Pilo classmates increases the likelihood of donating to this cause by 8.1–13.3 percentage points (see Table A.9 in the Appendix).

Does exposure to poor students affect students' academic performance? Table 3.7 uses the GPA obtained in the first semester (expressed in standard deviations) for Spring freshmen students between 2010 and 2015. Each row represents a separate regression, and both the observed and predicted shares of Pilos are expressed in standard deviations of the overall pooled sample of Spring 2010–2015 freshmen students. Columns (1) and (2), which exclude year fixed effects, suggest that the 2015 cohort is performing weakly better than older cohorts: a one standard deviation in the share of observed Pilos increases GPA by 0.0331 standard deviations and 0.0318 standard deviations for rich students. Columns (3) and (4) include year fixed effects, and suggest that GPA is not significantly affected by the presence of Pilos.

Does exposure to poor students affect dropout rates? Table 3.8 presents the results of a linear probability model where the dependent variable is an indicator that equals 1 if a freshman student dropped out the following semester.[45] Each row represents a separate regression. Restricting the sample to 2015 freshmen students, Column (1) suggests that a one standard deviation increase in the share of Pilo classmates raises the likelihood of dropping out by 2.1 percentage points. This coefficient is more than twice as large for rich students (Column (2)). Pooling together Spring freshmen from 2010–2015 cohorts, the coefficients become negative and only marginally significant (Columns (3) and (4)), while including time

[44] In fact, survey evidence suggests that rich students exposed to Pilos perceive they have received enough help by the government and the university. Specifically, Survey wave 2 (collected in Spring 2016) asked to what extent the participant agreed with the affirmation *"Pilos receive more than they deserve"*. Results suggest that a one percentage point increase in the share of students that are Pilo increases the likelihood of agreeing with this statement by 0.9–1.3 percentage points (not reported). Moreover, having at least 5% of Pilo classmates increases this sentiment by 11.1–13.5 percentage points. On a mean of 0.11-0.14, this means that exposure to Pilos doubles the likelihood of perceiving they receive more than they deserve. This suggests that there is some polarization among rich students, a phenomenon that future research should study.

[45] On average, almost 17% of college dropouts drop out after their first semester, according to statistics from the Ministry of Education (available here).

fixed effects makes the coefficients become positive and non-significant. This suggests that exposure to Pilos had little effect on students' dropout rates.

Finally, has the greater competition for college places at this selective university induced the most recent entering cohorts to exert more effort in preparing for the SABER 11 exam? To explore changes in test preparation behavior, survey wave 2 collected information about the study methods used to prepare for SABER 11, the amount of time and pesos spent preparing for this exam, and the list of universities they applied to, among others.

Individual i's SABER 11 score can be modeled as a function of innate ability θ_i, exam prepping p_i, and luck ϵ_i, $s_i = \Phi(\theta_i, p_i, \epsilon_i)$. Because SPP was announced two months *after* high school seniors took SABER 11 (recall Figure 3.1), differences in s_i between Spring 2015 freshmen and older cohorts can be attributed to a higher θ_i, but not a higher p_i. In contrast, cohorts scheduled to take SABER 11 *after* SPP was announced could prepare for the more competitive admissions process by exerting more effort in preparing for the exam. Thus, differences between them and older cohorts will be a result of a combination of θ_i and p_i.

To test whether rich students' preparation for SABER 11 increased after the announcement of SPP, I use the following linear probability model:

$$p_{it} = \alpha + \beta \cdot \text{SPP}_{t-1} + X'_{it}\gamma + \theta_{m(i)} + \epsilon_{it} \tag{3.2}$$

where p_{it} is a measure of the extent to which student i at time t prepared for SABER 11, SPP_{t-1} is an indicator for that equals 1 SABER 11 was taken *after* the announcement of SPP, $X'_{it}\gamma$ is a vector of individual controls, $\theta_{m(i)}$ is a major fixed effect, and ϵ_{it} is the error term.[46] A positive β coefficient would suggest that cohorts taking the exam after SPP was announced prepare more for SABER 11 than previous cohorts.

Table 3.9 presents the results from an OLS regression of specification 3.2. Column (1) suggests that students taking SABER 11 after the announcement of SPP are 8.7 percentage points more likely to prepare for the exam by either hiring a private tutor, using a test practice book or studying on their own. To deal with the fact that the higher admission cutoffs have raised average ability of more recent cohorts, Column (2) restricts the sample of older cohorts to students that would have been admitted under Spring 2016 thresholds: the coefficient is 7.1 and significant at the 5% level. Although younger cohorts do not seem to be spending more time studying for the exam (Columns (3) and (4)), they do, nevertheless, report investing a larger sum of money on exam prepping: younger cohorts are 9.8–10.3 percentage points more likely to have spent at least 500,000 pesos (US $160) to prepare for the exam.[47]

[46] I cluster standard errors at the major level.

[47] Students that applied to University X *after* SPP was announced did not apply to a larger number of universities than those who applied *before* SPP was announced. However, compared to older students that

Recall that Figure 3.4 provided evidence that SPP skyrocketed the number of poor applicants at this elite university, thus raising admission thresholds (see Figure 3.6) and reducing admission rates (see Figure A.5 in the Appendix). Moreover, given that the number of poor applicants varies considerably across majors (see Figure A.8 in the Appendix), then the degree to which SPP increased competition is major-specific and can be exploited to test for behavioral responses in exam prepping as a result of higher competition in admissions:

$$p_{imt} = \alpha + \beta \cdot \text{Number of Poor Applicants}_{m,t-1} + X_{it}'\gamma + \theta_m + \delta_t + \epsilon_{imt} \tag{3.3}$$

where p_{imt} is a measure of the extent to which student i applying to major m at time t prepared for SABER 11, Number of Poor Applicants$_{m,t-1}$ is the number of poor applicants at major m and time $t-1$, $X_{it}'\gamma$ is a vector of individual controls, θ_m is a major fixed effect, δ_t is a time fixed effect and ϵ_{it} is the error term.[48]

Table 3.10 presents the results of regressing specification (3.3) by OLS.[49] The results provide weak evidence that the higher competition in admissions as a result of SPP has raised effort exerted by rich students for SABER 11. Column (1) suggests that a one standard deviation in the lagged number of poor applicants increases the probability of preparing for the exam by 3.9 percentage points. The coefficient looses significance in Column (2), which restricts the sample to students that would have been admitted under 2016 major-specific thresholds to compare preparation among students of similar ability. Column (3) suggests that a one standard deviation increase in the lagged number of poor applicants increases the probability of spending at least 75 hours studying for SABER 11 by 4.6 percentage points, but the coefficient looses significance when restricting the sample in Column (4). Finally, the coefficient for the sum spent on prepping is positive but not statistically significant (Columns (5) and (6)).

While encouraging, the results from Tables 3.9 and 3.10 are somewhat difficult to interpret. Indeed, given that the sample represents enrolled students at this University, it constitutes by construction a truncated sample restricted to students who are admitted and enrolled at this university. Yet the admission requirement has become more selective in the last years (recall Figure 3.3) and cohorts taking SABER 11 after SPP was announced can alter their test score s_i by modifying their test preparation, p_i. Thus, the fact that the incen-

would have been admitted under 2016 admission cutoffs, they are 4.8 percentage points *less* likely to have been admitted at their top choice – but this coefficient is only significant at the 10% level (not reported).

[48] I cluster standard errors at the major level.

[49] Two features of this specification are worth discussing. First, I used the *lagged* and not the current number of poor applicants because, if anything, the result of a behavioral change in exam prepping will be evidenced in the period *after* the event. Second, I cannot replace the number of *poor* applicants by the number of *Pilo* applicants because the university cannot identify applicants that are Pilo, given that the "Pilo" condition depends on getting admitted to an accredited university. Thus, the university only records Pilos that are enrolled in its system.

tives to modify p_i are not held constant throughout the years makes the comparison between younger and older cohorts problematic: it is not sufficient to simply drop older students that would have been admitted under the Spring 2016 thresholds, as in Columns (2), (4), and (6) of Tables 3.9 and 3.10. Given the lack of a convincing counter-factual, the results from Tables 3.9 and 3.10 remain merely suggestive. A more rigorous test of the general equilibrium effects of a large shock in admissions processes at selective private universities would require expanding the survey sample to the universe of applicants, an area that should be explored in future research.

3.5 Conclusion

I exploit the plausibly exogenous timing of a financial aid program that generated an unprecedented and discontinuous jump in the presence of poor students at an elite university in Colombia. This setting allows me to causally identify the effect of exposure to poor peers on rich individuals' perceptions of inequality and poverty, beliefs of social justice, redistributive preferences, and academic outcomes. Using survey and administrative data, I find a significant positive effect of exposure to poor students on interactions among students with heterogeneous family backgrounds. Exposure to more heterogeneous peers reduces biases generated by rich reference groups in the perception of the income distribution and the poverty rate. In addition, rich students exposed to poor peers perceive more upward social mobility, and a higher perception of meritocracy in college admissions. They also show a stronger support for redistribution. Moreover, I find no significant effect on dropout rates and GPA.

Figures and Tables

Figure 3.1: Timeline of Events

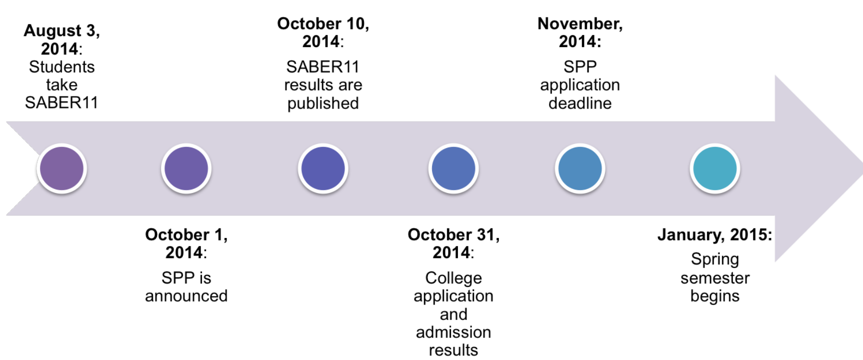


Figure 3.2: SPP Raised the Share of Poor Students at Elite College

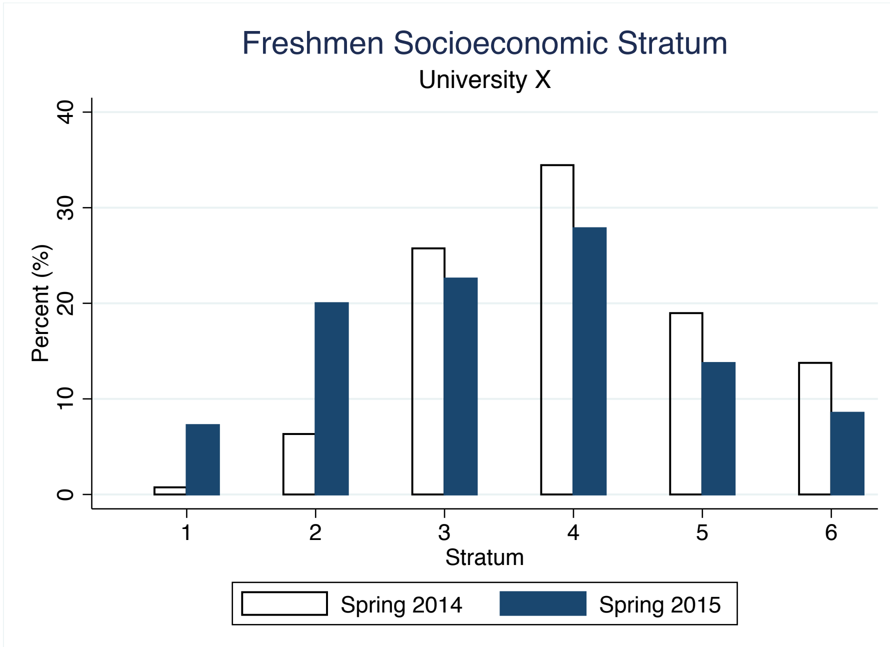
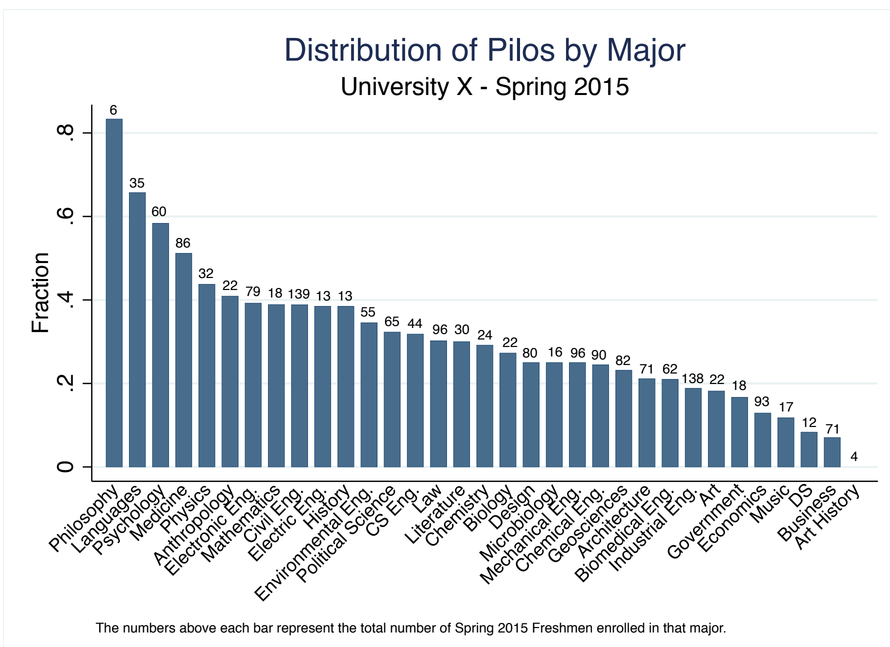
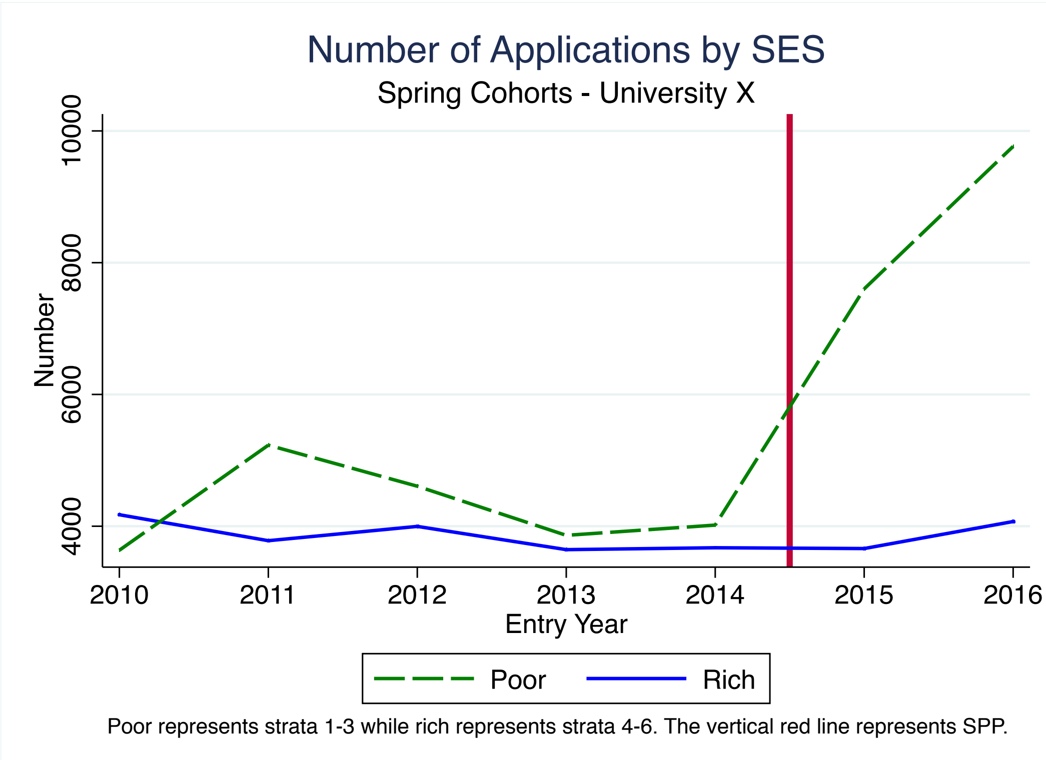


Figure 3.3: There is Wide Variation in the Share of Pilos Across Majors



Note: The numbers above each bar represent the total number of Spring 2015 Freshmen enrolled in that major.

Figure 3.4: SPP Increased the Number of Relatively Poor Applicants



Note: This graph plots the number of applications for Spring admissions by self-reported socioeconomic stratum. "Poor" refers to strata 1–3, while "rich" refers to strata 4–6. The vertical red line represents SPP.

Figure 3.5: Cohort Size Remained Constant

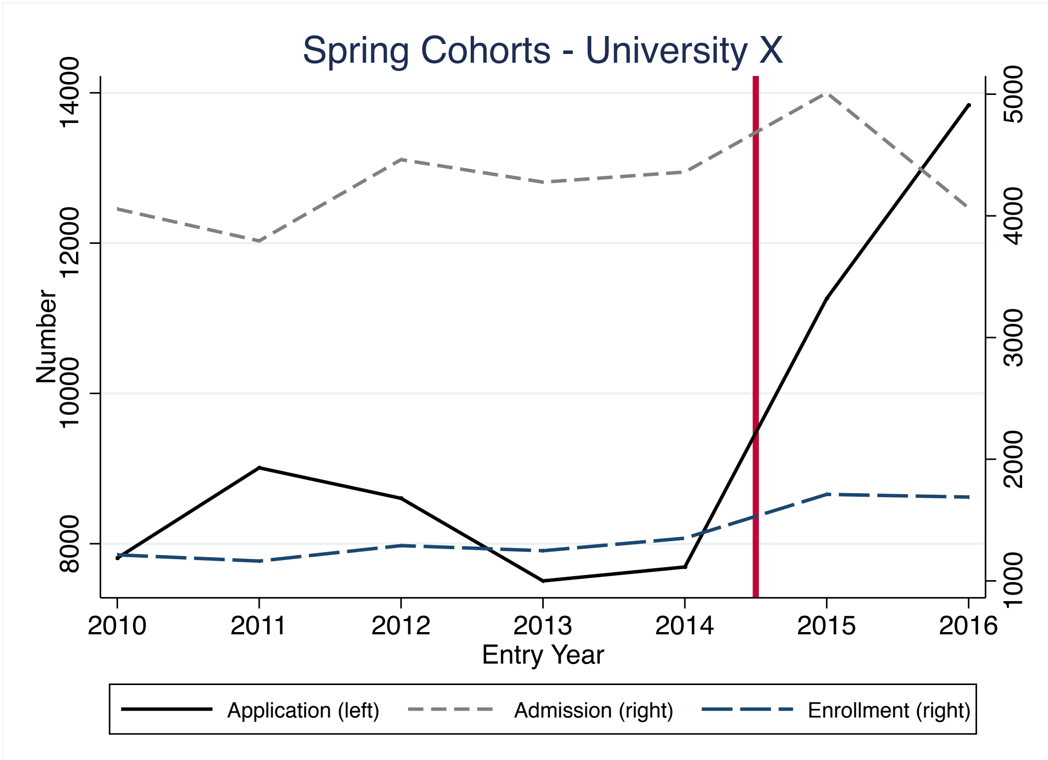


Figure 3.6: SPP Raised Admission Thresholds

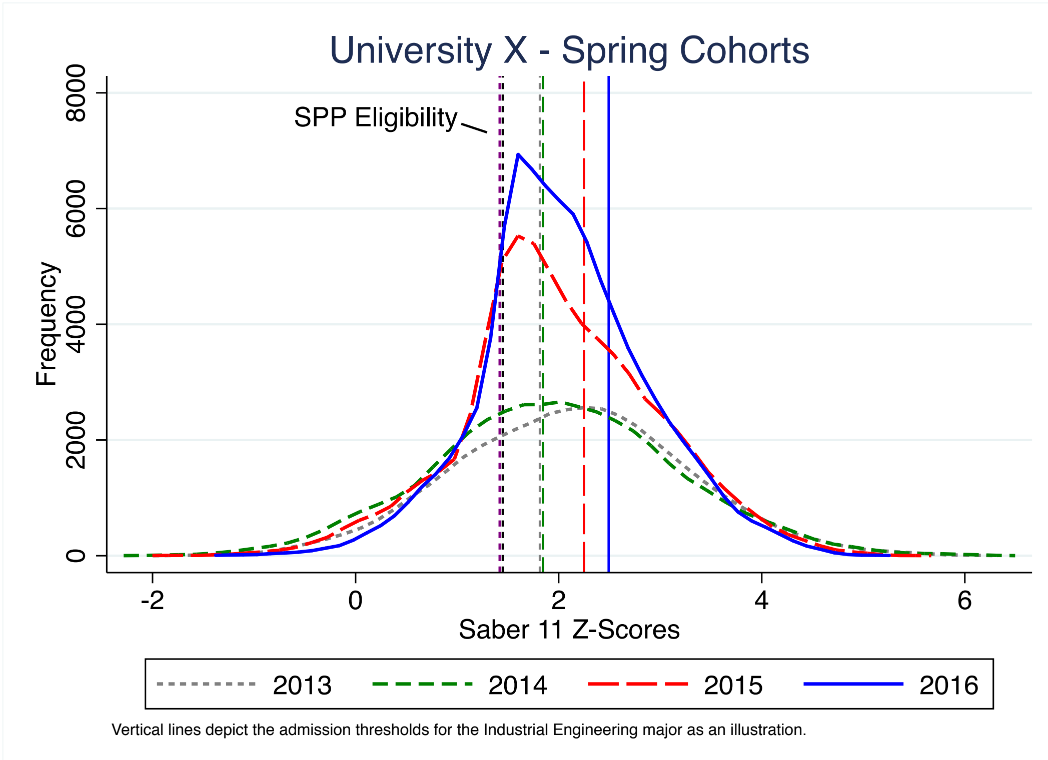
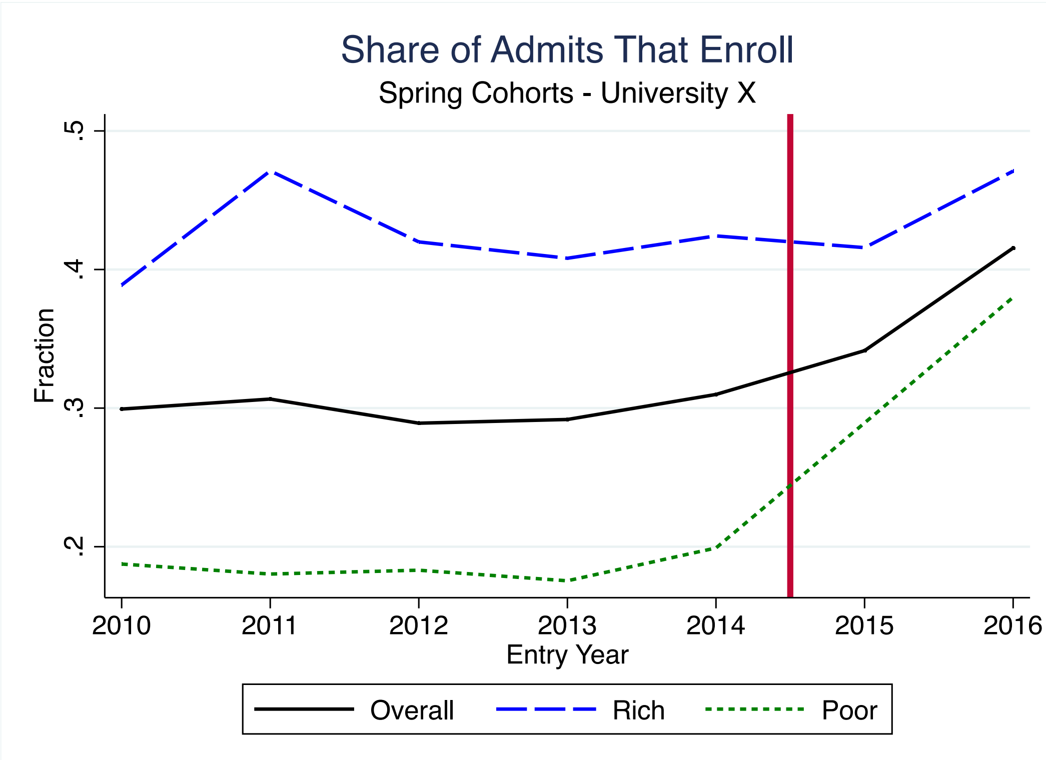
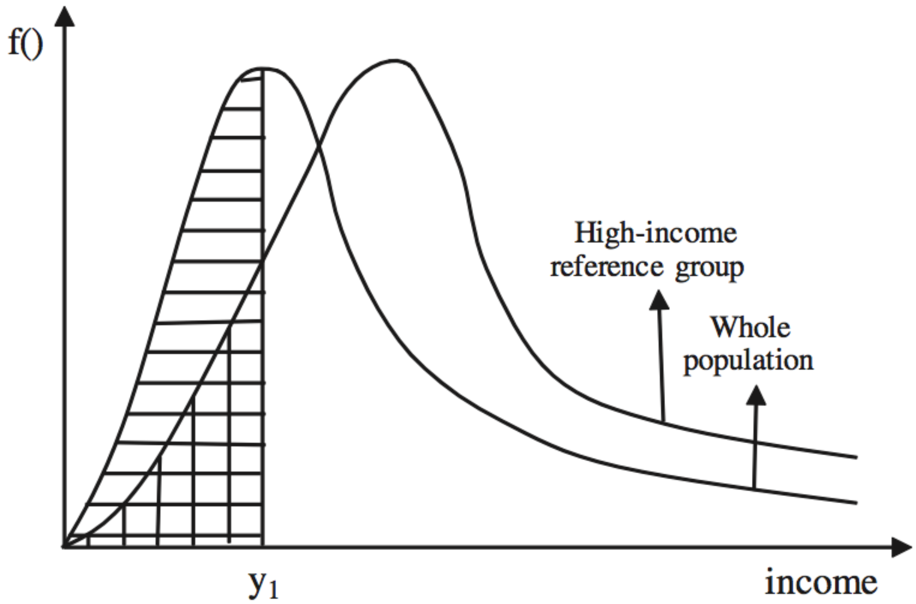


Figure 3.7: SPP Raised the Share of Poor Admits that Enroll



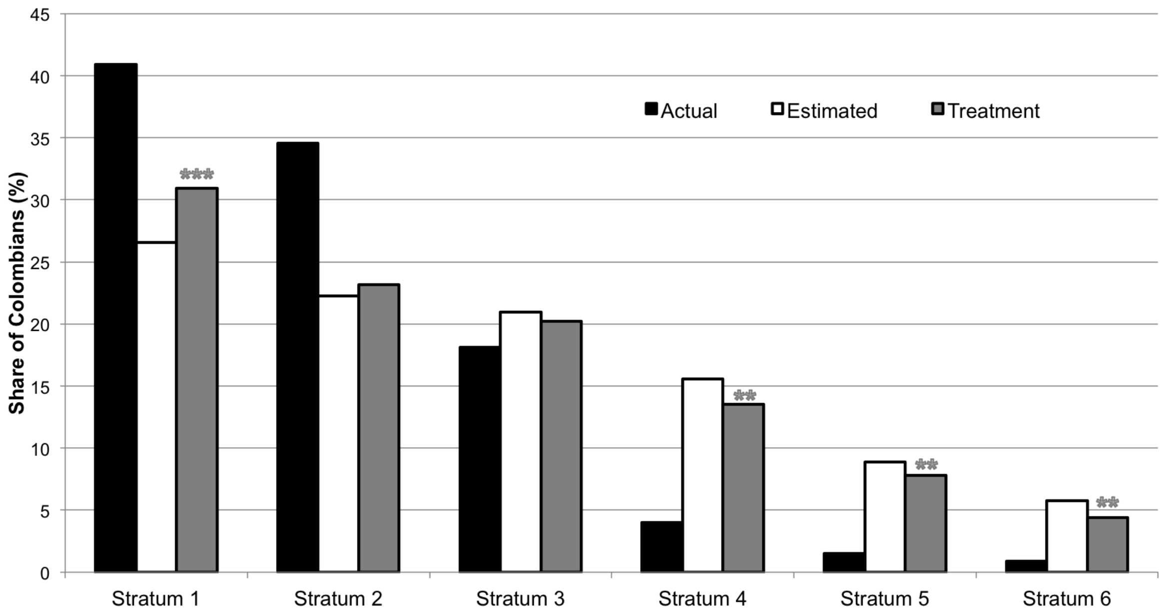
Note: This graph plots the fraction of Spring applicants that enroll by self-reported socioeconomic stratum. "Poor" refers to strata 1–3, while "rich" refers to strata 4–6. The vertical red line represents SPP.

Figure 3.8: Illustration of Biases with a Rich Reference Group



Source: Figure 1a in Cruces et al. (2013).

Figure 3.9: The Shift in the Reference Group Reduces the Bias



Note: The "Actual" bar represents the distribution of (self-reported) stratum for the universe of SABER 11 test takers in 2014 in Colombia. The "Estimated" bar represents the average response in survey, while the "Treatment" bar adds the estimated coefficient from the regression using a dummy for at least 5% Pilo classmates that controls for the increase in ability as a result of SPP by dropping older students that would not have been admitted under 2015 thresholds using Spring cohorts only. Controls include age, age squared, gender, migrant status, risk aversion, socioeconomic stratum, parental education, financial aid received, high school calendar, and SABER 11 test score. $***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.1: SPP generates rich-poor interactions

	Cohort		
	Spring 2015 (1)	Fall 2014 (2)	Spring 2014 (3)
Perceived % Pilo classmates	33.61	15.91	13.42
	(19.41)	(13.71)	(14.74)
Pilo among 5 closest friends	28.91	7.12	1.45
	(45.44)	(25.92)	(11.99)
Pilo among 5 regular study partners	33.18	6.54	3.62
	(47.20)	(24.80)	(18.75)
No. times worked with Pilo	3.05	1.18	1.00
	(3.13)	(2.38)	(2.08)

Note: This table presents means (and standard deviations in parentheses) by cohort, i.e., the semester in which they began their studies at University X.

Table 3.2: The Shift in the Reference Group Raises Perception of Poverty Rate

	Share of Colombians that are poor					
	(1)	(2)	(3)	(4)	(5)	(6)
Observed Share of Pilos	0.349***	0.406***	0.347**			
	(0.119)	(0.136)	(0.140)			
Predicted Share of Pilos				0.165*	0.278**	0.203
				(0.092)	(0.122)	(0.123)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Major FE	Yes	Yes	Yes	Yes	Yes	Yes
Spring Cohort Only	No	Yes	Yes	No	Yes	Yes
Spring '15 Threshold Only	No	No	Yes	No	No	Yes
N	493	342	309	493	342	309
R^2	0.16	0.2	0.2	0.15	0.2	0.2
Dep Mean	36.75	36.06	35.65	36.75	36.06	36.65
Dep SD	19.4	19.24	19.08	19.4	19.24	19.08

Notes: Each row represents a separate regression. Controls include age, age squared, gender, migrant status, risk aversion, socioeconomic stratum, parental education, financial aid received, high school calendar, and SABER 11 test score. Columns (2) and (5) restrict the sample to Spring cohorts only. Columns (3) and (6) control for higher positive selection on ability of admitted students due to SPP by dropping pre-Spring 2015 students with test scores below the Spring 2015 cutoffs. Clustered standard errors at the semester-by-major level in parentheses.
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.3: Upward social mobility among the poor

	Upward Stratum 1				**Upward Stratum 2**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Observed Share of Pilos	-0.002	0.004			0.007*	0.011**		
	(0.004)	(0.005)			(0.004)	(0.005)		
Predicted Share of Pilos			-0.004	0			0.003	0.006
			(0.003)	(0.004)			(0.003)	(0.004)
Major FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Spring Cohort Only	No	Yes	No	Yes	No	Yes	No	Yes
Spring '15 Threshold	No	Yes	No	Yes	No	Yes	No	Yes
N	493	309	493	309	493	309	493	342
R^2	0.12	0.16	0.13	0.16	0.13	0.21	0.13	0.2
Dep Mean	0.53	0.56	0.53	0.56	0.57	0.59	0.57	0.59
Dep SD	0.5	0.5	0.5	0.5	0.5	0.49	0.5	0.49

Notes: Each row represents a separate regression. Controls include age, age squared, gender, migrant status, risk aversion, socioeconomic stratum, parental education, financial aid received, high school calendar, and SABER 11 test score. Clustered standard errors at the semester-by-major level in parentheses.
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.4: A Higher Perception of Meritocracy in College Admissions

	Most talented students are admitted at best colleges					
	(1)	(2)	(3)	(4)	(5)	(6)
Observed Share of Pilos	0.010***	0.010***	0.010***			
	(0.003)	(0.003)	(0.003)			
Predicted Share of Pilos				0.005**	0.008***	0.007**
				(0.002)	(0.003)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Major FE	Yes	Yes	Yes	Yes	Yes	Yes
Spring Cohort Only	No	Yes	Yes	No	Yes	Yes
Spring '15 Threshold	No	No	Yes	No	No	Yes
N	493	342	309	493	342	309
R^2	0.11	0.15	0.19	0.1	0.15	0.18
Dep Mean	0.33	0.35	0.36	0.33	0.35	0.36
Dep SD	0.47	0.48	0.48	0.47	0.48	0.48

Notes: Each row represents a separate regression. Controls include age, age squared, gender, migrant status, risk aversion, socioeconomic stratum, parental education, financial aid received, high school calendar, and SABER 11 test score. Columns (3) and (6) control for higher positive selection on ability of admitted students due to SPP by dropping pre-Spring 2015 students with test scores below the Spring 2015 cutoffs. Clustered standard errors at the semester-by-major level in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.5: Preferences for redistribution

	State should tax the rich				**State should subsidize the poor**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Observed Share of Pilos	0.009***	0.012***			0.004	0.008*		
	(0.003)	(0.004)			(0.003)	(0.004)		
Predicted Share of Pilos			0.006**	0.007**			0.004	0.006
			(0.002)	(0.003)			(0.003)	(0.004)
Major FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Spring Cohort Only	No	Yes	No	Yes	No	Yes	No	Yes
Spring '15 Threshold	No	Yes	No	Yes	No	Yes	No	Yes
N	493	309	493	309	493	309	493	309
R^2	0.13	0.18	0.12	0.17	0.10	0.16	0.10	0.16
Dep Mean	0.72	0.71	0.72	0.71	0.48	0.46	0.48	0.46
Dep SD	0.45	0.45	0.45	0.45	0.5	0.5	0.5	0.5

Notes: Each row represents a separate regression. Controls include age, age squared, gender, migrant status, risk aversion, socioeconomic stratum, parental education, financial aid received, high school calendar, and SABER 11 test score. Clustered standard errors at the semester-by-major level in parentheses.
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.6: No effect on donation to SPP

	Share of Compensation Donated to SPP (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Observed Share of Pilos	-0.43	-0.244	-0.212			
	(0.271)	(0.348)	(0.373)			
Predicted Share of Pilos				-0.262	-0.285	-0.229
				(0.230)	(0.271)	(0.302)
Major FE	Yes	Yes	Yes	Yes	Yes	Yes
Spring Cohort Only	No	Yes	Yes	No	Yes	Yes
Spring ‘15 Threshold	No	No	Yes	No	No	Yes
N	493	342	309	493	342	309
R^2	0.14	0.16	0.17	0.14	0.16	0.17
Dep Mean	38.01	37.56	38.27	38.01	37.56	38.27
Dep SD	44.24	44.11	44.52	44.24	44.11	44.52

Notes: Each row represents a separate regression. Controls include age, age squared, gender, migrant status, risk aversion, socioeconomic stratum, parental education, financial aid received, high school calendar, and SABER 11 test score. Columns (3) and (6) control for higher positive selection on ability of admitted students due to SPP by dropping pre-Spring 2015 students with test scores below the Spring 2015 cutoffs. Clustered standard errors at the semester-by-major level in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.7: SPP did not significantly affect student grades

	First Semester GPA			
	All	Rich Only	All	Rich Only
	(1)	(2)	(3)	(4)
Observed Share of Pilos	0.0331*	0.0318*	0.027	0.0285
	(0.0172)	(0.0185)	(0.0438)	(0.0574)
Predicted Share of Pilos	0.0231	0.0311	-0.0247	0.0101
	(0.0164)	(0.0190)	(0.0351)	(0.0473)
Controls	Yes	Yes	Yes	Yes
Major FE	Yes	Yes	Yes	Yes
Cohort FE	No	No	Yes	Yes
N	7455	4729	7455	4729
R^2	0.17	0.17	0.18	0.18

Notes: The dependent variable is GPA obtained in the first semester, expressed in standard deviations. Each row represents a separate regression, and both the observed and predicted shares of Pilos are expressed in standard deviations of the overall sample of Spring 2010–2015 Freshmen students. To test whether the Spring 2015 cohort is performing differently relative to previous cohorts, columns (1)-(4) pool together freshmen grades from years 2010–2015. Columns (1) and (2) do not include year fixed effects, and suggest that the 2015 cohort is performing weakly better than older cohorts. Columns (3) and (4) include year fixed effects to gain better precision, and suggest that GPA is not significantly affected by the presence of Pilos. Columns (2) and (4) present results for rich students only (i.e., strata 4, 5, and 6). All columns include controls for SABER 11 z-score, and dummies for sex and socioeconomic stratum. Standard errors are in parentheses and are clustered at the major level.
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.8: SPP had no effect on freshmen dropout rates

	Dropped Out After First Semester			
	All	Rich Only	All	Rich Only
	(1)	(2)	(3)	(4)
Observed Share of Pilos	-0.00549*	-0.00072	0.00481	0.02073
	(0.00305)	(0.00406)	(0.00541)	(0.01268)
Predicted Share of Pilos	-0.00521	-0.00329	0.00433	-0.00377
	(0.00342)	(0.00368)	(0.00723)	(0.00802)
Controls	Yes	Yes	Yes	Yes
Major FE	Yes	Yes	Yes	Yes
Time FE	No	No	Yes	Yes
N	7533	4773	7533	4773
R2	0.02	0.022	0.021	0.023
Dep Mean	0.074	0.073	0.074	0.073
Dep SD	0.262	0.26	0.262	0.26

Notes: The sample pools together Spring Freshmen students from 2010 to 2015. Each row represents a separate regression. Both observed and predicted share of Pilo classmates are expressed in standard deviations. Controls include SABER 11 z-score, and dummies for sex and socioeconomic stratum. Columns (2), and (4), restrict the sample to students from strata 4, 5, and 6. Clustered standard errors at the major level in parentheses.
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 3.9: Preparation for SABER 11 After SPP is Announced

	Prepared for Exam		**Time Spent Preparing**		**Cost of Preparing**	
	(1)	(2)	(3)	(4)	(5)	(6)
SPP_{t-1}	0.0870***	0.0711**	0.0395	-0.0003	0.0983**	0.1033**
	(0.0229)	(0.0311)	(0.0489)	(0.0415)	(0.0361)	(0.0382)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Major FE	Yes	Yes	Yes	Yes	Yes	Yes
Spring '16 Threshold	No	Yes	No	Yes	No	Yes
N	1237	766	1237	766	1237	766
R^2	0.08	0.08	0.09	0.12	0.06	0.1

Note: "Prepared for exam" is an indicator variable that equals 1 if the student reported preparing for SABER 11 by either hiring a private tutor, using a practice book or through self-study (this is the case for 50.7% of the sample); 'Time to Prepare" is an indicator variable that equals 1 if the student reported spending more than 75 hours studying for the exam (this is the case for 50.3% of the sample); and "Cost to Prepare" is an indicator variable that equals 1 if the student reported spending more than 500,000 pesos (US $160) to prepare for the exam (this is the case for 38.1% of the sample). Columns (2), (4), and (6) restrict the sample to students that would have been admitted under the higher Spring 2016 thresholds. Controls include gender, migrant status, socioeconomic stratum, and a dummy for Spring cohort. Standard errors clustered at the major level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.10: Difference-in-Differences Estimation: Preparation for SABER 11

	Prepared for Exam		**Time Spent Preparing**		**Cost of Preparing**	
	(1)	(2)	(3)	(4)	(5)	(6)
Number of Poor $\text{Applicants}_{m,t-1}$	0.039*	0.04	0.046**	0.011	0.041	0.028
	(0.021)	(0.031)	(0.022)	(0.030)	(0.029)	(0.027)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Major FE	Yes	Yes	Yes	Yes	Yes	Yes
Spring '16 Threshold	No	Yes	No	Yes	No	Yes
N	1237	766	1237	766	1237	766
R^2	0.09	0.1	0.09	0.13	0.07	0.1

Note: "Prepared for exam" is an indicator variable that equals 1 if the student reported preparing for SABER 11 by either hiring a private tutor, using a practice book or through self-study (this is the case for 50.7% of the sample); 'Time to Prepare" is an indicator variable that equals 1 if the student reported spending more than 75 hours studying for the exam (this is the case for 50.3% of the sample); and "Cost to Prepare" is an indicator variable that equals 1 if the student reported spending more than 500,000 pesos (US $160) to prepare for the exam (this is the case for 38.1% of the sample). The independent variable of interest is the lagged number of poor applicants, expressed in standard deviations. Poor applicants are defined as those from strata 1, 2, and 3. Columns (2), (4), and (6) restrict the sample to students that would have been admitted under the higher Spring 2016 thresholds. Controls include gender, migrant status, socioeconomic stratum, and a dummy for Spring cohort. Standard errors clustered at the major level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Bibliography

A. Abadie. Bootstrap tests for distributional treatment effects in instrumental variables models. *Journal of the American Statistical Association*, 97(457):284–292, 2002.

A. Abdulkadiroglu, P. Pathak, and C. Walters. Free to choose: Can school choice reduce student achievement? *American Economic Journal: Applied Economics*, 10(1):175–206, 2018.

A. Alesina and G. Angeletos. Fairness and redistribution: US versus Europe. *American Economic Review*, 95(4):960–980, 2005.

A. Alesina and E. La Ferrara. Preferences for redistribution in the land of opportunies. *Journal of Public Economics*, 89:897–931, 2005.

A. Alesina, E. Glaeser, and B. Sacerdote. Why doesn't the United States have a European-style welfare state. *Brookings Papers on Economic Activity*, 2:1–69, 2001.

M. Almunia and D. Lopez-Rodriguez. Under the radar: The effects of monitoring firms on tax compliance. *American Economic Journal: Economic Policy*, 10(1):1–38, 2018.

S. Alon and O. Malamud. The impact of israel's class-based affirmative action policy on admission and academic outcomes. *Economics of Education Review*, 40:123–139, 2014.

A. Alstadsater, N. Johannesen, and G. Zucman. Who owns the wealth in tax havens? macro evidence and implications for global inequality. 162:89–100, June 2018a. Journal of Public Economics.

A. Alstadsater, N. Johannesen, and G. Zucman. Tax evasion and tax avoidance. 2018b.

A. Alstadsater, N. Johannesen, and G. Zucman. Tax evasion and inequality. forthcoming. American Economic Review.

J.G. Altonji and D. Card. *The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives*, chapter Immigration, Trade, and the Labor Market, pages 201–234. University of Chicago Press, 1991.

F. Alvaredo and J. Londono-Velez. High incomes and personal taxation in a developing economy: Colombia 1993–2010. Commitment to Equity Working Paper No. 12, March 2013.

F. Alvaredo and J. Londoño-Vélez. Altos ingresos e impuesto de renta en colombia, 1993–2010. *Revista de Economía Institucional*, 16(31):157–194, 2014.

F. Alvaredo, L. Chancel, T. Piketty, E. Saez, and G. Zucman. *World Inequality Report.* 2018.

J. Angrist and K. Lang. Does school integration generate peer effects? Evidence from Boston's Metco program. *The American Economic Review*, 94(5):1613–1634, 2004.

J. Angrist and V. Lavy. The effect of high stakes high school achievement awards: Evidence from a randomized trial. *American Economic Review*, 94(4):301–331, 2009.

J. Angrist, E. Bettinger, E. Bloom, E. King, and M. Kremer. Vouchers for private schooling in colombia: Evidence from randomized natural experiments. *American Economic Review*, 92(5):1535–1558, 2002.

J. Angrist, D. Autor, S. Hudson, and A. Pallais. Leveling up: Early results from a randomized evaluation of post-secondary aid. NBER Working Paper No. 20800, December 2014.

J. Angrist, D. Autor, S. Hudson, and A. Pallais. Evaluating post-secondary aid: Enrollment, persistence, and projected completion effects. December 2016.

D. Ariely and M. Norton. Building a better America – One wealth quintile at a time. *Perspectives on Psychological Science*, 6(1):9–12, 2011.

T. Aronsson, K. Jenderny, and G. Lanot. The quality of the estimators of the eti. 2017.

D.H. Autor. Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, 344(6186):843–851, 2014.

C. Avery, C. Hoxby, C. Jackson, K. Burek, G. Pope, and M. Raman. Cost should be no barrier: An evaluation of the first year of harvard's financial aid initiative. NBER Working Paper No. 12029, February 2006.

P. Bachas and M. Soto. Not(ch) your average tax system: Corporate taxation under weak enforcement. World Bank Policy Research Working Paper 8524, July 2018.

C. Balestra and R. Tonkin. Inequalities in household wealth across oecd countries: Evidence from the oecd wealth distribution database. *OECD Statistics Working Papers*, 2018. OECD Publishing, Paris.

F. Barrera-Osorio and H. Bayona-Rodriguez. The causal effect of college attendance on educational trajectory and labor market outcomes: Empirical evidence from Colombia. Documentos CEDE 2015-27, August 2015.

F. Barrera-Osorio and D. Filmer. Incentivizing schooling for learning: Evidence on the impact of alternative targeting approaches. *Journal of Human Resources*, 51(2):461–499, 2016.

R.C. Bayer, H. Oberhofer, and H. Winner. The occurence of tax amnesties: Theory and evidence. *Journal of Public Economics*, 125:70–82, 2015.

R. Benabou and E. Ok. Social mobility and the demand for redistribution: The POUM hypothesis. *The Quarterly Journal of Economics*, 116(2):447–487, 2001.

M. Bérgolo, R. Ceni, G. Cruces, M. Giaccobasso, and R. Perez-Truglia. Misperceptions about tax audits. *American Economic Association Papers and Proceedings*, 108:83–87, 2018.

E. Bettinger, O. Gurantz, L. Kawano, and B. Sacerdote. The long run impacts of merit aid: Evidence from california's cal grant. *American Economic Journal: Economic Policy*, 11 (1):64–94, 2019.

E.P. Bettinger, B.T. Long, P. Oreopoulos, and L. Sanbonmatsu. The role of application assistance and information in college decisions: Results from the h&r block fafsa experiment. *Quarterly Journal of Economics*, 127(3):1205–1242, 2012.

S. Blomquist and W. Newey. The bunching estimator cannot identify the taxable income elasticity. NBER Working Paper No. 24136, December 2017.

A. Brockmeyer and M. Hernandez. Taxation, information, and withholding: Evidence from costa rica. Working Paper, World Bank, 2016.

M. Brulhart, J. Gruber, M. Krapf, and M. Schmidheiny. The elasticity of taxable wealth: Evidence from switzerland. Working Paper, 2017.

A. Camacho, J. Messina, and J.P. Uribe. The expansion of higher education in colombia: Bad students or bad programs? IDB Discussion Paper 452, April 2016.

B. Castleman and J. Goodman. Intensive college counseling and the enrollment and persistence of low income students. *Education Finance and Policy*, 2017.

B.L. Castleman and B.T Long. Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation. NBER Working Paper No. 19306, September 2015.

M.D. Cattaneo, S. Calonico, and R. Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326, 2014.

M.D. Cattaneo, M. Jansson, and X. Ma. Manipulation testing based on density discontinuity. January 2016a.

M.D. Cattaneo, M. Jansson, and X. Ma. Simple local regression distribution estimators with an application to manipulation testing. Working Paper, University of Michigan, January 2016b.

J. Charite, R. Fisman, and I. Kuziemko. Reference points and redistributive preferences: Experimental evidence. NBER Working Paper 21009, March 2015.

R. Chetty, J. Friedman, T. Olsen, and L. Pistaferri. Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. *Quarterly Journal of Economics*, 126(2):749–804, 2011.

R. Chetty, J. Friedman, and E. Saez. Using differences in knowledge across neighborhoods to uncover the impacts of the eitc on earnings. *American Economic Review*, 103(7):2683–2721, 2013.

R. Chetty, J.N. Friedman, E. Saez, N. Turner, and D. Yagan. Mobility report cards: The role of colleges in intergenerational mobility. January 2017.

S.R. Cohodes and J.S. Goodman. Merit aid, college quality, and college completion: Massachusetts' adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics*, 6(4):251–285, 2014.

CollegeBoard. Trends in student aid, 2017.

G. Cruces, R. Perez-Truglia, and M. Tetaz. Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics*, 98:100–112, 2013.

R. Diamond and P. Persson. The long-term consequences of teacher discretion in grading of high-stakes tests. NBER Working Paper No. 22207, June 2016.

E.W. Dillon and J.A. Smith. Determinants of the match between student ability and college quality. *Journal of Labor Economics*, 35(1):45–66, 2017.

DNP, CNC, and Universidad de Los-Andes. Evaluación de impacto de corto plazo del programa ser pilo paga: http://sinergiapp.dnp.gov.co/#evaluaciones/evalfin, July 2016.

J.M. Durán-Cabé, A. Esteller-Moré, and M. Mas-Montserrat. Behavioral responses to the (re)introduction of wealth taxes: Evidence from spain. Universitat de Barcelona Working Paper, 2017.

S. Dynarski. Does aid matter? measuring the effect of student aid on college attendance and college completion. *American Economic Review*, 93(1):279–288, 2003.

S. Dynarski and J. Scott-Clayton. Financial aid policy: Lessons from research. *The Future of Children*, 23(1):67–91, 2013.

D. Epple, R.E. Romano, and M. Urquiola. School vouchers: A survey of the economics literature. *Journal of Economic Literature*, 55:441–492, 2017.

G. Fack and J. Grenet. Improving college access and success for low-income students: Evidence from a large need-based grant program. *American Economic Journal: Applied Economics*, 7(2):1–34, 2015.

M.M. Ferreyra, C. Avitabile, J. Botero, F. Haimovich, and S. Urzúa. *At a Crossroads: Higher Education in Latin America and the Caribbean.* The World Bank, 2017. doi: 10.1596/978-1-4648-1014-5. URL `https://elibrary.worldbank.org/doi/abs/10.1596/978-1-4648-1014-5`.

G.A. Flores-Macías. Financing security through elite taxation: The case of colombia's "democratic security taxes". *Studies in Comparative International Development*, 49(4):477–500, 2014.

J. Goodman, M. Hurwitz, and J. Smith. Access to 4-year public colleges and degree completion. *Journal of Labor Economics*, 35(3):829–867, 2017.

S. Goodman. Learning from the test: Raising selective college enrollment by providing information. *The Review of Economics and Statistics*, 98(4):671–684, October 2016.

B. Harrington. *Capital Without Borders: Wealth Management and the One Percent.* Harvard University Press, 2016.

J. Hastings, C. Neilson, and S. Zimmerman. The effects of earnings disclosure on college enrollment decisions. NBER Working Paper No. 21300, June 2015.

D. He, L. Peng, and X. Wang. Understanding the elasticity of taxable income: A tale of two approaches. 2018.

M. Hoekstra. The effect of attending the flagship state university on earnings: A discontinuity-based approach. *Journal of Economics and Statistics*, 91(4):717–724, 2009.

C. Hoxby and C. Avery. The missing "one-offs": The hidden supply of high-achieving, low-income students. *Brookings Papers on Economic Activity*, 2013.

C. Hoxby and S. Turner. Expanding college opportunities for high-achieving, low income students. SIEPR Discussion Paper No. 12-014, 2014.

G. Imbens and K. Kalyanaraman. Optimal bandwidth choice for the regression discontinuity estimators. *Review of Economic Studies*, 79(3):933–959, November 2012.

G.W. Imbens and D.B. Rubin. Estimating outcome distributions for compliers in instrumental variables models. *Review of Economic Studies*, 64:555–574, 1997.

K. Jakobsen, K. Jakobsen, H. Kleven, and G. Zucman. Wealth taxation and wealth accumulation: Theory and evidence from denmark. NBER Working Paper No. 24371, March 2018.

N. Johannesen and D.T. Larsen. The power of financial transparency: An event study of country-by-country reporting standards. *Economic Letters*, 145:120–122, 2016.

N. Johannesen and G. Zucman. The end of bank secrecy? an evaluation of the g20 tax haven crackdown. *American Economic Journal: Economic Policy*, 6(1):65–91, 2014.

N. Johannesen, P. Langetieg, D. Reck, M. Risch, and J. Slemrod. Taxing hidden wealth: The consequences of u.s. enforcement initiatives on evasive foreign accounts. NBER Working Paper No. 24366, March 2018.

I. Joumard and J. Londono-Velez. Income inequality and poverty in Colombia. Part 1: The role of the labor market. OECD Economics Department Working Paper No. 1036, April 2013.

R.D. Kahlenberg. *The Future of Affirmative Action: New Paths to Higher Education Diversity After Fisher V. University of Texas*. Century Foundation, 2014.

T.J. Kane. A quasi-experimental estimate of the impact of financial aid on college-going. NBER Working Paper No. 9703, May 2003.

H. Kleven and Schultz. Estimating taxable income responses using danish tax reforms. *American Economic Journal: Economic Policy*, 6(4):271–301, 2014.

H. Kleven and M. Waseem. Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *The Quarterly Journal of Economics*, 128(2):669–723, 2013.

M. Kremer, E. Miguel, and R. Thornton. Incentives to learn. *The Review of Economics and Statistics*, 91(3):437–456, 2009.

I. Kuziemko, M. Norton, E. Saez, and S. Stantcheva. How elastic are preferences for redistribution? Evidence from randomized survey experiments. *American Economic Review*, 105(4):1478–1508, 2015.

R. Laajaj, A. Moya, and F. Sanchez. Motivational effects of a nation-wide merit scholarship for low-income students: Quasi-experimental evidence from colombia. Documento CEDE No. 26, May 2018.

D. Langenmayr. Voluntary disclosure of evaded taxes—increasing revenue, or increasing incentives to evade? *Journal of Public Economics*, 171(7):110–125, 2017.

D.S. Lee. Training, wages and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76:1071–1102, 2009.

B. Lockwood. Malas notches. March 2018.

J. Londoño-Vélez. Diversity and redistributive preferences: Evidence from a quasi-experiment in Colombia. September 2016.

W.B. MacLeod and M. Urquiola. Reputation and school competition. *American Economic Review*, 105(11):3471–88, 2015.

W.B. MacLeod, E. Riehl, J.E. Saavedra, and M. Urquiola. The big sort: College reputation and labor market outcomes. *American Economic Journal: Applied Economics*, 9(3):223–261, 2017.

B.M. Marx and L.J. Turner. Borrowing trouble? Student loans, the cost of borrowing, and implications for the effectiveness of need-based grant aid. NBER Working Paper No. 20850, January 2015.

T. Melguizo, F. Sanchez, and T. Velasco. Credit for low-income students and access to and academic performance in higher education in Colombia: A regression discontinuity approach. *World Development*, 80:61–77, 2016.

A. Meltzer and S. Richard. A rational theory of the size of government. *Journal of Political Economy*, 89:914–927, 1981.

A. Montenegro and M. Meléndez, editors. *Equidad y movilidad social: Diagnósticos y propuestas para la transformación de la sociedad colombiana*. Universidad de los Andes, Facultad de Economía, CEDE, Ediciones Uniandes: Departamento Nacional de Planeación DNP, 2014.

J. Naritomi. Consumers as tax auditors. Working Paper, London School of Economics, 2016.

OECD. *Offshore Voluntary Disclosure: Comparative Analysis, Guidance, and Policy Advice*. OECD Publishing, September 2010.

OECD. *Update on Voluntary Disclosure Programmes: A Pathway to Tax Compliance*. OECD Publishing, August 2015.

OECD. *Education in Colombia*. 2016. doi: https://doi.org/https://doi.org/10.1787/9789264250604-en. URL `https://www.oecd-ilibrary.org/content/publication/9789264250604-en`.

OECD. *Tax Administration 2017: Comparative Information on OECD and Other Advanced and Emerging Economies*. OECD Publishing, 2017.

OECD. *The Role and Design of Net Wealth Taxes in the OECD*. OECD Publishing, 2018. OECD Tax Policy Studies No. 26.

OECD and The World Bank. *Reviews of National Policies for Education: Tertiary Education in Colombia*. OECD Publishing, 2012.

A. Pallais. Taking a chance on college: Is the tennessee education lottery scholarship program a winner? *Journal of Human Resources*, 44(1):199–222, 2009.

A. Pallais. Small differences that matter: Mistakes in applying to college. *Journal of Labor Economics*, 33(2):493–520, 2015.

T. Piketty. Social mobility and redistributive politics. *The Quartlery Journal of Economics*, 111:551–584, 1995.

T. Piketty. *Capital in the Twenty-First Century*. Harvard University Press, 2014.

D. Pomeranz. No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review*, 185(8):2538–2569, 2015.

G. Rao. Familiarity does not breed contempt: Diversity, discrimination, and generosity in Delhi Schools. Job Market Paper, December 2013.

E. Riehl. Assortative matching and complementarity in college markets. Job Market Paper, November 2016.

E. Riehl, J.E. Saavedra, and M. Urquiola. Learning and earning: An approximation to college value added in two dimensions. NBER Working Paper No. 22725, October 2016.

J.E. Saavedra. The learning and early labor market effects of college quality: A regression discontinuity analysis. Job Market Paper, 2009.

E. Saez. Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3):180–212, 2010.

E. Saez and G. Zucman. Wealth inequality in the united states since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics*, 131(2):519–578, 2016.

E. Saez, J. Slemrod, and S.H. Giertz. The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, 50(1):3–50, 2012.

F. Sanchez and O. Alvarez. La informalidad laboral y los costos laborales en colombia 1984-2009: Diagnostico y propuestas de politica. 2011.

F. Sanchez and J. Nunez. A dynamic analysis of human capital, female workforce participation, returns to education and changes in household structure in urban colombia, 1976-1998. *Colombian Economic Journal*, 1(1):109–149, 2003.

F. Sanchez and T. Velasco. Do loans for higher education lead to better salaries? Evidence from a regression discontinuity approach for Colombia. , October 2014.

J. Scott-Clayton and B. Zafar. Financial aid, debt management, and socioeconomic outcomes: Post-college effects of merit-based aid. NBER Working Paper 22574, August 2016.

D. Seim. Behavioral responses to wealth taxes: Evidence from sweden. *American Economic Journal: Economic Policy*, 9(4):395–421, 2017.

J. Slemrod. *Tax Compliance and Enforcement: New Research and its Policy Implications*, pages 81–102. Oxford University Press, 2017.

A. Solis. Credit access and college enrollment. *Journal of Political Economy*, 125(2):562–622, 2017.

M. Urquiola. In E. Hanushek, S. Machin, and L. Woessmann, editors, *Handbook of the Economics of Education*, volume 5, chapter Competition Among Schools: Traditional Public and Private Schools, pages 209–237. 2016.

V.C. Wong, P.M. Steiner, and T.D. Cook. Analyzing regression-discontinuity designs with multiple assignment variables: A comparative study of four estimation methods. *Journal of Educational and Behavioral Statistics*, 38(2):107–141, 2013.

S. Zimmerman. The returns to college admission for academically marginal students. *Journal of Labor Economics*, 32(4):711–754, 2014.

S. Zimmerman. Making the one percent: The role of elite universities and elite peers. NBER Working Paper No. 22900, December 2016.

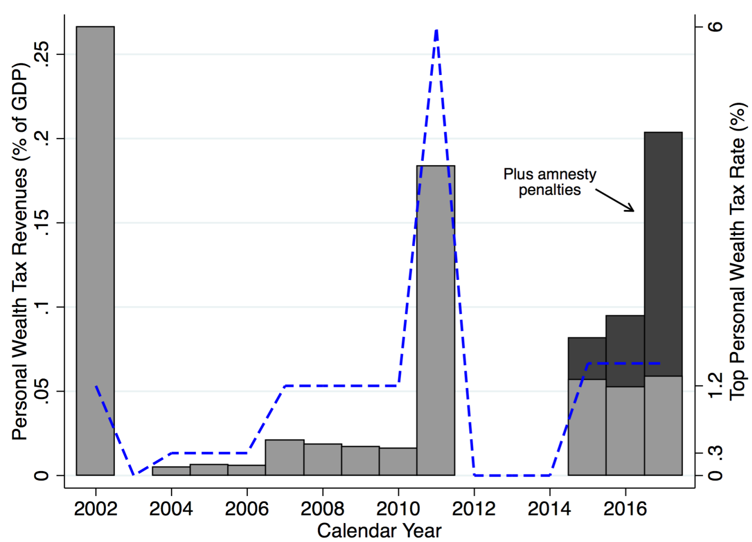
F.T. Zoutman. The effect of capital taxation on household savings. Working Paper, 2015.

G. Zucman. *The Hidden Wealth of Nations: The Scourge of Tax Havens.* University of Chicago, 2015.

Appendix A

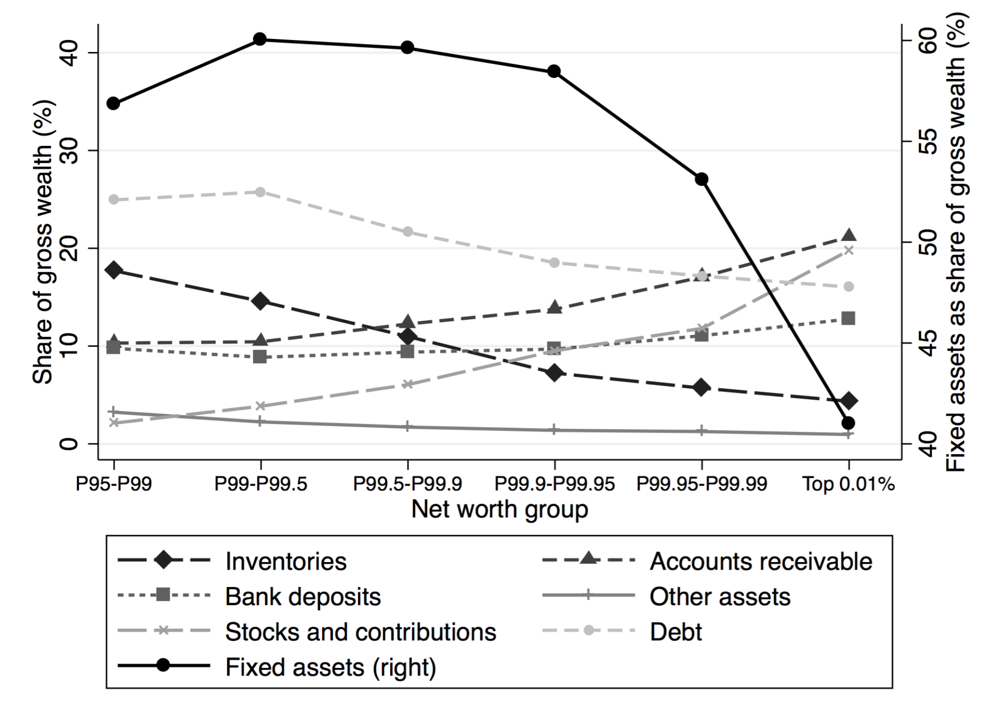
A.1 Appendix Figures and Tables from Chapter 1

Figure A.1: Personal Wealth Tax and Voluntary Disclosure Program Revenues



Notes: This figure plots personal wealth tax revenue (gray bars) plus disclosure penalty revenue (black bars) relative to GDP, and the statutory top personal wealth tax rate (dashed blue line) between calendar years 2002 and 2017. The figure shows that tax revenues on recurrent personal wealth taxes represented between 0 and 0.27% of GDP in Colombia during this period. As a comparison, in 2016 the equivalent share was 0.18% in Spain, 0.22% in France, 0.43% in Norway, and 1.0% in Switzerland (OECD, 2018). Total penalty revenues collected between 2015 and 2017 represent 0.21% of GDP. Wealth tax revenues do not systematically increase in 2016 and 2017 for non-disclosers because (i) eligibility for FY 2015–2018 is determined by net wealth held January 1, 2015 only, and (ii) rules regarding year-to-year changes in tax base. For instance, if net worth is 1,000 million in FY 2015 and 2,000 million in FY 2016 (i.e., an increase), FY 2016 wealth tax base is 1,000 million $\times(1+0.25i)$ where i is 2015 inflation rate. If net worth is 2,000 million in FY 2015 and 1,000 million in FY 2016 (i.e., a decrease), FY 2016 wealth tax base is 2,000 million $\times(1-0.25i)$. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

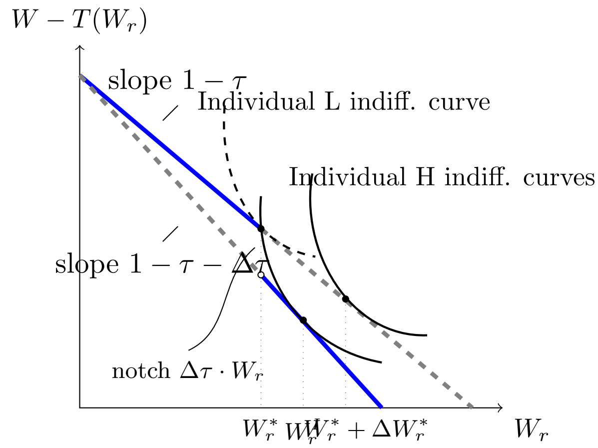
Figure A.2: Wealth Decomposition by Top Wealth Groups: Business Owners



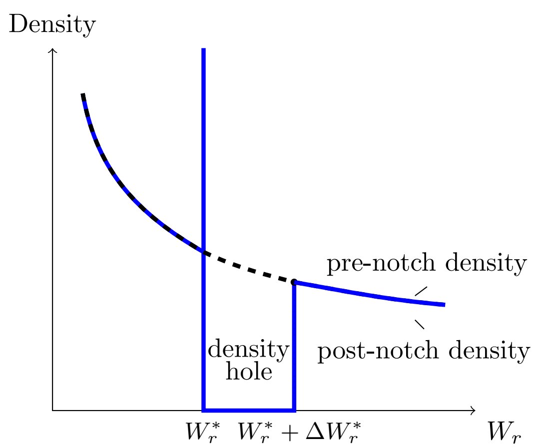
Notes: This figure plots mean wealth items as a share of gross wealth by top percentile groups. The figure suggests fixed assets (e.g., real estate, land) represent 56.9 percent of total wealth for individuals in P95–P99, but only 41.0 percent for the top 0.01 percent. In contrast, stocks and contributions, and accounts receivable are increasing in wealth, representing around 40 percent of wealth among the wealthiest 0.01 percent. Top wealth groups are generated ranking all income tax filers in 2016 by their reported net worth. The shares of wealth items are computing using the sample of business owners required to keep records and filing income tax form 110, because wealth decomposition is only available for these taxpayers.

Figure A.3: Behavioral Responses to a Tax Notch

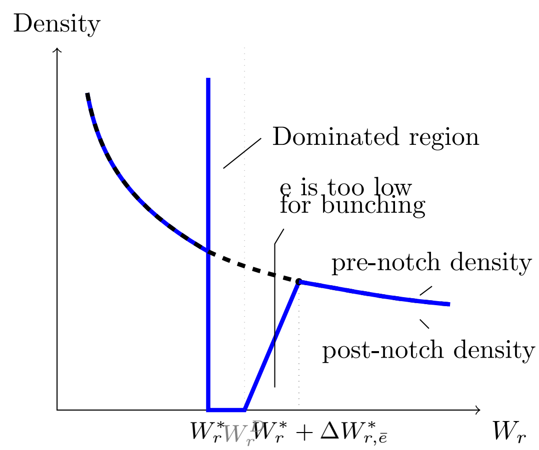
(a) Budget Sets



(b) Density Distributions in the Baseline Model (Homogeneous Elasticities and No Frictions)

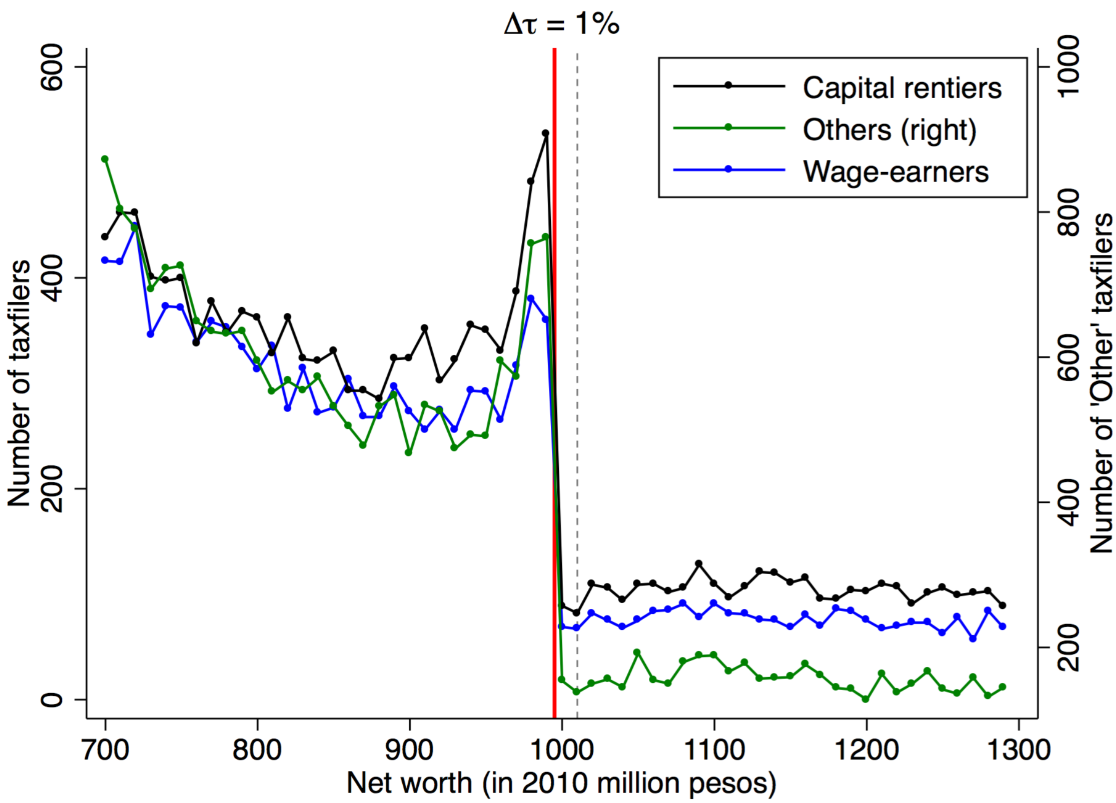


(c) Heterogeneity in Elasticities



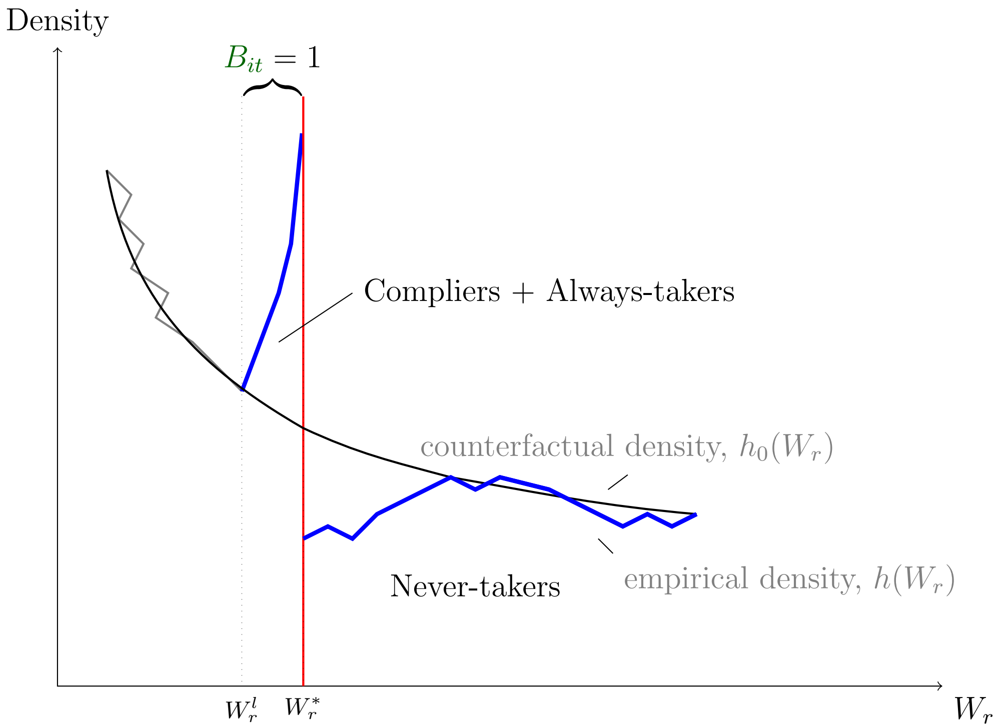
Notes: These figures illustrate the implications of a proportional tax notch ($\Delta\tau > 0$) in a budget diagram in Panel (a), and a density distribution diagram in Panels (b) and (c). Panel (a) shows the budget constraint of two individuals, L and H, assuming $W_r \leq W$. L has the lowest pre-notch reported wealth W_r (lowest true wealth W) among those who locate at the point; she chooses W_r^* both before and after the tax change. Individual H has the highest pre-notch reported wealth (highest true wealth) among those who locate at the notch point; she chooses reported wealth $W_r^* + \Delta W_r^*$ before the tax change and is exactly indifferent between the notch point W_r^* and the interior point W_r^I after the tax change. Panel (b) shows the corresponding distribution of net worth in the presence of such tax schedule in the baseline model, under homogeneous elasticities and no optimization frictions. There is bunching at the notch point by all individuals between L and H, i.e., who have reported wealth in an interval $(W_r^*, W_r^* + \Delta W_r^*]$. Panel (c) extends the baseline model to allow for heterogeneity in elasticities. Individual density is empty in the strictly dominated range $(W_r^*, W_r^D]$ and then increases gradually until it converges with the pre-notch density at $W_r^* + \Delta W_{r,\bar{e}}^*$. Tax filers with reported wealth in the range $(W_r^D, W_r^* + \Delta W_{r,\bar{e}}^*]$ do not bunch because their elasticity is too low. Figure 1.3 extends Panel (c) to incorporate optimization frictions.

Figure A.4: Wealth Bunching in 2010 by Economic Activity



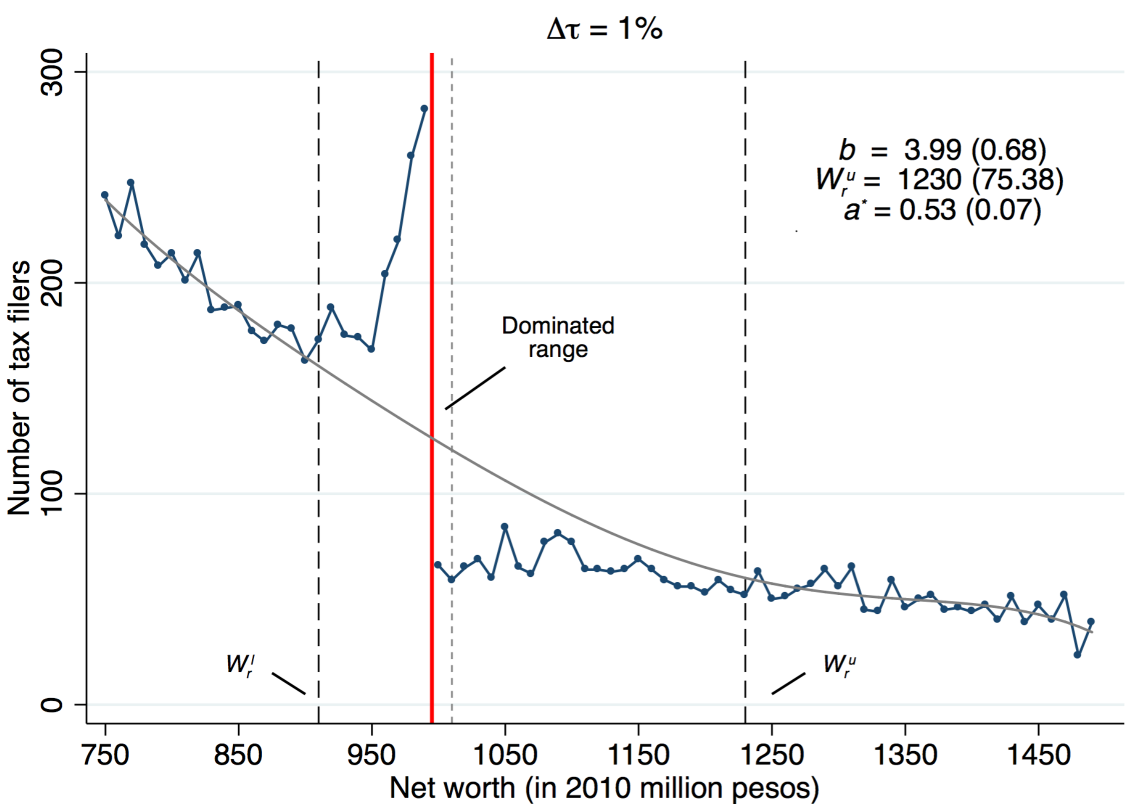
Notes: This figure shows heterogeneity in bunching among various samples of taxpayers that have different opportunities to shelter wealth from taxation. The three samples are built based on individuals' self-reported economic activity code: capital rentiers (in black), wage-earners (in blue), and others (the remainder, in green). The figure shows capital rentiers display more bunching behavior than other taxfilers, while wage-earners display the least bunching. Bin width is 2010 10,000,000 pesos (2010 USD 5,208.30 in 12/31/2010).
Source: Authors' calculations using administrative tax microdata from DIAN.

Figure A.5: Compliers (i.e., bunchers) analysis



Notes: This figure illustrates the analysis of compliers characteristics in the context of bunching in response to discontinuities in a tax schedule. The blue line represents the (hypothetical) empirical density of taxpayers by reported wealth W_r. The black solid line represents the counterfactual density. $B_{it} = 1$ if individual i is located in bunching range $[W_r^l, W_r^*)$ in year t. Tax filers located in this range will be a mix of compliers (i.e., those who react to tax notches by bunching below the cutoff) and always-takers (i.e., those who would be located in that range even in the absence of tax notches). Tax filers located above W_r^* are never-takers (i.e., those who will not or cannot bunch in response to the tax notch). A 2SLS-IV specification is as follows: $Y_{it} = \alpha_1 + \gamma_1 t + \beta_1 B_{it} + \epsilon_{it}$, where t is a time trend and B_{it} is instrumented with the post-reform dummy $Z_{it} = \mathbb{1}(t = 2010)$. The results of this analysis are presented in Table 1.2 for individuals around 2010's wealth tax eligibility cutoff.

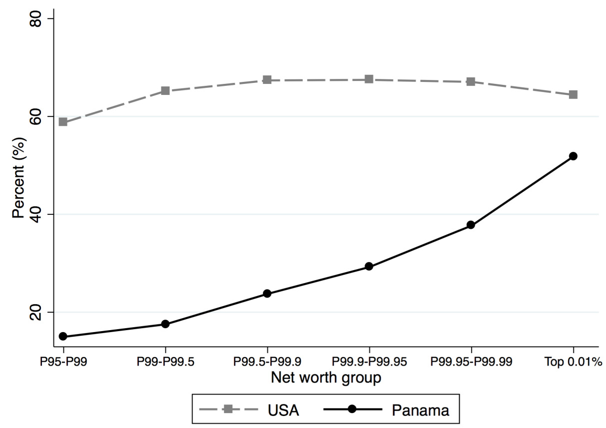
Figure A.6: Bunching Among Business Owners Required to Keep Records



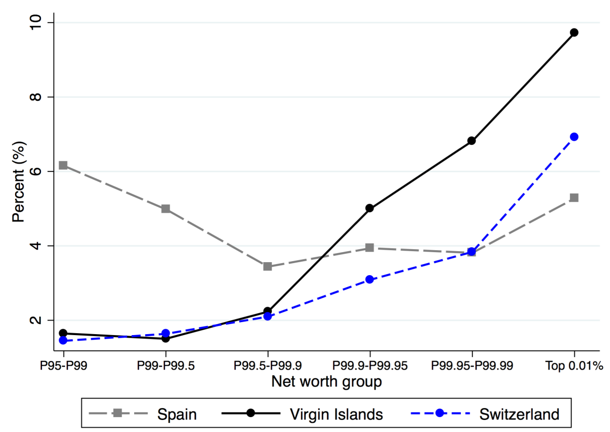
Notes: This figure displays the density of taxpayers required to keep records by their net worth in 2010, and fits the counterfactual distribution. b is the excess mass as a share of the counterfactual and W_r^u the net worth of the marginal buncher, obtained with the point of convergence method. The counterfactual is obtained from the regression of a polynomial of degree 5 on all data points outside the $[W_r^l, W_r^u]$ interval. The lower bound W_r^l is determined visually. The upper bound W_r^u is estimated from an iterative process: starting from $W_r^u = W_r^*$, we obtain the counterfactual and estimate the excess mass (B) below the threshold and missing mass (M) above the threshold. For low W_r^u, the excess mass is larger than the missing mass ($B >> M$). We iteratively increase W_r^u until the two masses are equal ($B = M$). a^* represents the share of individuals in the dominated range that do not bunch due to adjustment costs. Bootstrapped standard errors in parentheses are estimated from 1,000 bootstrap samples with replacement. Bin width is 10,000,000 pesos (2010 USD 5,208.30 in 12/31/2010). *Source:* Authors' calculations using administrative tax microdata from DIAN.

Figure A.7: Location of Foreign Assets according to Reports of Foreign Assets in 2017

(a) Two Most Popular Destinations

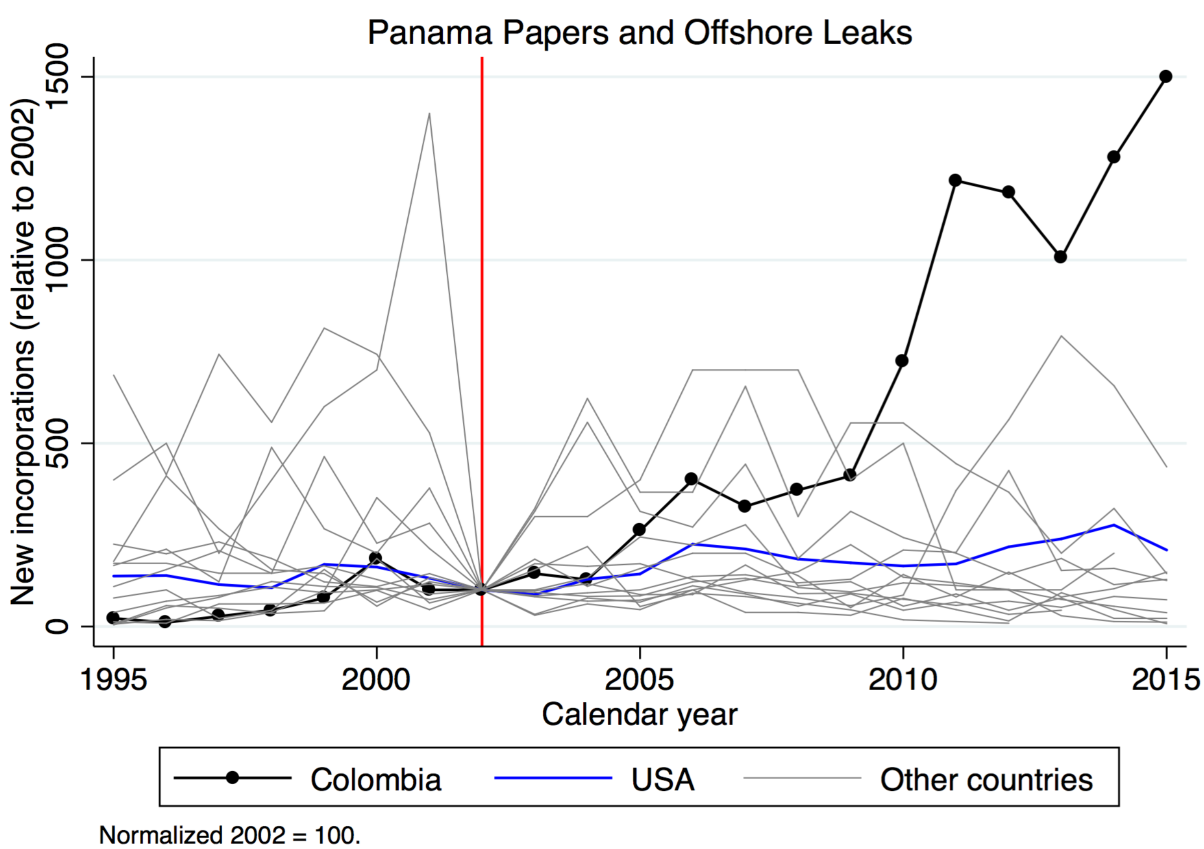


(b) Next Three Most Popular Destinations



Notes: These figures show the likelihood of reporting a foreign asset located in a given location for taxpayers filing a foreign asset return (form #160) in FY 2017. The sample is restricted to 2,076,685 individuals filing either the FY 2016 income tax return or FY 2017 a wealth tax return. This sample includes 29,183 taxpayers reporting foreign assets. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

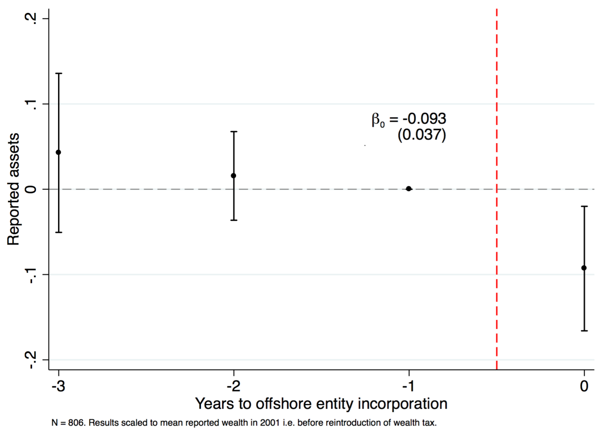
Figure A.8: The Use of Offshore Entities: Colombia versus Other Countries



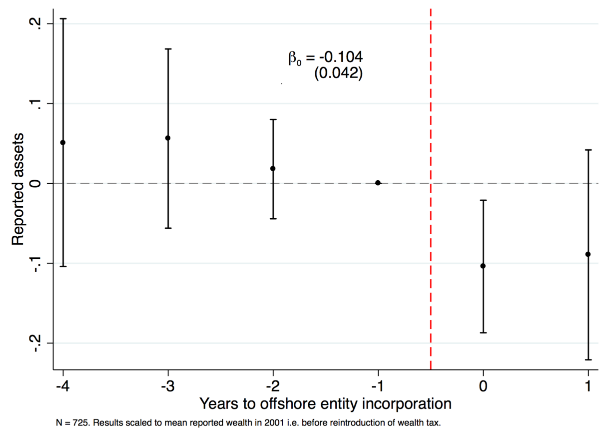
Note: This figure compares the dynamics of offshore entity incorporation between Colombia, the United States, and thirteen other countries. To facilitate the comparison, the number of offshore entities is expressed relative to 2002 (the year annual wealth taxation begins in Colombia). The comparison is restricted to countries that have not themselves been considered tax havens (e.g., Panama, British Virgin Islands, United Arab Emirates) and have at least 100 active or inactive offshore entities by December 31, 2002. The countries in gray are Brazil, Canada, Costa Rica, France, Germany, Greece, Ireland, Israel, Russia, Spain, United Kingdom, Uruguay, and Venezuela. Both active and inactive offshore entities are included. The high number of new offshore corporations in 2001 corresponds to Germany. *Source:* ICIJ. Accessed June 12, 2017.

Figure A.9: Robustness Check: The Effect of Opening an Offshore Entity on Assets Reported to the Tax Authority Using Different Event Time Windows

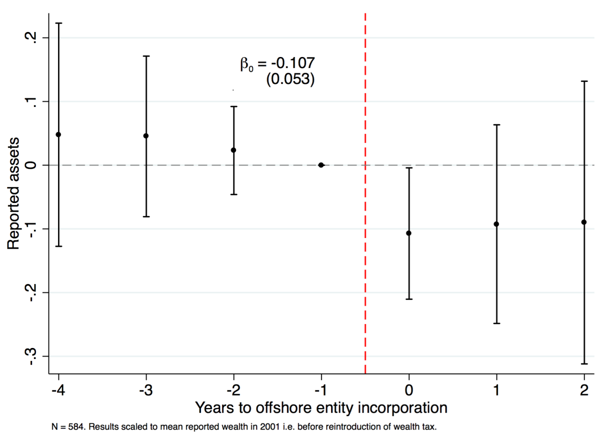
(a) Balanced in $k \in [-3, 0]$



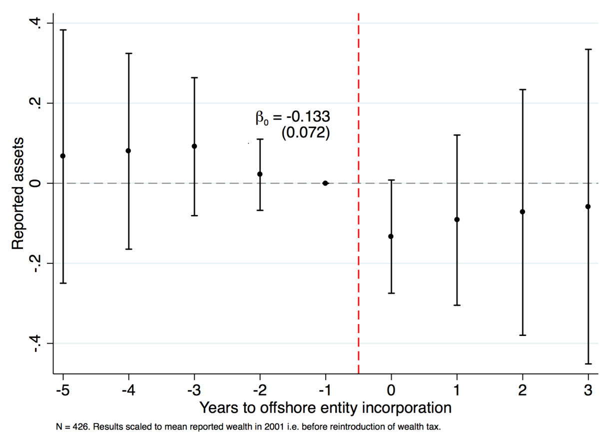
(b) Balanced in $k \in [-4, 1]$



(c) Balanced in $k \in [-4, 2]$

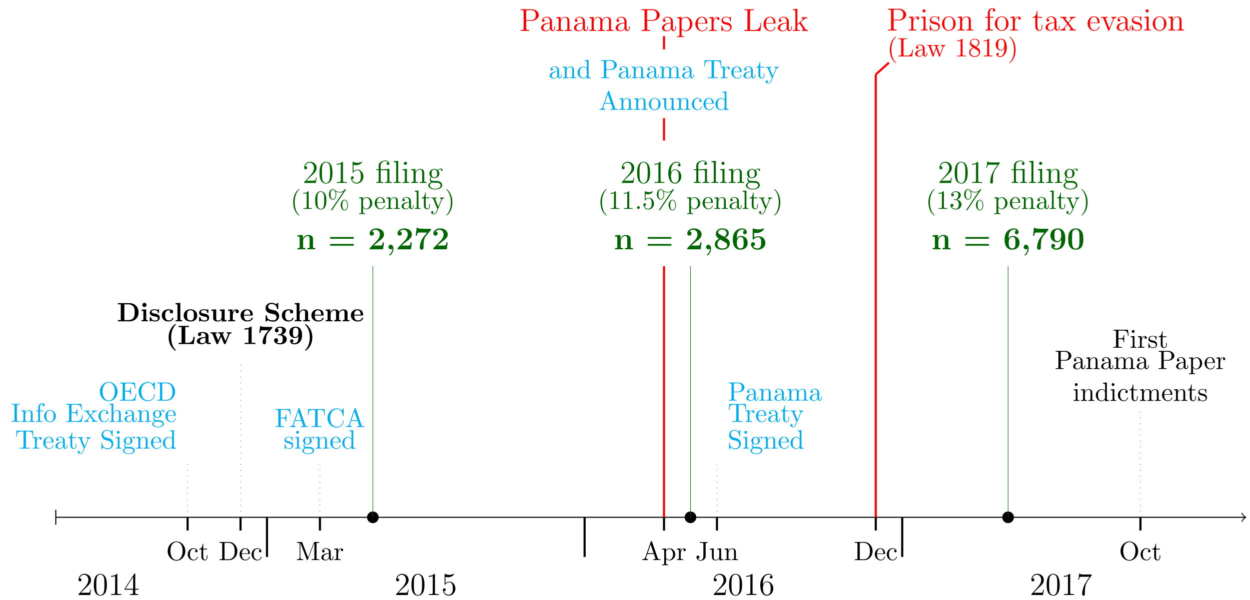


(d) Balanced in $k \in [-5, 3]$



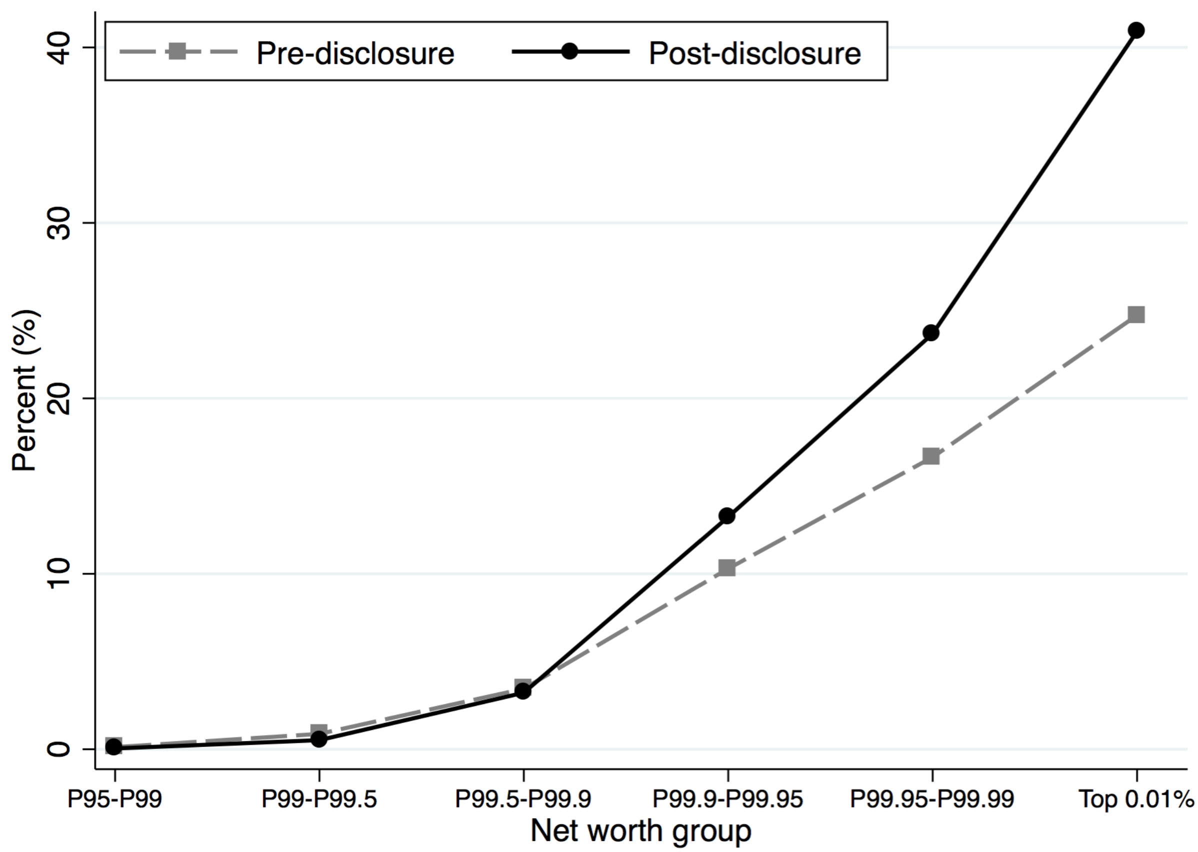
Notes: These figures present robustness checks on the β_k coefficients from event study specification (1.7), varying the event time window across balanced samples of taxpayers in the Panama Papers. The outcome variable is total assets reported to the Colombian tax authority, scaled to the mean wealth in 2001 (i.e., the year before annual taxation of net wealth is re-introduced in Colombia). An event is defined as the year a taxpayer incorporates an offshore entity through Mossack Fonseca, for taxpayers with only one offshore entity. *Sources:* Authors' calculations using administrative tax microdata from DIAN and ICIJ.

Figure A.10: Timeline of Events



Notes: This figure plots a timeline of events taking place around Colombia's voluntary disclosure scheme between 2014 and 2018. n (in green) represents the number of individuals first disclosing hidden wealth under the scheme.

Figure A.11: Probability of Participating in the 2015–17 Voluntary Disclosure Scheme, by Pre- and Post-Disclosure Wealth Group

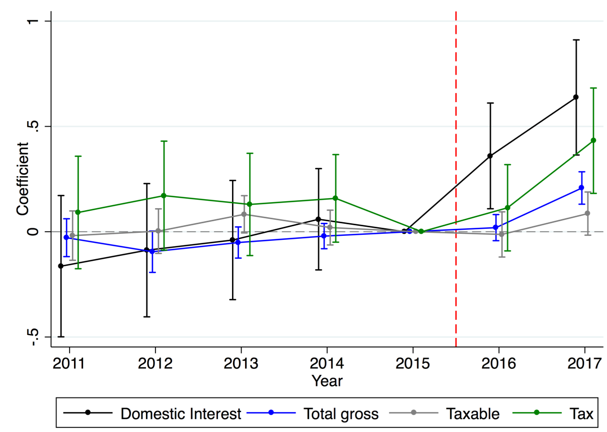


Notes: This figure plots the fraction of tax units in Colombia that participate in the 2015–17 voluntary disclosure program by bins of reported net worth. The figure ranks individuals by pre- and post-disclosure net worth, and shows that participation in the program is increasing in both measures of wealth. Ranking by pre-disclosure wealth, 24.7 percent of individuals in the wealthiest 0.01 percent disclosed hidden wealth (dashed gray line). Ranking by wealth including disclosures, 40.9 percent of individuals in the wealthiest 0.01 percent disclosed (solid black line). The sample is restricted to 1,633,383 individuals filing the income tax return in FY 2013 (they may or not file a wealth tax return in 2015–2017), and includes 11,210 disclosers and 1,085 individuals in the Panama Papers (of which 434 disclosed wealth). *Sources:* Authors' calculations using administrative tax microdata from DIAN.

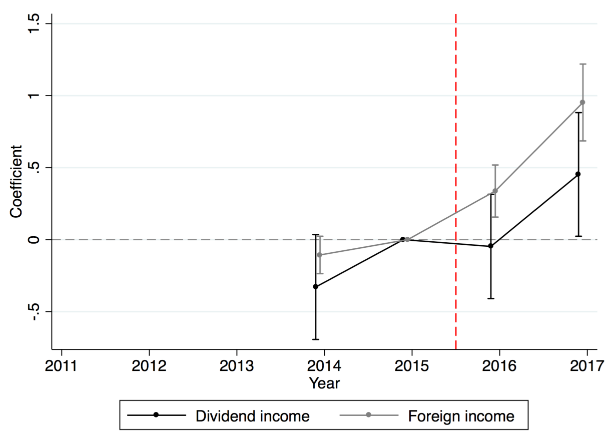
Figure A.12: The Impact of a Voluntary Disclosure Program on Reported Wealth and Income: 2016 Disclosers

(a) Wealth: Gross and Net

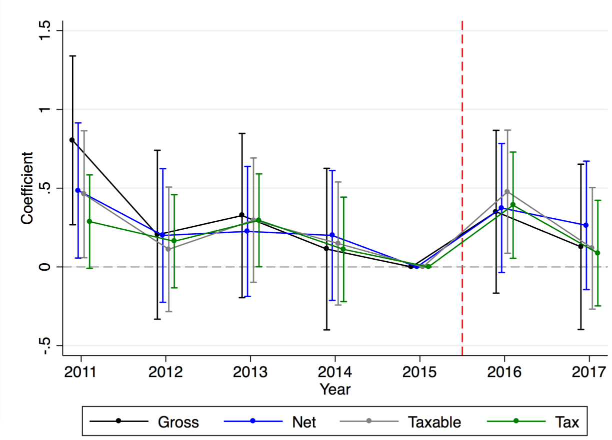
(b) Regular Income: Interest, Total Gross, Taxable, and Tax



(c) Regular Income: Dividend and Foreign



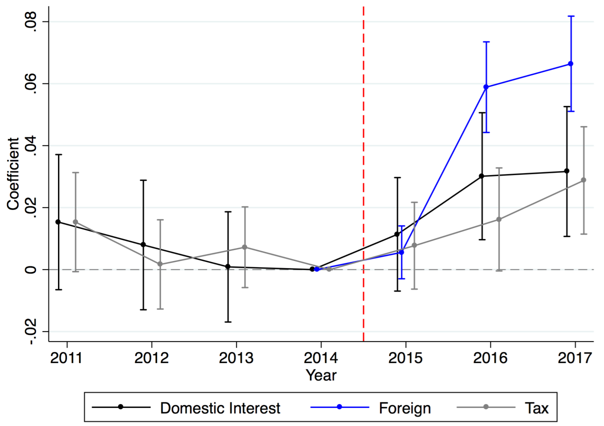
(d) Irregular Income (Includes Long-Rerm Realized Capital Gains): Gross, Net, Taxable, and Tax



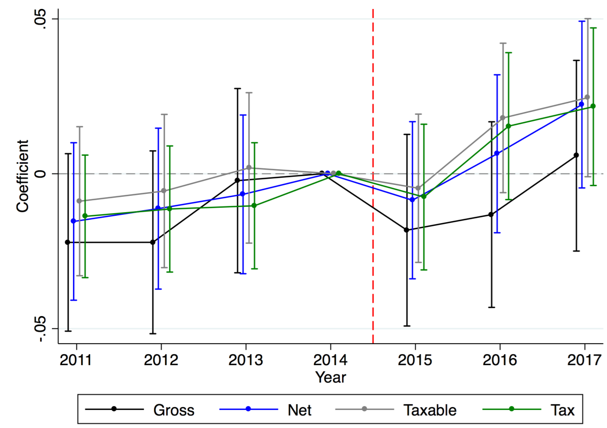
Notes: These figures compare outcomes between 2,074 taxpayers that first disclosed hidden assets or inexistent liabilities in 2016 and 43,181 that never disclosed between 2015 and 2017. The inverse hyperbolic sine transformation of a given outcome is regressed on individual fixed effects and a voluntary discloser dummy interacted with year fixed effects (2015 is the omitted category). The standard errors are clustered at the taxpayer level. The figures plot the coefficients on the interaction terms and 95 percent confidence intervals. The red vertical line marks the period taxpayers first disclosed their hidden assets and fake debts. The sample is a balanced panel of 45,255 individuals that both filed income taxes annually between 2011 and 2017, and filed the wealth tax in 2015, 2016, or 2017. Tax filers that first disclosed assets and liabilities in either 2015 or 2017 are excluded from the estimation sample. The corresponding difference-in-differences coefficients are presented in Table A.4. *Sources: Sources:* Authors' calculations using administrative tax microdata from DIAN.

Figure A.13: The Impact of a Voluntary Disclosure Program on the Probability of Reporting Strictly Positive Values of Capital Income

(a) Regular Income: Interest, Foreign, and Tax

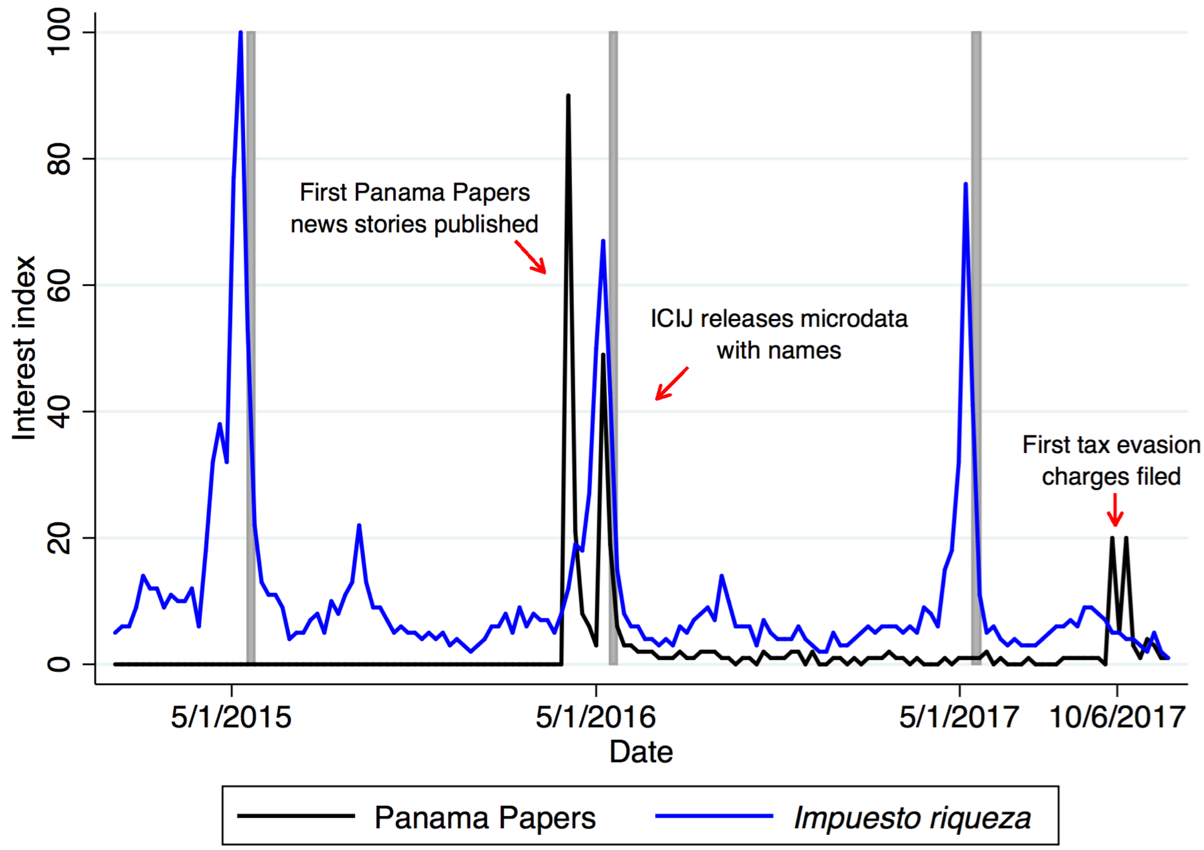


(b) Irregular Income (Includes Long-Rerm Realized Capital Gains): Gross, Net, Taxable, and Tax



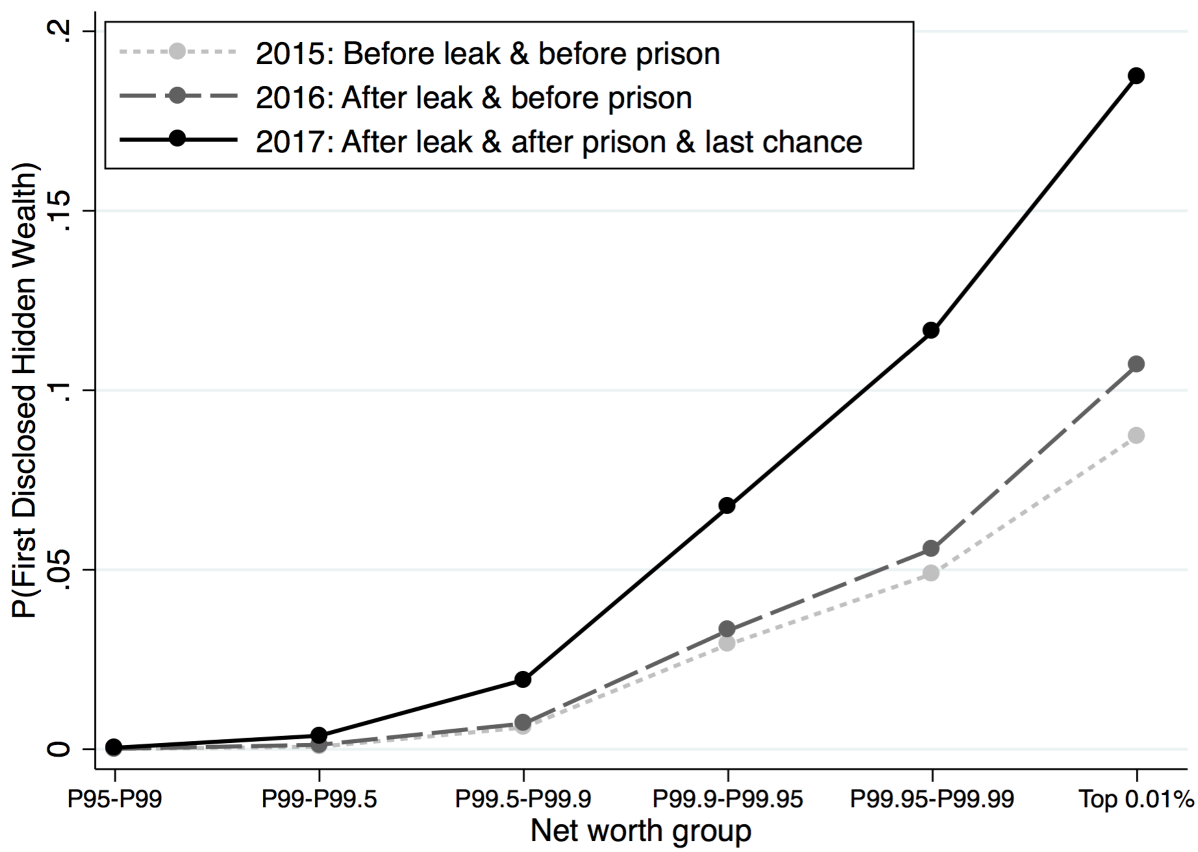
Notes: These figures present the effect of the 2015 voluntary disclosure program on the probability of reporting positive values of selected regular income categories in Panel (a) and irregular income categories in Panel (b). The figures compare outcomes between 1,777 taxpayers that voluntarily disclosed hidden assets and inexistent liabilities in 2015 and 43,181 that never disclose their assets and liabilities between 2015 and 2017. The outcome, which is a dummy for reporting strictly positive values, is regressed on individual fixed effects and a voluntary discloser dummy interacted with year fixed effects (2014 is the omitted category). The standard errors are clustered at the taxpayer level. The figures plot the coefficients on the interaction terms and 95 percent confidence intervals. The red vertical line marks the period taxpayers disclosed their hidden assets and fake debts. The sample is a balanced panel of 44,958 individuals that both filed income taxes annually between 2011 and 2017, and filed the wealth tax in 2015, 2016, or 2017. Tax filers that first disclosed hidden assets or fake liabilities *after* 2015 (i.e., in 2016 or 2017) are excluded from the estimation sample. The corresponding difference-in-differences coefficients are presented in Table A.3, Panel B. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Figure A.14: Comparison on Google Trends of Search Terms "Panama Papers" and "Impuesto Riqueza" (Wealth Tax) in Colombia



Notes: This figure plots the relative search interest in Colombia of the terms "Panama Papers" in black and "Impuesto riqueza" (wealth tax, in Spanish) in blue. The number in the y-axis represents search interest relative to the highest point on the chart for Colombia during the plotted period of time. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular, while a score of 0 means the term was less than 1 percent as popular as the peak. The gray bars represent the annual wealth tax filing season. The first Panama Papers news stories were published April 3, 2016. On May 9, 2016, the ICIJ released the database revealing the names and contact addresses of thousands of shareholders of offshore entities. The *fiscalía*—the Colombian equivalent of the district attorney's office—filed the first charges related to the Panama Papers on October 4, 2017. Nineteen individuals were charged for illicit enrichment, fraud, and money laundering, among others. *Source:* Google Trends, accessed November 30, 2017. Click here.

Figure A.15: Probability of Participating in the Voluntary Disclosure Program, by Post-Disclosure Wealth and Year of First Disclosure



Notes: This figure plots the fraction of tax units in Colombia that file taxes and participate in the voluntary disclosure program, by bins of net worth (including disclosures) and the year in which they first disclosed hidden wealth. The figure shows that (1) disclosing hidden wealth is increasing in net worth; and (2) the wealthiest taxpayers disclosed in 2017. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Table A.1: Compliers (i.e., bunchers) analysis for the fourth notch in 2010

	$E[Y(0)\|\text{Compliers}]$	$E[Y(1)-Y(0)\|\text{Compliers}]$	$E[Y\|\text{Always-takers}]$	$E[Y\|\text{Never-takers}]$
	(1)	(2)	(3)	(4)
Debt	0.112	0.061	0.1	0.096
	(0.047)	(0.039)	(0.012)	(0.015)
Inventories	-0.004	-0.043	0.078	0.043
	(0.033)	(0.049)	(0.021)	(0.015)
Bank deposits	0.133	-0.039	0.13	0.068
	(0.073)	(0.076)	(0.032)	(0.019)
Stock	0.01	-0.088	0.266	0.271
	(0.135)	(0.092)	(0.042)	(0.048)
Fixed assets	0.404	0.092	0.364	0.351
	(0.16)	(0.118)	(0.05)	(0.052)
Accounts receivable	0.458	0.085	0.158	0.261
	(0.161)	(0.117)	(0.041)	(0.049)
Other assets	-0.001	-0.008	0.004	0.005
	(0.013)	(0.015)	(0.002)	(0.005)

Notes: This table presents the results of a compliers analysis using the set-up illustrated by Figure A.5. In this setting, a complier refers to a taxpayer bunching below the exemption cutoff in response to the wealth tax. The sample is a balanced panel of 241 income tax filers reporting net wealth between $W_r^l = 4.5$ billion and $W_r^u = 5.8$ billion in 2008, 2009, and 2010. The endogenous variable is $B_{it} = 1$ if the individual has net wealth (in 2010 pesos) between W_r^l and W^*, i.e., the bunching region. Complier means in Column (1) are calculated as the coefficient on $1 - B_{it}$ in a 2SLS regression of $1 - B_{it}$ multiplied by Y_i and using 2010 as the instrument (Z_{it}). Always-taker and never-taker means are calculated in analogous 2SLS regressions of $B_{it}(1 - Z_{it})Y_{it}$ on $B_{it}(1 - Z_{it})$ and $(1 - B_{it})Z_{it}Y_{it}$ on $(1 - B_{it})Z_{it}$, respectively, again using 2010 as Z_{it}. The first stage coefficient is 0.139 (t-stat 3.88) for debt, and 0.19 (t-stat 2.06) for all others, as only business owners keeping records report asset types separately. Standard errors are clustered at the taxpayer level. *Source:* Authors' calculations using administrative tax microdata from DIAN.

Table A.2: Location and Type of Foreign Assets Reported in 2017: Amnesty Disclosers vs Non-disclosers

	Dependent variable					
	Tax Haven	Type of Asset				
		Bank Deposits	Portfolio Securities	Trusts	Real Estate	Cars
	(1)	(2)	(3)	(4)	(5)	(6)
Discloser	0.205***	0.031***	0.171***	0.028***	-0.081***	-0.018***
	(0.008)	(0.008)	(0.008)	(0.003)	(0.006)	(0.002)
Constant	0.323***	0.376***	0.401***	0.015***	0.187***	0.024***
	(0.005)	(0.006)	(0.006)	(0.001)	(0.005)	(0.002)
N	14,387	14,387	14,387	14,387	14,387	14,387
R^2	0.043	0.002	0.029	0.007	0.013	0.006

Notes: This table compares the location and type of foreign assets held by disclosers and non-disclosers for taxpayers reporting to own any foreign asset in 2017 (information tax return #160). Each column represents a separate regression with a different dependent variable obtained from the foreign asset information return. The dependent variable in Column (1) is an indicator for declaring a foreign asset located in a tax haven (Barbados, Bermuda, Cayman Islands, Curacao, Monaco, Panama, Switzerland, Uruguay, or the Virgin Islands). The outcomes in Columns (2)–(6) are indicator variables for reporting each type of foreign asset. This information is available only for taxpayers with foreign assets above approximately USD 40,000. Portfolio securities refer to portfolios of equities, bonds, and mutual fund shares owned by taxpayers on foreign accounts. The dependent variable is regressed on a dummy for having disclosed a foreign asset during the 2015–17 wealth disclosure program. Robust standard errors in parentheses. The sample is restricted to individuals having (1) filed a wealth tax return in either 2015, 2016, or 2017, and (2) filed a foreign asset information return in 2017. $*p < 0.1, **p < 0.05, ***p < 0.01$. *Source:* Authors' calculations using administrative tax microdata from DIAN.

Table A.3: Robustness Checks: The Effect of a Voluntary Wealth Disclosure Program on Reported Income and Wealth

	Wealth		Regular Income						Irregular income
	Gross	Net	Foreign	Dividend	Interest	Total gross	Taxable	Tax	Tax
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A. ArcSinh: no winsorizing								
DID	0.288***	0.400***	1.010***	0.061	0.416***	0.132***	0.132***	0.329***	0.339***
	(0.015)	(0.030)	(0.105)	(0.170)	(0.117)	(0.030)	(0.034)	(0.090)	(0.100)
C Mean	22.018	21.823	0.324	7.622	11.297	20.089	18.741	15.205	1.031
N	314,706	314,706	138,004	138,004	314,706	314,706	314,706	314,706	314,706
R^2	0.66	0.572	0.614	0.753	0.629	0.686	0.547	0.641	0.242
	Panel B. Dummy for strictly positive values								
DID	0	0.001	0.044***	-0.002	0.018***	0.001	0.001	0.012**	0.019***
	(0.000)	(0.001)	(0.005)	(0.008)	(0.007)	(0.001)	(0.002)	(0.005)	(0.006)
C Mean	1.000	0.999	0.0138	0.346	0.721	0.996	0.992	0.906	0.066
N	314,706	314,706	179,832	179,832	314,706	314,706	314,706	314,706	314,706
R^2	0.176	0.256	0.604	0.764	0.567	0.507	0.403	0.558	0.24
	Panel C. Levels (in million pesos)								
DID	2021.490***	1586.210***	8.298***	80.673***	5.844*	-60.53	24.642**	7.748**	-0.299
	(681.692)	(180.578)	(1.500)	(28.006)	(3.116)	(193.269)	(11.324)	(3.733)	(1.263)
C Mean	2489.714	2038.660	5.409	72.012	20.842	890.369	123.418	26.409	1.466
N	314,706	314,706	138,004	138,004	314,706	314,706	314,706	314,706	314,706
R^2	0.791	0.815	0.506	0.609	0.248	0.685	0.709	0.706	0.161
	Panel D. Levels: winsorizing at top 0.1% each year (in million pesos)								
DID	1326.350***	1413.511***	7.808***	42.251***	7.858***	43.151	15.431***	4.707***	0.727**
	(81.334)	(75.080)	(1.282)	(11.549)	(1.595)	(45.563)	(3.602)	(1.175)	(0.329)
C Mean	2432.978	1993.944	4.485	64.967	19.037	831.345	121.943	25.921	0.900
N	314,706	314,706	138,004	138,004	314,706	314,706	314,706	314,706	314,706
R^2	0.874	0.863	0.586	0.648	0.581	0.843	0.741	0.738	0.189
	Panel E. Levels: winsorizing at top 1% each year (in million pesos)								
DID	1116.518***	1176.559***	2.131***	21.910***	5.226***	41.387*	15.269***	4.642***	0.449***
	(54.471)	(49.964)	(0.235)	(5.390)	(0.940)	(22.634)	(2.613)	(0.843)	(0.109)
C Mean	2347.821	1925.880	0.543	53.249	16.750	732.548	116.873	24.245	0.438
N	314,706	314,706	138,004	138,004	314,706	314,706	314,706	314,706	314,706
R^2	0.874	0.866	0.592	0.667	0.63	0.845	0.772	0.768	0.208

Notes: See notes from Table 1.5. The outcome variables are in log-approximation form in Panel A (inverse hyperbolic sine function), a dummy for strictly positive values in Panel B, and in levels in Panels C–E. Panels D and E winsorize the outcome variables by replacing all values above the 99.9th and 99th percentile of the outcome variable by the 99.9th and 99th percentile value, respectively. *Source:* Authors' calculations using administrative tax microdata from DIAN.

Table A.4: The Impact of a Voluntary Disclosure Program on Reported Wealth and Income: 2016 Disclosers

	Wealth		Income						Capital gains and other irregular income			
	Gross (1)	Net (2)	Foreign (3)	Dividend (4)	Interest (5)	Total gross (6)	Taxable (7)	Tax (8)	Gross (9)	Net (10)	Taxable (11)	Tax (12)
DID	0.307***	0.343***	0.697***	0.366**	0.545***	0.153***	0.019	0.163*	-0.051	0.097	0.094	0.068
	(0.014)	(0.023)	(0.079)	(0.142)	(0.102)	(0.029)	(0.040)	(0.086)	(0.148)	(0.121)	(0.115)	(0.096)
N	316,785	316,785	139,725	139,725	316,785	316,785	316,785	316,785	316,785	316,785	316,785	316,785
R^2	0.66	0.576	0.605	0.754	0.628	0.685	0.551	0.643	0.264	0.245	0.241	0.24

Notes: This table presents the effects of the voluntary disclosure program on the logarithm of income and wealth reported to the Colombian tax authority. The dependent variables in columns (1) and (2) are taken from the wealth tax form 440, while those in columns (3)–(12) are taken from the individual income tax forms 110 and 210. Outcomes are expressed in log-approximation form using the inverse hyperbolic sine function. The table compares outcomes between 2,074 taxpayers that first disclosed hidden assets and inexistent liabilities in 2016 and 43,181 that did not disclose between 2015 and 2017. Each outcome is regressed on individual fixed effects, year fixed effects, and an interaction of the voluntary discloser dummy and post-disclosure years: $\log(y_{it}) = \alpha_i + \gamma_t + \beta \cdot \mathbb{1}(\text{Post} \times \text{Discloser}) + \nu_{it}$. The standard errors in parentheses are clustered at the taxpayer level. The sample is a balanced panel of 45,255 individuals that both filed income taxes annually between 2011 and 2017, and filed the wealth tax in 2015, 2016, or 2017. Tax filers that first disclosed assets and liabilities in 2015 or 2017 are excluded from the estimation sample. The number of observations with foreign income and dividend income is smaller than the rest because taxpayers report these two varibles as separate variables starting 2014. Wealth tax liability is not reported as an outcome because there is no wealth tax during most of the pre-program period. $*p < 0.1, **p < 0.05, ***p < 0.01$. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Table A.5: Top Wealth Shares in Colombia Using Tax and Survey Data, Including and Excluding Hidden Offshore Wealth

	Top 5% (1)	Top 1% (2)	Top 0.5% (3)	Top 0.1% (4)	Top 0.05% (5)	Top 0.01% (6)	Top 5% to 1% (7)	Top 1% to 0.5% (8)	Top 0.5% to 0.1% (9)	Top 0.1% to 0.05% (10)	Top 0.05% to 0.01% (11)
Survey Data	55.46	25.78	18.11	7.63	5.1	1.52	29.68	7.67	10.48	2.53	3.58
Before Amnesty	68.07	40.04	30.05	15.09	11.1	5.56	28.02	9.99	14.96	3.99	5.54
Baseline	68.39	40.64	30.75	15.85	11.81	6.01	27.74	9.9	14.89	4.05	5.79
Plus Offshore (LB)	69.74	43.17	33.66	19.06	14.8	8.23	26.57	9.51	14.6	4.26	6.57
Plus Offshore (UB)	71.47	46.4	37.39	23.17	18.64	11.08	25.06	9.01	14.22	4.53	7.57
Offshore 37.25%	75.07	53.14	45.17	31.73	26.64	17	21.92	7.98	13.43	5.09	9.64

Notes: This table presents top wealth shares in Colombia in 2017. The first row presents estimates using survey data only, while the rest combine tax and survey data. The second row presents estimates before disclosures of offshore hidden wealth during the 2015–2017 voluntary disclosure program. The third row is the baseline, without correcting for unreported hidden wealth. The last rows account for unreported offshore wealth, and make different assumptions about the size of hidden offshore wealth. Using data from Alstadsater et al. (2018a) and the Panama Papers leak, the lower bound assumes unreported offshore wealth today represents 6.2 percent of GDP, while the upper bound assumes it represents 15 percent. The last row augments unreported wealth by the observed increased in the stock of entities ever incorporated through Mossack Fonseca between 2007 and 2015. See Section A.5 for a descriptions of how these estimates were constructed. *Sources:* Authors' calculations using administrative tax microdata from DIAN and IEFIC from DANE.

A.2 A Brief Recount of Wealth Taxation in Colombia

Colombia began taxing the wealth of its wealthiest citizens in 1935 using a progressive schedule (Law 78/1935). After a series of reforms affecting its marginal tax rates and tax bases over the next five decades (Law 45/1942, Law 135/1944, Law 81/1960, Law 9/1983), the wealth tax was abolished in 1992, only to be re-established a decade later.

Colombia is unique in the extent to which it adopted multiple wealth tax reforms over the last two decades that have significantly changed both tax rates and base. Uribe was inaugurated president of Colombia on August 7, 2002 amid a precarious security situation and dismal economic conditions. Four days after inauguration, Uribe declared a state of emergency, enabling him to take extraordinary legislative measures to boost revenue to finance heightened military spending against illegal armed groups, including FARC (Decree 1837/2002).[1] Decrees 1838, 1885, and 1949 of August 2002 introduced a wealth tax dubbed "special tax for Democratic Security," in reference to Uribe's security policy.[2][3] Its rate was

[1] "*The adoption of temporary yet effective extraordinary measures is non-postponable to give Colombians their individual and collective security and to respond to the unprecedented challenge posed by criminal groups... every individual must make a significant tax effort to enable the State to ensure public security in vast parts of its territory*" (Decree 1837/2002, our translation).

[2] Importantly, this decree was announced a couple of months *after* the deadline to submit income tax returns for FY 2001.

[3] It should be noted that the exigencies of the war against drug cartels and illegal armed groups had previously led to the creation of forced investment bonds, in 1996 and 1998. Tax filers with net worth above 150 million

established at a flat rate of 1.2 percent of *all* net worth for individuals and firms whose *gross* wealth is at or above \$169.5 million pesos by August 31, 2002.[4] In other words, the *average* tax rate jumps for individuals with *gross* wealth at or above \$169.5 million pesos. This reform affected 48.05% of individual income tax filers: around 158,430 individual filers and 151,101 corporations were subject to this tax.[5] However, individual taxpayers contributed less than one-quarter of this tax revenue, with the bulk being paid for by corporations.

The following year, Uribe extended the "special" tax on wealth to continue funding the exigencies of war against illegal armed groups. Individuals with net worth of \$3 billion (thousand million) Colombian pesos or more (base year 2004) by January 1, 2004, January 1, 2005, and January 1, 2006 would be subject to this tax (Law 863/2003).[6] For these individuals, the tax was levied at a flat rate of 0.3 percent on *all* taxable wealth (i.e., net worth minus two allowances) for FY 2003-05 using tax form 420.[7][8] This reform thus generates a notch around 3 billion pesos, the threshold at which the average rate jumps from 0 to 0.3 percent. This reform annually affected only 1,420 individual taxpayers and some 4,850 firms, with the overwhelming majority (97%) of the tax burden falling on corporations.

After Uribe's re-election, the wealth tax was extended for tax years 2007-10 (Law 1111/2006), and its *average* rate was raised to 1.2 percent for net worth of \$3 billion (thousand million) pesos or more by January 1, 2007. To be clear, this is *not* a marginal tax rate of 1.2 percent; rather, the 1.2 percent is levied on *all* net worth after substracting two allowances, thus generating a notch around the 3 billion pesos threshold.[9] Importantly, even though the tax was levied on net worth held in tax years 2006, 2007, 2008, and 2009, only taxpayers with net worth of at least 3 billion pesos in 2007 were levied. That is, an individual reporting 2,999,999,999 pesos in FY 2006 and 3 billion pesos or more in 2007, 2008, or 2009 will *not* be subject to the wealth tax. This reform annually affected only some 1,800 individual filers and 5,690 corporations, with 97% of the tax burden falling on firms.

In 2009, Uribe introduced a one-off wealth tax with a increasing average rates of 2.4 percent for taxable wealth held by January 1, 2011 between \$3–5 billion pesos, and 4.8 percent for taxable wealth of \$5 billion pesos or more for 2011 (Law 1370/2009).[10] However, in De-

pesos in 1996 were required to invest 0.5 percent of taxable net worth in "Bonds for Security" (Law 345/1996), while taxpayers with net worth above 210 million pesos in 1998 were required to invest 0.6 percent of taxable net worth in "Solidarity Bonds for Peace" in 1999 and 2000 (Law 487/1998).

[4]Deductions are (i) net wealth value of assets in national businesses, and (ii) mandatory contributions to pension funds.

[5]Flores-Macías (2014) studies the factors behind the adoption of the "Democratic Security" tax by Uribe in 2002.

[6]Note that net worth possessed Jan 1, year t refers to the amount declared in the income tax in FY $t-1$, updated to valuations in year t.

[7]Allowances are (i) net wealth value of assets in national businesses, and (ii) the first 200 million pesos of the principal residence (increased to 212,200,000 pesos for FY 2005 and 222,959,000 for FY 2006).

[8]Importantly, the reform was announced and adopted several months *before* the deadline to submit returns for FY 2003.

[9]These allowances are (i) net wealth value of assets in national businesses, and (ii) the first 220 million pesos of the principal residence.

[10]Allowances are (i) net wealth value of assets in national businesses, and (ii) the first 319,215,000 pesos of

cember the following year, and allegedly to cover expenses to palliate the disastrous effects of the 2010 extreme weather conditions, newly-elected Santos reduced the filing thresholds to \$1 billion pesos, and introduced two additional rates: 1 percent for taxable wealth between \$1–2 billion pesos and 1.4 percent for wealth between \$2–3 billion pesos. In addition, Santos imposed a surcharge of 25 percent on taxpayers covered in Law 1370/2009. This set the previous rates to 3 percent for taxable wealth between \$3–5 billion pesos and 6 percent for wealth of \$5 billion pesos or more to be made in up to eight equal payments between 2011 and 2014 (Law 1430/2010). Note that, once again, these are *not* marginal tax rates but *average* tax rates on *all* taxable net worth. This affected 31,690 individual filers and 21,512 firms, with 94% of the tax revenue being collected from corporations.

On September 10, 2014, the Minister of Finance announced a bill to establish a permanent and progressive wealth tax for individuals with net worth of at least 1 billion pesos. The bill proposed the following tax schedule: an average rate of 0.4 percent for net worth \$1–3 billion, a marginal rate of 1.1 percent for net worth \$3–5\$ billion, 2 percent for net worth \$5–8 billion, and 2.25 percent for net worth of \$8 billion and above. Importantly, this announcement was made *before* individual taxpayers' deadline to submit their income tax return for FY 2013; in fact, some taxpayers could file their return up to 22 October, 2014 (Decree 2972/2013). Insofar as some taxpayers may have submitted their 2013 tax return expecting wealth taxation, we may see reporting responses starting FY 2013.

The law creating a permanent wealth tax was adopted in December 2014, albeit with a different tax schedule and significantly lower marginal rates that the initial proposed bill (Law 1739/2014). Individuals and corporations with net worth of 1 billion pesos and above on January 1, 2015 would be required to file a wealth tax return. A tax would be levied on net worth held on January 1, 2015, 2016, 2017 and 2018 for individuals using form 440. Importantly, even though the tax was levied on net worth held in tax years 2014–2017, only taxpayers with net worth of 1 billion pesos or more in FY 2014 were levied. That is, an individual reporting 999,999,999 pesos in FY 2014 and 1 billion pesos or more in 2015, 2016, or 2017 will *not* be subject to the wealth tax. For individuals, the tax rate is an *average* rate of 0.125 percent for taxable wealth below 2 billion, and a *marginal* rate of 0.35 percent for taxable wealth between 2 and 3 billion, 0.75 percent for taxable wealth between 3 and 5 billion, and 1.5 percent for taxable wealth of 5 billion and above.[11][12] The reform thus generates a notch around 1 billion pesos in net worth and kinks at 2, 3, and 5 billion pesos in taxable net worth. This reform affected 4.19 percent of filers, with again the bulk of the

---

the principal residence, and some other items.

[11]If the wealth tax base in either year t 2016, 2017 or 2018 is bigger (smaller) than that in 2015, the resulting tax base will be the minimum between 2015 tax base plus (minus) 25 percent of the inflation rate in year $t-1$ and the tax base in year t.

[12]For corporations, the wealth tax is phased out between 2015 and 2017 according to the following schedule: in 2015, the tax rates are an average rate of 0.2 percent (0.15 percent in 2016, and 0.05 percent in 2017) for taxable wealth below 2 billion pesos, and a marginal rate of 0.35 percent (0.25 percent in 2016, and 0.10 percent in 2017) for taxable wealth between 2 and 3 billion, 0.75 percent (0.5 percent in 2016, and 0.2 percent in 2017) for taxable wealth between 3 and 5 billion, and 1.15 percent (1 percent in 2016, and 0.4 percent in 2017) for taxable wealth of 5 billion and above.

burden falling on corporations.[13][14]

Improvements in Third-Party Reporting: The number of third-party reporting institutions, as well as the coverage of reported items, has been subject to significant changes over the past years. In 2006, Colombia established the list of public and private institutions required to provide third-party reports, and allowed uploading these reports through its newly-created online web portal, *sistema Muisca*. In 2012, the tax authority sought to further improve tax technology (e.g., Resolutions 111–118 from October 31 2012, Law 1607/2012), expanding the coverage of third-party reports and requiring this information be submitted online. Since then, taxpayers have been granted online access to all their third-party reported information. Note, however, that there is no return pre-filling and that taxpayers are not currently required to neither file nor pay their taxes electronically (OECD, 2017).

The Issue of Valuation: Regarding the issue of valuation, the value of some assets reported in tax records is often below its corresponding market values, defined as the price at which an asset would be traded in a competitive market. For instance, real estate is reported at cadastral values. In Colombia, as in many developing countries, cadastres are typically outdated; updating cadastres requires massive fieldwork and labor-intensive operations, as Colombia—unlike OECD countries—does not use values from transactions to estimate property values.[15] As a result, cadastral values today represent between 60 and 70 percent of market values.[16] Moreover, unlisted equities are recorded at the price at which they are bought rather than their market price. For other illiquid assets that are infrequently traded and therefore hard-to-value, including artwork and high-value jewellery, insured values could be used instead of market values; in practice, underreporting is rampant.

A.3 Bunching Theory: Extensions

Heterogeneity in Elasticities But Not in Cost Function. If there is heterogeneity in elasticities e, the tax notch creates different incentives to bunch for individuals with the same latent wealth W. Behavioral responses can be characterized as in the baseline model detailed above at each elasticity level: the bunching segment at elasticity e is given by $(W_r^*, W_r^* + \Delta W_{r,e}^*)$, where $\Delta W_{r,e}^*$ is increasing in e and equals ΔW_r^D for $e = 0$.[17] If $e > 0$, the bunching interval will be larger than the region of strictly dominated choice ($\Delta W_r^* > \Delta W_r^D$).

[13] Allowances are (i) net wealth value of assets in national businesses, and (ii) the first 12,200 UVT of the principal residence, and some other items.

[14] Note that, in both cases, the first marginal tax rate applies to *all* taxable wealth below 2 billion, *not* only taxable wealth between 1 and 2 billion.

[15] Property transaction data is limited and, when available (e.g., real estate sales are reported by notaries), it usually suffers from underreporting due to tax avoidance.

[16] By law, cadastral values must represent at least 60 percent of market values (Law 1450/2011). Historically, Bogota has updated cadastral values closer to market values more systematically than other cities in Colombia.

[17] See proof in Appendix A.3.

The dominated range therefore represents a lower bound on the wealth reporting response to tax notches under any compensated elasticity in this frictionless model.

The post-notch density is empty in the strictly dominated range, as depicted in Figure A.3, Panel (c). It then increases gradually—the elasticity being too low for some individuals to bunch—until converging with the pre-notch density at $W_r^* + \Delta W_{r,\bar{e}}^*$. With heterogeneity, the bunching method estimates the *average* response in the population $E\left[\Delta W_{r,e}^*\right]$. Excess bunching at the notch is then

$$B = \int_e \int_{W_r^*}^{W_r^* + \Delta W_{r,e}^*} \tilde{h}_0(W_r, e) dW_r de \approx h_0(W_r^*) E\left[\Delta W_{r,e}^*\right] \tag{A.1}$$

where $\tilde{h}_0(W_r, e)$ represents the joint reported wealth-elasticity distribution in the baseline without a notch, and $h_0(W_r) \equiv \int_e \tilde{h}_0(W_r, e) de$ represents the unconditional reported wealth distribution in the baseline case.[18]

Framework for Recovering the Structural Bunching Elasticity

This section presents the conceptual framework used to identify the structural elasticity of reported wealth with respect to the net-of-tax rate. Unlike the reduced-form approach developed in Section 1.3, this section builds a parametrized model and assumes a specific utility functional form.

Consider with a utility function of the form

$$u(W, W_r) = W - T(W_r) - W\left[\frac{1}{1+e} - \frac{W_r}{W} + \frac{1}{1+1/e}\left(\frac{W_r}{W}\right)^{1+1/e}\right] \tag{A.2}$$

where $T(W_r)$ represents wealth tax liability and the convex cost function $C(1 - W_r/W)$ is parametrized by $\frac{1}{1+e} - \frac{W_r}{W} + \frac{1}{1+1/e}\left(\frac{W_r}{W}\right)^{1+1/e}$.[19]

If tax liability implies a proportional (average and marginal) tax rate on reported wealth, $T(W_r) = \tau W_r$, then the individual maximization problem leads to the first order condition

$$\tilde{W_r} = W\left(1 - \tau\right)^e \tag{A.3}$$

The optimality condition (A.3) indicates that a positive tax rate depresses $\tilde{W_r}$ below W, with the strength of the effect determined by e, the parameter of interest. If $e \to 0$, then individuals report their true wealth ($\tilde{W_r} = W$), while if $e \to \infty$, individuals report no wealth at all ($\tilde{W_r} = 0$).

The combination of the wealth distribution and the reported wealth function (A.3) yields a reported wealth distribution associated with the baseline linear tax system. We denote

[18] The approximation assumes that the counterfactual density is locally constant in reported wealth (but not elasticities).

[19] Note that $C(0) = 0$, i.e., there is no cost of underreporting when taxpayers do not underreport. Moreover, $C(\cdot)$ is convex: if $u = 1 - W_r/W$, $C'(u) = W[1 - (1-u)^{1/e}] \geq 0$ and $C''(u) = W(1-u)^{1/e-1}/e \geq 0$.

$H_0(W_r)$ and $h_0(W_r)$ the distribution and density functions for reported wealth associated with this baseline. Using the optimality condition (A.3), we obtain $H_0(W_r) = F\left(\frac{W_r}{(1-\tau)^e}\right)$ and hence $h_0(W_r) = H_0'(W_r) = f\left(\frac{W_r}{(1-\tau)^e}\right)/(1-\tau)^e$. Therefore, given a smooth tax system (i.e., no notches and no kinks), the smooth wealth distribution converts into a smooth reported wealth distribution.

Consider the marginal buncher H located at $W_r^* + \Delta W_r^*$ before the reform, whose wealth is $W^* + \Delta W^*$, and is indifferent between the notch point W_r^* and the best interior point W_r^I after the tax change. At notch point W_r^*, her utility level is given by

$$\begin{aligned} u^N &= W^* + \Delta W^* - \tau W_r^* \\ &\quad - (W^* + \Delta W^*) \cdot \left[\frac{1}{1+e} - \frac{W_r^*}{W^* + \Delta W^*} + \frac{1}{1+1/e}\left(\frac{W_r^*}{W^* + \Delta W^*}\right)^{1+\frac{1}{e}}\right] \end{aligned} \tag{A.4}$$

Using the first order condition $W_r^I = (W^* + \Delta W^*)(1-\tau-\Delta\tau)^e$, the utility level obtained at the best interior location can be written as

$$u^I = (W^* + \Delta W^*) \cdot \left[\frac{(1-\tau-\Delta\tau)^{1+e}}{1+e} + \frac{1}{1+1/e}\right] \tag{A.5}$$

From the condition $u^I = u^N$ and using the relationship $W^* + \Delta W^* = (W_r^* + \Delta W_r^*)/(1-\tau)^e$, we can rearrange the terms so as to obtain

$$\frac{1}{1+\Delta W_r^*/W_r^*} - \frac{1}{1+1/e}\left(\frac{1}{1+\Delta W_r^*/W_r^*}\right)^{1+1/e} - \frac{1}{1+e}\left(1 - \frac{\Delta\tau}{1-\tau}\right)^{1+e} = 0 \tag{A.6}$$

This condition characterizes the relationship between the percentage reporting response $\Delta W_r^*/W_r^*$, the percentage change in the average net-of-tax rate created by the notch, $\Delta\tau/(1-\tau)$, and the structural elasticity e.[20] Although it is not possible to obtain an analytical solution for e, this can be solved numerically given an estimate for ΔW_r^* and the observed value of the other arguments.

The disadvantage of this approach is that it relies on a functional form for utility. While the wealth reporting response ΔW_r^* can be non-parametrically identified, the underlying structural elasticity e from equation (A.6) that could be used for out-of-sample prediction cannot. It is therefore useful to develop a reduced-form approach that does not rely on the specific functional form for individuals' utility, as we do in Section 1.3. As discussed in Kleven and Waseem (2013), under some assumptions equation (1.3) for this reduced-form elasticity represents an upper bound on the structural elasticity e from equation (A.6).

Finally, note that, as the compensated elasticity e converges to zero (L-shaped Leontief preferences), equation (A.6) implies

$$\lim_{e\to 0} \Delta W_r^* = \frac{\Delta\tau \cdot W_r^*}{1-\tau-\Delta\tau} \equiv \Delta W_r^D \tag{A.7}$$

[20] See proofs and derivations in Section A.3.

This means that, as $e \to 0$, the bunching interval ΔW_r^* converges to the strictly dominated range ΔW_r^D. The dominated range therefore represents a lower bound on the wealth reporting response to tax notches under any compensated elasticity in a frictionless model (Kleven and Waseem, 2013).

Table A.6 presents the estimated parameters. For each reform year, the table shows the notch point (column 2), whether this notch also defined eligibility for the wealth tax (column 3), the average tax rate jump (column 4), the size of the dominated range (column 5), the share of taxpayers in dominated ranges that are unresponsive to the tax notch (column 6), the lower and upper bounds on the reporting responses (columns 7 and 8, respectively), and the bounds on the elasticities based on either the parametric equation (A.6) (columns 9 and 10) or the reduced-form formula (1.3) (column 11 and 12).

The table shows that the structural elasticities driving the large wealth reporting responses for the first notch in 2010 range between 0.47 to 1.42. The reduced-form elasticities, which represent an upper bound if the uncompensated reported wealth elasticity is not too strongly negative, are between 0.6 and 2.0.[21] These elasticities obtained for the first notch are all statistically significantly different from zero at the 1 percent level. In contrast, elasticities obtained from the second notch are smaller and often less precisely estimated: the structural elasticities between 0.32 to 0.84, and the reduced-form elasticities are between 0.37 and 1.0, but the upper bounds using the convergence method are not statistically significantly different from zero.

[21] Reduced-form elasticities are somewhat larger than structural elasticities, as the former provide an approximation (upper bound) of the true structural elasticity. Kleven and Waseem (2013) show that, given the size of the notch $\Delta\tau/(1-\tau)$ and a true functional form for utility, the bias of the reduced-form approach is determined by the percentage reporting response $\Delta W_r^*/W_r^*$.

Table A.6: Summary of Notches, Responses, and Elasticities

Year of Reform (1)	Notch Point (mill. pesos) (2)	Exemption Cutoff (3)	ATR Jump $\Delta\tau$ (%) (4)	Dominated Range ΔW_r^D (mill. pesos) (5)	Frictions a^* using ΔW_r^D (6)	Response ΔW_r^* (mill. pesos): Bunching-Hole Method (7)	Response ΔW_r^* (mill. pesos): Convergence Method (8)	Structural Elasticity e: Bunching-Hole Method (9)	Structural Elasticity e: Convergence Method (10)	Reduced-Form Elasticity e_R: Bunching-Hole Method (11)	Reduced-Form Elasticity e_R: Convergence Method (12)
2003	3,000	✓	0.3	9	0.74	120	180	0.24	0.53	0.27	0.60
					(0.18)	(80.09)	(137.93)	(0.52)	(1.27)	(0.62)	(1.76)
2006	3,000	✓	1.2×4	151	0.41	340	560	0.07	0.21	0.13	0.36
					(0.04)	(52.04)	(109.62)	(0.03)	(0.10)	(0.04)	(0.16)
2010	1,000	✓	1.0	10	0.43	110	200	0.47	1.42	0.60	2.00
					(0.02)	(7.91)	(16.86)	(0.07)	(0.22)	(0.09)	(0.35)
2010	2,000		0.4	8	0.57	110	180	0.32	0.84	0.37	1.00
					(0.07)	(24.10)	(64.63)	(0.14)	(0.73)	(0.19)	(0.99)
2010	3,000		1.6	49	0.35	220	360	0.18	0.33	0.17	0.44
					(0.04)	(40.82)	(87.44)	(0.05)	(0.14)	(0.09)	(0.20)
2010	5,000		3.0	160	0.45	360	680	0.06	0.20	0.08	0.30
					(0.06)	(105.16)	(238.74)	(0.06)	(0.16)	(0.06)	(0.23)
2014	1,000	✓	0.0125×4	5	0.38	110	210	0.98	3.17	1.21	4.41
					(0.02)	(6.31)	(16.03)	(0.12)	(0.42)	(0.15)	(0.71)

Notes: This table presents elasticity estimates at different wealth levels exploiting four wealth tax reforms taking place in 2003, 2006, 2010, and 2014. Column (1) presents the year of the wealth tax reform. Column (2) indicates the bracket cutoff, expressed in current million pesos. Column (3) indicates whether this cutoff also marks the eligibility threshold, below which taxpayers are exempt from the wealth tax. Column (4) presents the size of the wealth tax notch. Column (5) presents the dominated range in current million pesos, defined as $\Delta\tau \cdot W_r^*/(1-\tau-\Delta\tau)$. Column (6) presents the estimate of frictions (the fraction of individuals in dominated ranges who are unresponsive). Columns (7)–(8) present the reporting responses in current million pesos using bunching-hole and convergence methods, respectively. Columns (9)–(12) present elasticities based on either the parametric equation (A.6) in columns (9)–(10) or the reduced-form formula (1.3) in columns (11)–(12). *Source:* Authors' calculations using administrative tax microdata from DIAN.

Proofs and Derivations

Structural elasticity formula, equation (A.6): Recall from Figure A.3, Panel (a), that the marginal buncher H is located at $W_r^* + \Delta W_r^*$ before the reform and has (latent) wealth $W^* + \Delta W^*$. This individual is indifferent between the notch point W_r^* and the best interior point W_r^I after the tax change. Her utility level at notch point W_r^* is given by

$$u^N = W^* + \Delta W^* - \tau W_r^* - (W^* + \Delta W^*) \cdot \left[\frac{1}{1+e} - \frac{W_r^*}{W^* + \Delta W^*} + \frac{1}{1+1/e}\left(\frac{W_r^*}{W^* + \Delta W^*}\right)^{1+\frac{1}{e}}\right]$$

Using the first order condition $W_r^I = (W^* + \Delta W^*)(1 - \tau - \Delta\tau)^e$, the utility level obtained at the best interior location is

$$\begin{aligned}
u^I &= W^* + \Delta W^* - (\tau + \Delta\tau) W_r^I \\
&\quad - (W^* + \Delta W^*) \cdot \left[\frac{1}{1+e} - \frac{W_r^I}{W^* + \Delta W^*} + \frac{1}{1+1/e}\left(\frac{W_r^I}{W^* + \Delta W^*}\right)^{1+\frac{1}{e}}\right] \\
&= W^* + \Delta W^* - (W^* + \Delta W^*)(\tau + \Delta\tau)(1 - \tau - \Delta\tau)^e \\
&\quad - (W^* + \Delta W^*) \cdot \left[\frac{1}{1+e} - \frac{W^* + \Delta W^*}{W^* + \Delta W^*} \cdot (1 - \tau - \Delta\tau)^e \right. \\
&\quad \left. + \frac{1}{1+1/e}\left(\frac{W^* + \Delta W^*}{W^* + \Delta W^*} \cdot (1 - \tau - \Delta\tau)^e\right)^{1+\frac{1}{e}}\right] \\
&= (W^* + \Delta W^*)\left[1 - (\tau + \Delta\tau)(1 - \tau - \Delta\tau)^e - \frac{1}{1+e} + (1 - \tau - \Delta\tau)^e \right. \\
&\quad \left. - \frac{1}{1+1/e}(1 - \tau - \Delta\tau)^{1+e}\right] \\
&= (W^* + \Delta W^*)\left[1 + (1 - \tau - \Delta\tau)^{1+e} - \frac{1}{1+1/e}(1 - \tau - \Delta\tau)^{1+e} - \frac{1}{1+e}\right] \\
&= (W^* + \Delta W^*) \cdot \left[\frac{(1 - \tau - \Delta\tau)^{1+e}}{1+e} + \frac{1}{1+1/e}\right]
\end{aligned}$$

From the condition $u^I = u^N$ and using the relationship $W^* + \Delta W^* = (W_r^* + \Delta W_r^*)/(1 - \tau)^e$, we can obtain

$$W^* + \Delta W^* - \tau W_r^* - (W^* + \Delta W^*) \cdot \left[\frac{1}{1+e} - \frac{W_r^*}{W^* + \Delta W^*} + \frac{1}{1+1/e} \left(\frac{W_r^*}{W^* + \Delta W^*} \right)^{1+\frac{1}{e}} \right]$$
$$-(W^* + \Delta W^*) \cdot \left[\frac{(1-\tau-\Delta\tau)^{1+e}}{1+e} + \frac{1}{1+1/e} \right] = 0$$
$$\Leftrightarrow \frac{W_r^* + \Delta W_r^*}{(1-\tau)^e} - \tau W_r^* - \frac{W_r^* + \Delta W_r^*}{(1-\tau)^e} \cdot \left[\frac{1}{1+e} - \frac{W_r^*(1-\tau)^e}{W_r^* + \Delta W_r^*} + \frac{1}{1+1/e} \left(\frac{W_r^*(1-\tau)^e}{W_r^* + \Delta W_r^*} \right)^{1+\frac{1}{e}} \right]$$
$$- \frac{W_r^* + \Delta W_r^*}{(1-\tau)^e} \cdot \left[\frac{(1-\tau-\Delta\tau)^{1+e}}{1+e} + \frac{1}{1+1/e} \right] = 0$$

Dividing by $\frac{W_r^*+\Delta W_r^*}{(1-\tau)^e}$:

$$\frac{e}{1+e} - \frac{\tau W_r^*(1-\tau)^e}{W_r^* + \Delta W_r^*} + \frac{W_r^*(1-\tau)^e}{W_r^* + \Delta W_r^*} - \frac{1}{1+1/e} \left(\frac{W_r^*(1-\tau)^e}{W_r^* + \Delta W_r^*} \right)^{1+\frac{1}{e}} - \frac{(1-\tau-\Delta\tau)^{1+e}}{1+e}$$
$$- \frac{1}{1+1/e} = 0$$
$$\frac{W_r^*(1-\tau)^{1+e}}{W_r^* + \Delta W_r^*} - \frac{1}{1+1/e} \left(\frac{W_r^*(1-\tau)^e}{W_r^* + \Delta W_r^*} \right)^{1+\frac{1}{e}} - \frac{(1-\tau-\Delta\tau)^{1+e}}{1+e} = 0$$
$$\frac{(1-\tau)^{1+e}}{1+\Delta W_r^*/W_r^*} - \frac{1}{1+1/e} \left(\frac{(1-\tau)^e}{1+\Delta W_r^*/W_r^*} \right)^{1+\frac{1}{e}} - \frac{(1-\tau-\Delta\tau)^{1+e}}{1+e} = 0$$

Dividing by $(1-\tau)^{1+e}$:

$$\frac{1}{1+\Delta W_r^*/W_r^*} - \frac{1}{1+1/e} \left(\frac{1}{1+\Delta W_r^*/W_r^*} \right)^{1+1/e} - \frac{1}{1+e} \left(1 - \frac{\Delta\tau}{1-\tau} \right)^{1+e} = 0$$

□

Equivalence between wealth and capital income elasticities: Let τ_W be the wealth tax rate and τ_K the equivalent tax rate on capital income so that $\tau_W = r \cdot \tau_K$ where r is the rate of return on wealth. For instance, if τ_W is 1 percent, and r is 5 percent, then τ_K is 20 percent.

$$d(1-\tau_W) = d(1-\tau_K) \cdot r$$
$$\frac{d(1-\tau_W)}{(1-\tau_W)} = \frac{d(1-\tau_K)}{(1-\tau_K)} \cdot r \left(\frac{1-\tau_K}{1-\tau_W} \right)$$
$$\left[\frac{d(1-\tau_W)}{(1-\tau_W)} \cdot \frac{W}{dW} \right]^{-1} = \left[\frac{d(1-\tau_K)}{(1-\tau_K)} \cdot r \left(\frac{1-\tau_K}{1-\tau_W} \right) \cdot \frac{W}{dW} \right]^{-1}$$
$$\epsilon_W = \epsilon_K \cdot \left[\frac{1-\tau_W}{r(1-\tau_K)} \right]$$

In our example, $\frac{1-\tau_W}{r(1-\tau_K)} = 24.75$. Therefore, to translate our estimated elasticities of reported wealth into the equivalent elasticities of capital income, we divide ϵ_W by a factor of 24.75.

A.4 Difference-in-Differences Comparing Taxpayers Close to the Bracket Cutoffs

Section 1.3 showed that taxpayers respond to wealth taxation by immediately lowering their reported wealth below the notch points. In this section, we complement the bunching analysis by examining responses over a longer time horizon, leveraging the panel microdata and variation from four wealth tax reforms in 2003, 2006, 2010, and 2014.

We run two types of analyses. First, we compare the number of individuals reporting wealth above and below the bracket cutoffs across time. Figure D.1 plots these series separately for each reform and bracket cutoff, normalizing each series to equal 0 in the year before a given wealth tax reform. Three key insights emerge from the figure. First, the two series are on parallel trends in the years leading up to the reform. Second, consistent with the immediate bunching responses documented in Section 1.3, there is a sudden drop in the relative number of taxpayers reporting wealth just above the cutoff and a corresponding rise in those reporting just below it. Third, this gap persists over time, even when reporting above the cutoff no longer triggers wealth tax liability. For instance, while wealth taxes owed between 2007 and 2009 depend only on wealth reported in 2006, Panel (b) shows that the gap between the two series remains large throughout this period. Similarly, even though wealth is not taxed in 2011–2013, Panel (c) shows that the gap in induced by the 2010 wealth tax persists across time. However, the effect is not growing over time. In fact, it can also revert back, as shown in Panels (d) and (e).

Second, we consider a difference-in-differences approach in which we compare wealth reported by treatment and control groups in a balanced panel of taxpayers. To define the groups that were differentially affected by wealth taxation, we exploit the notched tax schedule whereby wealth tax liability jumps discontinuously at the bracket cutoffs. We assign taxpayers to treatment and control groups based on their reported wealth in the year immediately before a reform. We then compare reported wealth between taxpayers below (control) and above (treated) the cutoffs using the following specification:

$$\log(W_{it}) = \alpha_i + \sum_{j \neq K} \beta_j^C 1(Year)_{j=t} + \sum_{j \neq K} \beta_j^T 1(Year)_{j=t} \times 1(\text{Treated})_i + u_{it} \tag{A.8}$$

where W_{it} denotes the wealth reported by individual i in year t, α_i is an individual fixed effect, $1(Year)_{j=t}$ equals 1 when the year equals t, $1(\text{Treated})_i$ equals 1 for treated taxpayers, and u_{it} is the error term. We choose K such as to normalize with respect to the year immediately before the reform.

In Figures D.2 through D.6, Panel (a) shows the time series of log reported wealth in the control group (black dots) and treated group (blue triangles) in the years before and after a wealth tax reform, i.e., estimates $\hat{\beta}_t^C$ and $\hat{\beta}_t^C + \hat{\beta}_t^T$. Panel (b) shows the differences between the control and treated groups in year t ($\hat{\beta}_t^T$), i.e., our reduced-form estimate of the effect of the wealth tax reform. Several insights emerge from these figures. First, while the two series evolve in tandem in the years leading up to the reform, they immediately diverge afterwards. Second, most of the overall effect is concentrated in the year of the reform, consistent with a one-time avoidance adjustment driving the results. Third, the gap between the two series—which conflates mechanical and behavioral responses—persists several years after a reform. This is expected if wealth is taxed more than once, as is the case between 2003 and 2005 (Figure D.2) or between 2006 and 2009 (Figure D.3), with the mechanical effect—governed by the large change in the average after-tax return—driving the result. Yet, interestingly, the gap persists *even when wealth is no longer taxed*: a one-off wealth tax in 2010 generates a persistent gap in Figure D.4. This is partly due to the peristence of bunching documented in Section 1.3 and Figure 1.5.[22]

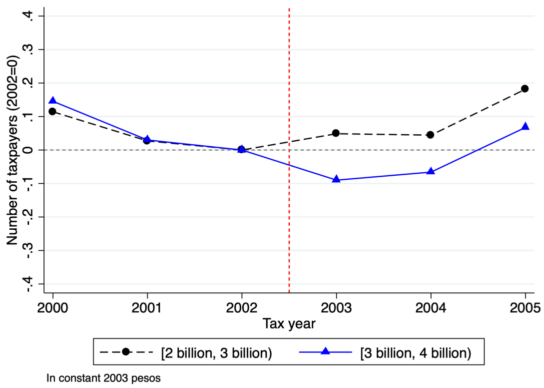
Because the treatment assignment is based on *pre-reform* reported wealth, it may not systematically determine whether a taxpayer actually owed wealth taxes. For instance, some taxpayers reporting above the cutoff in $t-1$ (treated) may immediately bunch below the cutoff in year t to avoid the wealth tax. Indeed, one-fifth of treated taxpayers based on wealth reported in 2009 would belong to the control group based on wealth reported in 2010. This means β is a reduced-form or ITT estimate. Moreover, some taxpayers in the control group who would have reported above the cutoff may remain below it in response to the reform. This implies that control taxpayers also lower their reported wealth in response to wealth taxation, driving β towards zero. Third, as in other studies using difference-in-differences approaches with panel data (Saez et al., 2012), the possibility of mean reversion implies that treated wealth in base year $t-1$ is likely to decrease because some individuals are part of the treated group in $t-1$ due to having large positive transitory wealth shocks.

To examine how sensitive our estimates are to the response of taxpayers around the notch point, Panels (c) and (d) in Figures D.2 through D.6 re-estimate the results excluding taxpayers in the bunching area around the bracket cutoff. Although the results vary somewhat depending on the reform year and notch point, the qualitative findings remain very similar.

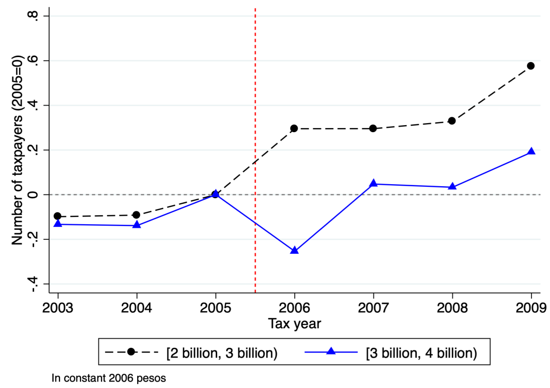
[22] This effect is less precisely identified and estimated for taxpayers further up the distribution (Figures D.5 and D.6), partly because these taxpayers were already subject to wealth taxes prior to the 2010 reform. Further, we do not present results for two comparisons: (a), the impact of the reintroduction of the wealth tax in 2014, because of persistent bunching from the reform in 2010; and (b) the second notch in 2010 at 2 billion pesos, because it is too close to the other two notches at 1 and 3 billion pesos.

Figure D.1: Number of Taxpayers Above and Below Wealth Tax Bracket Cutoffs

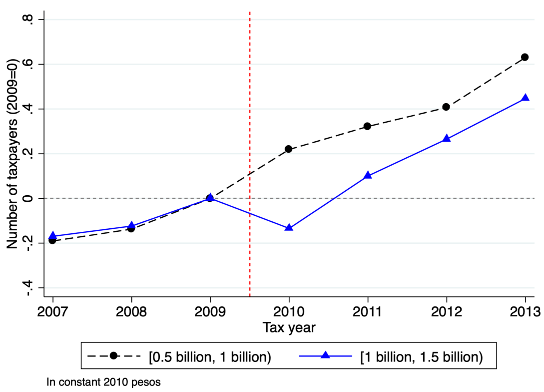
(a) 2003–05: $\Delta\tau = 0.3\%$ and $W_r^* = 3$ Billion



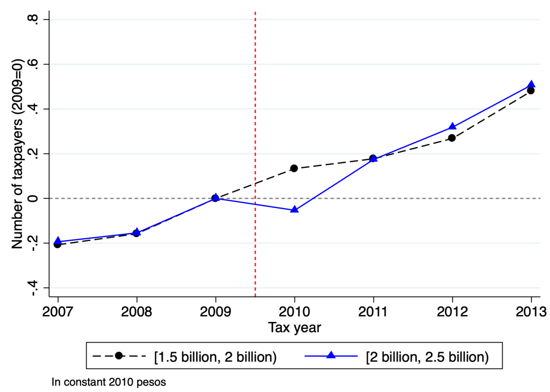
(b) 2006: $\Delta\tau = 1.2\% \times 4$ and $W_r^* = 3$ Billion



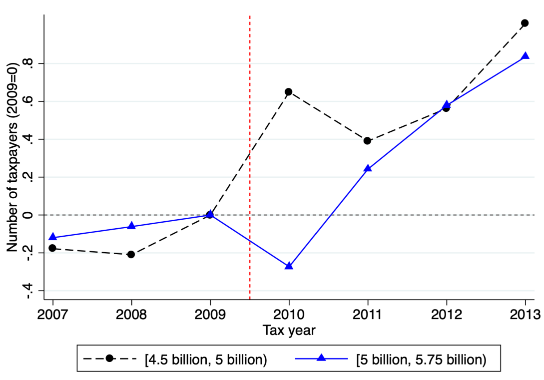
(c) 2010: $\Delta\tau = 1\%$ and $W_r^* = 1$ Billion



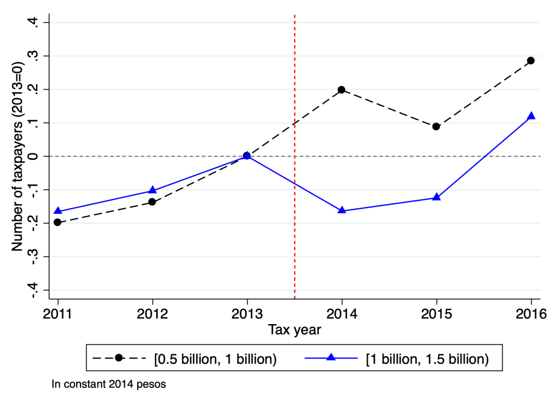
(d) 2010: $\Delta\tau = 0.4\%$ and $W_r^* = 2$ Billion



(e) 2010: $\Delta\tau = 3\%$ and $W_r^* = 5$ Billion



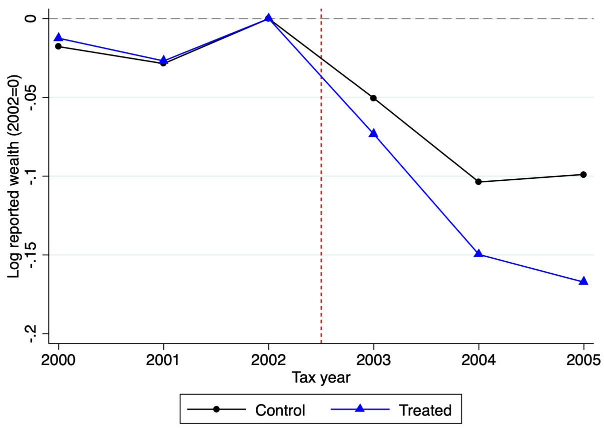
(f) 2014: $\Delta\tau = 0.125\% \times 4$ and $W_r^* = 1$ Billion



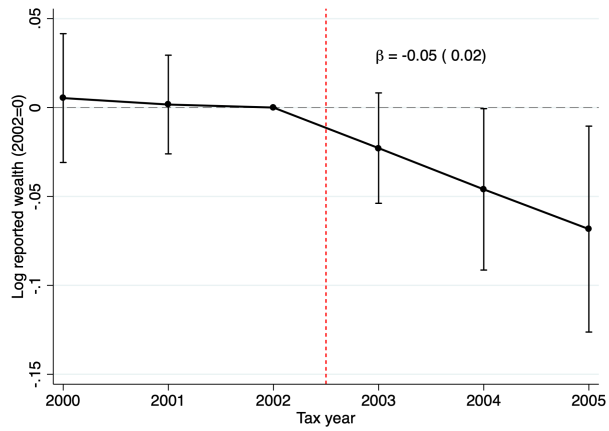
Notes: The figure shows the number of taxpayers reporting wealth in a given range below (control, dashed black line) and above (treated, blue solid line) bracket cutoff W_r^*, normalized to zero in the pre-reform year. We do not plot the series around 2010's third notch of $\Delta\tau = 1.6$ percent at $W_r^* = 3$ billion pesos, as it interferes with the wealth tax put in place between 2006 and 2009 at a similar cutoff.

Figure D.2: 2003–05 Notch with $\Delta\tau = 0.3\%$ and $W_r^* = 3$ Billion (2003 pesos)

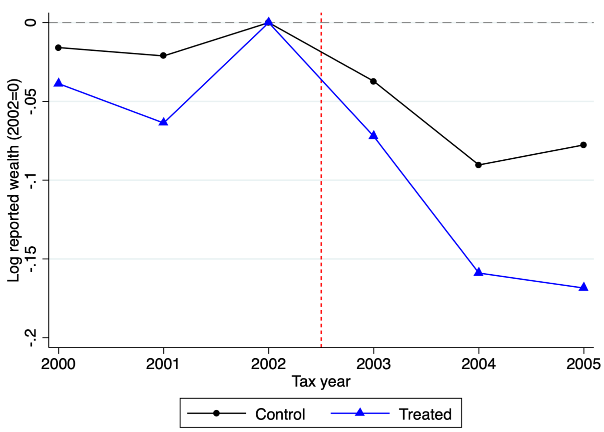
(a) $W_r^{2002} : C = [2, 3)$ vs $T = [3, 4)$



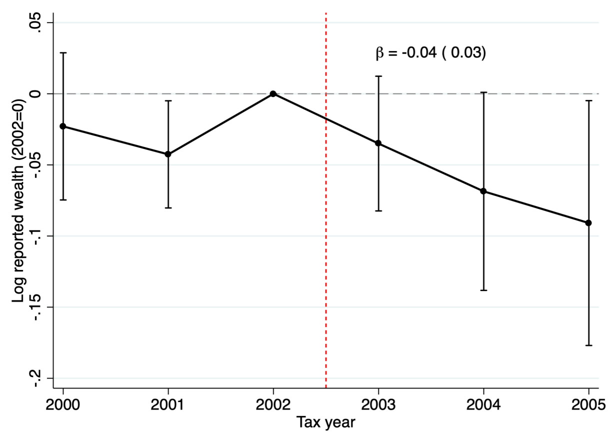
(b) $W_r^{2002} : C = [2, 3)$ vs $T = [3, 4)$



(c) $W_r^{2002} : C = [2, 2.5)$ vs $T = [3.5, 4)$



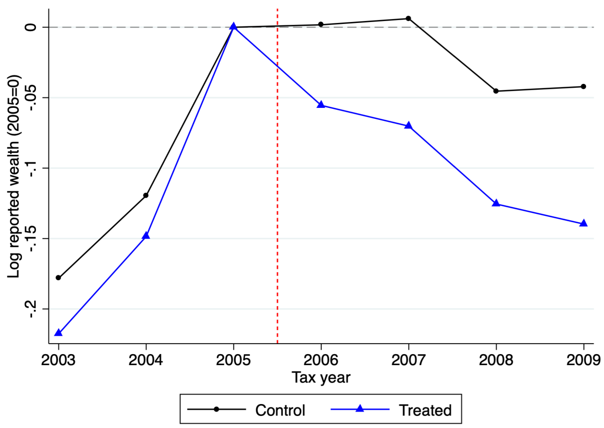
(d) $W_r^{2002} : C = [2, 2.5)$ vs $T = [3.5, 4)$



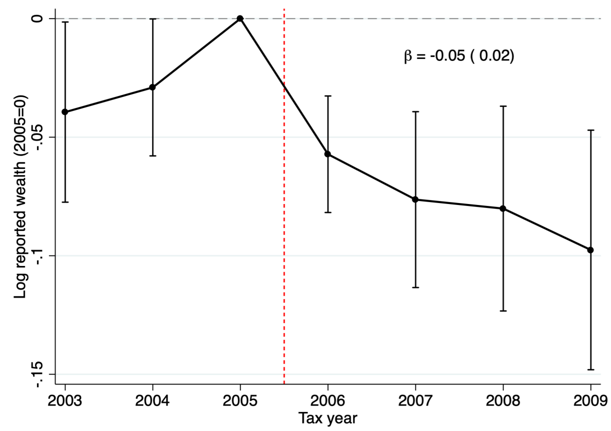
Notes: The figure shows the (intention-to-treat) effects of a wealth tax reform on reported wealth based on a difference-in-differences comparison of taxpayers around the eligibility cutoff. Taxpayers reporting 3 billion pesos (in 2003 pesos) or more in wealth owed 0.3 percent of taxable wealth in 2003, 2004, and 2005. The estimation sample is a balanced panel of taxpayers observed in all years plotted in the figure. Panel (a) shows the evolution of reported wealth for taxpayers reporting wealth between 2003 2 and 3 billion pesos in 2002 (control) and between 2003 3 and 4 billion pesos in 2002 (treated), normalized to zero in 2002 and using specification (A.8). Panel (b) shows the differences between these two series, i.e., our difference-in-differences estimates. The 95 percent confidence intervals are based on robust standard errors clustered at the individual level. Panels (c) and (d) examine how these estimates vary when excluding taxpayers in the bunching area, that is, comparing taxpayers owning between 2 and 2.5 billion versus 3.5 and 4 billion pesos in 2002.

Figure D.3: 2006 Notch with $\Delta\tau = 1.2\% \times 4$ and $W_r^* = 3$ Billion

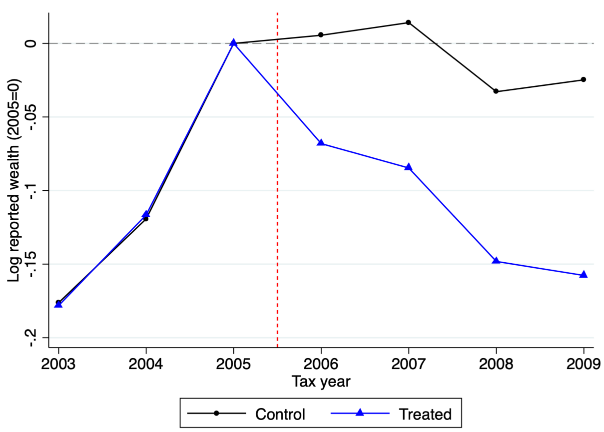
(a) $W_r^{2005}: C = [2, 3)$ vs $T = [3, 4)$



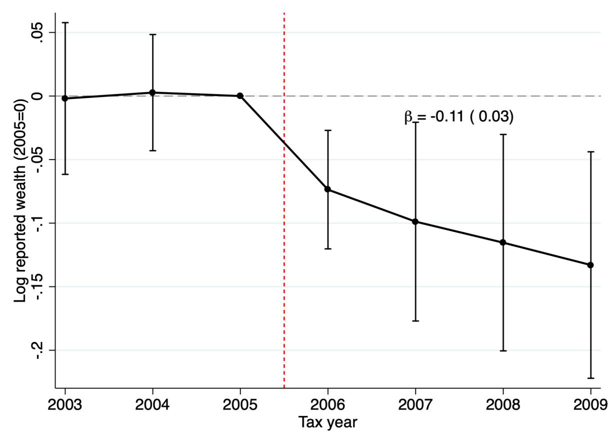
(b) $W_r^{2005}: C = [2, 3)$ vs $T = [3, 4)$



(c) $W_r^{2005}: C = [2, 2.5)$ vs $T = [3.5, 4)$



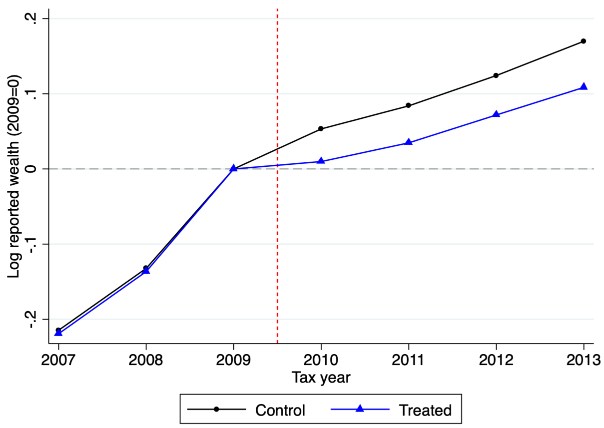
(d) $W_r^{2005}: C = [2, 2.5)$ vs $T = [3.5, 4)$



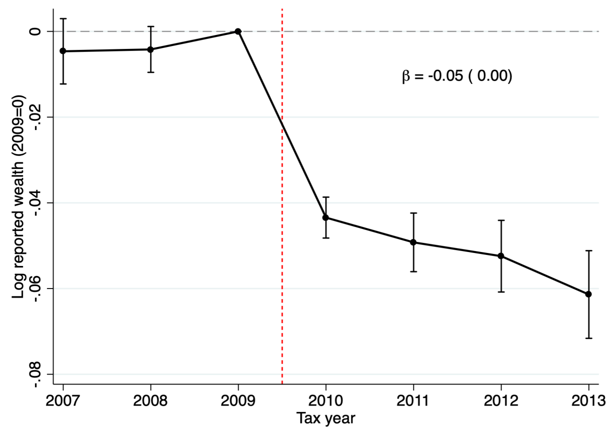
Notes: The figure shows the (intention-to-treat) effects of a wealth tax reform on reported wealth based on a difference-in-differences comparison of taxpayers around the eligibility cutoff. Taxpayers reporting 3 billion pesos or more in wealth in 2006 owed 1.2 percent of taxable wealth in 2006, 2007, 2008, and 2009. The estimation sample is a balanced panel of taxpayers observed in all years plotted in the figure. Panel (a) shows the evolution of reported wealth for taxpayers reporting wealth between 2006 2 and 3 billion pesos in 2005 (control) and between 2006 3 and 4 billion pesos in 2005 (treated), normalized to zero in 2005 and using specification (A.8). Panel (b) shows the differences between these two series, i.e., our difference-in-differences estimates. The 95 percent confidence intervals are based on robust standard errors clustered at the individual level. Panels (c) and (d) examine how these estimates vary when excluding taxpayers in the bunching area, that is, comparing taxpayers owning between 2 and 2.5 billion versus 3.5 and 4 billion pesos in 2005.

Figure D.4: 2010 First Notch with $\Delta\tau = 1\%$ and $W_r^* = 1$ Billion

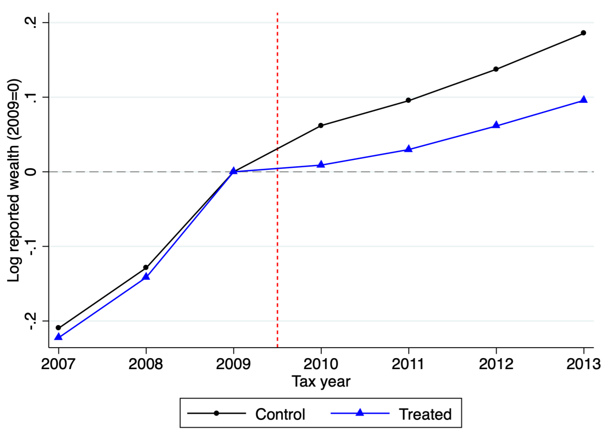
(a) $W_r^{2009} : C = [0.5, 1)$ vs $T = [1, 1.5)$



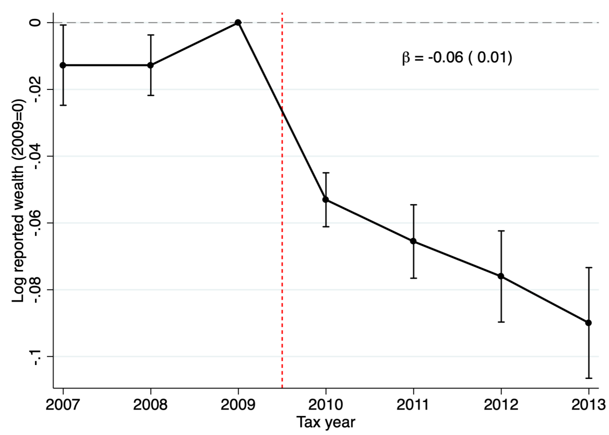
(b) $W_r^{2009} : C = [0.5, 1)$ vs $T = [1, 1.5)$



(c) $W_r^{2009} : C = [0.5, 0.75)$ vs $T = [1.25, 1.5)$



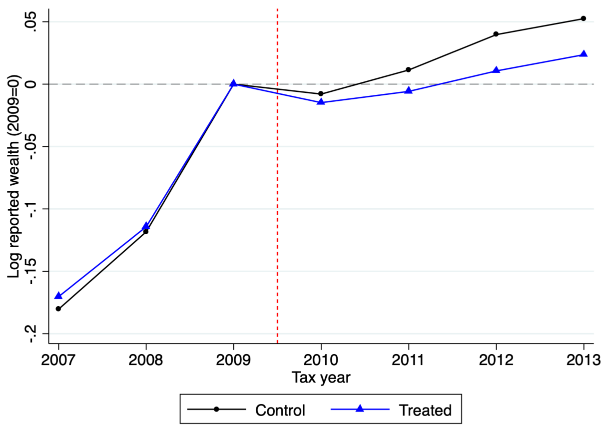
(d) $W_r^{2009} : C = [0.5, 0.75)$ vs $T = [1.25, 1.5)$



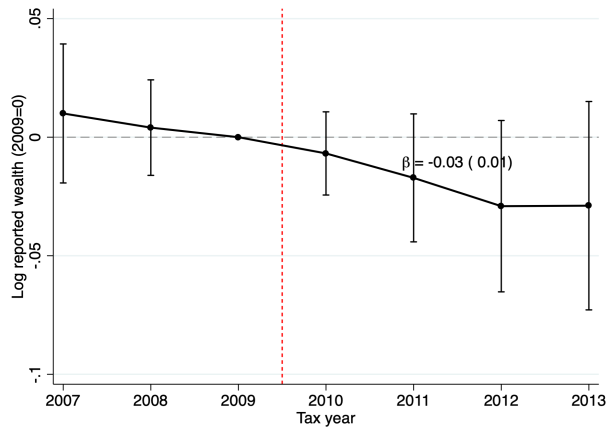
Notes: The figure shows the (intention-to-treat) effects of a wealth tax reform on reported wealth based on a difference-in-differences comparison of taxpayers around the eligibility cutoff. Taxpayers reporting between 1 and 2 billion pesos in 2010 owed 1 percent of taxable wealth that year. The estimation sample is a balanced panel of taxpayers observed in all years plotted in the figure. Panel (a) shows the evolution of reported wealth for taxpayers reporting wealth between 2010 0.5 and 1 billion pesos in 2009 (control) and between 2010 1 and 1.5 billion pesos in 2009 (treated), normalized to zero in 2009 and using specification (A.8). Panel (b) shows the differences between these two series, i.e., our difference-in-differences estimates. The 95 percent confidence intervals are based on robust standard errors clustered at the individual level. Panels (c) and (d) examine how these estimates vary when excluding taxpayers in the bunching area, that is, comparing taxpayers owning between 0.5 and 0.75 billion versus 1.25 and 1.5 billion pesos in 2009.

Figure D.5: 2010 Third Notch with $\Delta\tau = 1.6\%$ and $W_r^* = 3$ Billion

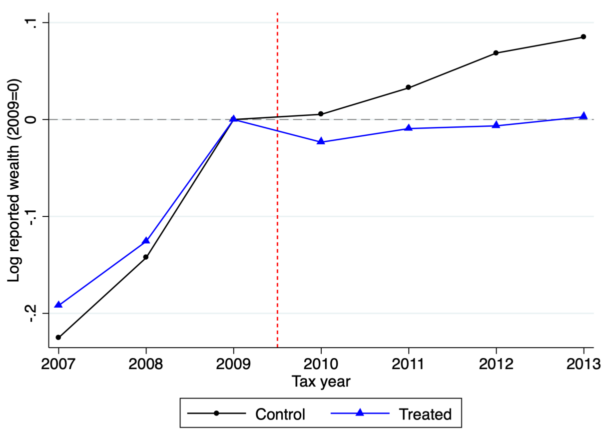
(a) $W_r^{2009} : C = [2.5, 3)$ vs $T = [3, 4)$



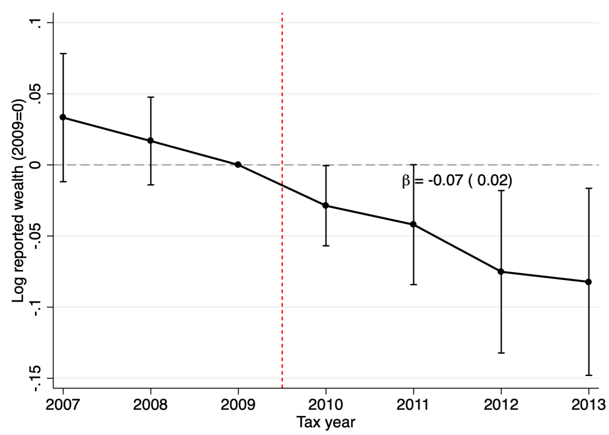
(b) $W_r^{2009} : C = [2.5, 3)$ vs $T = [3, 4)$



(c) $W_r^{2009} : C = [2.25, 2.5)$ vs $T = [3.5, 4)$



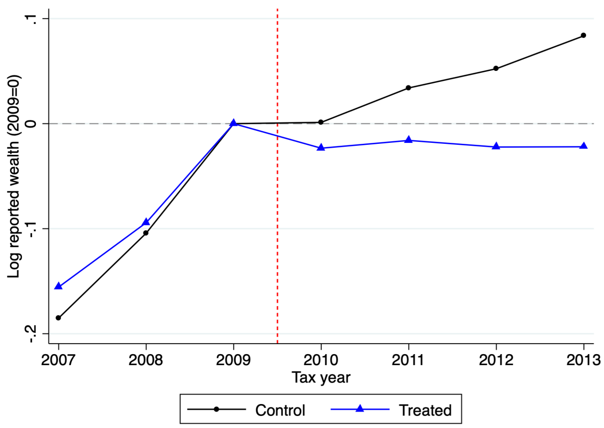
(d) $W_r^{2009} : C = [2.25, 2.5)$ vs $T = [3.5, 4)$



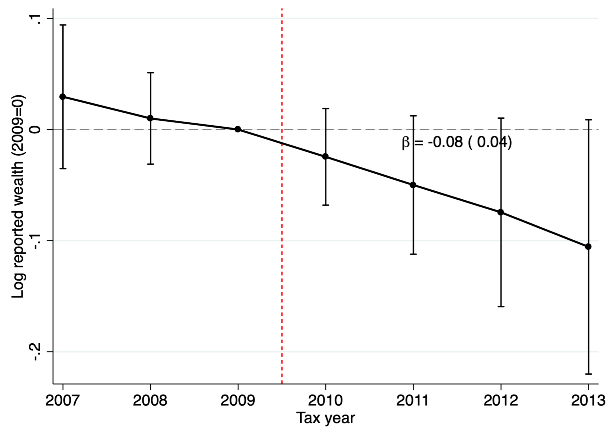
Notes: The figure shows the (intention-to-treat) effects of a wealth tax reform on reported wealth based on a difference-in-differences comparison of taxpayers around the eligibility cutoff. Taxpayers reporting between between 2 and 3 billion pesos in 2010 owed 1.4 percent of taxable wealth that year, whereas those reporting between 3 and 5 billion pesos owed 3 percent. The estimation sample is a balanced panel of taxpayers observed in all years plotted in the figure. Panel (a) shows the evolution of reported wealth for taxpayers reporting wealth between 2010 2.5 and 3 billion pesos in 2009 (control) and between 2010 3 and 4 billion pesos in 2009 (treated), normalized to zero in 2009 and using specification (A.8). Panel (b) shows the differences between these two series, i.e., our difference-in-differences estimates. The 95 percent confidence intervals are based on robust standard errors clustered at the individual level. Panels (c) and (d) examine how these estimates vary when excluding taxpayers in the bunching area, that is, comparing taxpayers owning between 2.25 and 2.5 billion versus 3.5 and 4 billion pesos in 2009.

Figure D.6: 2010 Fourth Notch with $\Delta\tau = 3\%$ and $W_r^* = 5$ Billion

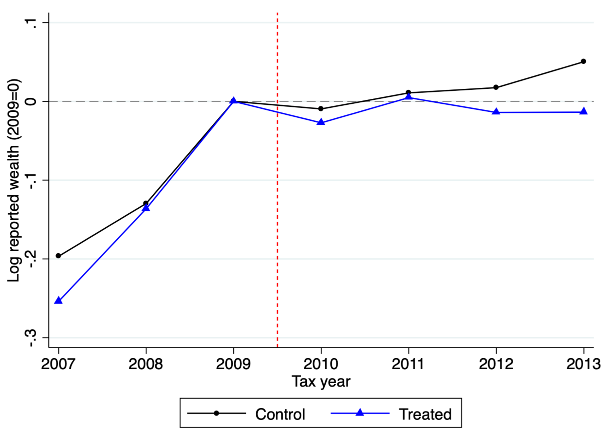
(a) $W_r^{2009}: C = [4.5, 5)$ vs $T = [5, 5.75)$



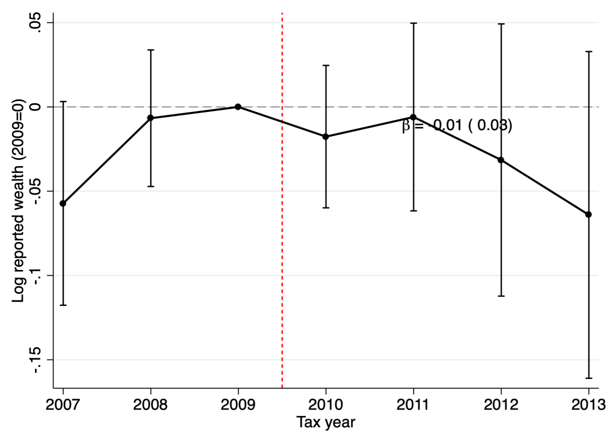
(b) $W_r^{2009}: C = [4.5, 5)$ vs $T = [5, 5.75)$



(c) $W_r^{2009}: C = [4, 4.5)$ vs $T = [6, 8)$



(d) $W_r^{2009}: C = [4, 4.5)$ vs $T = [6, 8)$



Notes: The figure shows the (intention-to-treat) effects of a wealth tax reform on reported wealth based on a difference-in-differences comparison of taxpayers around the eligibility cutoff. Taxpayers reporting between between 3 and 5 billion pesos in 2010 owed 3 percent of taxable wealth that year, whereas those reporting between 5 billion pesos or more owed 6 percent. The estimation sample is a balanced panel of taxpayers observed in all years plotted in the figure. Panel (a) shows the evolution of reported wealth for taxpayers reporting wealth between 2010 4.5 and 5 billion pesos in 2009 (control) and between 2010 5 and 5.75 billion pesos in 2009 (treated), normalized to zero in 2009 and using specification (A.8). Panel (b) shows the differences between these two series, i.e., our difference-in-differences estimates. The 95 percent confidence intervals are based on robust standard errors clustered at the individual level. Panels (c) and (d) examine how these estimates vary when excluding taxpayers in the bunching area, that is, comparing taxpayers owning between 4 and 4.5 billion versus 6 and 8 billion pesos in 2009.

A.5 Measuring Wealth Inequality in Colombia

Measuring top wealth shares (e.g., the fraction of total wealth held by the top 1 percent) faces challenges due to severe data limitations in Colombia. These data limitations affect both our measure of the amount of wealth held by wealthy individuals (the numerator) and the total amount of wealth held by individuals (the denominator). This section discusses these limitations and describes how we deal with each one of them to estimate top wealth shares in Colombia.

Total Wealth of Non-Filers

Unlike in many developed countries, there is no aggregate wealth measure to construct the denominator in Colombia. National accounts do not report personal wealth estimates and personal financial wealth, as reported by the Central Bank, appears significantly underestimated. Moreover, we cannot compute total wealth as wealth reported in the tax records because only a fraction of tax units file taxes in Colombia. For instance, in FY 2016, taxpayers with gross wealth below 133,889,000 pesos (USD 46,780) did not have to file income taxes. This excluded the bottom 94% of tax units (adults aged 20+) from filing income taxes, which means we do not observe wealth holdings for most tax units. As a second-best alternative, we refer to survey data to capture wealth for non-filers.

For this purpose, we use *Encuesta de Carga Financiera y Educación Financiera de los Hogares* (IEFIC). IEFIC surveys a representative sample of households with formal financial services from three largest urban areas (Bogota, Medellin, and Cali). In 2017, 28,114 households were surveyed from Colombia's main household survey, *Gran Encuesta Integrada de Hogares* (GEIH). Among these surveyed households, 19,419 households reported to have access to financial services and are thus included in IEFIC. Therefore, our initial survey sample comes from 47,347 individuals aged 18 and above. Before any corrections, monthly individual income ranges from 0 to 100 million pesos (USD 0 to 33,500), and household net wealth ranges from 0 to 10 billion pesos (USD 0 to 3,605,362). 45.13% of households are self-reported home-owners.

Using household survey data from IEFIC to estimate the wealth of non-filers has three main issues. The first issue is that the unit of observation is the tax unit in our study (individuals aged 20 and above) while individuals aged 18 and above are included in the survey. We thus drop survey respondents aged below 20 from the sample. Further, some assets and debts are reported at the family-level by the head of household in the survey (real estate, business assets, vehicles, and livestock; and the outstanding debt of each asset), while others are reported at the individual-level (financial assets, consumption debt). This implies that we must make assumptions about the intra-family distribution of assets and debts reported at the family level in the survey. We proceed as follows:

- For family size $n = 1$, we attribute 100% of assets and debts to head of household ($w_h = 1$)

- For family size $n = 2$ with head of household and spouse/partner, we split assets and debts equally ($w_h = w_s = .5$)
- For family size $n \geq 2$ with head of household but no spouse/partner, we attribute 80% to head of household and the remaining 20% split equally across other members ($w_h = .8, w_{j \neq h} = .2/(n-1)$)
- For family size $n > 2$ with head of household and spouse/partner, we attribute 40% to each spouse/partner and the remaining 20% split equally across other members ($w_h = w_s = .4, w_{j \neq h,s} = .2/(n-2)$)

The second issue is valuation. At face value, wealth items reported in the survey are similar to those in the tax records: primary and secondary housing, business assets, real estate properties (e.g., industrial buildings, land, offices, warehouses, parking lots, hotels and lodgings), livestock, vehicles (e.g., motorcycles, private vehicles, boats, planes), inventories, financial assets (e.g., savings accounts, mutual investment funds, shares, swaps), shares and contributions, and voluntary pension contributions are all included in the survey. None of these items are top-coded. However, the survey asks respondents to self-assess their wealth at "market" values. The questionnaire reads as follows: "*If you wanted to sell this asset, what would be the minimum price at which you would sell it?*" Survey respondents are encouraged to use bank account statements to answer questions regarding debts. Nevertheless, it is clear that values reported by survey respondents are not systematically the same as values reported in tax records.

The direction of the bias of wealth items in the survey relative to tax records could go in either direction. On the one hand, survey respondents are more likely to self-report their assets at market values, which are larger than cadastral values. Moreover, given incentives for underreporting wealth in tax records, survey respondents are also likely to overstate their wealth in surveys compared to what they would report to the tax authority. On the other hand, wealthy individuals with financial assets poorly covered in the survey questionnaire will underreport their wealth. Because we focus on potential non-filers in the survey to capture wealth at the bottom of the distribution, it is more likely that our estimates of wealth for this population suffers from upward bias, thus artificially deflating top wealth shares.

The third and last issue is the representativeness of the survey. IEFIC is representative of households in Bogota, Medellin, and Cali that have access to financial services. It is therefore *not* representative of all Colombian adults.[23] Because urban household with access to financial services are likely to be wealthier than other households, this again implies that our estimates of wealth for non-filers will likely suffer from upward bias. Given are wealth denominator will be biased upward, our top wealth shares will be biased downwards. We thus interpret our top wealth shares as conservative estimates of wealth inequality in Colombia.

In the survey data, we find that the wealthiest 10 percent of individuals own 71 percent of all wealth reported in the survey. This is very close to the equivalent top share of 75.3

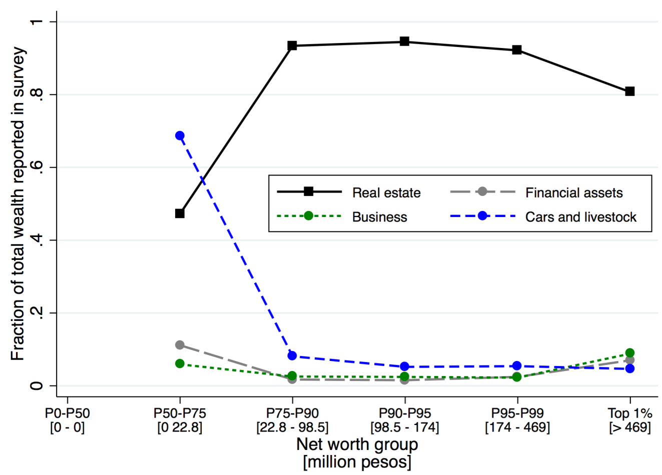
[23]In fact, the sum of survey weights add up to 6,719,291, i.e., 21 percent of all tax units.

percent in the United States, based on data from the Survey of Consumer Finances from 2013 (Saez and Zucman, 2016). Moreover, the wealthiest 1 percent in Colombia own 25.8 percent of total wealth according to IEFIC survey, which is significantly less than the United States' 35.8 percent estimate for households using the Survey of Consumer Finances. Figure E.7 plots wealth decomposition by net wealth groups. The figure shows that 50 percent of individuals have 0 net wealth. This is not surprising, given recent evidence that one-quarter of households in OECD countries have negative net wealth (Balestra and Tonkin, 2018). Individuals in the third quartile (P50–P75) have less than 22.8 million pesos, that is, less than USD 7,641. For these individuals, most of their wealth comes from vehicle and livestock ownership. For middle and upper-middle class individuals (e.g., P75–P95), real estate represents more than 90 percent of wealth. Finally, for individuals in the top 1 percent, the share of wealth belonging to real estate falls to 80.7 percent while the shares of financial and business assets increase. However, it is clear from Figure E.7 that financial assets are underreported in the survey data, making this less than ideal to study wealth inequality at the top. Because our use of survey data is to focus on wealth at the bottom for non-filers, this issue is less of a concern for our purposes.

We impute average net worth for non-filers using net worth of surveyed units with gross wealth below the filing threshold of 133,889,000 pesos. For this group, average net worth is 12,763,333 pesos (USD 4,277). At baseline, non-filers have one-third of total wealth.

Forbes: According to the 2018 Forbes rich list, the fortune of Colombia's richest man alive, Luis Carlos Sarmiento, was worth US \$12.1 billion. Sarmiento ranked 123 on the list of the world's wealthiest individuals, and was followed by the Santo Domingo dynasty (Alejandro and Andrés ranked 449 with US \$4.3 billion; Julio Mario III ranked 1103 with US \$2.2 billion.), Jaime Gilinski Bacal (ranked 606 with US \$3.7 billion), and Carlos Ardila Lulle (ranked 859 with US \$2.8 billion).

Figure E.7: Wealth Decomposition in Survey Data



Notes: This figure plots wealth decomposition by asset types across wealth groups using household survey data from 2017. Individuals aged 20 and above are included. For assets reported at the family level (real estate, business assets, vehicles, and livestock—and the outstanding debt of each asset), we make the following assumptions about the intra-family distribution. For family size $n = 1$, we attribute 100% of assets and debts to head of household. For family size $n = 2$ with head of household and spouse/partner, we split assets and debts equally. For family size $n \geq 2$ with head of household but no spouse/partner, we attribute 80% to head of household and the remaining 20% split equally across other members. Finally, for family size $n > 2$ with head of household and spouse/partner, we attribute 40% to each spouse/partner and the remaining 20% split equally across other members. Individuals are ranked by their net worth. *Sources:* Authors' calculations using 2017 IEFIC from DANE.

Cadastral-to-Market Values

For most middle-class individuals, real estate represents the largest share of gross assets (Balestra and Tonkin, 2018). Yet in Colombia, real estate is reported in tax records at cadastral (not market) values and, as in other developing countries, cadasters are outdated. This implies that cadastral values represent a fraction of market values today. As a result, our measure of real estate in the tax records must be adjusted to obtain wealth at market values W^*:

$$W^* = [K \cdot (1 - \alpha) + \alpha \cdot K \cdot \delta] - L \tag{A.9}$$

where K represents gross wealth as reported in tax records, $\alpha \in [0, 1]$ represents real estate as a fraction of K, $\delta \in [0, 1]$ is the cadastral-to-market value conversion factor, and L represents liabilities. Equation (A.9) thus shows measuring W^* depends critically on accurate measures of α and δ. We discuss how we estimate each parameter next.

Unfortunately, since 2004, wealth is not decomposed by type of assets for most taxpayers, so it is impossible to know what share of assets α should be inflated to reflect market values.

Table E.1: Net wealth groups in survey data

Fractile (1)	Min (million pesos) (2)	Mean (million pesos) (3)	Mean P_0–P_1 (million pesos) (4)	Share of total wealth (%) (5)	Sum of survey weights (6)
P50	0.0	71.39	6.35	100	6,719,291
P75	22.8	136.43	52.00	95.6	6,719,291
P90	98.5	263.10	130.21	73.7	6,719,291
P95	174.0	396.01	264.70	55.5	6,719,291
P99	469.0	923.38	923.38	25.8	6,719,291

Notes: This table plots mean net worth across wealth groups, as well as the minimum wealth needed to belong to each group, using household survey data from Colombia. Individuals aged 20 and above are included. For assets reported at the family level (real estate, business assets, vehicles, and livestock—and the outstanding debt of each asset), we make the following assumptions about the intra-family distribution. For family size $n = 1$, we attribute 100% of assets and debts to head of household. For family size $n = 2$ with head of household and spouse/partner, we split assets and debts equally. For family size $n \geq 2$ with head of household but no spouse/partner, we attribute 80% to head of household and the remaining 20% split equally across other members. Finally, for family size $n > 2$ with head of household and spouse/partner, we attribute 40% to each spouse/partner and the remaining 20% split equally across other members. Individuals are ranked by their net worth. *Sources:* Authors' calculations using 2017 IEFIC from DANE.

To deal with this issue, we obtain α using data from taxpayers required to keep accounting books, which are mostly business owners. In FY 2016, these taxpayers represented 8% of all taxpayers. The tax return used by these taxpayers (income tax form #110) has a "fixed assets" category that includes real estate, land ownership, vehicles, and boats. We assume that the share of fixed assets is similar between individuals required and not required to keep accounting books, and impute estimated shares for all taxpayers. We estimate these shares separately by top wealth groups for FY 2016: $\alpha_{P0-P99} = 0.6, \alpha_{P99-P99.9} = 0.55, \alpha_{P99.9-P99.99} = 0.4$, and $\alpha_{P99.99} = 0.25$.

To inflate cadastral values to reflect market values, we account for the fact that Bogota has done a better job updating its cadasters than other cities in Colombia. We assume cadastral values represent 70 percent of market values in Bogota, and 60 percent in all other cities (hence $\delta = 1/.7 = 1.43$ in Bogota and $\delta = 1/.6 = 1.67$ elsewhere). Unfortunately, there is time and spatial variation in how outdates cadastral values are in Colombia. Legislation has been introduced to force regular updating of cadastres, such that cadastral values be at least 40 (Law 223/1995) or 60 (Law 1450/2011) percent of market values. However, compliance with this norm varies substantially across neighborhoods and time. We ignore these issues and assume δ is the same across individuals within a given city.

Unreported Offshore Wealth

As discussed in Sections 1.4 and 1.5, offshore wealth may be underreported in tax records for the purposes of reducing the tax burden. To the extent that wealthier individuals are disproportionately likely to hold foreign assets, our measures of top wealth shares will underestimate inequality if we do not account for unobserved offshore wealth. Indeed, while the 2015–2017 voluntary disclosure program incentivized some taxpayers to disclose (at least part of) their assets hidden in tax havens, it is likely that other taxpayers choose to continue evading and remain keeping their fortunes concealed from the tax authority. In this section, we place bounds on the total amount of offshore wealth that could potentially remain hidden abroad, and illustrate their implications for estimates of wealth inequality in Colombia.

We begin from the macro estimate for total offshore wealth by Colombians from Alstadsater et al. (2018a). Using fiduciary deposits data from the Central Bank of Switzerland in 2003–2004 as well as cross-border bank deposits data from offshore financial centers in 2007, Alstadsater et al. (2018a) estimate that total offshore wealth by Colombians is 9.0% of GDP. This places Colombia just below the world average of 9.8 percent of GDP kept offshore. How much of this is reported to the tax agency?

In FY 2017, total offshore wealth reported by individuals in tax return #160 for foreign assets amounts to 2.8% of GDP. That is, less than one-third of the baseline measure of offshore wealth is reported to the tax authorities. Half of this amount (1.4% of GDP) was disclosed thanks to the voluntary disclosure program.[24] This means 6.2% of GDP remains concealed offshore. Who holds this offshore wealth?

We assume that the distribution of unreported offshore wealth is similar to the distribution of offshore wealth disclosures made during the 2015–2017 voluntary disclosure program by each net wealth group. Figure E.8 shows the total amount of offshore assets disclosed by wealth group, ranking individuals by their pre- and post-disclosure wealth. We use the black solid line as our estimate of unreported wealth for each wealth group: 58% if P99.99; 24% if P99.95–P99.99, 9% if P99.9–P99.95, 8% if P99.5–P99.9, 0.8% if P99–P99.5, 0.2% if P95–P99, and 0% if P0–P95. We then re-rank individuals according to this augmented measure of wealth and re-compute total wealth accordingly.

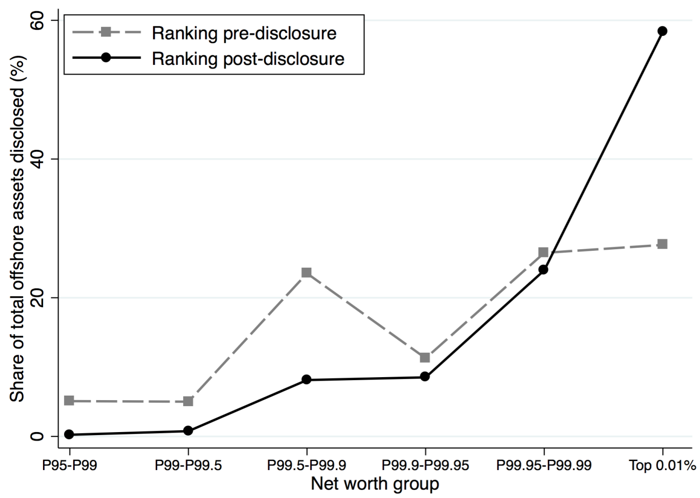
Note, however, that the estimates from (Alstadsater et al., 2018a) are based mostly from 2007, a year that predates high wealth taxation in Colombia. If individuals respond to higher wealth taxes by obscuring their wealth offshore, as suggested in Section 1.4, this implies that our baseline measure of unreported offshore wealth today may be underestimated. How much could overall offshore wealth have increased due to higher wealth taxes in Colombia?

We use the Panama Papers microdata to estimate increases in offshore wealth between 2007 and 2015 due to high wealth taxation in Colombia. The cumulative number of entities ever incorporated through Mossack Fonseca was 400 in 2007 and 1778 in 2015. This represents a 345 percent increase. If the increase in the use of offshore structures also reflects rises in assets held offshore, then offshore wealth today could reach 40 percent of GDP

[24] Total disclosures of hidden foreign and domestic assets and fake liabilities during the 2015–2017 voluntary disclosure program represent 1.73% GDP.

$(40\% = (1 + 3.45) \times 9\%)$ and 37.25 percent of GDP would be unreported to the tax agency. We use this estimate as an upper bound on the total amount of offshore wealth that is unreported to the tax authority today.

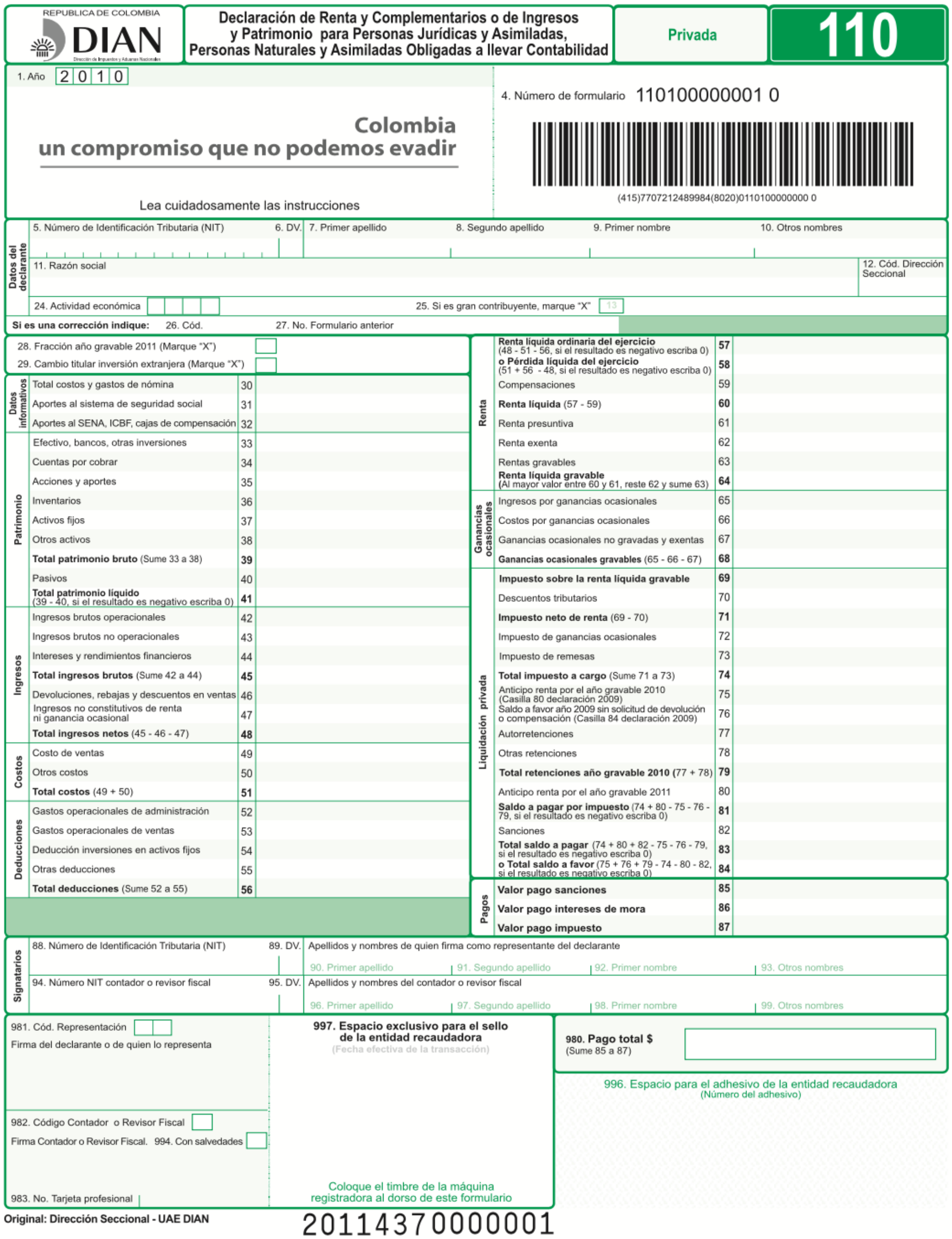
Figure E.8: Distribution of Hidden Offshore Assets in 2015–2017, by Pre/Post-Disclosure Top Wealth Group



Notes: This figure shows the fraction of total disclosures of hidden offshore assets during the 2015–2017 voluntary disclosure program for each wealth group, ranking by pre- and post-disclosure net worth. The figure shows that the volume of offshore assets disclosed in 2015–2017 is increasing in net worth. Tax filers in the wealthiest 0.01 percent post-disclosure disclosed 58 percent of all disclosures. The sample is restricted to 1,633,383 individuals filing the income tax return in FY 2013 (they may or not file a wealth tax return in 2015–2017). This sample includes 11,210 disclosers and 1,085 taxpayers in the Panama Papers (of which 434 disclosed wealth). *Sources:* Authors' calculations using administrative tax microdata from DIAN.

A.6 Income and Wealth Tax Returns in Colombia

Figure F.9: Income Tax Form 110 for Tax Filers Required to Keep Records (2010)



REPUBLICA DE COLOMBIA
DIAN
Dirección de Impuestos y Aduanas Nacionales

Declaración de Renta y Complementarios o de Ingresos y Patrimonio para Personas Jurídicas y Asimiladas, Personas Naturales y Asimiladas Obligadas a llevar Contabilidad

Privada

110

1. Año 2010

4. Número de formulario 110100000001 0

(415)7707212489984(8020)0110100000000 0

Colombia
un compromiso que no podemos evadir

Lea cuidadosamente las instrucciones

Datos del declarante: 5. Número de Identificación Tributaria (NIT) · 6. DV. · 7. Primer apellido · 8. Segundo apellido · 9. Primer nombre · 10. Otros nombres · 11. Razón social · 12. Cód. Dirección Seccional · 24. Actividad económica · 25. Si es gran contribuyente, marque "X" 13

Si es una corrección indique: 26. Cód. 27. No. Formulario anterior

28. Fracción año gravable 2011 (Marque "X")
29. Cambio titular inversión extranjera (Marque "X")

Datos informativos	Total costos y gastos de nómina	30
	Aportes al sistema de seguridad social	31
	Aportes al SENA, ICBF, cajas de compensación	32
Patrimonio	Efectivo, bancos, otras inversiones	33
	Cuentas por cobrar	34
	Acciones y aportes	35
	Inventarios	36
	Activos fijos	37
	Otros activos	38
	Total patrimonio bruto (Sume 33 a 38)	39
	Pasivos	40
	Total patrimonio líquido (39 - 40, si el resultado es negativo escriba 0)	41
Ingresos	Ingresos brutos operacionales	42
	Ingresos brutos no operacionales	43
	Intereses y rendimientos financieros	44
	Total ingresos brutos (Sume 42 a 44)	45
	Devoluciones, rebajas y descuentos en ventas	46
	Ingresos no constitutivos de renta ni ganancia ocasional	47
	Total ingresos netos (45 - 46 - 47)	48
Costos	Costo de ventas	49
	Otros costos	50
	Total costos (49 + 50)	51
Deducciones	Gastos operacionales de administración	52
	Gastos operacionales de ventas	53
	Deducción inversiones en activos fijos	54
	Otras deducciones	55
	Total deducciones (Sume 52 a 55)	56
Renta	**Renta líquida ordinaria del ejercicio** (48 - 51 - 56, si el resultado es negativo escriba 0)	57
	o Pérdida líquida del ejercicio (51 + 56 - 48, si el resultado es negativo escriba 0)	58
	Compensaciones	59
	Renta líquida (57 - 59)	60
	Renta presuntiva	61
	Renta exenta	62
	Rentas gravables	63
	Renta líquida gravable (Al mayor valor entre 60 y 61, reste 62 y sume 63)	64
Ganancias ocasionales	Ingresos por ganancias ocasionales	65
	Costos por ganancias ocasionales	66
	Ganancias ocasionales no gravadas y exentas	67
	Ganancias ocasionales gravables (65 - 66 - 67)	68
Liquidación privada	**Impuesto sobre la renta líquida gravable**	69
	Descuentos tributarios	70
	Impuesto neto de renta (69 - 70)	71
	Impuesto de ganancias ocasionales	72
	Impuesto de remesas	73
	Total impuesto a cargo (Sume 71 a 73)	74
	Anticipo renta por el año gravable 2010 (Casilla 80 declaración 2009)	75
	Saldo a favor año 2009 sin solicitud de devolución o compensación (Casilla 84 declaración 2009)	76
	Autorretenciones	77
	Otras retenciones	78
	Total retenciones año gravable 2010 (77 + 78)	79
	Anticipo renta por el año gravable 2011	80
	Saldo a pagar por impuesto (74 + 80 - 75 - 76 - 79, si el resultado es negativo escriba 0)	81
	Sanciones	82
	Total saldo a pagar (74 + 80 + 82 - 75 - 76 - 79, si el resultado es negativo escriba 0)	83
	o Total saldo a favor (75 + 76 + 79 - 74 - 80 - 82, si el resultado es negativo escriba 0)	84
Pagos	**Valor pago sanciones**	85
	Valor pago intereses de mora	86
	Valor pago impuesto	87

Signatarios: 88. Número de Identificación Tributaria (NIT) · 89. DV. · Apellidos y nombres de quien firma como representante del declarante · 90. Primer apellido · 91. Segundo apellido · 92. Primer nombre · 93. Otros nombres

94. Número NIT contador o revisor fiscal · 95. DV. · Apellidos y nombres del contador o revisor fiscal · 96. Primer apellido · 97. Segundo apellido · 98. Primer nombre · 99. Otros nombres

981. Cód. Representación
Firma del declarante o de quien lo representa

982. Código Contador o Revisor Fiscal
Firma Contador o Revisor Fiscal. 994. Con salvedades

983. No. Tarjeta profesional

997. Espacio exclusivo para el sello de la entidad recaudadora
(Fecha efectiva de la transacción)

Coloque el timbre de la máquina registradora al dorso de este formulario

980. Pago total $ (Sume 85 a 87)

996. Espacio para el adhesivo de la entidad recaudadora
(Número del adhesivo)

Original: Dirección Seccional - UAE DIAN

20114370000001

Figure F.10: Income Tax Form 210 for Tax Filers Not Required to Keep Records (2010)

REPUBLICA DE COLOMBIA
DIAN
Dirección de Impuestos y Aduanas Nacionales

Declaración de Renta y Complementarios Personas Naturales y Asimiladas No Obligadas a llevar Contabilidad

Privada

210

1. Año 2 0 1 0

4. Número de formulario 210100000001 0

(415)7707212489984(8020)02101 00000000 0

Colombia un compromiso que no podemos evadir

Lea cuidadosamente las instrucciones

Datos del declarante

5. Número de Identificación Tributaria (NIT) | 6. DV. | 7. Primer apellido | 8. Segundo apellido | 9. Primer nombre | 10. Otros nombres | 12. Cód. Dirección Seccional

24. Actividad económica

Si es una corrección indique: 25. Cód. 26. No. Formulario anterior

27. Fracción año gravable 2011 (Marque "X")

28. Cambio titular inversión extranjera (Marque "X")

Sección	Concepto	Casilla	Valor
Datos informativos	Total costos y gastos de nómina	29	
	Aportes al sistema de seguridad social	30	
	Aportes al SENA, ICBF, cajas de compensación	31	
Patrimonio	Total patrimonio bruto	32	
	Deudas	33	
	Total patrimonio líquido (32 - 33, si el resultado es negativo escriba 0)	34	
Ingresos	Salarios y demás pagos laborales	35	
	Honorarios, comisiones y servicios	36	
	Intereses y rendimientos financieros	37	
	Otros ingresos (Arrendamientos, etc.)	38	
	Total ingresos recibidos por concepto de renta (Sume 35 a 38)	39	
	Ingresos no constitutivos de renta	40	
	Total ingresos netos (39 - 40)	41	
Costos y deducciones	Deducción inversión en activos fijos	42	
	Otros costos y deducciones	43	
	Total costos y deducciones (42 + 43)	44	
Renta	**Renta líquida ordinaria del ejercicio** (41 - 44, si el resultado es negativo escriba 0)	45	
	o Pérdida líquida del ejercicio (44 - 41, si el resultado es negativo escriba 0)	46	
	Compensaciones (Por exceso de renta presuntiva)	47	
	Renta líquida (45 - 47)	48	
	Renta presuntiva	49	
	Renta exenta	50	
	Rentas gravables	51	
	Renta líquida gravable (Al mayor valor entre 48 y 49, reste 50 y sume 51, si el resultado es negativo escriba 0)	52	
Ganancias ocasionales	Ingresos por ganancias ocasionales	53	
	Costos por ganancias ocasionales	54	
	Ganancias ocasionales no gravadas y exentas	55	
	Ganancias ocasionales gravables (53 - 54 - 55)	56	
Liquidación privada	**Impuesto sobre la renta líquida gravable**	57	
	Descuentos tributarios	58	
	Impuesto neto de renta (57 - 58)	59	
	Impuesto de ganancias ocasionales	60	
	Total impuesto a cargo (Sume 59 a 60)	61	
	Anticipo renta por el año gravable 2010 (Casilla 65 declaración 2009)	62	
	Saldo a favor año 2009 sin solicitud de devolución o compensación (Casilla 69 declaración 2009)	63	
	Total retenciones año gravable 2010	64	
	Anticipo renta por el año gravable 2011	65	
	Saldo a pagar por impuesto (61 + 65 - 62 - 63 - 64, si el resultado es negativo escriba 0)	66	
	Sanciones	67	
	Total saldo a pagar (61 + 65 + 67 - 64 - 63 - 62, si el resultado es negativo escriba 0)	68	
	o Total saldo a favor (62 + 63 + 64 - 61 - 65 - 67, si el resultado es negativo escriba 0)	69	
Pagos	**Valor pago sanciones**	70	
	Valor pago intereses de mora	71	
	Valor pago impuesto	72	

Servicios Informáticos Electrónicos - Más formas de servirle!

Este formulario también puede diligenciarlo ingresando a www.dian.gov.co
Asistido, sin errores y de manera gratuita

Signatario

73. Número de Identificación Tributaria (NIT) | 74. D.V. | Apellidos y nombres de quien firma como representante del declarante: 75. Primer apellido | 76. Segundo apellido | 77. Primer nombre | 78. Otros nombres

981. Cód. Representación

Firma del declarante o de quien lo representa

997. Espacio exclusivo para el sello de la entidad recaudadora
(Fecha efectiva de la transacción)

Coloque el timbre de la máquina registradora al dorso de este formulario

980. Pago total $ (Sume 70 a 72)

996. Espacio para el adhesivo de la entidad recaudadora
(Número del adhesivo)

Original: Dirección seccional - UAE DIAN

20114380000001

Figure F.11: Wealth Tax Form 420 (2004–2011)

REPUBLICA DE COLOMBIA
DIAN
Dirección de Impuestos y Aduanas Nacionales

Declaración y Pago del Impuesto al Patrimonio | **Privada** | **420**

1. Año

4. Número de formulario

Colombia
un compromiso que no podemos evadir

Lea cuidadosamente las instrucciones

Datos del declarante

5. Número de Identificación Tributaria (NIT) | 6. DV. | 7. Primer apellido | 8. Segundo apellido | 9. Primer nombre | 10. Otros nombres

11. Razón social | 12. Cód. Dirección Seccional

24. Si es gran contribuyente, marque "X" 13

25. Contribuyente, marque "X":
1 Si es contribuyente del impuesto al patrimonio creado por la Ley 1370 de 2009 y de la sobretasa creada por el Decreto Legislativo 4825 de 2010
2 Si es contribuyente del impuesto al patrimonio creado por el Decreto Legislativo 4825 de 2010

Si es una corrección indique: 26. Cód. 27. No. Formulario anterior

Sección	Concepto	Casilla	Valor
Base gravable	Total patrimonio bruto	28	
	Menos: Total pasivos	29	
	Patrimonio líquido (28 - 29)	30	
Liquidación privada	Menos: Valor patrimonial neto de los bienes excluidos	31	
	Menos: Valor casa o apartamento de habitación	32	
	Base para el impuesto (30 - 31 - 32)	33	
	Impuesto al patrimonio (Casilla 33 x tarifa)	34	
	Sobretasa	35	
	Sanciones	36	
	Total saldo a pagar (34 + 35 + 36)	37	
Pagos	**Valor pago sanciones**	38	
	Valor pago intereses de mora	39	
	Valor pago impuesto	40	

Servicios Informáticos Electrónicos - Más formas de servirle!

Este formulario también puede diligenciarlo ingresando a www.dian.gov.co
Asistido, sin errores y de manera gratuita

Signatarios

41. Número de Identificación Tributaria (NIT) | 42. DV. | Apellidos y nombres de quien firma como representante del declarante
43. Primer apellido | 44. Segundo apellido | 45. Primer nombre | 46. Otros nombres

47. Número NIT contador o revisor fiscal | 48. DV. | Apellidos y nombres del contador o revisor fiscal
49. Primer apellido | 50. Segundo apellido | 51. Primer nombre | 52. Otros nombres

981. Cód. Representación
Firma del declarante o de quien lo representa

997. Espacio exclusivo para el sello de la entidad recaudadora
(Fecha efectiva de la transacción)

980. Pago total (Sume 38 a 40) $

996. Espacio para el adhesivo de la entidad recaudadora
(Número del adhesivo)

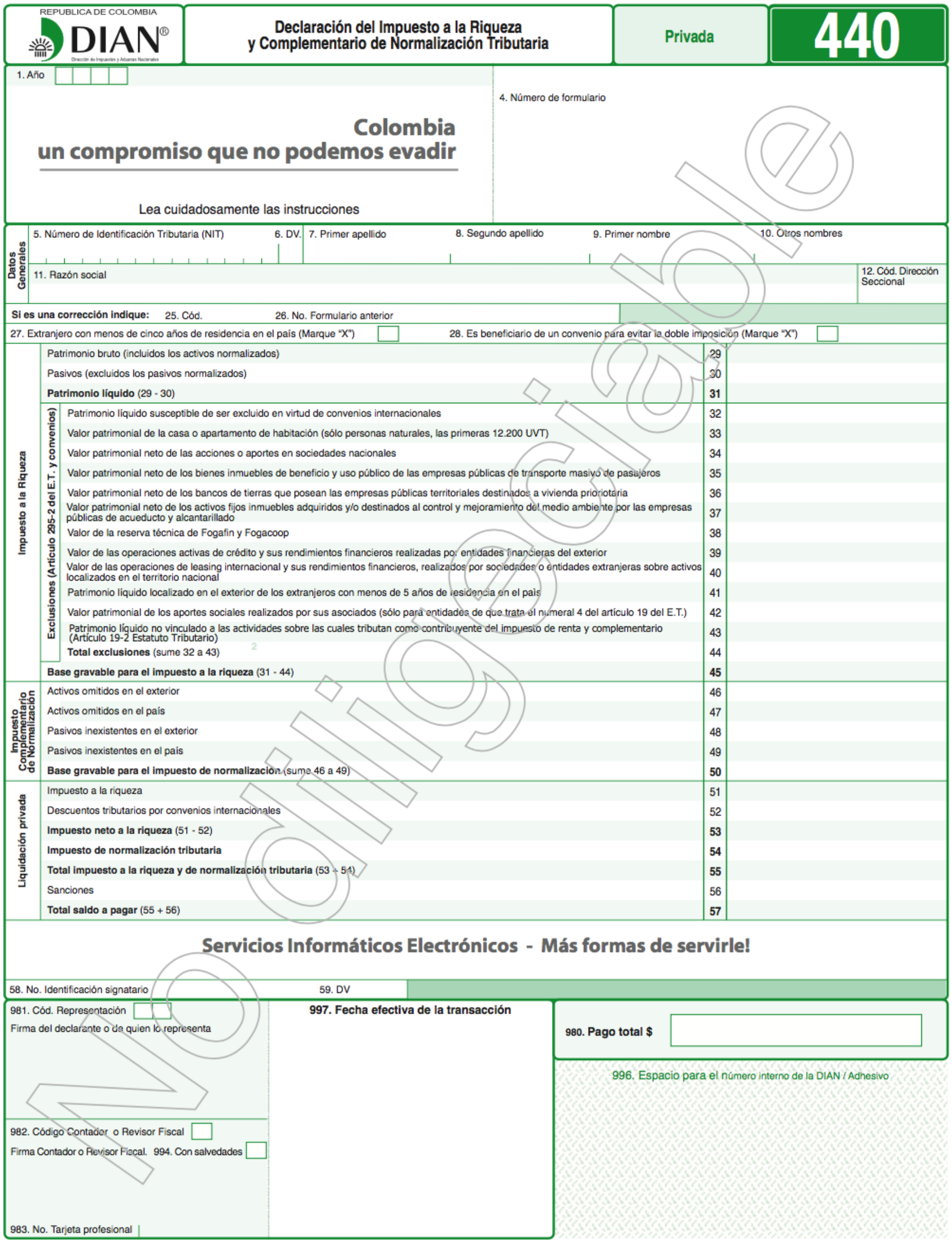
982. Código Contador o Revisor Fiscal
Firma Contador o Revisor Fiscal

983. No. Tarjeta profesional

Coloque el timbre de la máquina registradora al dorso de este formulario

2011423

Figure F.12: Wealth Tax Form 440 (2015–2018)



REPUBLICA DE COLOMBIA

DIAN®

Dirección de Impuestos y Aduanas Nacionales

Declaración del Impuesto a la Riqueza
y Complementario de Normalización Tributaria

Privada

440

1. Año

4. Número de formulario

Colombia
un compromiso que no podemos evadir

Lea cuidadosamente las instrucciones

Datos Generales

5. Número de Identificación Tributaria (NIT) | 6. DV. | 7. Primer apellido | 8. Segundo apellido | 9. Primer nombre | 10. Otros nombres

11. Razón social | 12. Cód. Dirección Seccional

Si es una corrección indique: 25. Cód. 26. No. Formulario anterior

27. Extranjero con menos de cinco años de residencia en el país (Marque "X")

28. Es beneficiario de un convenio para evitar la doble imposición (Marque "X")

Impuesto a la Riqueza	Patrimonio bruto (incluidos los activos normalizados)	29	
	Pasivos (excluidos los pasivos normalizados)	30	
	Patrimonio líquido (29 - 30)	31	
Exclusiones (Artículo 295-2 del E.T. y convenios)	Patrimonio líquido susceptible de ser excluido en virtud de convenios internacionales	32	
	Valor patrimonial de la casa o apartamento de habitación (sólo personas naturales, las primeras 12.200 UVT)	33	
	Valor patrimonial neto de las acciones o aportes en sociedades nacionales	34	
	Valor patrimonial neto de los bienes inmuebles de beneficio y uso público de las empresas públicas de transporte masivo de pasajeros	35	
	Valor patrimonial neto de los bancos de tierras que posean las empresas públicas territoriales destinados a vivienda prioritaria	36	
	Valor patrimonial neto de los activos fijos inmuebles adquiridos y/o destinados al control y mejoramiento del medio ambiente por las empresas públicas de acueducto y alcantarillado	37	
	Valor de la reserva técnica de Fogafin y Fogacoop	38	
	Valor de las operaciones activas de crédito y sus rendimientos financieros realizadas por entidades financieras del exterior	39	
	Valor de las operaciones de leasing internacional y sus rendimientos financieros, realizados por sociedades o entidades extranjeras sobre activos localizados en el territorio nacional	40	
	Patrimonio líquido localizado en el exterior de los extranjeros con menos de 5 años de residencia en el país	41	
	Valor patrimonial de los aportes sociales realizados por sus asociados (sólo para entidades de que trata el numeral 4 del artículo 19 del E.T.)	42	
	Patrimonio líquido no vinculado a las actividades sobre las cuales tributan como contribuyente del impuesto de renta y complementario (Artículo 19-2 Estatuto Tributario)	43	
	Total exclusiones (sume 32 a 43)	44	
	Base gravable para el impuesto a la riqueza (31 - 44)	45	
Impuesto Complementario de Normalización	Activos omitidos en el exterior	46	
	Activos omitidos en el país	47	
	Pasivos inexistentes en el exterior	48	
	Pasivos inexistentes en el país	49	
	Base gravable para el impuesto de normalización (sume 46 a 49)	50	
Liquidación privada	Impuesto a la riqueza	51	
	Descuentos tributarios por convenios internacionales	52	
	Impuesto neto a la riqueza (51 - 52)	53	
	Impuesto de normalización tributaria	54	
	Total impuesto a la riqueza y de normalización tributaria (53 + 54)	55	
	Sanciones	56	
	Total saldo a pagar (55 + 56)	57	

Servicios Informáticos Electrónicos - Más formas de servirle!

58. No. Identificación signatario | 59. DV

981. Cód. Representación

Firma del declarante o de quien lo representa

997. Fecha efectiva de la transacción

980. Pago total $

996. Espacio para el número interno de la DIAN / Adhesivo

982. Código Contador o Revisor Fiscal

Firma Contador o Revisor Fiscal. 994. Con salvedades

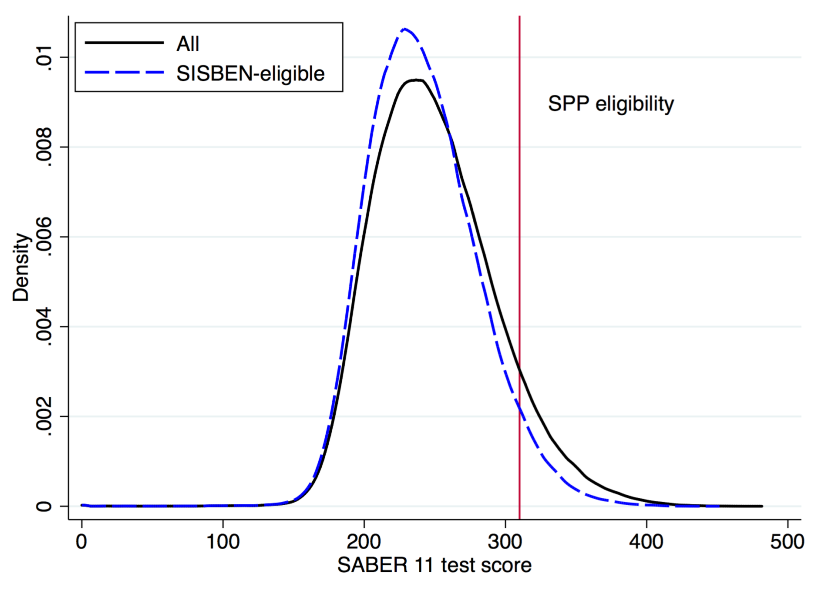
983. No. Tarjeta profesional

No diligenciable

Appendix B

B.1 Appendix Figures and Tables from Chapter 2

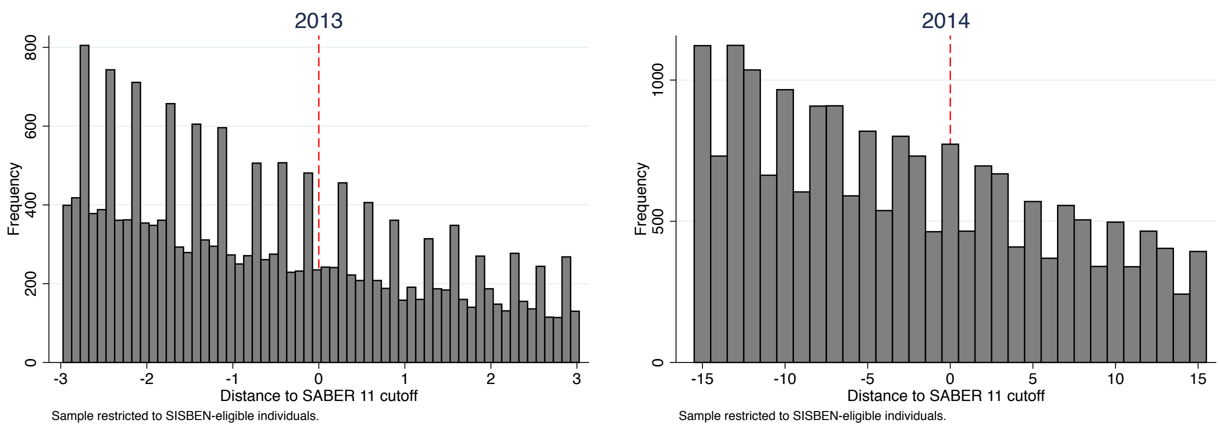
Figure A.1: Distribution of SABER 11 Scores by SISBEN-Eligibility Status



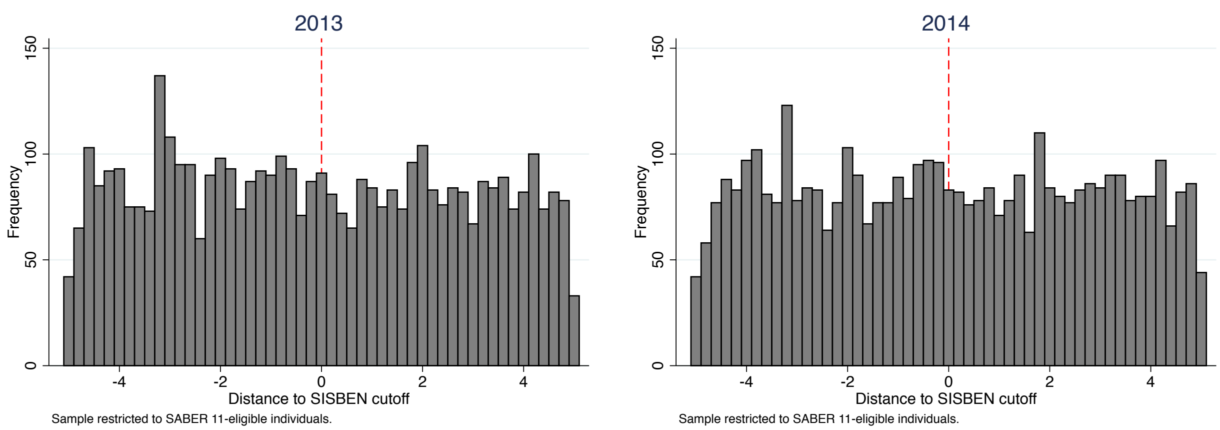
Notes: This figure plots the distribution of SABER 11 scores among Fall 2014 test-takers by SISBEN-eligibility status. The black solid line represents all test-takers, while the blue dashed line represents the subsample of test-takers that are SISBEN-eligible. The red vertical line represents the SPP eligibility cutoff. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure A.2: Histograms of SABER 11 and SISBEN scores in the Fall 2013 and 2014

(a) SABER 11



(b) SISBEN



Note: The SABER 11 score plotted above is the product of a linear combination of exam category sub-scores. Before Fall 2014, the formula is

$$score_i^{pre-20142} = \frac{N_{i,chem} + N_{i,bio} + N_{i,phys} + 2N_{i,sosci} + N_{i,philo} + 3N_{i,lang} + 3N_{i,math} + N_{i,eng}}{13}$$

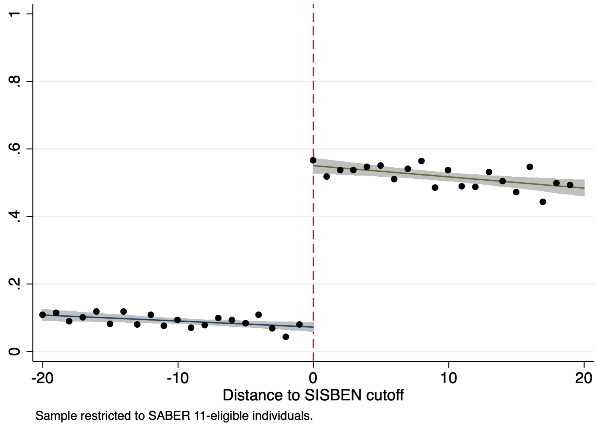
Starting Fall 2014, the formula is

$$score_i^{post-20142} = \frac{500}{1300} \times (3N_{i,reading} + 3N_{i,maths} + 3N_{i,natsci} + 3N_{i,sosci} + N_{i,eng})$$

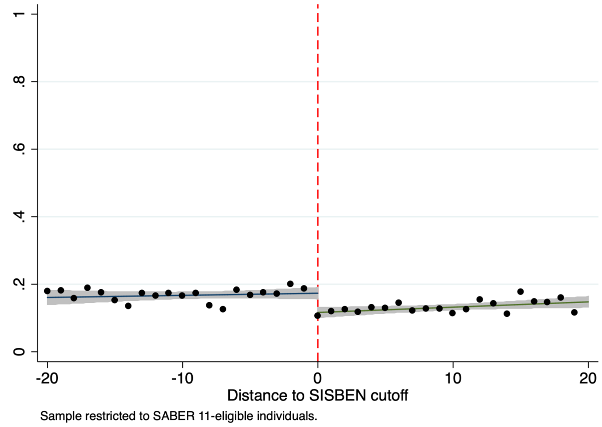
Sources: Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure A.3: Immediate Postsecondary Enrollment: High- vs. Low-Quality, Private vs. Public Institutions (R_i = SISBEN wealth index)

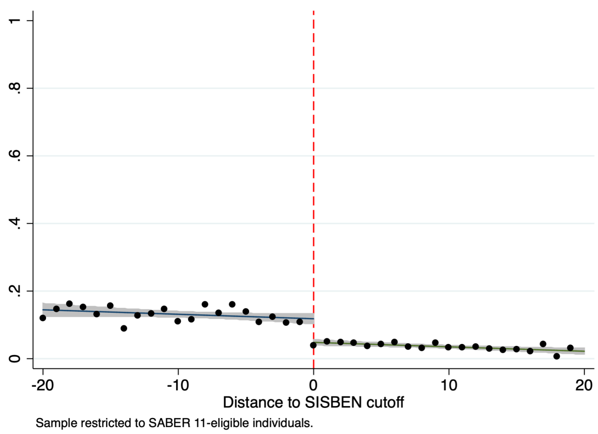
(a) Private High Quality



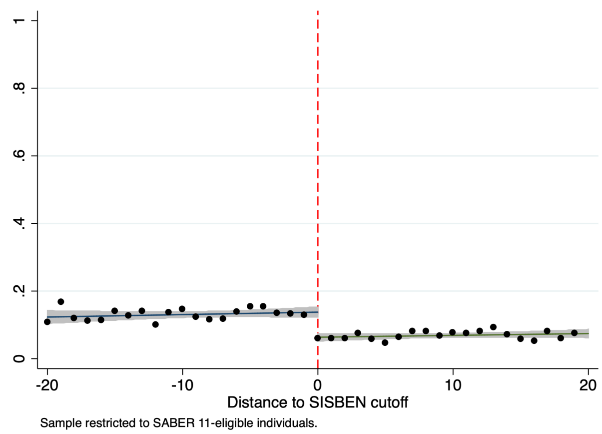
(b) Public High Quality



(c) Private Low Quality



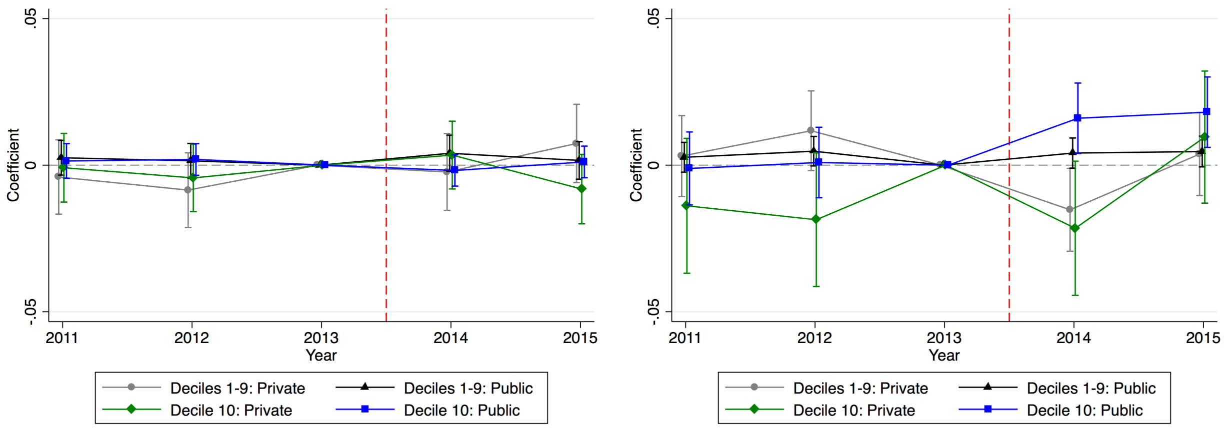
(d) Public Low Quality



Note: The figures plot the probability of immediate enrollment in a private or public, high- or low-quality HEI as a function of the distance to the SISBEN wealth eligibility cutoff. The sample is restricted to SABER 11-eligible students. The figures suggest the likelihood of immediately attending a private, high-quality HEI rose 47.7 percentage points, while the probability of attending a public, high-quality institution decreased by 7.9 percentage points. The likelihood of attending a low-quality institution decreased by 5.2 and 7.6 percentage points for private and public institutions, respectively. See reduced-form estimates in Table 2.2. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure A.4: High-Income Students: Enrollment by Test Score Decile and HEI Type

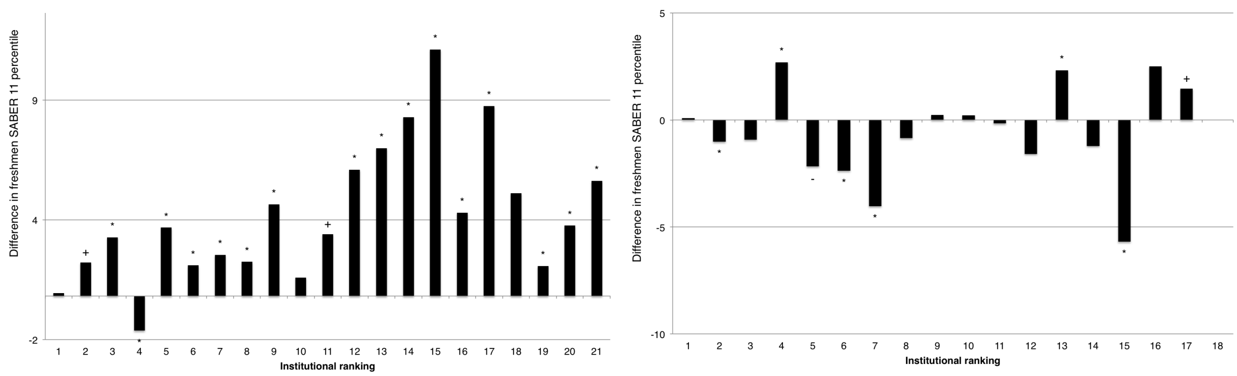
(a) Low-Quality HEIs: Private vs Public (b) High-Quality HEIs: Private vs Public



Notes: These figures plot, for high-income students (strata 4–6), the difference in immediate enrollment probabilities separately by test score decile and HEI quality between Fall (treatment) and Spring (control) test-takers before and after SPP financial aid is introduced (red vertical line) using specification (2.1). Beside a temporary displacement effect from private high-quality HEIs (Panel B, green line), financial aid expansion had no statistically significant effect on enrollment of high-income students. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

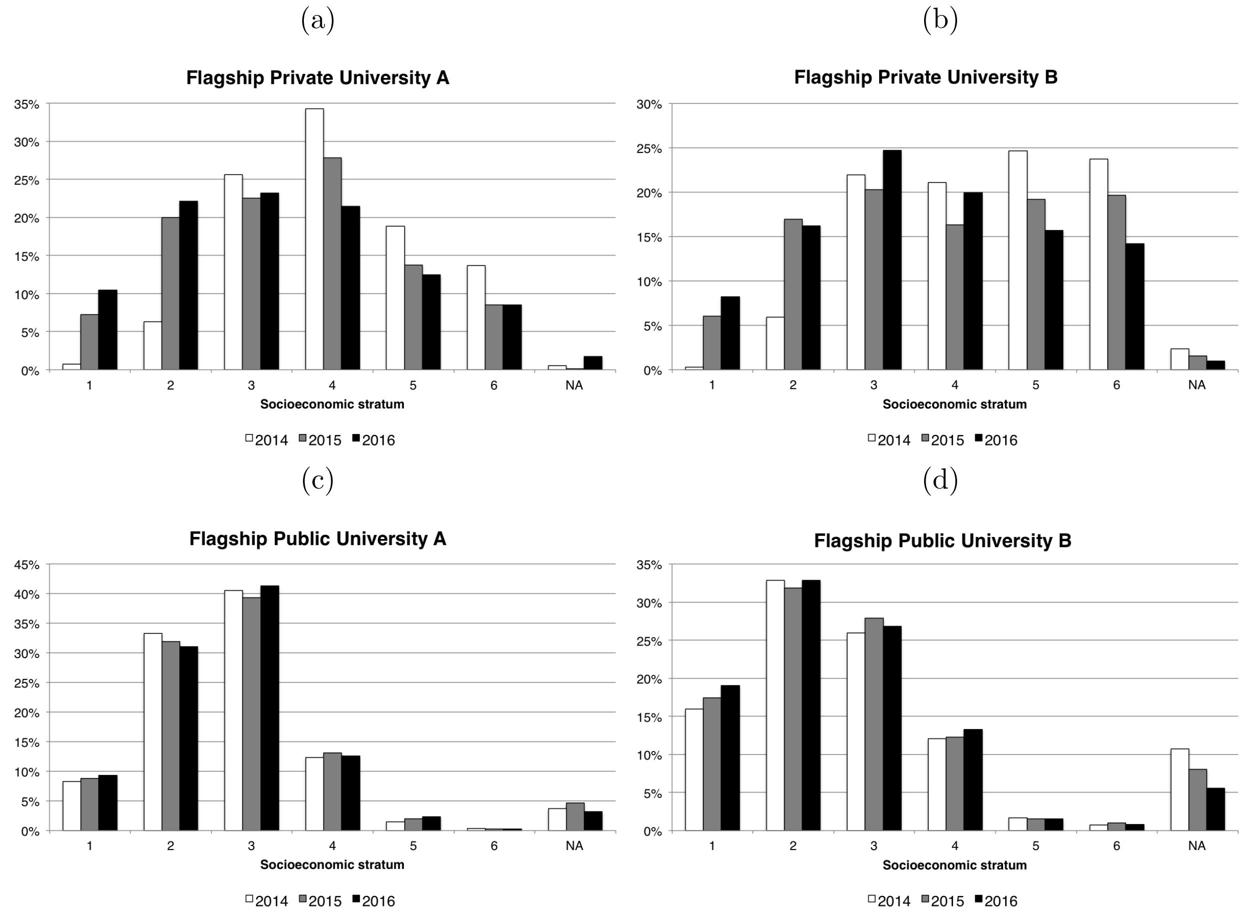
Figure A.5: Difference in Mean Percentile of Entering Students

(a) Private High-Quality HEIs (b) Public High-Quality HEIs



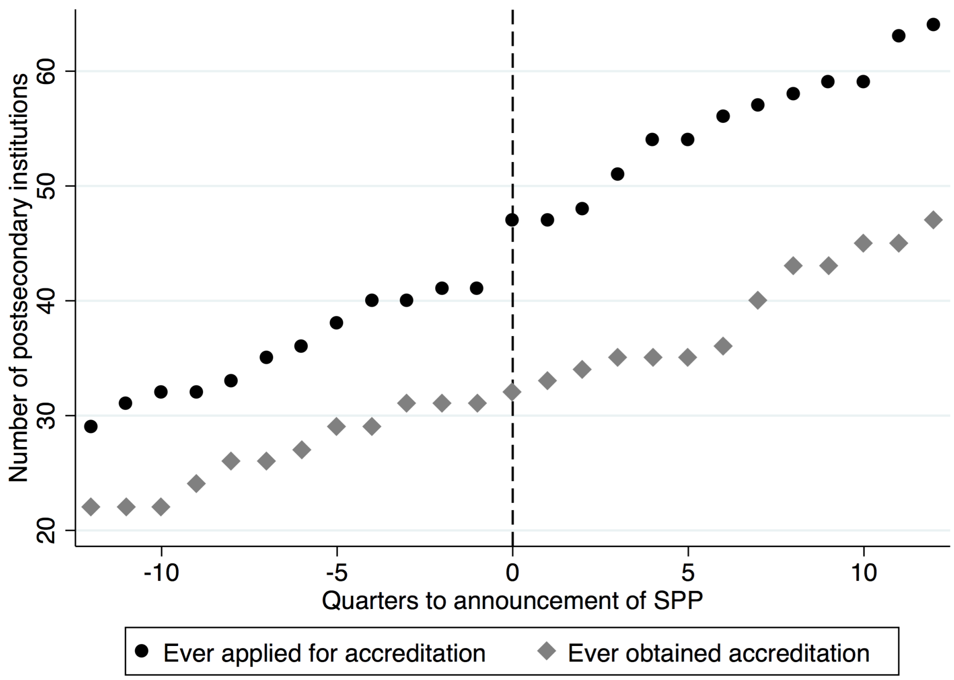
Note: The figures plot the difference between Spring 2014 and 2015 in the mean SABER 11 percentile of the entering cohort for students taking SABER 11 in Fall 2013 and 2014, respectively, aged 14 to 23. Institutions are ordered by mean SABER 11 score of their Spring 2014 entering cohort. Panel A suggests that the average student quality significantly increased for most private, high-quality institutions, with the magnitude of the impact being inversely proportional to the institutional ranking (as measured by the average quality of admitted students before SPP). Panel B suggests that average student quality decreased for many public, high-quality institutions. $-p < 0.1, +p < 0.05, *p < 0.01$. *Sources:* Authors' calculations based on ICFES, DNP, MEN, and SPADIES (2016).

Figure A.6: Freshmen Socioeconomic Stratum at Selected High-Quality HEIs



Notes: These figures plot the distribution of first-semester students by socioeconomic stratum in 2014–2016 Spring semesters at selected accredited institutions. At the flagship private institution in Bogota, the share of first-semester students in the bottom two strata increased from 7.0 percent to 32.6 between 2014 and 2016 (Panel A), while the share of students from the bottom two strata decreased by 3.0 percent at a top-ranked public institution in Bogota (Panel C).
Sources: Authors' calculations based on SPADIES (2016) and administrative record from universities.

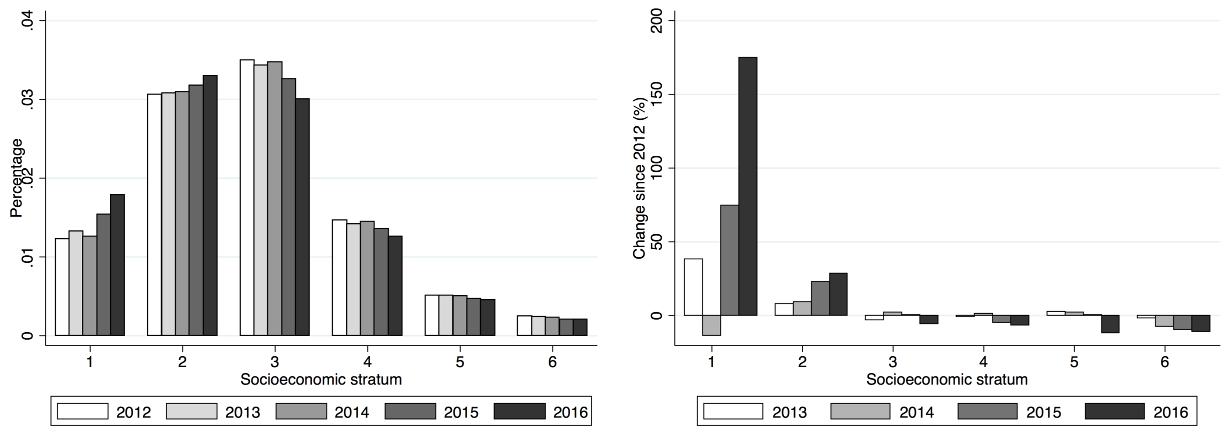
Figure A.7: Evolution of the Number of HEIs with High Quality Accreditation



Note: The figure plots the number of HEIs ever having requested institutional *High Quality Accreditation* status (in black) and the number of HEIs ever having received this status (in gray) in Colombia. The number of HEIs requesting this status increased the moment SPP was announced, while the number of HEIs having obtained this status did not. *Sources:* Authors' calculations based on MEN and SPADIES (2016).

Figure A.8: Gains in Test Performance for Low-Income Students

(a) Distribution of top decile by strata (b) Change in share of strata in top percentile



Note: Panel A plots the distribution of socioeconomic strata for the SABER 11 top 10% of performing students in the Fall semesters between 2012 and 2016, and suggests that the share of top performers from the bottom stratum increased by 45.8 percent between 2012 and 2016. Panel B focuses on the top percentile and plots the percentage change between in Fall 2013 through 2016, using 2012 as baseline. The sample in all figures is restricted test-takers aged 14–23. *Sources:* Authors' calculations based on ICFES (2016).

Table A.1: Manipulation Testing based on Density Discontinuity

	Running variable	
	SABER 11 (1)	SISBEN (2)
Robust Bias-Corrected p-value	0.2179786	0.4364368
Robust Bias-Corrected SE	0.0001818	0.0017997
Number of obs - left	284052	7709
Number of obs - right	15423	15423
Eff. number of obs - left	6042	2731
Eff. number of obs - right	6122	3692
Order loc. poly. (p)	2	2
Order BC (q)	3	3
Bandwidth values - left	8.705387	6.21869
Bandwidth values - right	10.00451	8.619373

Notes: This table tests for manipulation of the running variable based on density discontinuity. All results are estimated with package rddensity (Cattaneo et al., 2016a) using an unrestricted model and a triangular kernel function, and employ the jackknife standard errors estimator. Column (1) restricts the sample to SISBEN-eligible individuals. Column (2) restricts the sample to SABER 11-eligible individuals. The table suggests we cannot statistically detect manipulation in either variable. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Table A.2: Baseline Covariate Balance Test around SABER 11 and SISBEN Cutoffs

	Running Variable			
	SABER 11		SISBEN	
	Coef/SE	Mean Control	Coef/SE	Mean Control
	(1)	(2)	(3)	(4)
Female	-0.017 (0.013)	0.473	-0.004 (0.026)	0.442
Age	-0.008 (0.051)	16.627	0.143 (0.108)	16.393
Ethnic minority	0 (0.006)	0.079	-0.007 (0.011)	0.066
Employed	-0.006 (0.006)	0.085	-0.002 (0.012)	0.073
Family size	-0.042 (0.033)	4.601	-0.132 (0.064)	4.378
Mother education: Primary	-0.009 (0.012)	0.252	0.01 (0.016)	0.130
Mother education: Secondary	-0.011 (0.012)	0.504	-0.073 (0.026)	0.486
Mother education: T&T	0.001 (0.007)	0.134	0.005 (0.019)	0.178
Mother education: Higher	0.019 (0.008)	0.110	0.054 (0.021)	0.207
Father education: Primary	0 (0.011)	0.343	0.014 (0.02)	0.184
Father education: Secondary	-0.003 (0.011)	0.429	-0.052 (0.022)	0.444
Father education: T&T	0.002 (0.008)	0.103	-0.011 (0.018)	0.173
Father education: Higher	0.004 (0.008)	0.122	0.051 (0.022)	0.198
Household SES: Stratum 1	-0.002 (0.01)	0.343	-0.013 (0.015)	0.131
Household SES: Stratum 2	-0.013 (0.011)	0.459	-0.003 (0.021)	0.516
Household SES: Stratum 3	0.01 (0.009)	0.184	0.008 (0.02)	0.326
Household SES: Stratum 4	0.007 (0.003)	0.010	0.011 (0.007)	0.020
Household SES: Stratum 5	0 (0.001)	0.003	-0.003 (0.003)	0.007
Household SES: Stratum 6	-0.001 (0.001)	0.001	0 (0.001)	0.001
School hours: Full day	-0.002 (0.009)	0.197	0.035 (0.023)	0.286
School hours: Morning	-0.002 (0.011)	0.614	-0.049 (0.029)	0.548
School hours: Evening	0.003 (0.002)	0.009	0.001 (0.003)	0.006
School hours: Afternoon	-0.001 (0.008)	0.172	0.019 (0.021)	0.154
School hours: Weekends	0.003 (0.002)	0.008	-0.006 (0.003)	0.008
Private school	-0.002 (0.008)	0.172	0.067 (0.025)	0.301
Joint F-Stat (*p*-value, LB on b-width)	0.347		0.082	
Joint F-Stat (*p*-value, UB on b-width)	0.475		0.196	

Notes: This table plots the reduced-form coefficient from a RD specification where the outcome is a baseline characteristic and the running variable is either SABER 11 test scores in columns (1) and (2) or SISBEN poverty index in columns (3) and (4). The sample is restricted to SISBEN-eligible individuals in columns (1) and (2), and SABER 11-eligible individuals in columns (3) and (4). Columns (1) and (3) present conventional coefficients and standard errors in parentheses, while columns (2) and (4) present control means. The last two rows report the p-value from a joint significance test using all baseline characteristics, and small or large bandwidths ($\pm$ 20 or 40 test score units in column (1); and $\pm$ 7 or 15 household wealth units in column (3)). All results are estimated with package rdrobust (Cattaneo et al., 2014). *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Table A.3: Robustness Check: Predicted Access by Type of Postsecondary Institution

	Any	High Quality			Low Quality		
		Any	Private	Public	Any	Private	Public
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: SABER 11 as the Running Variable						
RD Coefficient	0.006	0.004	0.002	0.001	0.002	0.002	0.001
	(0.004)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
Realized Access	0.32	0.465	0.466	0	-0.154	-0.063	-0.087
Mean Control	0.37	0.109	0.033	0.075	0.267	0.105	0.159
	Panel B: SISBEN as the Running Variable						
RD Coefficient	0.02	0.01	0.009	0.002	0.01	0.01	0
	(0.008)	(0.004)	(0.003)	(0.001)	(0.004)	(0.003)	(0.001)
Realized Access	0.274	0.396	0.477	-0.079	-0.12	-0.052	-0.076
Mean Control	0.485	0.261	0.067	0.194	0.225	0.097	0.134

Note: The table plots the reduced-form coefficient from a RD specification where the dependent variable is predicted access to postsecondary education using all 25 baseline characteristics from the universe of test-takers in Fall 2013. Panel A uses SABER 11 test scores as the running variable, while Panel B uses SISBEN wealth index as the running variable. The first row in each panel presents the RD estimate with robust standard errors. The third and forth rows reproduce the estimates and control means from Table 2.2. The table shows that controls cannot predict the magnitude of the effects we document of financial on immediate postsecondary enrollment. Bias-corrected RD results estimated with package rdrobust (Cattaneo et al., 2014). Robust standard errors in parentheses.
Sources: Authors' calculations based on ICFES, DNP, and MEN (2016).

Table A.4: Immediate Postsecondary Enrollment, by Type of Institution (With and Without Controls)

	Any	High Quality			Low Quality		
		Any	Private	Public	Any	Private	Public
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: SABER 11 as the Running Variable						
RF (without controls)	0.32	0.465	0.466	0	-0.154	-0.063	-0.087
	(0.012)	(0.012)	(0.011)	(0.007)	(0.011)	(0.007)	(0.009)
RF (with controls)	0.329	0.479	0.484	-0.004	-0.155	-0.066	-0.088
	(0.012)	(0.012)	(0.012)	(0.007)	(0.011)	(0.007)	(0.008)
Mean Control	0.495	0.164	0.044	0.118	0.341	0.08	0.265
Observations	273,361	273,361	273,361	273,361	273,361	273,361	273,361
BW Loc. Poly.	28.317	24.143	22.777	27.582	26.247	30.453	27.345
Effect Obs Control	27,382	21,516	19,087	25,728	24,109	30,166	25,728
Effect Obs Treat	10,500	9,654	9,199	10,304	10,072	10,834	10,304
	Panel B: SISBEN as the Running Variable						
RF (without controls)	0.274	0.396	0.477	-0.079	-0.12	-0.052	-0.076
	(0.027)	(0.024)	(0.02)	(0.018)	(0.022)	(0.015)	(0.016)
RF (with controls)	0.253	0.392	0.479	-0.085	-0.135	-0.059	-0.082
	(0.026)	(0.025)	(0.021)	(0.019)	(0.022)	(0.015)	(0.016)
Mean Control	0.69	0.3	-0.009	0.319	0.389	0.148	0.239
Observations	21,071	21,071	21,071	21,071	21,071	21,071	21,071
BW Loc. Poly.	10.422	10.516	11.775	11.255	9.347	10.613	11.021
Effect Obs Control	4,085	4,117	4,536	4,363	3,730	4,140	4,284
Effect Obs Treat	4,116	4,154	4,612	4,417	3,708	4,182	4,339

Note: This table compares how the exclusion or inclusion of baseline covariates affects the reduced-form coefficient on the effect of financial aid eligibility on immediate postsecondary enrollment using a regression discontinuity design. The dependent variable is immediate enrollment by type of postsecondary institution (e.g., high-quality, low-quality, private, public). Panel A uses SABER 11 test score as the running variable, restricting the sample to SISBEN-eligible students. Panel B uses SISBEN wealth index as the running variable, restricting the sample to SABER 11-eligible students. The first row reproduces the main reduced-form estimates from Table 2.2, while the next rows include baseline covariates in the regression. Including controls, the reduced-form coefficient in Column (1) of Panel A suggests that, for individuals below a certain level of poverty, financial aid eligibility raises immediate postsecondary enrollment by 32.9 percentage points. On a basis of 49.5 percent, this implies a 66.5 percent increase in immediate enrollment. Bias-corrected RD results estimated with package rdrobust (Cattaneo et al., 2014). Robust standard errors in parentheses. *Sources:* Authors' calculations based on ICFES, DNP, MEN, and SPADIES (2016).

Table A.5: Reason Student Chose to Attend her HEI, by Type

	Type of Postsecondary Institution		
	Any	Private	Public
It offers the major I want	58.1%	62.0%	52.4%
It's prestigious	47.4%	55.7%	36.8%
It's better academically	35.2%	42.4%	27.0%
It offers better job prospects	28.5%	36.1%	18.1%
It's public	24.1%	1.4%	62.7%
It's affordable	20.3%	9.1%	31.2%
I liked the curriculum there better	17.1%	13.5%	8.6%
It's private	6.9%	10.7%	1.1%
I have friends or relatives there	6.5%	6.3%	5.8%
It provides better contacts	5.5%	6.7%	4.5%
It's the only one that admitted me	4.5%	4.9%	3.9%
It's the only one available in my region	3.5%	3.0%	3.9%
N	1012	571	359

Note: The data is based on a survey collected from 1,487 SISBEN-eligible individuals that took SABER 11 in Fall 2015 and scored just below or above the SABER 11 eligibility cutoff. 82 observations did not match with an institution.
Source: Authors' calculations based on survey data.

Table A.6: Difference-in-Differences Outcomes by Type of Institution

	Any	High Quality			Low Quality		
		Any	Private	Public	Any	Private	Public
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: *Enrollment by Socioeconomic Strata and Test Score Decile*						
	Panel A.1: Strata 1–3 Only						
All Deciles							
DiD	0.0384	0.0232	0.0171	0.0061	0.0152	0.0089	0.0063
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean Control	0.1789	0.0556	0.0194	0.0363	0.1233	0.073	0.0503
Deciles 1-9							
DiD	0.0237	0.0061	0.0026	0.0035	0.0176	0.0098	0.0078
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean Control	0.1688	0.0452	0.0157	0.0295	0.1236	0.0741	0.0495
Decile 10							
DiD	0.1819	0.1983	0.1573	0.041	-0.0164	-0.0049	-0.0115
	(0.007)	(0.007)	(0.004)	(0.006)	(0.005)	(0.003)	(0.004)
Mean Control	0.4182	0.2748	0.1022	0.1725	0.1435	0.056	0.0875
	Panel A.2: Strata 4–6 Only						
All Deciles							
DiD	0.0036	-0.0016	-0.009	0.0074	0.0051	0.0045	0.0007
	(0.005)	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)	(0.001)
Mean Control	0.4201	0.3135	0.2723	0.0413	0.1065	0.0862	0.0203
Deciles 1-9							
DiD	-0.0009	-0.009	-0.0109	0.0019	0.008	0.0065	0.0015
	(0.006)	(0.005)	(0.005)	(0.002)	(0.005)	(0.004)	(0.002)
Mean Control	0.3413	0.2023	0.1792	0.0232	0.139	0.113	0.026
Decile 10							
DiD	0.0188	0.0214	0.0046	0.0168	-0.0026	-0.0011	-0.0015
	(0.008)	(0.008)	(0.007)	(0.004)	(0.004)	(0.004)	(0.002)
Mean Control	0.5481	0.4948	0.4206	0.0742	0.0533	0.043	0.0103
	Panel B: *Share of Entering Students Scoring in Top Decile*						
DiD	-0.0057	0.0762	0.1268	0.0057	-0.0424	-0.0295	-0.0662
	(0.009)	(0.019)	(0.021)	(0.016)	(0.006)	(0.005)	(0.012)
Mean Control	0.1443	0.3627	0.39	0.3242	0.0438	0.0442	0.0426
	Panel C: *Share of Entering Students From Strata 1–3*						
DiD	0.0313	0.0904	0.1367	-0.0031	0.005	0.0082	0.0012
	(0.008)	(0.018)	(0.019)	(0.008)	(0.004)	(0.006)	(0.004)
Mean Control	0.7393	0.5258	0.2987	0.8395	0.8724	0.8328	0.9388

Note: This table presents the difference-in-differences coefficients on a series of outcomes. Panel A plots the results from specification (2.1) while Panels B and C plot the results from specification (2.2). Robust standard errors in parentheses. These regressions exclude baseline covariates. *Sources:* Authors' calculations based on ICFES, DNP, MEN, and SPADIES (2016).

B.2 Alternative Measures of College Quality

This section investigates how the RD results presented and discussed in Section 2.4 compare when using other metrics of college quality. This helps further understand how financial aid, by shifting students to HEIs with "high-quality" accreditation, affects the peer composition and institutional resources students are exposed to. For this purpose, we use alternative measures of resource and quality differences across universities.

Specifically, we use six proxies of institutional quality in 2014 (i.e., before SPP was introduced):

1. *Mean SABER 11 scores*: Mean standardized high school exit scores of first-year students enrolling in that university for the first time in Spring 2014.

2. *Mean SABER PRO scores*: Mean standardized score in the national college exit exam, SABER PRO, of students taking that exam in Spring 2014. Since 2009, all college students are required to present this test to graduate from an undergraduate program in Colombia. This allows us to compare the value added of colleges in terms of student performance at the time of graduation.

3. *Graduation Rate*: Percentage of students graduating from that HEI within 14 semesters from initial enrollment.

4. *Faculty with a Doctorate*: Percentage of faculty members with a doctoral degree.

5. *Log Spending per Student*: Logarithm of total operating expenses divided by the number of undergraduate students enrolled.

6. *Log Research Spending per Faculty*: Logarithm of the ratio of operating expenses assigned to research and the number of faculty members.

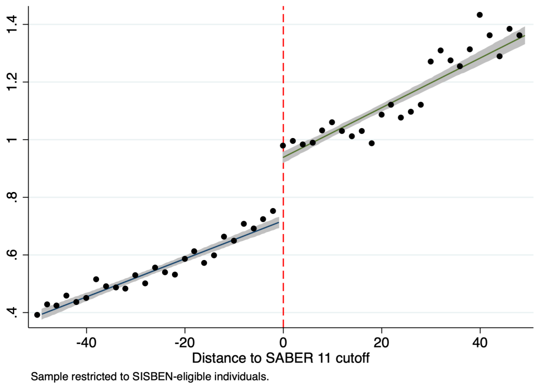
We restrict the RD estimation sample to Fall 2014 test-takers who immediately enroll in any university after having taken SABER 11 exam. Hence, we are interested in the impact of financial aid on HEI quality along the *intensive* margin, comparing different metrics of institutional quality. As always, we present results for the two populations of compliers separately. Figure B.9 and Panel A in Table B.7 use SABER 11 as the running variable, while Figure B.10 and Panel B in Table B.7 use SISBEN as the running variable.

Using SABER 11 as the running variable, the results suggest financial aid significantly improved the average quality of the HEI attended by high school seniors immediately after graduation, as measured through our six indicators above described. The results are large and precisely estimated (Table B.7, Panel A). Financial aid eligibility raised peer quality (mean high school test scores), college quality (college exit test scores, graduation rate, faculty with a doctorate), and the resources students are exposed to (spending per student, research spending per faculty member).

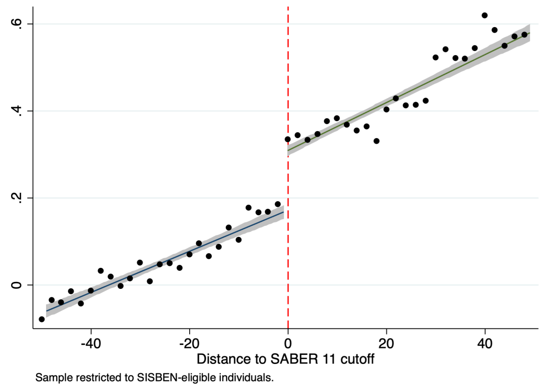
The results are qualitatively similar, although at times less precisely estimated, using SISBEN as the running variable. Figure B.10 and Panel B in Table B.7 show that, among attendees, the quality of the university they attend is falling as the SISBEN score approaches the eligibility cutoff: students are progressively poorer and thus enroll in worse-quality universities. Aid eligibility increases enrollment in higher-quality universities. Aid eligibility improved access to universities with higher peer quality, measured by mean SABER 11 scores, as well as the graduation rate. However, financial aid had no statistically significant impact on the quality of universities attended measured through resources available to students (columns 4–6).

Figure B.9: Alternative Measures of University Quality (R_i = SABER 11 test score)

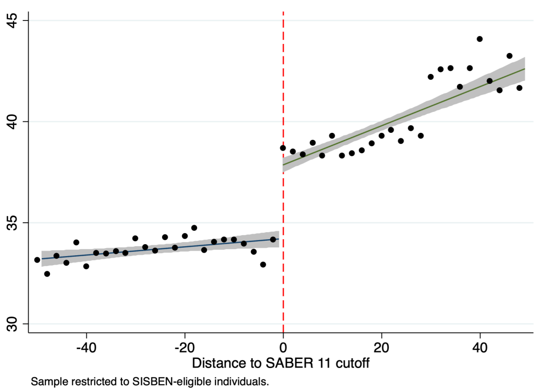
(a) Mean (Standardized) SABER 11 scores



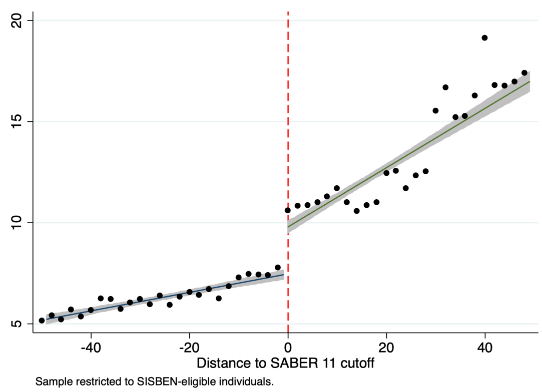
(b) Mean (Standardized) SABER PRO scores



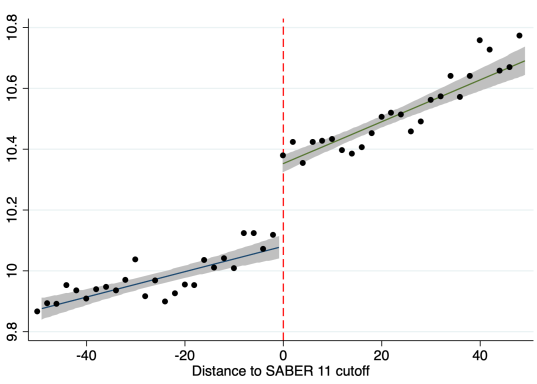
(c) Graduation Rate (%)



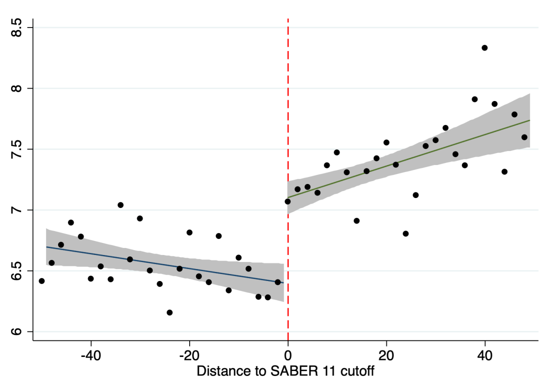
(d) Faculty with a Doctorate (%)



(e) Log Spending per Student



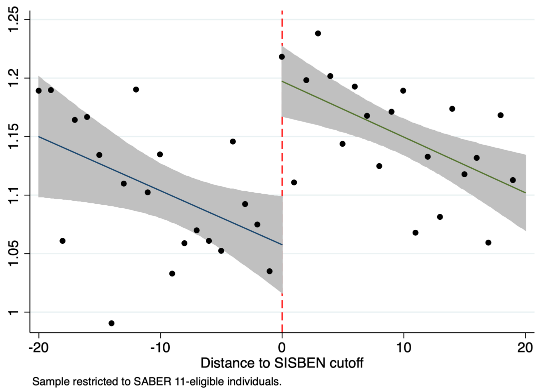
(f) Log Research Spending per Faculty



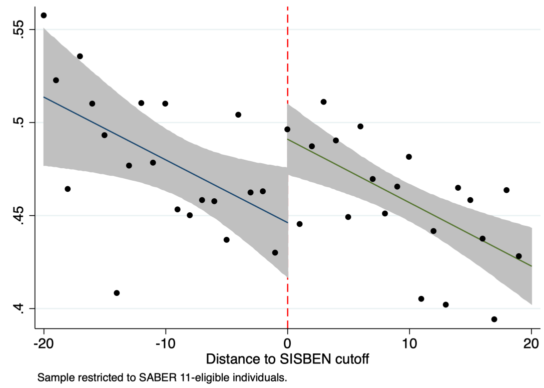
Note: The figures plot six different measures of the quality of the college students attend as a function of the distance to SABER 11 SPP eligibility cutoffs. The sample is restricted to Fall 2014 test-takers who immediately enrolled in a university the semester after having taken SABER 11 exam and are SISBEN-eligible. See reduced-form estimates in Table B.7. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure B.10: Alternative Measures of University Quality (R_i = SISBEN score)

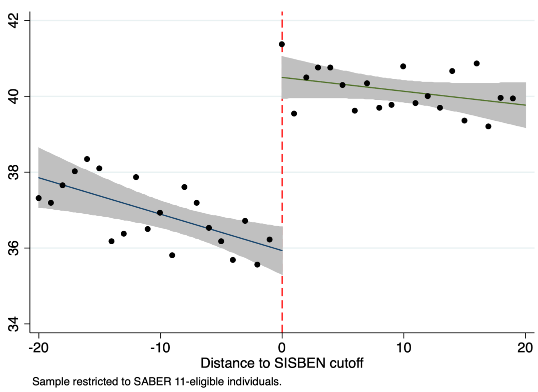
(a) Mean (Standardized) SABER 11 scores



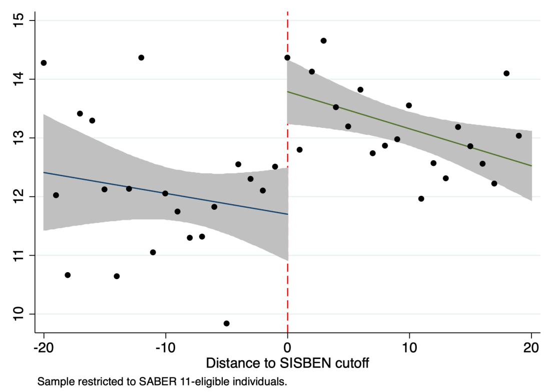
(b) Mean (Standardized) SABER PRO scores



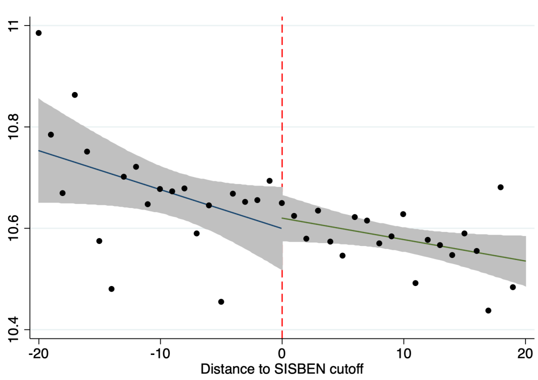
(c) Graduation Rate (%)



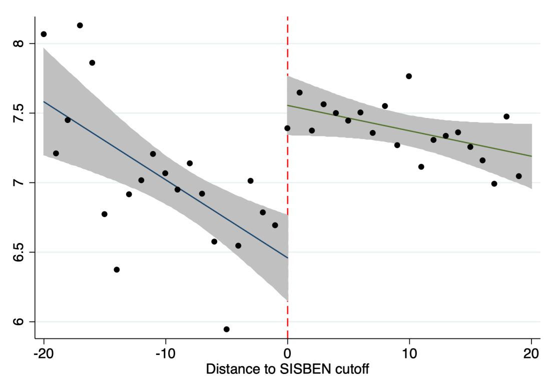
(d) Faculty with a Doctorate (%)



(e) Log Spending per Student



(f) Log Research Spending per Faculty



Note: The figures plot six different measures of the quality of the college students attend as a function of the distance to SISBEN SPP eligibility cutoffs. The sample is restricted to Fall 2014 test-takers who immediately enrolled in a university the semester after having taken SABER 11 exam and are SABER 11-eligible. See reduced-form estimates in Table B.7. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Table B.7: The Impact of Financial Aid on University Quality, by Measure of Quality

	Mean SABER 11 scores (1)	Mean SABER PRO scores (2)	Graduation Rate (%) (3)	Faculty with a Doctorate (%) (4)	Log Spending per Student (5)	Log Research Spending per Faculty (6)
	Panel A: SABER 11 as the Running Variable					
ITT	0.222	0.136	5.211	2.893	0.24	0.776
	(0.025)	(0.018)	(0.535)	(0.364)	(0.04)	(0.178)
LATE	0.273	0.167	6.433	3.562	0.294	0.954
	(0.03)	(0.022)	(0.656)	(0.445)	(0.05)	(0.22)
Mean Control	0.765	0.204	33.454	7.918	10.156	6.36
First Stage	0.813	0.814	0.81	0.812	0.814	0.814
	(0.01)	(0.01)	(0.011)	(0.01)	(0.01)	(0.009)
Observations	30,373	30,373	30,373	30,373	30,373	30,373
BW Loc. Poly.	22.549	23.402	17.397	21.129	26.688	29.829
Effect Obs Control	4,589	4,876	3,409	4,309	5,527	6,286
Effect Obs Treat	6,928	7,165	5,943	6,695	7,651	8,085
	Panel B: SISBEN as the Running Variable					
ITT	0.118	0.029	4.981	0.847	-0.047	0.662
	(0.047)	(0.035)	(0.794)	(1.034)	(0.078)	(0.367)
LATE	0.144	0.036	6.154	1.049	-0.057	0.817
	(0.057)	(0.043)	(0.967)	(1.269)	(0.096)	(0.449)
Mean Control	1.07	0.448	35.906	12.954	10.675	6.77
First Stage	0.814	0.81	0.809	0.809	0.822	0.811
	(0.018)	(0.02)	(0.02)	(0.021)	(0.017)	(0.019)
Observations	14,418	14,418	14,418	14,418	14,418	14,418
BW Loc. Poly.	10.241	9.002	8.736	7.374	12.068	8.958
Effect Obs Control	1,885	1,668	1,632	1,401	2,157	1,663
Effect Obs Treat	3,262	2,889	2,791	2,319	3,796	2,871

Note: This table presents the effect of financial aid on six different measures of the quality of the university students attend using a RD design. Panel A uses SABER 11 test score as the running variable, restricting the sample to SISBEN-eligible students. Panel B uses SISBEN wealth index as the running variable, restricting the sample to SABER 11-eligible students. No controls are included in these regressions. The LATE estimates scale the ITT estimates by the first stage. In the absence of always-takers, the LATE is equal to the TOT. Bias-corrected RD results estimated with package rdrobust (Cattaneo et al., 2014). Robust standard errors in parentheses. *Sources:* Authors' calculations based on ICFES, DNP, MEN, and SPADIES (2016).

B.3 Medium-Term Enrollment, Persistence, and Academic Performance

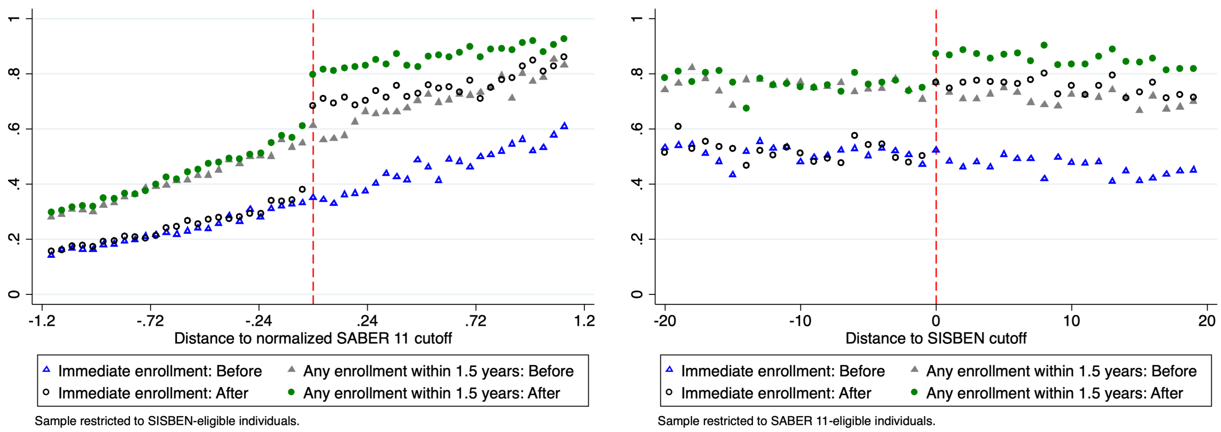
This section analyzes the effect of financial aid on medium-term postsecondary enrollment, persistence, and performance.

Figure C.11 compares immediate and medium-term enrollment probabilities before and after SPP. Specifically, the figure compares (i) immediate enrollment and (ii) any enrollment within 1.5 years from taking SABER 11 as a function of the distance to the normalized SABER 11 (Panel A) and SISBEN (Panel B) SPP eligibility cutoffs for Fall 2013 test-takers ("Before") and Fall 2014 test-takers ("After"). Three lessons emerge from this figure. First, as shown in Section 2.4, financial aid significantly improved immediate enrollment: while this probability is the same before SPP (in blue), it increases significantly above the eligibility cutoff after SPP (in black). Second, any enrollment within 1.5 years increases for eligible students after SPP is rolled out (in green). Third, given control students have somewhat caught up over time (although not disproportionately more than before SPP, in gray), the magnitude of the enrollment gains within 1.5 years diminishes relative to the large immediate enrollment result.

Column (1) in Table C.8 presents the reduced-form estimates on any enrollment within 1.5 years using the regression discontinuity design. The table shows that this probability increases by 19.1 percentage points or 31 percent at the SABER 11 cutoff (Panel A) and by 12.9 percentage points or 17.3 percent at the SISBEN cutoff (Panel B). While these coefficients are smaller than those on immediate enrollment from Table 2.2, the results are consistent with enrollment effects of aid remaining positive and significant even over a longer time horizon.

Figure C.11: Immediate Enrollment and Any Enrollment 1.5 Years After Taking SABER 11

(a) R_i = SABER 11 test score (b) R_i = SISBEN wealth index



Note: This figure compares immediate enrollment and any enrollment within 1.5 years from taking SABER 11 as a function of the distance to the normalized SABER 11 (Panel A) and SISBEN (Panel B) SPP eligibility cutoffs for Fall 2013 test-takers ("Before") and Fall 2014 test-takers ("After). The likelihood of any postsecondary enrollment within 1.5 years increases by 19.1 percentage points (31 percent) using SABER 11 as the running variable (Panel A) and by 12.9 percentage points (17.3 percent) using SISBEN as the running variable (Panel B). See reduced-form estimates in Table C.8, Column (1). *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

The first cohort of SPP recipients (i.e., Fall 2014 test-takers) began their undergraduate studies in Spring 2015. The average length of a program being 4.5 years, degree completion outcomes will only become available once they begin graduating, around 2021–2022. Instead, we observe two shorter-run measures of academic performance and persistence: the cumulative fraction of courses passed by Spring 2016 (included), and whether the student dropped out by that term. It should be noted that persistence during the first year is a strong predictor of completion: one in two college dropouts does so during freshman year in university (MEN, 2016).

We use two empirical strategies to compare college persistence between financial aid recipients and non-recipients. First, we use a RD design where the outcome is a dummy for being enrolled in a HEI in Spring 2016, i.e., 1.5 years after taking the high school exit exam. Columns (2) through (8) in Table C.8 present these results by type of HEI. Panel A uses SABER 11 test score as the running variable, while Panel B uses SISBEN wealth index as the running variable. The reduced-form estimate in Column (2) suggests the likelihood of being enrolled in any HEI in Spring 2016 increased by 15.3 percentage points. On a base of 54.2 percent, this implies an increase of 28.2 percent. That is, the effects of financial aid on enrollment 1.5 years after taking SABER 11 are roughly a third of the immediate enrollment results documented in Table 2.2. This is in large part because barely-ineligible students have somewhat caught up, as discussed above. However, enrollment at high-quality HEIs remains 38.5 percentage points (227.8 percent) higher.

Table C.8: Medium-Term Enrollment in Postsecondary Education, by Type of Institution

	Any	*In Spring 2016*						
		Any	High Quality			Low Quality		
			Any	Private	Public	Any	Private	Public
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: SABER 11 as the Running Variable							
RF	0.191	0.153	0.385	0.406	-0.027	-0.231	-0.094	-0.131
	(0.012)	(0.015)	(0.012)	(0.012)	(0.008)	(0.014)	(0.009)	(0.01)
Mean Control	0.609	0.542	0.169	0.049	0.123	0.372	0.147	0.22
Observations	299,475	299,475	299,475	299,475	299,475	299,475	299,475	299,475
BW Loc. Poly.	25.31	19.778	25.457	20.771	28.761	17.782	21.235	23.777
Effect Obs C	25,256	16,955	25,256	18,379	29,988	14,717	19,382	22,473
Effect Obs T	10,927	9,280	10,927	9,632	11,581	8,724	9,808	10,442
	Panel B: SISBEN as the Running Variable							
RF	0.129	0.117	0.311	0.438	-0.127	-0.191	-0.074	-0.117
	(0.023)	(0.022)	(0.026)	(0.019)	(0.021)	(0.022)	(0.017)	(0.017)
Mean Control	0.745	0.688	0.344	0.08	0.267	0.342	0.147	0.195
Observations	23,132	23,132	23,132	23,132	23,132	23,132	23,132	23,132
BW Loc. Poly.	9.063	12.549	10.77	14.382	10.636	10.568	9.257	12.145
Effect Obs C	3,883	5,133	4,526	5,644	4,490	4,469	4,001	5,025
Effect Obs T	3,918	5,320	4,638	5,984	4,573	4,548	4,011	5,195

Note: This table presents the effect of financial aid on medium-term postsecondary enrollment using a regression discontinuity design. The dependent variable is any postsecondary enrollment by Spring 2016—1.5 years having taken SABER 11—in Column (1). Columns (2)–(8) restrict attending to enrollment in the Spring 2016 term. Their dependent variable is enrollment by type of postsecondary institution (e.g., high-quality, low-quality, private, public) in Spring 2016. Panel A uses SABER 11 test score as the running variable restricts, restricting the sample to SISBEN-eligible students. Panel B uses SISBEN wealth index as the running variable, restricting the sample to SABER 11-eligible students. The RF coefficient in Column (1) of Panel A suggests that, for individuals below a certain level of poverty, financial aid eligibility raises any postsecondary enrollment 1.5 years after graduating high school by 19.1 percentage points. On a base of 60.9 percent, this implies a 31.4 percent increase in any postsecondary enrollment with 1.5 years. No controls are included in these regressions. Bias-corrected RD results estimated with package rdrobust (Cattaneo et al., 2014). Robust standard errors in parentheses. *Sources:* Authors' calculations based on ICFES, DNP, MEN, and SPADIES (2016).

The results from Table C.8 conflate persistence with the positive enrollment results and a shift in the quality of institutions attended. For this reason, it is worth comparing persistence and other academic outcomes across financial aid recipients and their classmates using the following OLS regression, controlling for a set of relevant observables. Our second empirical strategy is as follows:

$$y_{imj} = \alpha + \beta 1(\text{SPP recipient})_i + \delta_{mj} + \boldsymbol{X}^{'}_{imj}\Gamma + \varepsilon_{imj} \tag{B.1}$$

where y_{imj} is outcome y for student i in major m in postsecondary institution j, 1(SPP recipient) is a dummy variable that turns 1 if a student is SPP beneficiary and 0 otherwise, δ_{mj} are major-by-institution fixed effects, $\boldsymbol{X}_{imj}$ is a vector of 25 baseline characteristics, and ε_{imj} is a student-specific error term. Standard errors are clustered at the institution-by-major level.

Columns (1)–(4) in Table C.9 show that, on average 22.6 percent of students who enrolled in Spring 2015 are absent by Spring 2016.[1] SPP recipients are 12.1 percentage points less likely to be absent than their classmates (Column 1). Because retention rates are higher in high-quality HEIs attended by SPP recipients, Column 2 includes major-by-institution fixed effects. This reduces the magnitude of the coefficient, but not its statistical significance. Since dropout rates tend to decrease with pre-collegiate ability, controlling for SABER 11 score further reduces the magnitude of the coefficient (Column 3). Finally, including baseline covariates—our preferred specification—suggests SPP recipients are 3.8 percentage points less likely to drop out than their classmates (Column 4). On a mean of 22.1 percent, SPP recipients are 17.2 percent (3.8/22.1) less likely to drop out. This is partly a result of the program design, which requires beneficiaries to graduate from their program for the loan to become forgivable.

Columns (5)–(8) in Table C.9 use as dependent variable the cumulative share of courses passed by Spring 2016 (inclusive). On average, students passed 85.2 percent of their courses in their first three semesters in college. This share is 0.6 percentage points higher for SPP recipients than for their classmates (Column 5). Including major-by-college fixed effects reduces the magnitude of this coefficient (Column 6). Further, controlling for SABER 11 makes this coefficient switch sign and become significant (Column 7). Including a rich set of baseline controls suggests SPP recipients are 1.9 percentage points more likely to fail a course during their freshman year (Column 8). On a base of 85.3 percent, this implies a 2.2 percent drop in performance.

[1] The first-year retention rate in Colombia varies significantly by type of HEI, and is lower than the average retention rate in the United States. In 2015, data from SPADIES shows the percentage of first-time, degree-seeking undergraduates retained was 77.1 percent at 4-year institutions in Colombia, 82.2 percent at high-quality institutions, 75.4 percent at low-quality institutions, 84.5 percent at private, high-quality institutions, and 78.6 percent at public, high-quality institutions. In the U.S., data from the U.S. Department of Education suggest the percentage of first-time, full-time degree-seeking undergraduates retained at 4-year institutions was 80 percent from 2013 to 2014. For selective institutions (i.e., acceptance rate of 25 percent or below), this share is 96 percent at both public and private non-profit institutions.

The second row in Table C.9 treats SPP recipiency status as endogenous and instruments it using program eligibility (i.e., scoring above 310/500 in SABER 11 and having a SISBEN score below the cutoff). The resulting 2SLS estimates suggest that selection on unobservables does not drive our main results, namely that SPP recipients are less likely to drop out but more likely to fail courses. However, by construction, we observe course-passing information only for retained students. The differential non-random attrition from postsecondary education documented in Columns (1)–(4) may thus compromise the course-passing comparability of SPP recipients and non-recipients, possibly generating selection bias.

To deal with this issue, we follow Abdulkadiroglu, Pathak, and Walters (2018) and formally assess the robustness of our results to selected attrition by constructing non-parametric bounds on the 2SLS estimate under a monotonicity assumption on the attrition process, in the spirit of Lee (2009).[2] This monotonicity assumption in response to SPP eligibility is as follows:

$$D_i(1) \neq D_i(0) \implies D_i(1) = 1$$

This restriction implies that any student who changes behavior in response to SPP eligibility does so to receive financial aid. That is, no student declines SPP in response to eligibility, and no one drops out of college (and hence exits our estimation sample) in response to SPP eligibility. Intuitively, if SPP eligibility reduces the likelihood of dropping out for students, the usual LATE framework must be augmented with a set of "at risk" compliers who drop out of college (exit the sample) when ineligible for financial aid: $D_i(1) = 1, D_i(0) = a$, where a represents attrition from the sample. This prevents identification of the mean treated outcome for the subgroup of compliers who remain in the sample, but these means can be bounded using observed response probabilities and quantiles of the outcome distribution.[3] We then obtain standard errors by conducting 100 bootstrap replications of the entire procedure. In practice, risk set indicators and baseline covariates are included in all regressions used to estimate group shares, CDFs, and mean potential outcomes.

The third row in Table C.9 displays estimated bounds on 2SLS estimates for compliers. These bounds are tight around 2SLS estimates and suggest that adjustments for differential attrition do not overturn the conclusion that SPP recipients are 2 percent more likely to fail courses in their first year of college. The seemingly contradicting evidence on performance outcomes is likely also a result of program design. As mentioned above, SPP requires beneficiaries to graduate from their undergraduate program for the loan to become forgivable. SPP does not, however, provide achievement incentives; beneficiaries are not required to maintain a minimum grade point average in college. This partly explains why SPP recipients are slightly more likely to fail courses than counterfactual students. Ultimately, while this small negative effect on performance may potentially affect time to degree, we interpret the

[2] We thank Christopher Walters for sharing the Stata code to perform this estimation.

[3] A full description of the methods for bounding local average treatment effects in the presence of differential attrition is available in Abdulkadiroglu et al. (2018), Appendix A.2.

generally positive results displayed in Table C.9 as evidence against mismatch documented in other settings (Goodman et al., 2017; Cohodes and Goodman, 2014; Hoxby and Avery, 2013; Dillon and Smith, 2017).

Table C.9: Persistence and Performance by Spring 2016: Cumulative Dropout and Course-Rassing Rates

	Dependent variable							
	Absent by Spring 2016				Cumulative share of courses passed by Spring 2016			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OLS	-0.121	-0.049	-0.019	-0.038	0.006	0.002	-0.017	-0.019
	(0.018)	(0.005)	(0.005)	(0.006)	(0.011)	(0.003)	(0.002)	(0.003)
2SLS	-0.139	-0.060	-0.007	-0.035	0.014	0.013	-0.020	-0.023
	(0.019)	(0.007)	(0.006)	(0.007)	(0.011)	(0.003)	(0.003)	(0.003)
Bounds (LB; UB)					-0.052 0.015	0.003 0.013	-0.020 -0.020	-0.023 -0.023
					(0.004) (0.003)	(0.004) (0.002)	(0.003) (0.002)	(0.003) (0.003)
Major-by-college FE		Yes	Yes	Yes		Yes	Yes	Yes
SABER 11 score			Yes	Yes			Yes	Yes
Other controls				Yes				Yes
N	109,735	109,595	109,595	101,157	84,654	84,451	84,451	78,368
R2	0.006	0.208	0.224	0.239	0.000	0.492	0.524	0.535
Dep Mean	0.226	0.226	0.226	0.221	0.852	0.852	0.852	0.853
Dep SD	0.418	0.418	0.418	0.415	0.199	0.199	0.199	0.198

Note: Sample restricted to Fall 2014 SABER 11 test-takers who immediately enrolled in any postsecondary institution in Spring 2015. The dependent variable in Columns (1)–(4) is a dummy for whether a student registered in Spring 2015 is absent by Spring 2016, and in Columns (5)–(8) it is the cumulative share of courses passed by Spring 2016. Each row represents a separate regression. Robust standard errors in parentheses clustered at the first major-by-college pair in rows 1 and 2. Because the share of courses passed are observed for retained students, we assess the robustness of our results in row 2 to selective attrition by constructing non-parametric bounds on local average treatment effects under a monotonicity assumption on the attrition process, in the spirit of Abdulkadiroglu et al. (2018) and Lee (2009). Row 3 reports these non-parametric bounds and standard errors using 100 bootstrap replications of the entire procedure. Controls include 25 baseline characteristics. $*p < 0.1, **p < 0.05, ***p < 0.01$.
Sources: Authors' calculations based on ICFES, DNP, MEN, and SPADIES (2016).

B.4 The Second Cohort of SPP Recipients

Having documented the enrollment effects of financial aid on the first cohort of SPP beneficiaries (henceforth SPP1), this section estimates these effects for the second cohort of students (henceforth SPP2). Unlike SPP1, students from SPP2 were high school juniors when the government announced SPP. They therefore had one year to prepare their college application. Expecting an increasing take-up of SPP over time, the government raised the SABER 11 eligibility cutoff from 310 to 318 for SPP2, while keeping the SISBEN cutoff unchanged. Therefore, relative to SPP1, SPP2 beneficiaries will be more selected on ability (top 8 percent versus top 9 percent of test scores), but a priori equally wealthy. As before, we exploit the SABER 11 and SISBEN eligibility cutoffs using a RD design separately for the two population of compliers.

First, we test whether the key identification assumptions required in a RD strategy still hold for this second cohort. We find that 77 students received SPP without having the required SISBEN score in SPP2. That is, unlike in SPP1, there are some always-takers in SPP2. As discussed in Section 2.4, students from cohorts 2 and above reacted to the expansion of financial aid by requesting an evaluation from local authorities to be included in SISBEN. They also reacted by requesting a SISBEN re-evaluation, in the hopes that their new score would fall below the eligibility cutoff. This generates a concern that, unlike SPP1, there might be some manipulation of SISBEN scores around the cutoffs in this second cohort (and those that followed).

Interestingly though, we do not statistically detect any manipulation of either running variable. The robust bias-corrected p-value for the RD density test from Cattaneo et al. (2016a) is 0.218 for R_i = SABER 11 and 0.436 for R_i = SISBEN. Moreover, Figure D.12 further suggests there is no observable bunching around either eligibility cutoff. Lastly, we do not detect any imbalance in baseline covariates across the SABER 11 cutoff for SISBEN-eligible students (see Column (1) in Table D.10). However, we do indeed detect some small differences among SABER 11-eligible students around the SISBEN cutoff (Column (3)). Specifically, merit-eligible students who are barely need-eligible are more likely to be female, low-income, and have more educated fathers than barely ineligible students. For narrow bandwidths, we can reject the joint null hypothesis of balance across all baseline covariates for this population of compliers.

Table D.11 shows that take-up rate of SPP significantly increased by 12.3 percentage points for this second cohort, from 59.4 percent in SPP1 to 71.7 percent in SPP2. The higher take-up rate of SPP should not be surprising: as SPP becomes more well-known and students have more time to prepare their college application, a higher fraction of eligible students are able to receive the scholarship-loan.

Third, when compared with SPP1 outcomes, SPP2 enrollment impacts vary depending on the population of compliers one observes. Figure D.13 plots immediate enrollment probabilities Fall 2014 test-takers (SPP1) and Fall 2015 test-takers (SPP2) as a function of the two running variables. Raising the SABER 11 cutoff improved enrollment for those above the new threshold; but not those below it. The reduced-form coefficient for SPP2 (Table D.11) is 0.308 and thus very similar to the 0.32 coefficient for SPP1 (Table 2.2). For individuals below the poverty cutoff, aid eligibility raises immediate enrollment by 30.8 percentage points in SPP2 versus 32.0 percentage points in SPP1. Critically, however, raising the SABER 11 cutoff also raised the control mean for SPP2 (45.6 percent) relative to SPP1 (37 percent). This implies that the enrollment effect in percentage terms falls from 86.5 percent in SPP1 to 67.5 percent in SPP2.

In contrast, we much smaller differences in enrollment outcomes between SPP1 and SPP2 using SISBEN as the running variable. For individuals scoring above the test score threshold, aid eligibility raises immediate enrollment by 24.5 percentage points in SPP2 versus 27.4 percentage points in SPP1. Given the control mean increased from 48.5 percent to 56.8 percent, the enrollment effect falls from 56.5 percent in SPP1 to 43.1 percent in SPP2.

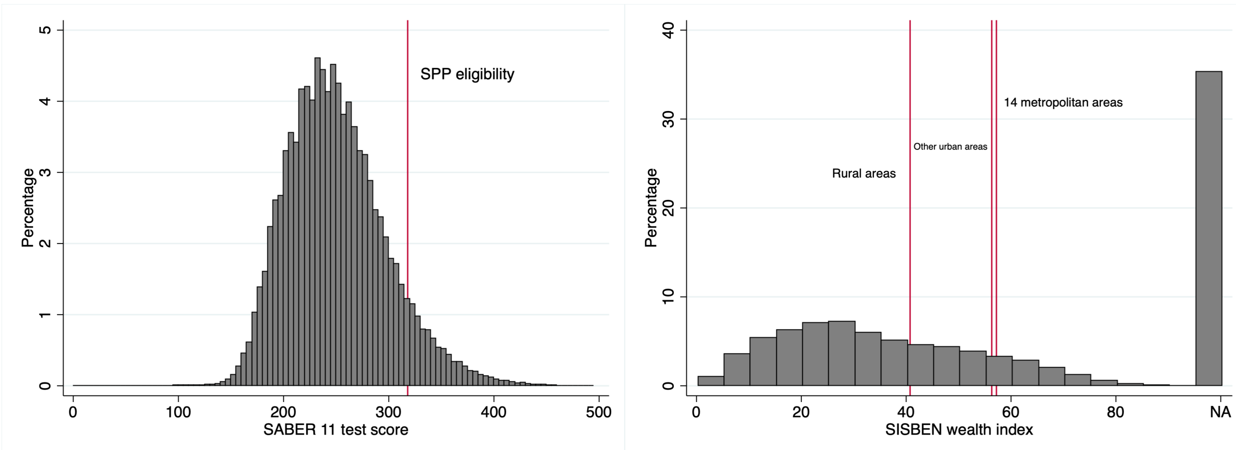
Finally, students eligible for SPP2 shifted away from public, high-quality HEIs—a result that is statistically significant and large in percentage terms (29.5 to 34.7 percent) for both populations of compliers. The shift away from low-quality HEIs is somewhat less accentuated in SPP2 relative to SPP1: negative 42–51 percent versus negative 53–58 percent, respectively. This is both a result of a lower reduced-form coefficient using SISBEN as the running variable as well as a higher control means in both comparisons in SPP2 relative to SPP1.

A mixture of demand and supply responses to financial aid expansion are potentially at play in producing these positive results for control students, as discussed in Section 2.5. It is worth noting that, in May 2015, the Colombian government implemented a separate student loan program called "Tú Eliges." The objective of this loan program was to expand financial aid to low- and middle-income (albeit lower-performing) students.[4] As a result, loans were often awarded to lower-performing, middle-income students, which might drive part of the positive low-quality enrollment effects for documented in Table D.11. Although these loans could explain part of the increase in immediate enrollment for the control population in SPP2, they cannot explain most of the large enrollment gains among students barely ineligible for SPP: Section 2.5 documented that low-income, lower-performing Fall 2014 already had large enrollment gains relative to Spring 2014 test-takers (the control group in the difference-in-differences approach) before "Tú Eliges" was introduced.

[4]Unlike SPP2, loans from "Tú Eliges" (1) were not forgivable upon graduation, requiring beneficiaries to pay back the full amount of the loan; (2) did not require beneficiaries to graduate from high school by a specific date; (3) could be used to attend both low- and high- quality HEIs; and (4) were less restrictive in their eligibility criteria than SPP2.

Figure D.12: SPP Eligibility Conditions for the Second Cohort of SPP

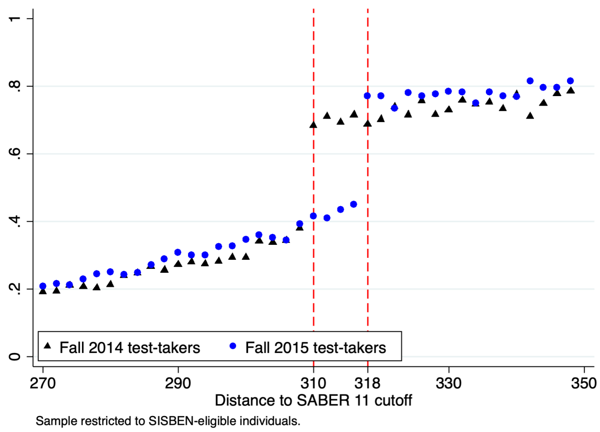
(a) Merit: SABER 11 test score $\geq 318/500$ (b) Need: SISBEN wealth index $<$ threshold



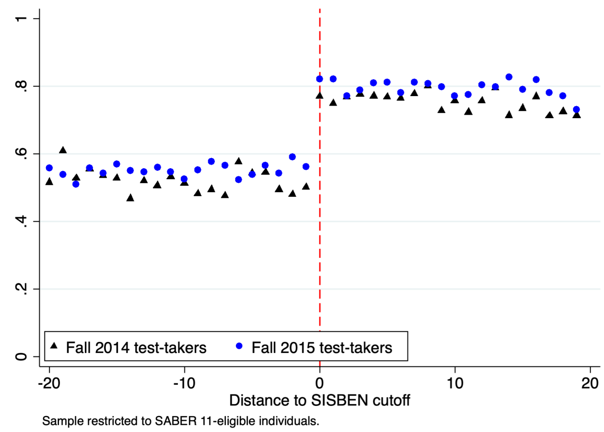
Notes: These figures show the distribution of SABER 11 test scores (Panel A) and SISBEN poverty index (Panel B) for Fall 2015 test-takers. The red vertical lines represent the SPP eligibility cutoffs. The figures suggest both variables are distributed smoothly around the eligibility cutoffs. In Panel B, the SISBEN eligibility cutoff varies by the applicant's geographic location. Test-takers not included in SISBEN (e.g., individuals that do not receive welfare) do not have a SISBEN score and appear in Panel B as "N/A". *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure D.13: Immediate Postsecondary Enrollment: SPP1 versus SPP2

(a) R_i = SABER 11 test score



(b) R_i = SISBEN wealth index



Note: These figures plot the probability of immediately accessing any postsecondary institution for test-takers in Fall 2014 (SPP1, in black) and Fall 2015 (SPP2, in blue) as a function of the SABER 11 score (Panel A) and SISBEN (Panel B) eligibility cutoffs. The SABER 11 cutoff increased from 310 to 318 for SPP2 (i.e., from top 9 percent to top 8 percent of test scores), while the SISBEN cutoff remained the same. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Table D.10: Baseline Covariate Balance Test for the Second Cohort of SPP

	Running Variable			
	SABER 11		SISBEN	
	Coef/SE	Mean Control	Coef/SE	Mean Control
	(1)	(2)	(3)	(4)
Female	0.004 (0.011)	0.452	0.049 (0.024)	0.387
Age	-0.005 (0.047)	16.631	0.05 (0.092)	16.533
Ethnic minority	0.001 (0.002)	0.009	-0.003 (0.003)	0.006
Employed	0.007 (0.007)	0.094	-0.014 (0.014)	0.089
Family size	0.023 (0.037)	4.476	0.034 (0.064)	4.311
Mother education: Primary	-0.004 (0.009)	0.226	-0.002 (0.017)	0.136
Mother education: Secondary	-0.002 (0.011)	0.484	0.004 (0.023)	0.431
Mother education: T&T	0.002 (0.007)	0.148	-0.007 (0.017)	0.192
Mother education: Higher	0.007 (0.007)	0.138	0.005 (0.02)	0.243
Father education: Primary	0.007 (0.009)	0.303	0.016 (0.018)	0.176
Father education: Secondary	-0.015 (0.011)	0.444	-0.075 (0.026)	0.469
Father education: T&T	0.001 (0.007)	0.113	0.011 (0.018)	0.147
Father education: Higher	0.012 (0.008)	0.134	0.045 (0.022)	0.210
Household SES: Stratum 1	-0.012 (0.011)	0.361	0.006 (0.015)	0.119
Household SES: Stratum 2	0.006 (0.011)	0.448	0.04 (0.023)	0.497
Household SES: Stratum 3	0.007 (0.008)	0.171	-0.046 (0.021)	0.336
Household SES: Stratum 4	0 (0.003)	0.017	0.005 (0.009)	0.037
Household SES: Stratum 5	-0.001 (0.001)	0.002	-0.003 (0.004)	0.010
Household SES: Stratum 6	-0.001 (0.001)	0.001	0 (0.001)	0.000
School hours: Full day	-0.01 (0.01)	0.207	-0.001 (0.024)	0.348
School hours: Morning	0.002 (0.011)	0.620	0.022 (0.024)	0.491
School hours: Evening	0 (0.002)	0.007	0.002 (0.004)	0.006
School hours: Afternoon	0.009 (0.007)	0.159	-0.022 (0.016)	0.148
School hours: Weekends	-0.001 (0.002)	0.007	0.003 (0.003)	0.003
Private school	-0.003 (0.009)	0.190	-0.019 (0.021)	0.352
Joint F-Stat (p-value, LB on b-width)	0.247		0.000	
Joint F-Stat (p-value, UB on b-width)	0.699		0.184	

Notes: This table plots the reduced-form coefficient from a RD specification where the outcome is a baseline characteristic and the running variable is either SABER 11 test scores in columns (1) and (2) or SISBEN poverty index in columns (3) and (4). The sample is restricted to Fall 2015 test-takers who are SISBEN-eligible in columns (1) and (2), or SABER 11-eligible in columns (3) and (4). Columns (1) and (3) present conventional coefficients and standard errors in parentheses, while columns (2) and (4) present control means. There is no statistically significant imbalance in Column (1), and we cannot reject the joint null hypothesis. There is imbalance in 5 of 25 covariates in Column (3) and, while imbalance in these observable characteristics using SISBEN as the running variable are relatively small, we can reject the joint null hypothesis for small bandwidths. All results are estimated with package rdrobust (Cattaneo et al., 2014). *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Table D.11: Immediate Postsecondary Enrollment for the Second Cohort of SPP, by Type of Institution

	Any	High Quality			Low Quality		
		Any	Private	Public	Any	Private	Public
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: SABER 11 as the Running Variable						
RF	0.308	0.444	0.473	-0.028	-0.129	-0.02	-0.108
	(0.013)	(0.013)	(0.012)	(0.008)	(0.01)	(0.007)	(0.008)
Mean Control	0.456	0.15	0.055	0.095	0.304	0.129	0.172
Observations	312,863	312,863	312,863	312,863	312,863	312,863	312,863
BW Loc. Poly.	28.578	27.713	26.655	26.442	38.145	44.042	38.615
Effect Obs Control	25,378	23,874	22,943	22,943	40,952	51,629	40,952
Effect Obs Treat	10,841	10,678	10,408	10,408	12,769	13,584	12,769
	Panel B: SISBEN as the Running Variable						
RF	0.245	0.397	0.451	-0.06	-0.151	-0.06	-0.086
	(0.024)	(0.032)	(0.024)	(0.023)	(0.028)	(0.02)	(0.02)
Mean Control	0.568	0.278	0.106	0.173	0.297	0.144	0.146
Observations	23,135	23,135	23,135	23,135	23,135	23,135	23,135
BW Loc. Poly.	12.917	7.321	11.667	8.224	7.011	9.594	7.877
Effect Obs Control	4,459	2,650	4,111	2,977	2,546	3,443	2,831
Effect Obs Treat	4,804	2,720	4,342	3,104	2,588	3,618	2,960

Note: This table presents the effect of financial aid on immediate postsecondary enrollment for the second cohort of SPP using a regression discontinuity design. The dependent variable is immediate enrollment by type of postsecondary institution (e.g., high-quality, low-quality, private, public). Panel A uses SABER 11 test score as the running variable restricts, restricting the sample to SISBEN-eligible students. Panel B uses SISBEN wealth index as the running variable, restricting the sample to SABER 11-eligible students. No controls are included in these regressions. The reduced-form coefficient in Column (1) of Panel A suggests that, for individuals below a certain level of poverty, financial aid eligibility raises immediate postsecondary enrollment by 30.8 percentage points, which is very similar to 32.0 result from Table 2.2 from the first cohort of SPP. Critically, however, the basis for this second cohort of SPP has increased substantially, from 37 percent to 45.6 percent. This implies that the enrollment effect in percentage terms falls from 86.5 percent to 67.5 percent. Bias-corrected RD results estimated with package rdrobust (Cattaneo et al., 2014). Robust standard errors in parentheses. *Sources:* Authors' calculations based on ICFES, DNP, MEN, and SPADIES (2016).

B.5 Where Did Displaced Applicants Enroll?

The greater entry competition at elite private universities induced by financial aid may have displaced some applicants from attending top schools. This section uses admission records microdata from one of the country's top-ranked universities to identify and characterize displaced applicants, and track them to the colleges where they end up enrolling around the country.

In this top-ranked university, 900 SPP-*in*eligible applicants whose test scores would have granted them admission in Spring 2014 were rejected the following year as a result of the more selective cutoffs.[5] Figure E.15 tracks where these displaced students enrolled in Spring 2015, i.e., the semester immediately after presenting SABER 11. Panel A shows that around 42 percent of displaced students did not immediately enroll in any HEI in Colombia, with the remainder enrolling at institutions with significantly lower-quality students, according to their mean SABER 11 score before SPP.[6] Specifically, 22 percent of displaced students attended Bogota's second-best private HEI, whose mean normalized SABER 11 score is 1.94 (i.e., a loss in peer quality of 1.1 points). Roughly 9.3 percent of displaced students attended an institution with a score of 1.54 (i.e., a loss in peer quality of 1.5 points). One-third of all displaced students belong to relatively wealthy households, i.e., strata 4, 5, or 6. Panel B restricts the sample to these relatively wealthy students and shows that one-third of them opted not to immediately enroll in any HEI in Colombia, while 26.7 percent opted for the 8th ranked institution in the country; a private, high-quality college in Bogota.

How does this compare to successful applicants who would not have been admitted the following year due to the increase in cutoffs, that is, the would-be marginal applicants? Indeed, comparing the two cohorts is important because not all those who are admitted enroll. Specifically, Londoño-Vélez (2016) suggests that on average roughly 30 percent of admitted applicants enrolled at this elite private university before SPP, and that those from strata 4–6 were twice as likely to enroll as those from strata 1–3. Thus, Figure E.16 presents the difference in the frequency of marginal and would-be marginal applicants enrolled at each institution before and after SPP.[7] The figure suggests an increase in the number of students that did not immediately enroll in any HEI and that enrolled in lower-ranked institutions.[8]

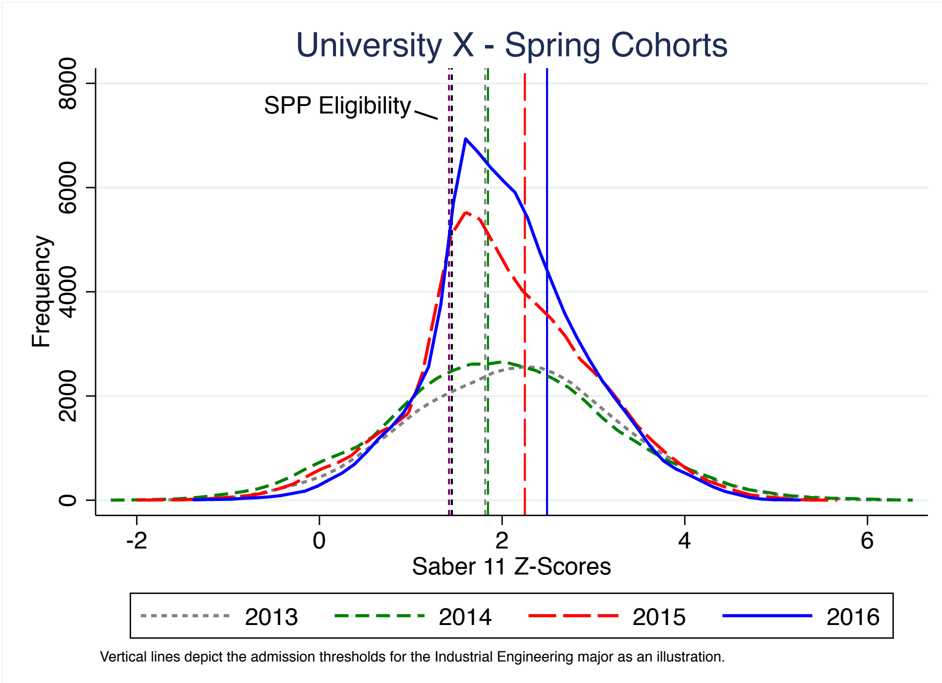
[5] Figure E.14 illustrates the rightward shift in admission cutoffs at this institution produced by SPP.

[6] Roughly one-third of those (immediately) crowded out from higher education re-applied to this same university to no avail. By Fall 2016 (i.e., 1.5 years later), 70 percent were attending higher education.

[7] There are 317 such would-be marginal applicants in Spring 2014.

[8] Note that almost 42 percent of would-be marginal applicants come from relatively wealthy households (i.e., strata 4-6).

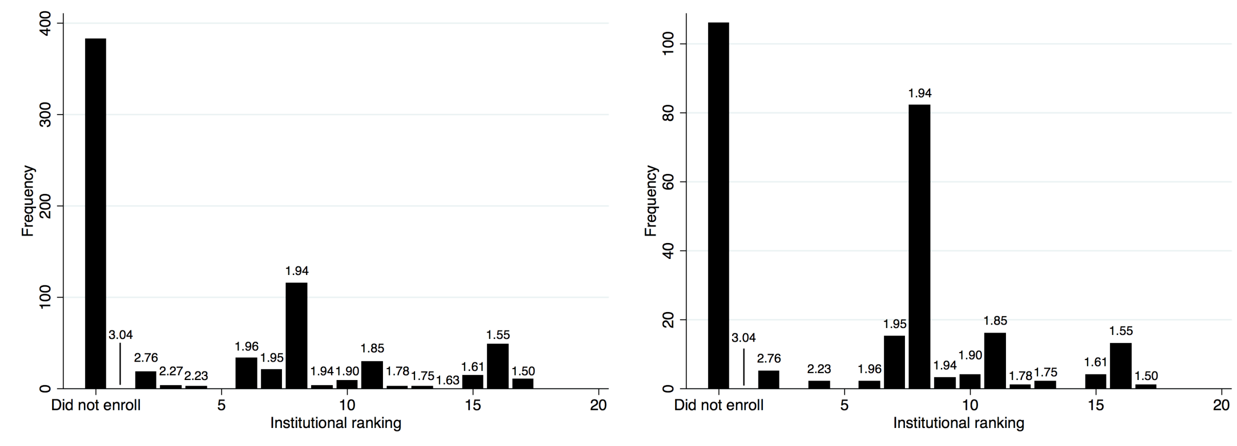
Figure E.14: SPP displaced students from an elite private university



Notes: This figure plots the distribution of SABER 11 scores among applicants wishing to enroll in an elite private university in Bogota in the Spring semesters of 2013, 2014, 2015, and 2016. The vertical lines depict the admission cutoffs for a representative undergraduate program (Industrial Engineering). The figure shows that the increase in the number of applicants following the announcement of SPP induced a rightward shift in the admission threshold, thus raising admission selectivity. *Source:* Londoño-Vélez (2016).

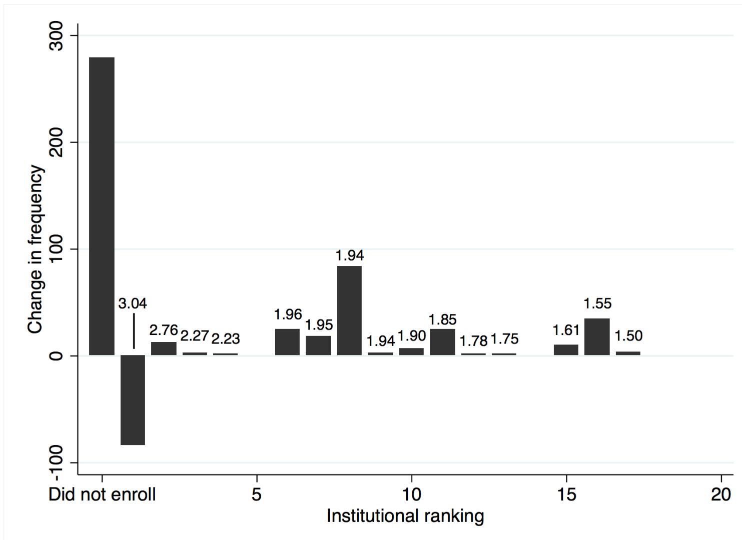
Figure E.15: Institutions where displaced applicants immediately enrolled

(a) All displaced applicants (b) Relatively wealthy displaced applicants



Notes: These figures tracks the postsecondary institution in which displaced applicants from a flagship private university immediately enrolled. A Spring 2015 applicant is considered “displaced” if he/she is *not* eligible for SPP, has standardized SABER 11 test score above the previous year’s admission cutoff, and was not admitted. The values above the bars represent institutional mean normalized SABER 11 test score of Spring 2014 entering cohort. Panel A suggests around 42% of displaced students did not immediately enroll in any postsecondary institution in Colombia, with the remainder enrolling at institutions with significantly lower-quality students. Panel B restricts the sample to displaced applicants from strata 4, 5, and 6. *Sources:* Authors’ calculations based on university administrative data, ICFES, DNP, MEN, and SPADIES (2016).

Figure E.16: Difference in frequency of displaced and would-be displaced applicants before and after SPP



Note: This figure presents the difference in the frequency of displaced and would-be displaced applicants enrolled at each institution before and after SPP. A Spring 2014 applicant is considered "displaced" if he/she is *not* eligible for SPP, was admitted the year he/she applied, and has standardized SABER 11 test score below the following year's admission cutoff. A Spring 2015 applicant is considered "displaced" if he/she is *not* eligible for SPP, has standardized SABER 11 test score above the previous year's admission cutoff, and was not admitted. Values represent institutional mean normalized SABER 11 test score of Spring 2014 entering cohort. The figure suggests there was an increase in the number of students that did not immediately enroll in postsecondary education and that enrolled in a lower-ranked institution. *Sources:* Authors' calculations based on university administrative data, ICFES, DNP, MEN, and SPADIES (2016).

B.6 Shifts in College Demand using College Admission Records

This section studies demand responses to financial aid using college admission records. First, we use data on undergraduate applications received by HEIs between 2012 and 2016 from SNIES, the Ministry of Education's information database on postsecondary education. We employ a difference-in-difference approach comparing the relative difference in the number of undergraduate applications received by HEI j in year t for Spring (treatment) and Fall (control) enrollment across time using specification (2.2). The coefficients of interest represent the difference in the number of undergraduate applicants with a given HEI in the Spring (treatment) and Fall (control) terms across time.

Figure F.17 plots the β_k coefficients from specification (2.2). The figure shows the number of applicants increased at high-quality private HEIs but remained unchanged at all other

HEIs. High-quality private universities reacted to the higher demand by expanding their cohort size. Indeed Figure F.18, which re-estimates specification (2.2) using the number of incoming students as the outcome variable, shows that entering cohort size significantly increased at these institutions, while remaining relatively unaffected at other institutions.[9] Critically, these impacts cannot be explained by changes in HEIs' decision to re-allocate spaces from Spring to Fall semester following scholarship rollout. Figure F.19, Panel A shows that, while the number of Fall test-takers (treated) accessing private high-quality HEIs immediately after high school increased significantly after SPP, the number of Spring test-takers (control) remained stable.

Unfortunately, inconsistent application reporting by some HEIs affects SNIES's data quality. For instance, the SNIES data misses undergraduate admission records for Colombia's flagship public and high-quality HEI, UNAL, in several years. For this reason, we complement SNIES data with admissions records provided specially to us by selected private high-quality HEIs. Figure F.20, Panel A, plots the number undergraduate applicants at several private high-quality colleges from Spring 2011 to Spring 2016 (normalized to equal 1 in Spring 2014) in the four largest metropolitan areas: Los Andes, PUJ, Sabana, and Jorge Tadeo Lozano, in Bogota; EAFIT and EIA, in Medellin; ICESI, in Cali; and Norte, in Barranquilla. The number of applicants increased between 19 percent (EAFIT) and 94 percent (ICESI) between 2014 and 2015. One year later, this fraction had further increased to 80 percent (Los Andes, PUJ – Bogota), 97 percent (Jorge Tadeo Lozano), and even 150 percent (Sabana, ICESI).[10] Importantly, this rise in college applications at private high-quality HEIs stands in stark contrast with their public counterparts. For instance, at UNAL, Colombia's flagship public university, the number of applicants stayed constant between 2014 and 2016.

Consequently, admission rates plummeted at most high-quality HEIs, with the fall being particularly severe at top-ranked universities: at Los Andes and Javeriana, admission rates fell by almost 50 percent and 35 percent in just two years, respectively (Figure F.20, Panel C).[11] This phenomenon is not restricted to private colleges in Bogota: the admission rate

[9] It should be noted that private colleges are relatively indifferent between admitting a SPP recipient and a non-recipient regarding tuition payments: they charge the same tuition and thus receive the same payment from parents or the government for either student type. In contrast, public colleges do receive more payments for enrolling SPP recipients starting 2016: in addition to the usual annual subsidies from the central government, SPP transfers to these institutions the students' average *cost*—not the tuition fee, which is artificially lower thanks to heavy government subsidies. This implies that, all else equal, public, high-quality colleges do have an incentive to enroll a SPP beneficiary over another student. While we do not observe such behavior, our finding that SPP beneficiaries sort *out* of—not *into*—public colleges suggests that, if public colleges are in fact actively changing their admission practices to favor SPP beneficiaries, their efforts are being met with limited success.

[10] Using admissions microdata from a private, high-quality university in Bogota, Londoño-Vélez (2016) shows that the documented rise in applications after SPP is driven exclusively by low-SES applicants from the bottom socioeconomic strata.

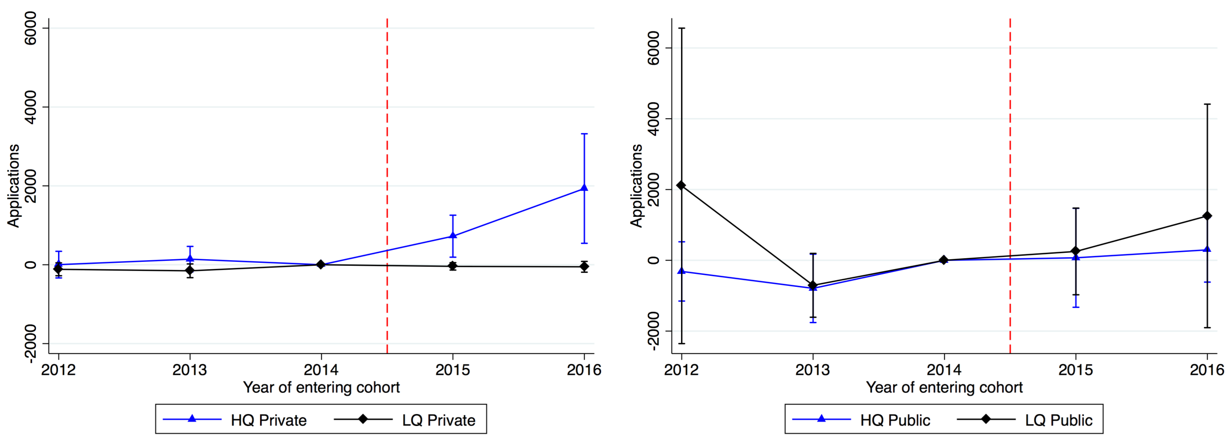
[11] Note that average admission rates are significantly lower in public versus private universities. Moreover,

fell by 20 percent in Barranquilla's Norte University, 15 percent in Medellin's EIA, and 15 percent in Cali's ICESI. In contrast, the undergraduate admission rate at UNAL, Colombia's top public HEIs, *increased* from 8.3 percent in 2014 to 9.5 percent in 2016.

Importantly, again, private colleges expanded supply as a response to this rise in demand. Panel B in Figure F.20 shows that the increase between 2014 and 2016 in the size of entering cohorts varies significantly across institutions, from 8 percent in EAFIT to 100 percent in ICESI. Note, however, that the slope of the curves flattens after 2015, suggesting a possible conscientious attempt on behalf of some elite colleges to restrict this increase in supply (see MacLeod and Urquiola, 2015).[12]

Figure F.17: Applications by Quality and Type of HEI (SNIES data)

(a) Private HEIs (b) Public HEIs



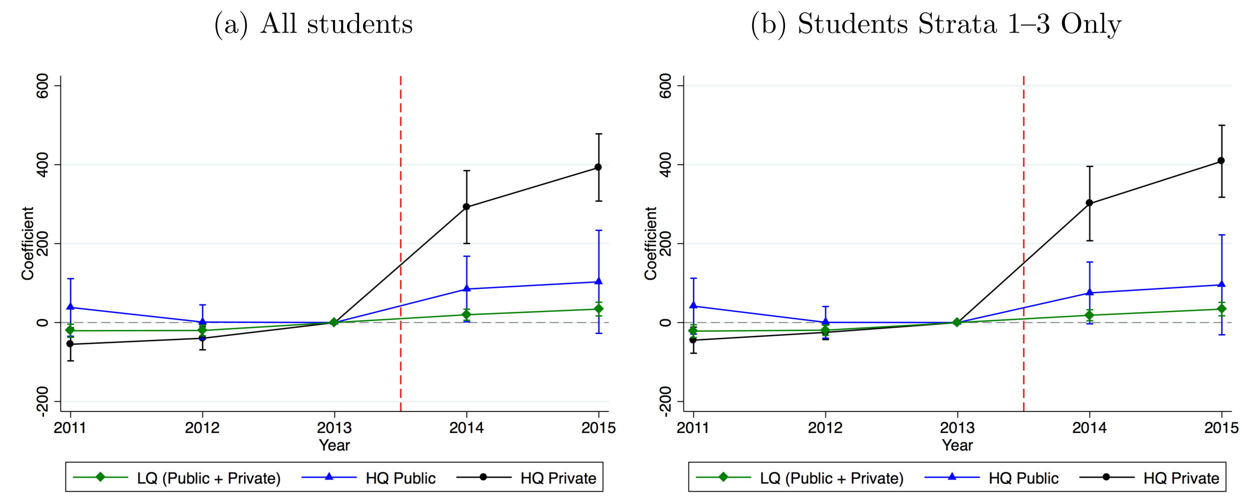
Notes: Panel A plots the difference in the number of undergraduate applicants at HEIs in the Spring (treatment) and Fall (control) terms, using specification (2.2). The figure suggests that the number of applicants increased only at high-quality private HEIs. The sample is restricted to a balanced sample of HEIs that reported annually between 2012 and 2016. Standard errors are clustered at the HEI level.
Source: Authors' calculations using SNIES.

---

universities are generally more selective in Bogota than elsewhere in the country: the admission rate in Spring 2014 was 57 percent in the University of Los Andes and Sabana, and 58 percent in PUJ, versus 92 percent in EAFIT and 96 percent in EIA.

[12] MacLeod and Urquiola (2015) offer a model to explain why elite colleges have restricted supply despite increasing college applications. In their setup, asymmetric information about individual innate ability leads firms to set wages to expected skill conditional upon college reputation and an individual-specific measure of skill. As a response, college applicants display an endogenous taste for abler peers and colleges with good reputation. This provides schools with incentives to be selective and remain small.

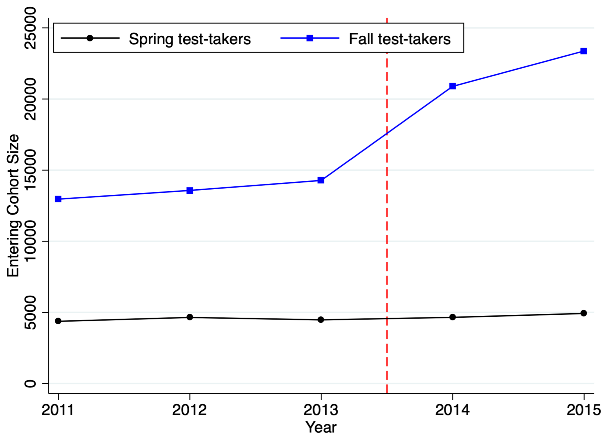
Figure F.18: Entering Cohort Size by Quality and Type of HEI (SPADIES data)



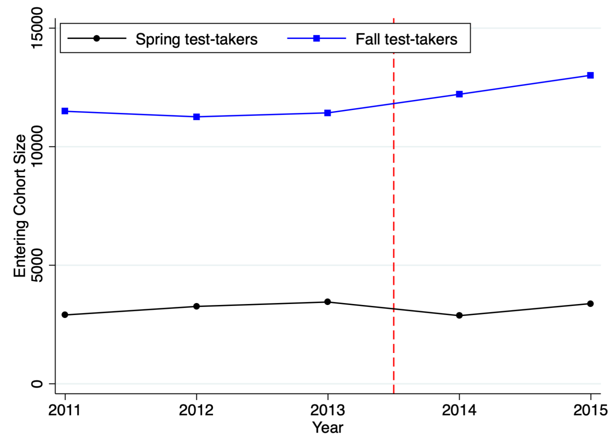
Notes: These figures plot the difference in the number of students who immediately enroll at a HEI after presenting SABER 11 in Spring (control) or Fall (treatment), using specification (2.2). Panel A uses information for students, while Panel B restricts the sample to students from strata 1–3. The figures suggest that entering cohort size increased only at high-quality private HEIs. This rise was driven by low-income students. The sample is restricted to a balanced sample of HEIs that reported annually between 2011 and 2015. Standard errors are clustered at the HEI level. *Source:* Authors' calculations using ICFES and SPADIES.

Figure F.19: Cohort Size by HEI Type Before and After SPP

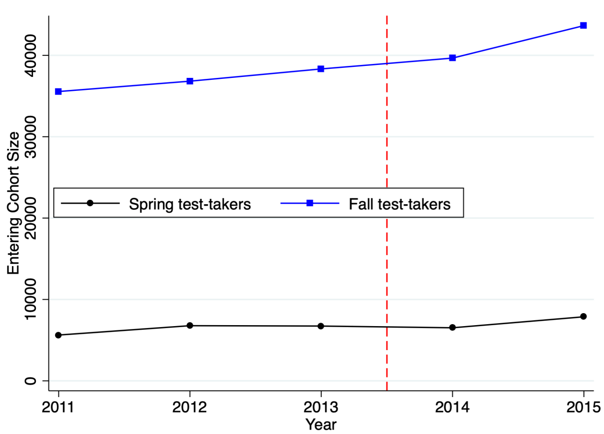
(a) Private High Quality HEIs



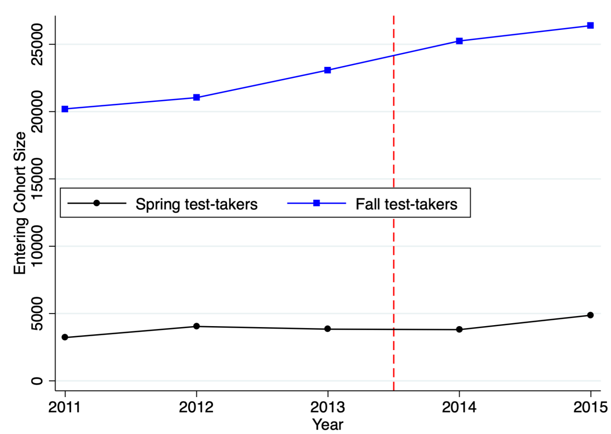
(b) Public High Quality HEIs



(c) Private Low Quality HEIs



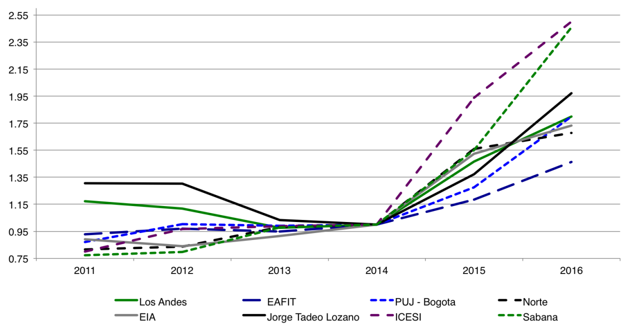
(d) Public Low Quality HEIs



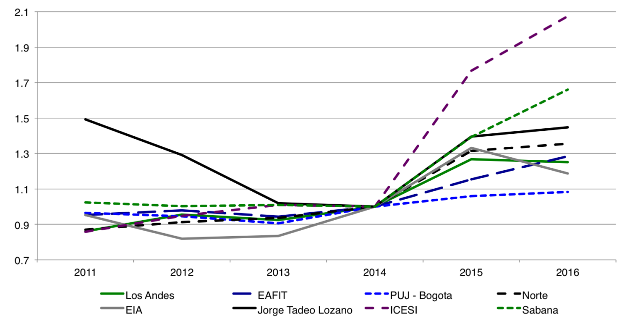
Note: This figure plots the number of students who access HEIs immediately after graduating high school between 2011 and 2015, separately for Fall (treated, in blue) and Spring (control, in black) test-takers, by type of HEI. *Sources:* Authors' calculations based on ICFES, DNP, and MEN (2016).

Figure F.20: Admissions at Selected Private High-Quality HEIs

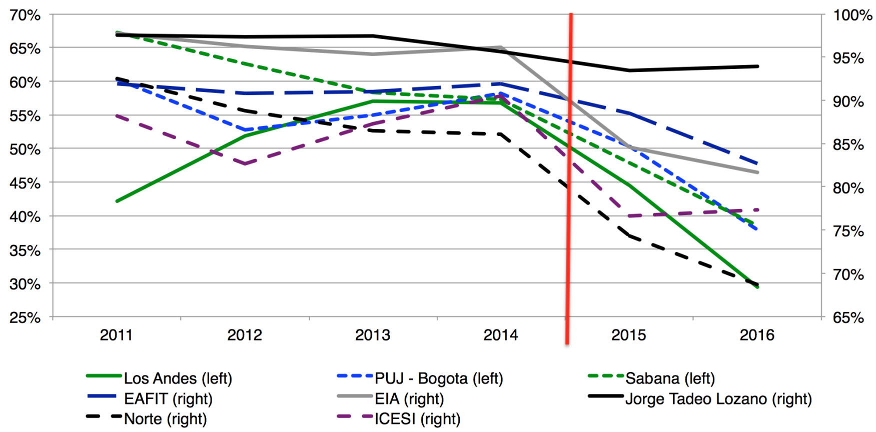
(a) Number of Spring Applicants



(b) Size of Entering Cohort



(c) Admission Rates



Notes: Panel A plots the number of undergraduate applicants at selected private high-quality institutions from Spring 2011 to Spring 2016 (normalized for each institution to equal 1 in Spring 2014), in the four largest metropolitan areas in Colombia. The figure suggests that the number of applicants increased between 46 percent (EAFIT) and 150 percent (ICESI) between 2014 and 2016. Panel B plots the size of the entering cohort at these same institutions (normalized for each institution to equal 1 in Spring 2014), and suggests that cohort size increased between 8 and 100 percent between 2014 and 2016. Panel C plots the admission rate, that is, the share of applicants that are admitted, and suggests that this fraction decreased significantly across institutions. *Source:* Authors' calculations using administrative data.

B.7 Institutional Responses: Tuition Fees

A final concern is whether the expansion of government-provided financial aid and the resulting higher demand for education has been captured (partly or entirely) by schools via tuition hikes, as predicted by the so-called Bennett Hypothesis.[13] Private colleges would be especially prone to have this behavioral response, as they have significantly more flexibility

[13] The Bennett Hypothesis is named after former Secretary of Education William J. Bennett who, in a New York Times opinion piece published in February 18, 1987, argued that "increases in financial aid in recent years

in setting their own tuition fees than their public counterparts. Interestingly, however, real tuition fees for entering, full-time undergraduate students at private high-quality HEIs have remained stable in 2014 through 2017 and not statistically different than private low-quality HEIs between 2014 and 2017.[14][15] We therefore do not yet find empirical evidence that aid increases feed tuition increases.[16]

---

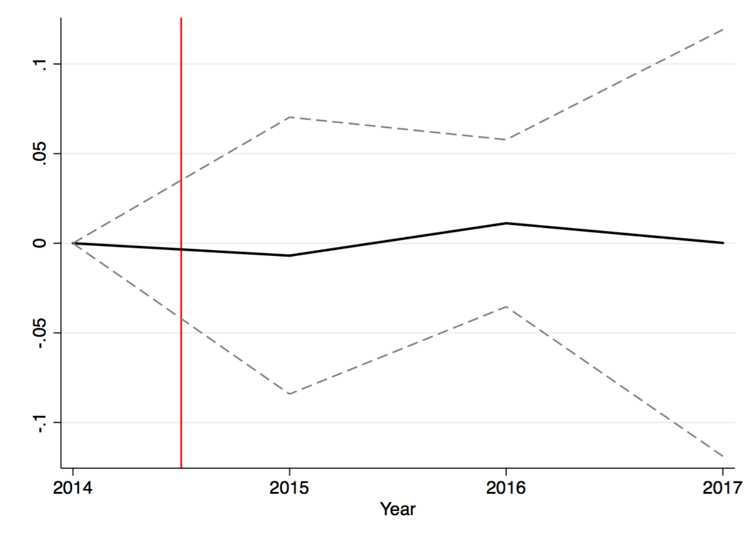
have enabled colleges and universities blithely to raise their tuitions, confident that Federal loan subsidies would help cushion the increase."

[14]Appendix Figure G.21 presents the results from running a regression of a full interaction of year fixed effects and a dummy for a HEI having received High Quality Accreditation by 2014 on a balanced sample of HEIs. The figure suggests there is no statistically significant difference in relative tuition fees between accredited and non-accredited private institutions, despite the rightward shift in the demand for the former type of education. Results using an unbalanced sample of HEIs are qualitatively similar. It should be noted that public HEIs are excluded from this comparison because these institutions set their tuition fees by student household income.

[15]In Colombia, private HEIs are forbidden by law to have for-profit status. Their tuition fees are regulated, and may be increased annually based on the maximum annual inflation rate as of December of the previous year.

[16]On a related note, after SPP, many high-quality universities have incorporated new academic and financial aid complementary programs available to all their student body with the aim of reducing dropout rates and providing a smoother transition into postsecondary education (DNP et al., 2016).

Figure G.21: Difference in the relative tuition fees between high- and low-quality private HEIs



Note: This figure presents the difference in relative tuition fees between high-quality and low-quality private HEIs in Colombia in 2014–2017. The figure compares tuition fees for first-time, undergraduate students from a balanced sample of 40 private HEIs, 5 of which are high-quality and 35 of which are low-quality. The average HEI tuition relative to 2014 is regressed on a full interaction of year fixed effects and a dummy that equals 1 if the HEI has received High Quality Accreditation status by 2014 (2014 is the omitted category). The black solid line plots the coefficient on the interaction terms. The gray dashed lines represent the 95 percent confidence intervals. The red vertical line marks the period SPP is announced. The figure shows that tuition did not increase at high-quality private HEIs relative to low-quality private HEIs despite the rightward shift in demand for the former type of education. Results using an unbalanced sample of HEIs are qualitatively similar. *Sources:* Authors' calculations based on MEN (2018).

B.8 A Stylized Cost-Benefit Discussion

Ultimately, when assessing the impact of SPP on students, we care about how their long-run life outcomes are affected, such as college exit test scores, graduation rates, earnings, and the likelihood of employment. These long-run outcomes also enable cost-benefit assessments of the program. For instance, SPP geared students towards private high-quality universities, which have the highest returns to schooling but are also more costly to both the taxpayer and the student (if he or she were to drop out) relative to their public and low-quality counterparts. Unfortunately, having begun their undergraduate studies in 2015, SPP recipients will only begin to graduate *circa* 2020. Therefore we currently lack complete information on the full costs and benefits of the program.[17]

[17] In addition, the ultimate returns of SPP requires observing a more diverse set of outcomes, such as how the program affected, *inter alia*, labor force decisions, mobility, health, and family formation. This too requires

In addition to these timing and data availability constraints, there are a number of other complications in our analysis. First, SPP recipients are very different from other graduates of high-quality HEIs, both in terms of SES and test scores. It is not clear that their labor market outcomes will resemble those of other graduates or similar graduates from prior cohorts, as some evidence suggests SES and college quality are complementary educational inputs (Riehl, 2016) and elite colleges produce high-end labor market outcomes only for students from wealthy households (Zimmerman, 2016). Moreover, we ignore potential market-wide earnings losses from graduates of high-quality colleges resulting from low-SES student entry at these schools (Riehl, 2016). If SPP indeed changes the education production function, an ideal cost-benefit analysis should take these changes into account. Second, we find evidence that SPP also indirectly positively impacted college access for aid-ineligible students, which should also be incorporated in a comprehensive cost-benefit analysis. Third, the initial shock in college admissions caused by SPP at elite private colleges, documented in Section 2.5, displaced some applicants to lower-ranked universities the first year SPP was introduced. These displaced applicants could potentially incurr some earnings losses. Unfortunately, we do not observe application behavior nor college preference rank orderings for all high school students, so we cannot identify all displaced applicants to incorporate these "losers" into our cost-benefit analysis.[18] Finally, a more careful analysis would also include, *inter alia*, the potential societal gains from promoting class diversity at elite colleges, and the positive peer effects at private high-quality HEIs (and potentially negative peer effects at public low-quality HEIs). We leave that exercise for future work.

These complications imply that our cost-benefit discussion is therefore necessarily speculative, as the only way to be sure of SPP's long-run effects is to directly measure long-run outcomes for SPP participants. It is imperative that future research evaluate the impact this program will have on social mobility and explore its effects on longer-term outcomes when these data become available.

Notwithstanding these limitations, we proceed with the cost-benefit analysis as follows. Using data on total payments in tuition fees to universities and maintenance subsidies to beneficiaries, Appendix Table H.12 presents the annual cost for the first cohort of SPP re-

following students over a much longer time-frame than is currently possible.

[18]Indeed, Section 2.5 documented how a SPP recipient who is admitted to and attends a top school instead of a lower-ranked college reaps a lifetime benefit, but that student displaces another student from attending that top school, who then attends a slightly less prestigious school. That displaced student in turn displaces another student at that school, and so on. The empirical literature suggests that the long-term effects of not attending a top-ranked institution are substantial. Exploiting discontinuities in entry into the University of Los Andes, Saavedra (2009) estimates admitted applicants are 18 percent more likely to be formally employed and earn 20 percent more one year after graduation. Moreover, admitted applicants who enroll are 30 percent more likely to have a formal job and earn 35 percent more after graduation. Importantly, the effects on exit exam scores and formal employment are larger for low-SES students than for high-SES students.

cipients in 2015, which amounted to 135.4 billion Colombian pesos (USD 42.7 million) or 14.8 million pesos per capita (USD 4,657). This is around 4.2 percent of annual central government spending on higher education. Assuming the annual dropout rate stabilizes at 2.32 percent after 2016, and given the average program is 9.5 semesters long, the cost of the full 4.5 years of study for the first cohort of SPP recipients reaches 598.9 billion pesos (USD 188.7 million), or 78.4 million (USD 24,691) per beneficiary. After this first cohort, and as a response to complaints that resources were disproportionately benefitting private institutions, the government appeased critics by paying public universities the marginal student cost instead of their tuition fee (artificially lower after public subsidies). This shrank the per student resource gap between public and private institutions, but raised the cost of the program to 913.8 billion (USD 287.9 million) pesos for this second cohort, that is, 94.7 million pesos per beneficiary (USD 29,845).

Since earnings are unavailable for our sample because SPP recipients have not yet graduated college, we project earnings impacts using income data from household surveys and social security contributions. We caution that our projections will produce reliable estimates of earnings impacts if the relationship between (type of) college attendance and earnings is causal, a strong assumption that is not likely to hold due to selection into (type of) college. We complement our analysis with literature estimates that relate the *type* of college attended with later life outcomes, but this still leaves the important caveat that the treated populations from SPP are very different from that of previous studies, both in terms of SES and test scores. This becomes particularly relevant in contexts where elite colleges produce high-end labor market outcomes only for students from wealthy households (Zimmerman, 2016), although evidence from Colombia is more nuanced.[19] Moreover, the magnitude of the impacts documented in Section 2.5 could potentially also affect labor market outcomes for these cohorts (Riehl, 2016). This implies that our cost-benefit discussion will be necessarily speculative, as the only way to be sure of SPP's long-run effects is to directly measure long-run outcomes for SPP participants.

Our reduced form point estimate is that SPP eligibility raises enrollment in postsecondary education by 25.4 to 32.3 percentage points. Consider the strong assumption that the only impact of the program is to raise 30 of each 100 students from "no postsecondary education" to "high-quality postsecondary education", thus ignoring any graduate degree or other unobserved effects.[20] We estimate that the average college graduate from a high-quality HEI in Colombia will earn 1.37 billion in 2015 pesos (USD 432,625) over the course

[19]Appendix Figure H.24, which plots earnings profiles for graduates by socioeconomic stratum, suggests that, even controlling for a degree from a private, high-quality HEI, there is a positive correlation between baseline SES and earnings after college. Moreover, the wage gap between high- and low-SES graduates roughly doubles six years after graduation. Yet controlling for selection into college type yields different results: (Saavedra, 2009), exploiting a sharp discontinuity in the likelihood of admission into college, shows that low-SES students benefit disproportionately from attending the flagship private selective university in Colombia.

[20]See Appendix B.8 for the list of assumptions made to obtain our estimates of earnings gain.

of his/her lifetime, while the average high school graduate will earn 448,891,226 pesos (USD 141,428).[21] To calculate the earnings gain of SPP recipients, we take into account the foregone earnings of attending a 4.5-year college (the average duration of a program attended by SPP recipients), as well as the monetary subsidies received by SPP beneficiaries over the course of their studies.[22] We thus estimate that SPP might raise earnings by an annuity of 16,309,969 pesos (USD 5,139) for forty years for a net present value of around 382,703,013 pesos (USD 120,574) at a 5 percent interest rate. This significant premium reflects the fact that Colombia has the largest return to college education in Latin America, and the second largest wage gap between college- and high school-educated workers in the region.[23]

Moreover, we documented how students sorted into high-quality HEIs, and especially top private schools. Previous literature has shown that graduates from top private HEIs in Colombia enjoy a significant wage premium over both low-quality HEIs and selective public universities (Camacho et al., 2016; MacLeod and Urquiola, 2015; Riehl et al., 2016; Saavedra, 2009). To estimate how much more are SPP recipients likely to earn by switching to top private high-quality schools, we turn to administrative OLE data for 2008–2013 formal employees' wages for 2001–2013 college graduates.[24]

Figure H.22 plots the earnings profiles of college graduates by type of HEI. Panel A plots mean log monthly real earnings separately for graduates from high- and low-quality HEIs, while Panel B plots earnings profiles separately for college graduates from private and public high-quality HEIs.[25] Panel A illustrates the magnitude of the earnings gap between graduates from high- and low-quality HEIs. It also shows that the slope of workers'

[21] The average wage of a college graduate from a high-quality institution is taken from the average monthly wage in 2013 of an individual graduating college in 2012, as recorded in OLE, weighted by the frequency of the institutions where SPP recipients enrolled and converted to 2015 pesos.

[22] Our calculations likely underestimate the total amount of subsidies received by SPP recipients. First, the government's *Departamento para la Prosperidad Social* provided additional subsidies upon completion of the academic semester as well as for good academic performance. Second, some private institutions provided in-kind subsidies (e.g., food, photocopies, transportation).

[23] The coefficient from the college dummy in a Mincer regression for men is 1.053 in Colombia compared to 1.004 in Costa Rica, 0.857 in Brazil, 0.851 in Chile, 0.719 in Mexico, 0.559 in Peru, and 0.417 in Argentina. Also, the ratio of hourly wages between workers with tertiary and secondary education is 2.82 in Brazil, 2.72 in Colombia, 2.61 in Mexico, 2.60 in Chile, 2.28 in Costa Rica, 1.7 in Peru, and 1.49 in Argentina (SEDLAC – CEDLAS and The World Bank).

[24] We restrict our sample by excluding the following observations: (i) self-employed workers (whose wages we do not observe) and other non-employees; (ii) non-university degrees (e.g., technical or technological, specialization, masters, PhD); (iii) distance learning degrees; (iv) earnings after obtaining a second undergraduate degree or a graduate degree; (v) ages below 18 or above 35; (vi) more than 9 years of experience (there are very few individuals in our data with 10 or more years of experience, making these estimates very noisy); (vii) graduates from the National Appenticeship Service (SENA), a public vocational training school; and (viii) graduates from military or police academies, as they would typically be outside of the traditional labor market.

[25] We define high-quality institutions as HEIs having received High Quality Accreditation by October 1, 2014, i.e., the day SPP was announced.

earnings-experience profiles is higher for high-quality than for low-quality HEIs, with the earnings gap more than doubling over the first nine years of experience. Moreover, there is a significant gap between earnings from private and public high-quality college graduates that stays roughly constant across time (Panel B).[26]

Although the correlation depicted in Figure H.22 cannot be interpreted as the causal effect of college type on earnings, we can use it to feed our back-of-the-envelope calculation of the benefits of SPP for beneficiaries. Thus, moving a college graduate from low- to high-quality education might raise earnings by an annuity of 2,455,573 pesos (USD 774) for forty years for a net present value of around 127.3 million pesos (USD 21,172) at a 5 percent interest rate.[27] Moreover, moving a college graduate from low-quality to *private* high-quality education raises earnings by an annuity of 3,941,820 pesos (USD 1,242) for forty years for a net present value of around 174.4 million pesos (USD 29,482) at a 5 percent interest rate.[28] Note that these returns reflect not only higher wages, but also lower unemployment and informality rates.

These back-of-the-envelope calculations suggest that SPP's costs are more than offset by the increased earnings of beneficiaries. The private internal rate of return of SPP is 44.5 percent, while the social rate of return—which takes into account the opportunity cost of other social investments, an expanded income tax base, positive externalities from an educated workforce, lower public spending on subsidized healthcare, and lower crime rates (see Appendix B.8)—is 16.1 percent. Moreover, by subsidizing students to enroll in the universities with the highest returns, SPP has the potential to reshape intergenerational mobility in a country where social mobility has historically been low (Montenegro and Meléndez, 2014).

An argument, often voiced by critics of SPP, sustains that resources for SPP could have been used to benet a larger number of students at public, high-quality universities. Our estimations reveal that the biannual marginal student cost is roughly similar in high-quality private and public universities (see Table H.12). However, tuition fees at public HEIs are significantly lower because they are heavily subsidized. A direct implication is that such an argument is not entirely accurate.

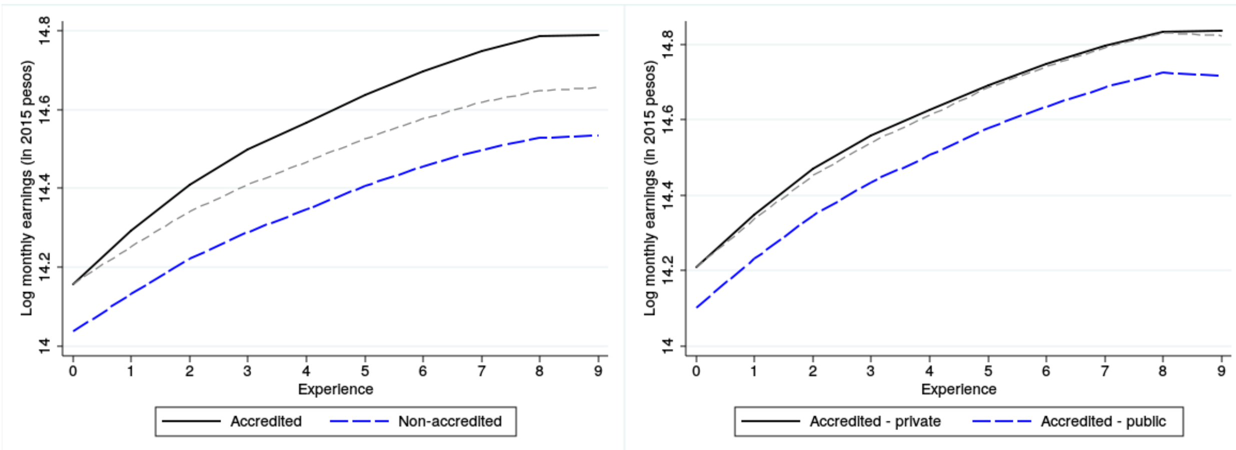
[26] Appendix Figure H.23 compares the average formal earnings profiles plotted in Figure H.22 with those of college graduates that are similar to SPP recipients in two observable characteristics, namely SES (strata 1–3) and SABER 11 score (top decile). These students enjoy a wage premium compared to an average graduate.

[27] We multiply OLE's monthly averages by twelve to obtain annual estimates.

[28] These calculations assume wages stabilize after the ninth year of experience.

Figure H.22: Earnings profiles of college graduates

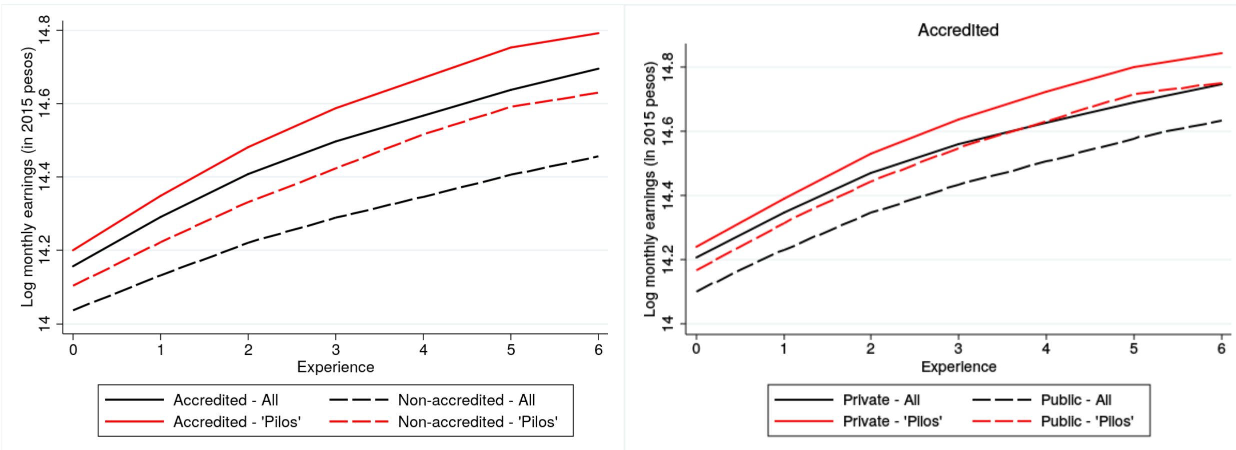
(a) High- versus low-quality (b) Accredited: private versus public



Notes: The figures include 2001–2013 graduates from 234 different college campuses with earnings observations in 2008–2013. Lines depict the mean log monthly real earnings (in 2015 Colombian pesos) for university graduates from high- and low-quality tertiary institutions (Panel A) and high-quality private and public institutions (Panel B). We define experience as earnings year minus graduation year. Wages with a value of zero (the unemployed) are excluded. The dashed gray line parrallels the low-quality (Panel A) and public high-quality (Panel B) profiles starting from the high-quality (Panel A) and private high-quality (Panel B) intercept. *Source:* Author's calculations using OLE and ICFES (2016).

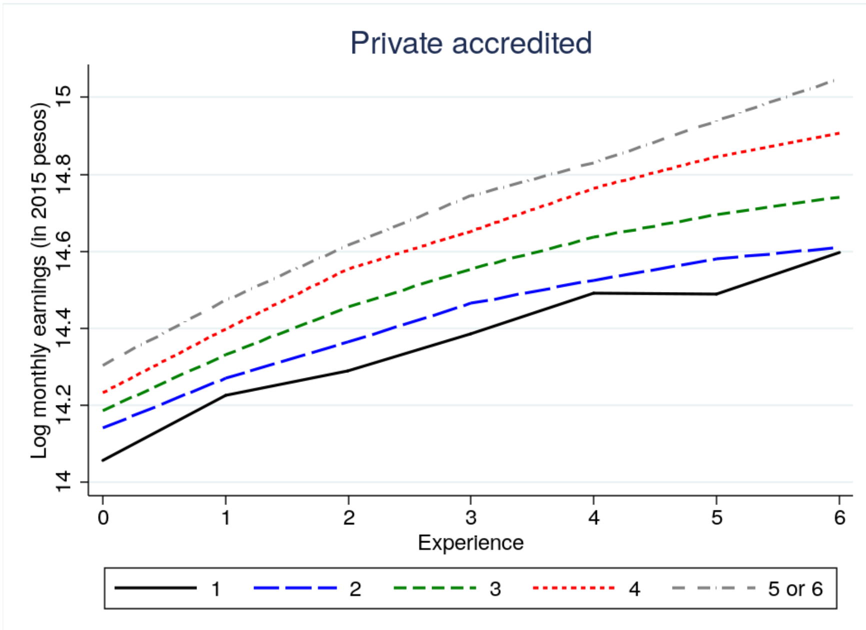
Figure H.23: Earnings profiles of college graduates: All versus SPP recipients

(a) High- versus low-quality (b) Accredited: Private versus public



Notes: The figures include 2002–2013 graduates from 234 different college campuses with earnings observations in 2008–2013. Lines depict the mean log monthly real earnings (in 2015 Colombian pesos) for university graduates from high- and low-quality tertiary institutions (Panel A) and high-quality private and public institutions (Panel B). The red lines restrict the sample to high school graduates from strata 1, 2, and 3 scoring in the top SABER 11 decile. We define experience as earnings year minus graduation year. Wages with a value of zero (the unemployed) are excluded. The dashed gray line is parrallel to the low-quality (Panel A) and public high-quality (Panel B) profiles starting from the high-quality (Panel A) and private high-quality (Panel B) intercept. *Source:* Author's calculations using OLE and ICFES (2016).

Figure H.24: Earnings profiles of private and high-quality college graduates by socioeconomic stratum



Notes: The figures include 2002–2013 graduates from 234 different college campuses with earnings observations in 2008–2013. Lines depict the mean log monthly real earnings (in 2015 Colombian pesos) for university graduates from high-quality private tertiary institutions by socioeconomic stratum. We define experience as earnings year minus graduation year. Wages with a value of zero (the unemployed) are excluded. *Source:* Author's calculations using OLE and ICFES (2016).

Table H.12: Cost of SPP tuition fees and maintenance subsidies (in Colombian pesos)

	First cohort				Second cohort		All 4 cohorts
	2015-1	2015-2	2016-1	Expected total	2016-1	Expected total	Expected total
Total tuition fees							
Total (in million pesos)	51,714	51,569	53,767	457,162	83,050	709,920	2,586,922.59
N	9,205	9,110	8,899	7,642	11,912	9,646.95	36,582.55
Per capita	5,618,051	5,660,747	6,041,968	59,824,651	6,971,921	73,590,110	280,594,981.45
Public institutions							
Subtotal (in million pesos)	3,421	3,423	3,651	30,876	11,426	97,675	323,900.15
N	2,018	1,941	1,906	1,637	1,879	1,521.71	6,201.84
Per capita	1,695,403	1,763,677	1,915,689	18,864,798	6,081,107	64,187,381	211,426,940.88
Private institutions							
Subtotal (in million pesos)	48,293	48,146	50,116	426,285	71,289	609,390	2,254,454.81
N	7,187	7,169	6,993	6,005	9,985	8,086.37	30,264.09
Per capita	6,719,471	6,715,875	7,166,619	70,988,599	7,139,612	75,360,128	297,068,984.36
Maintenance subsidies							
Total (in million pesos)	16,132	15,949	16,658	141,717	23,856	203,926	753,495.35
Per capita	1,752,485	1,750,744	1,871,884	18,545,258	2,002,698	21,138,909	81,961,985.22
TOTAL (in million pesos)	67,846	67,519	70,425	598,879	106,906	913,846	3,340,418.08
Per capita	7,370,536	7,411,490	7,913,852	78,369,909	8,974,619	94,729,024	91,311,785

Notes: 2015 USD 1 = COP 3174.

Assumptions in the cost-benefit calculation

Assumptions in the cost-benefit calculation

1. *Work life expectancy:* 40 years after graduating from university. We also assume that all college graduates are 25 years old at the time of graduation and experience no unemployment spells.

2. *Expected wages:* Mincerian regressions are used to derive wage profiles for high school and college graduates. The coefficient for an additional year of experience is 0.0386 and the curvature on age is -0.0008 (based on Sanchez and Nunez (2003), and Sanchez and Alvarez (2011)).

3. *Duration of study:* The average duration of study for SPP recipients is 9.5 semesters. We assume SPP recipients complete their undergraduate studies by the end of this period.

4. *Entry wage:* The entry wage for college graduates, taken from OLE, is 1,971,317 pesos. It is based on the observed wages in 2013 for students that graduated college in 2012, converted to 2015 pesos. To account for the fact that graduates from private and high-quality universities have higher average wages than those from public and low-quality institutions, we weight OLE observations by the frequency of SPP recipients enrolled in each institution. Wages for high school graduates are 612,152.48 pesos, based on household survey data (GEIH).

5. *Dropout rate:* Note that although the average drop out rate is 50 percent (SPADIES, 2016), it is much lower for students with financial aid. Based on first-year retention rates among SPP recipients, we estimate that the overall dropout rate for SPP recipients is 10 percent.

6. *Student cost for colleges:* Tuition fee before financial aid ("sticker price") roughly reflects average student cost at private high-quality universities in Colombia. Because 15.8 percent of SPP recipients enrolled in public institutions, whose student costs are significantly lower than sticker prices, we assume average student costs for colleges receiving SPP recipients are 80 percent.

7. *Maintenance subsidies provided by the government*: Average annual maintenance subsidies provided by the government to SPP recipients are 4,000,000 pesos. See Table H.12.

To calculate the social rate of return, we use the following additional assumptions:

1. *Opportunity cost:* We assume the opportunity cost for other social investments is 10 percent.

2. *Additional taxes collected:* The average additional income tax due to lower informality and higher wages is assumed to be 16 percent.

3. *Education externalities:* This is assumed to be 8 percent of the difference in average wages between high school and college graduates.

4. *Subsidized healthcare:* The rate of informality is 22 percent among college-educated workers and 55 percent among high school graduates (DANE, 2016), which thus translates into lower government healthcare subsidies (616,849.20 pesos per person, equivalent to the *UPC–Régimen Subsidiado* established by Resolution 5593 of 2015).

5. *Lower crime rate:* The annual cost of an incarcerated individual is 2015 pesos 13,122,078 (INPEC, 2012). We assume that the probability of committing a crime is 0 for college-educated workers, and 0.003171571 for high-school educated workers (this is the share of incarcerated individuals over population aged 16 and above in 2012).

Appendix C

C.1 Appendix Figures and Tables from Chapter 3

Figure A.1: There is no bunching at admission thresholds

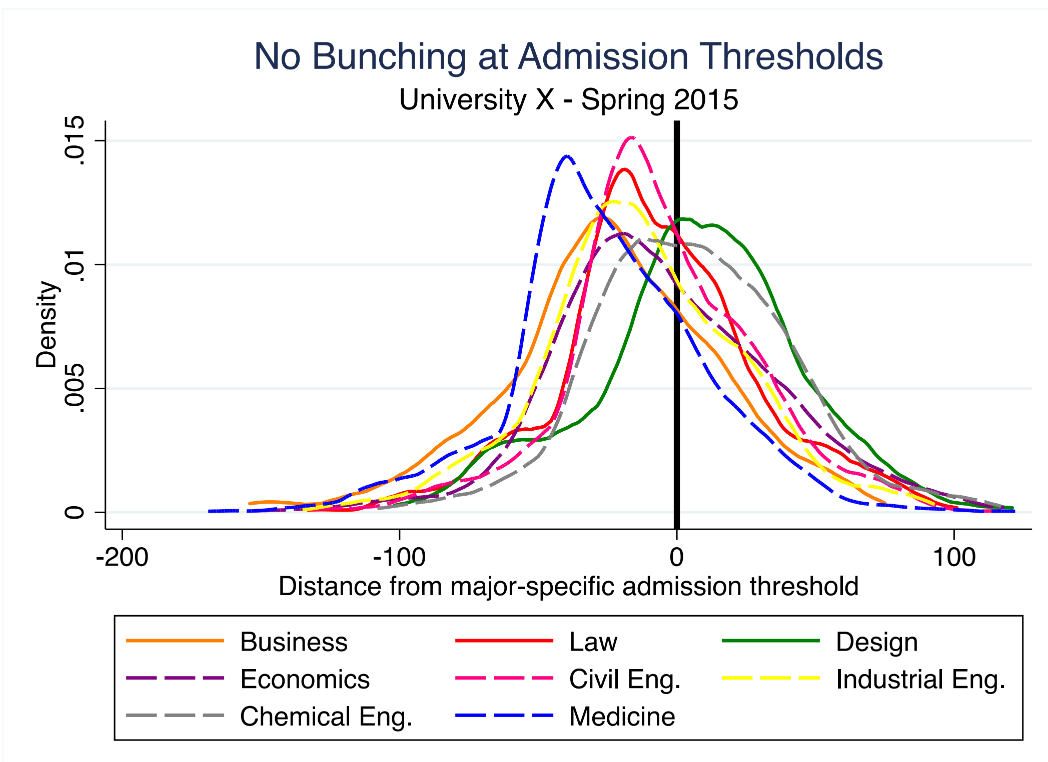


Figure A.2: SPP 'High-Achieving' Condition

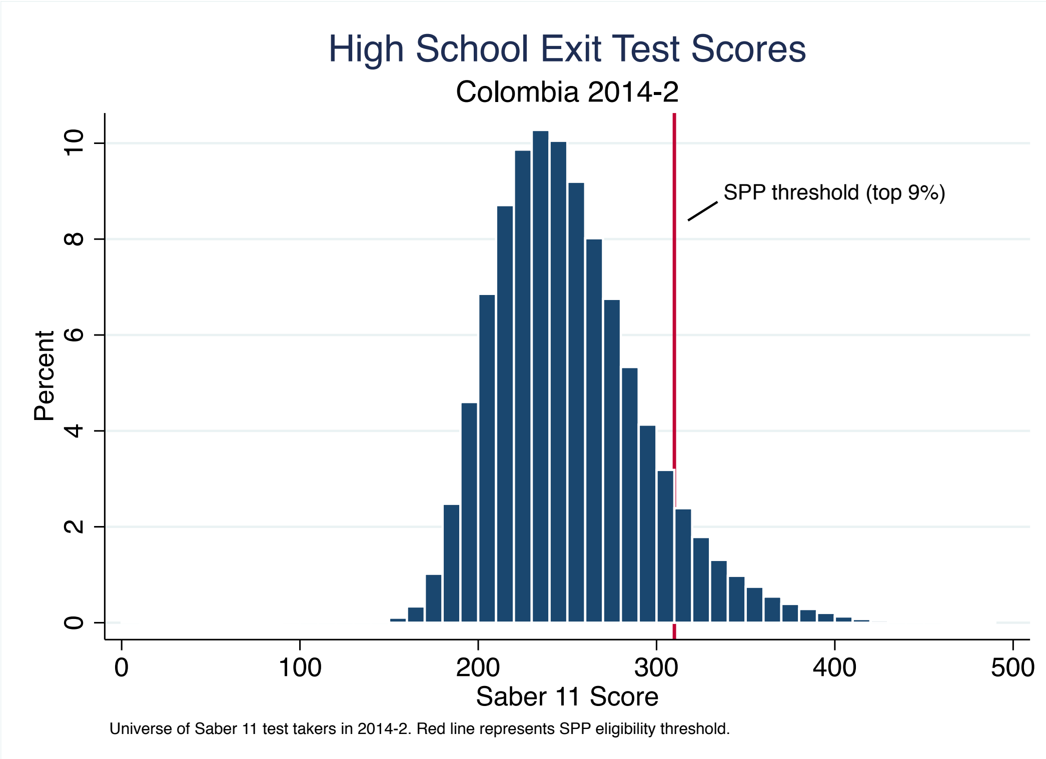


Figure A.3: SPP 'Poor' Condition

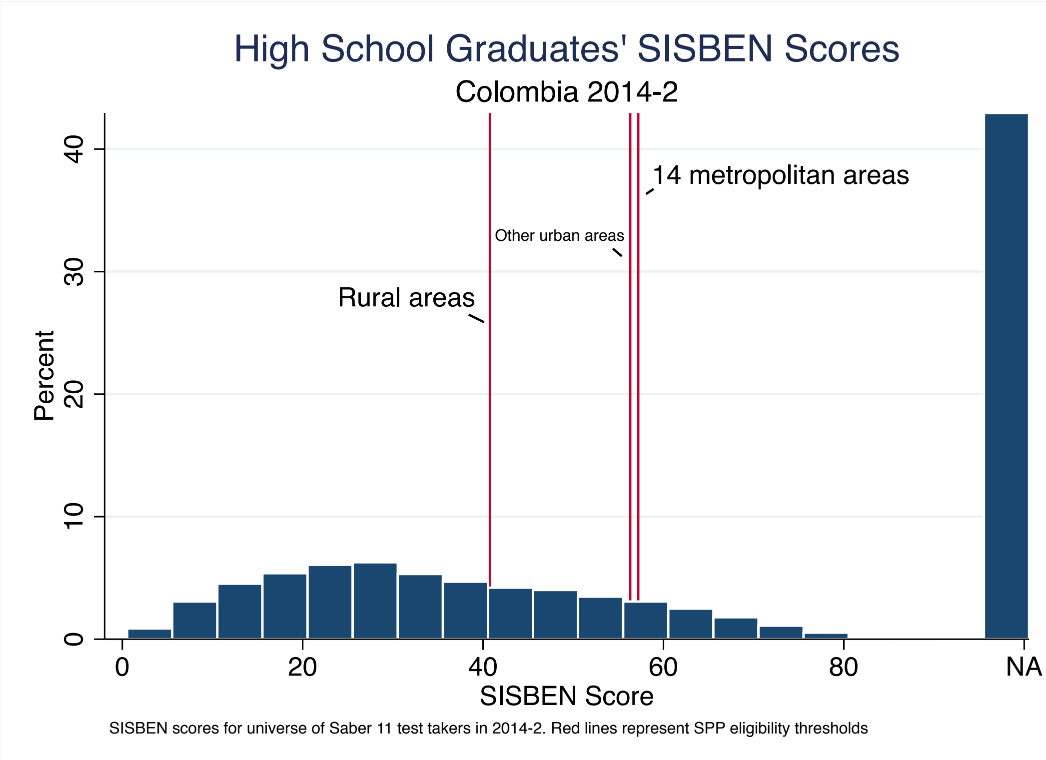


Figure A.4: SPP Continued Raising the Share of Poor Students in 2016

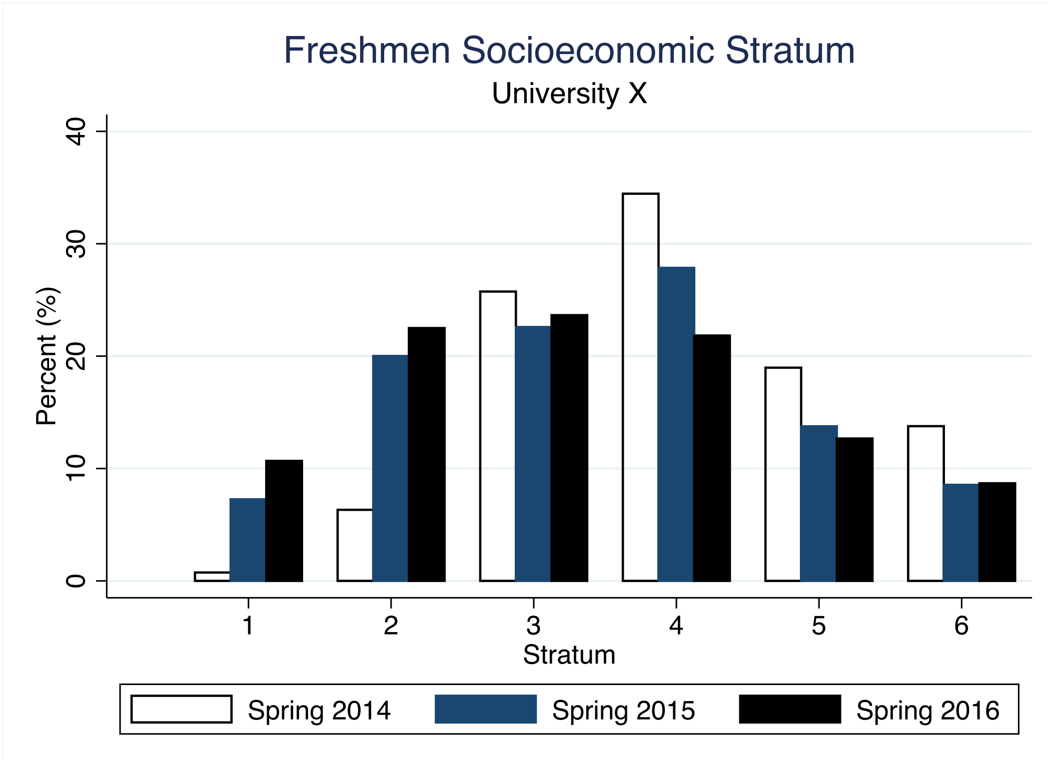
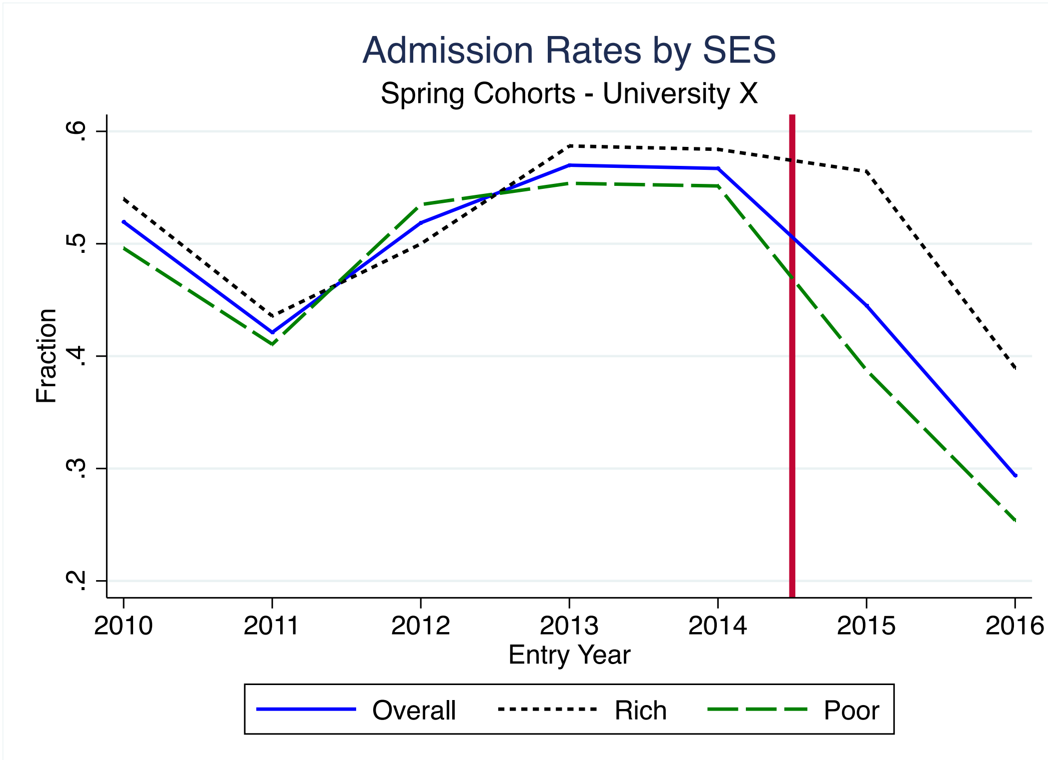


Figure A.5: Admission Rates Decreased as a Result of SPP



Note: "Poor" refers to strata 1–3, while "rich" refers to strata 4–6. The vertical red line represents SPP.

Table A.1: Survey Wave 1 Response Rates by Cohort

Entering Cohort	Response Rate (%)
Spring 2015	25.16
Fall 2014	18.72
Spring 2014	20.46

Note: A joint F test can be rejected at the 5% level (but not at the 1% level).

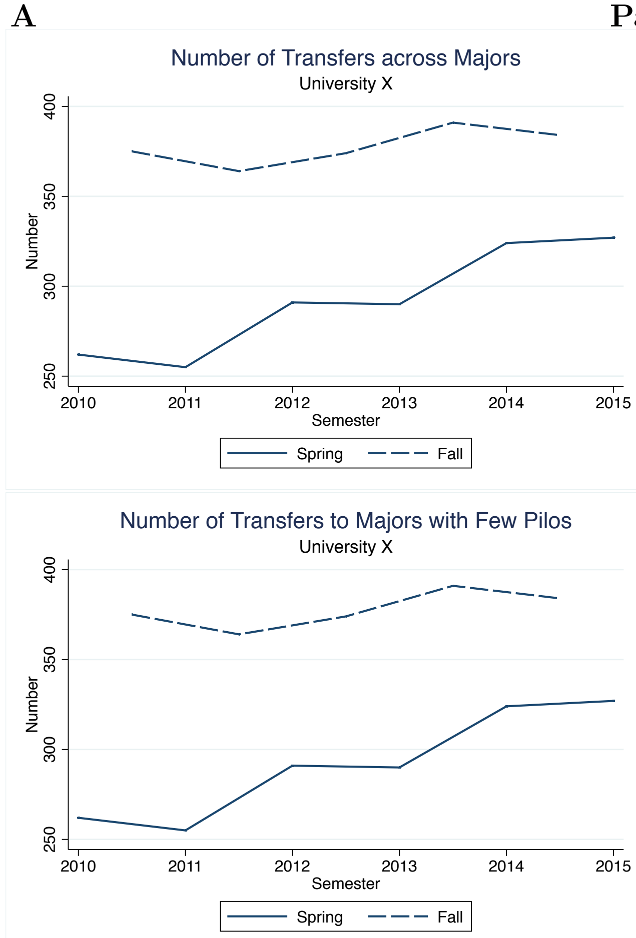
Table A.2: Survey Wave 2 Response Rates by Cohort

Entering Cohort	Response Rate (%)
Spring 2016	27.25
Fall 2015	22.50
Spring 2015	19.71
Fall 2014	17.69
Spring 2014	20.88
Fall 2013	17.94
Spring 2013	24.87

Note: A joint F test can be rejected at the 1% level.

Figure A.6: Students Are Not Avoiding Pilos by Switching Across Majors

Panel A **Panel B**



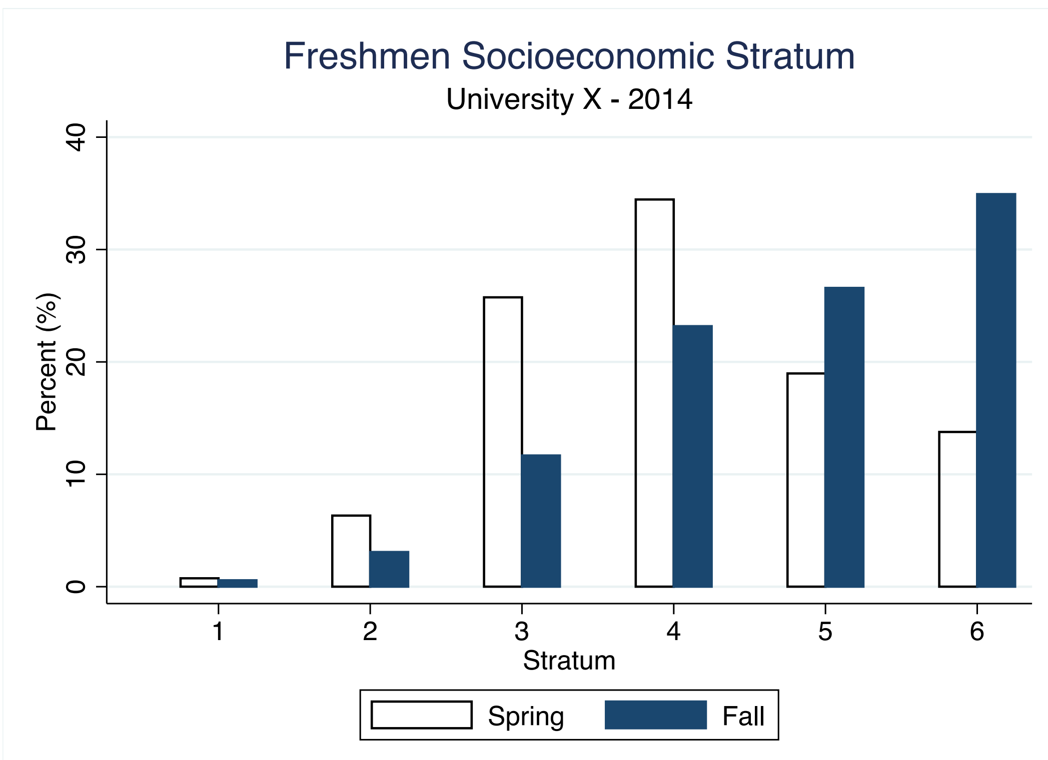
Note: Panel B restricts to students switching to majors where less than 20% of Freshmen in Spring 2015 are Pilo: Art History, Business, Directed Studies, Music, Economics, Government and Public Affairs, Art, and Industrial Engineering.

Table A.3: Rich-poor Interactions After One Year

	Cohort				
	Spring 2015	Fall 2014	Spring 2014	Fall 2013	Spring 2013
	(1)	(2)	(3)	(4)	(5)
Perceived % Pilo classmates	35.00	19.25	17.59	16.23	14.95
	(18.35)	(14.06)	(15.37)	(13.43)	(12.76)
Observed % Pilo classmates	12.58	4.20	2.58	1.66	1.40
	(4.63)	(2.41)	(1.72)	(1.75)	(1.21)
Pilo among 5 closest friends	17.37	4.91	1.35	2.38	0.96
	(37.99)	(21.67)	(11.59)	(15.29)	(9.81)
Pilo among 5 regular study partners	19.76	3.68	2.03	2.38	0.96
	(39.94)	(18.89)	(14.14)	(15.29)	(9.81)
No. times worked with Pilo	3.51	1.27	1.20	1.05	(0.97)
	(3.18)	(2.35)	(2.18)	(1.89)	(2.05)

Note: This table presents means (and standard deviations in parentheses) by cohort, i.e., the semester in which they began their studies at University X.

Figure A.7: Spring vs. Fall



Note: "Stratum" is a measure of socio-economic status designed to target public service subsidies in Colombia. The system classifies dwellings into 6 strata (1 being the poorest) according to their physical characteristics and surroundings. While correlation with income is clearly imperfect, one advantage of using the strata system is straightforwardness: most Colombians are well aware of their stratum, making this information easy to collect.

Table A.4: Perception of the Income Distribution

	Share of Colombians by Socioeconomic Stratum					
	Strat. 6	Strat. 5	Strat. 4	Strat. 3	Strat. 2	Strat. 1
	(1)	(2)	(3)	(4)	(5)	(6)
1[Observed % Pilos $\geq$ 5%]	-1.292**	-1.014**	-2.095**	-0.806	0.938	4.269***
	(0.531)	(0.485)	(1.002)	(1.050)	(0.817)	(1.420)
1[Predicted % Pilos $\geq$ 5%]	-0.584	-0.574	-1.4	-1.341	0.322	3.577**
	(0.519)	(0.549)	(0.853)	(0.952)	(0.716)	(1.437)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Major FE	Yes	Yes	Yes	Yes	Yes	Yes
Spring Only	Yes	Yes	Yes	Yes	Yes	Yes
Spring 2015 Threshold	Yes	Yes	Yes	Yes	Yes	Yes
N	309	309	309	309	309	309
R^2	0.2	0.22	0.17	0.17	0.19	0.2
Dep Mean	5.71	8.8	15.6	21	22.24	26.65
Dep SD	5.25	4.94	7.32	7.84	6.22	12.3

Notes: Each row represents a separate regression. Controls include age, age squared, gender, migrant status, risk aversion, socioeconomic stratum, parental education, financial aid received, high school calendar, and SABER 11 test score. The sample is restricted to Spring cohorts and control for higher positive selection on ability of admitted students due to SPP by dropping pre-Spring 2015 students with test scores below the Spring 2015 cutoffs. Clustered standard errors at the semester-by-major level in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table A.5: The Shift in the Reference Group Raises Perception of Poverty Rate

	Share of Colombians that are poor					
	(1)	(2)	(3)	(4)	(5)	(6)
1[Observed % Pilos $\geq$ 5%]	3.684**	6.599***	6.342***			
	(1.573)	(1.723)	(1.858)			
1[Predicted % Pilos $\geq$ 5%]				4.402***	7.543***	6.639***
				(1.561)	(1.752)	(1.769)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Major FE	Yes	Yes	Yes	Yes	Yes	Yes
Spring Only	No	Yes	Yes	No	Yes	Yes
Spring 2015 Threshold Only	No	No	Yes	No	No	Yes
N	493	342	309	493	342	309
R^2	0.15	0.2	0.21	0.16	0.21	0.21
Dep Mean	36.75	36.06	35.65	36.75	36.06	35.65
Dep SD	19.4	19.24	19.08	19.4	19.24	19.08

Notes: Each row represents a separate regression. Controls include age, age squared, gender, migrant status, risk aversion, socioeconomic stratum, parental education, financial aid received, high school calendar, and SABER 11 test score. Columns (2) and (5) restrict the sample to Spring cohorts only. Columns (3) and (6) further control for higher positive selection on ability of admitted students due to SPP by dropping pre-Spring 2015 students with test scores below the Spring 2015 cutoffs. Clustered standard errors at the semester-by-major level in parentheses.
$***p < 0.01, **p < 0.05, *p < 0.1$

Table A.6: Upward social mobility among the poor

	Upward Stratum 1				**Upward Stratum 2**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1[Observed % Pilos $\geq$ 5%]	-0.057	0.053			0.087**	0.125**		
	(0.046)	(0.063)			(0.040)	(0.054)		
1[Predicted % Pilos $\geq$ 5%]			-0.076	-0.024			0.098**	0.141**
			(0.050)	(0.061)			(0.048)	(0.060)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Major FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Spring Only	No	Yes	No	Yes	No	Yes	No	Yes
Spring 2015 Thresholds	No	Yes	No	Yes	No	Yes	No	Yes
N	493	309	493	309	493	309	493	309
R^2	0.13	0.16	0.13	0.16	0.13	0.2	0.13	0.21
Dep Mean	0.53	0.56	0.53	0.56	0.57	0.59	0.57	0.59
Dep SD	0.5	0.5	0.5	0.5	0.5	0.49	0.5	0.49

Notes: Each row represents a separate regression. Controls include age, age squared, gender, migrant status, risk aversion, socioeconomic stratum, parental education, financial aid received, high school calendar, and SABER 11 test score. Clustered standard errors at the semester-by-major level in parentheses.
$***p < 0.01, **p < 0.05, *p < 0.1$

Table A.7: A Higher Perception of Meritocracy in College Admissions

	Most talented students are admitted at best colleges					
	(1)	(2)	(3)	(4)	(5)	(6)
1[Observed % Pilos $\geq$ 5%]	0.140***	0.119***	0.122**			
	(0.043)	(0.042)	(0.049)			
1[Predicted % Pilos $\geq$ 5%]				0.075*	0.131**	0.129**
				(0.042)	(0.050)	(0.056)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Major FE	Yes	Yes	Yes	Yes	Yes	Yes
Spring Only	No	Yes	Yes	No	Yes	Yes
Spring 2015 Thresholds	No	No	Yes	No	No	Yes
N	493	342	309	493	342	309
R^2	0.11	0.15	0.18	0.1	0.15	0.18
Dep Mean	0.33	0.35	0.36	0.33	0.35	0.36
Dep SD	0.47	0.48	0.48	0.47	0.48	0.48

Notes: Each row represents a separate regression. Controls include age, age squared, gender, migrant status, risk aversion, socioeconomic stratum, parental education, financial aid received, high school calendar, and SABER 11 test score. Columns (3) and (6) control for higher positive selection on ability of admitted students due to SPP by dropping pre-Spring 2015 students with test scores below the Spring 2015 cutoffs. Clustered standard errors at the semester-by-major level in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table A.8: Preferences for redistribution

	State should tax the rich				State should subsidize the poor			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1[Observed % Pilos ≥ 5%]	0.116***	0.195***			0.048	0.111*		
	(0.039)	(0.054)			(0.045)	(0.064)		
1[Predicted % Pilos ≥ 5%]			0.111**	0.142**			0.072*	0.126**
			(0.046)	(0.057)			(0.039)	(0.056)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Major FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Spring Only	No	Yes	No	Yes	No	Yes	No	Yes
Spring 2015 Thresholds	No	Yes	No	Yes	No	Yes	No	Yes
N	493	309	493	309	493	309	493	309
R^2	0.13	0.19	0.12	0.17	0.1	0.16	0.1	0.16
Dep Mean	0.72	0.71	0.72	0.71	0.48	0.46	0.48	0.46
Dep SD	0.45	0.45	0.45	0.45	0.5	0.5	0.5	0.5

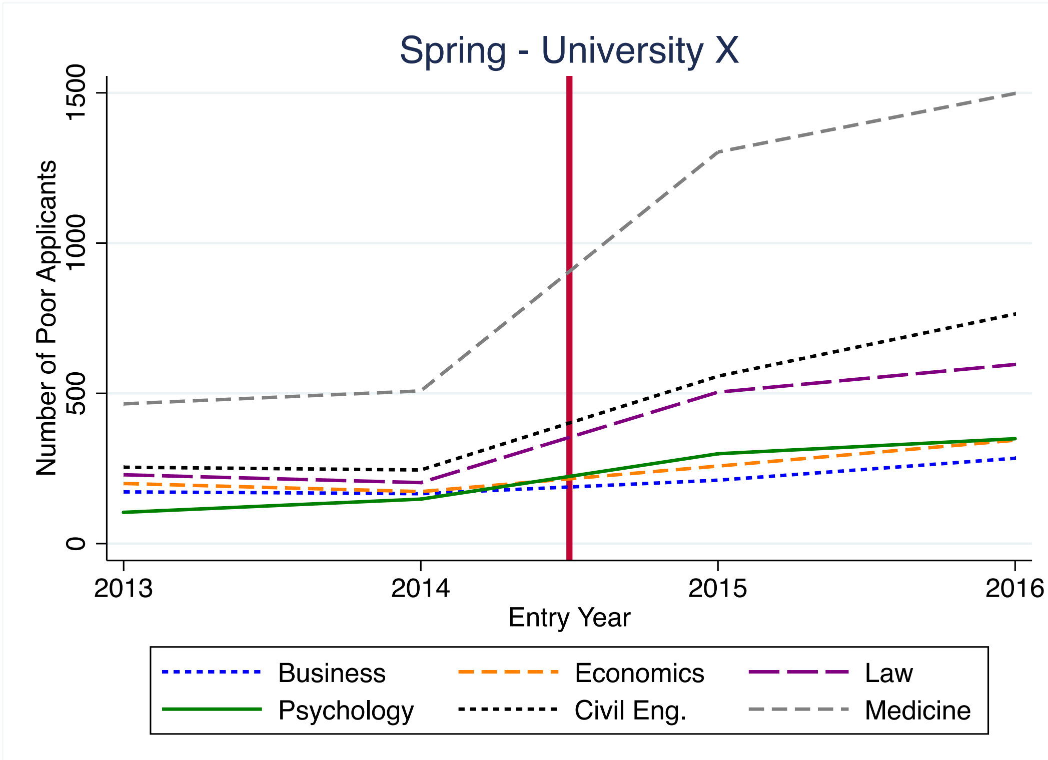
Notes: Each row represents a separate regression. Controls include age, age squared, gender, migrant status, risk aversion, socioeconomic stratum, parental education, financial aid received, high school calendar, and SABER 11 test score. Clustered standard errors at the semester-by-major level in parentheses.
$***p < 0.01, **p < 0.05, *p < 0.1$

Table A.9: Exposure to Pilos Increases Donations to GiveDirectly

	Donated to GiveDirectly		
	(1)	(2)	(3)
1[Observed % Pilos ≥ 5%]	0.08110**	0.09182*	0.13276**
	(0.03871)	(0.05261)	(0.05248)
Controls	Yes	Yes	Yes
Major FE	Yes	Yes	Yes
Spring Only	No	Yes	Yes
Spring 15 Threshold	No	No	Yes
N	729	400	328
R^2	0.1	0.16	0.19
Dep Mean	0.15	0.14	0.13
Dep SD	0.36	0.34	0.33

Notes: Controls include age, age squared, gender, migrant status, risk aversion, socioeconomic stratum, parental education, financial aid received, high school calendar, and SABER 11 test score. Clustered standard errors at the semester-by-major level in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Figure A.8: The Number of Poor Applicants Varies Considerably Across Majors



C.2 Predicting the Distribution of Pilo Classmates

Use distribution of students among classes in the year **prior** to SPP, i.e., Spring 2014 (see Altonji and Card (1991)).

Step 1 Estimate the share of students in every cohort-major pair enrolled in each class

- E.g., 100% of Freshmen Econ majors take ECON 1
- E.g., 90% of Sophomore Econ majors take MATH 2
- E.g., 5% of Freshmen Econ majors take MATH 2

Step 2 Multiply Spring 2014 Freshmen major-class shares by the fraction of Freshmen in each major that are Pilo in Spring 2015 (e.g., 12% in Econ): This gives share of Pilos in each classroom absent sorting due to SPP

- E.g., 12% (= 100% × 12%) of ECON 1 students are Pilo
- E.g., 0.6% (= 5% × 12%) of MATH 2 students are Pilo

Step 3 For the older cohorts, multiply share derived in Step 1 by that obtained in Step 2: This gives the share of Pilo classmates if students behaved similar to the year prior to SPP

- E.g., A Sophomore Econ major will have 0.54% (= 0.6% × 90%) Pilo classmates

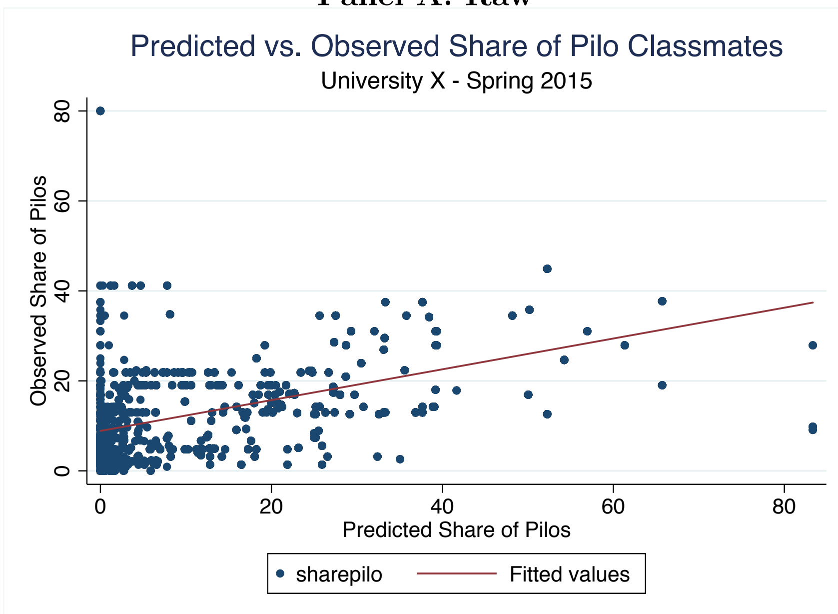
Table B.10: Predicting the distribution of Pilo classmates

	Share of Pilo Classmates in Spring 2015					
	Spring 2015 Cohort		Fall 2014 Cohort		Spring 2014 Cohort	
	Observed (1)	Predicted (2)	Observed (3)	Predicted (4)	Observed (5)	Predicted (6)
Mean	15.49	15.03	4.99	1.26	2.34	0.17
Median	14.12	13.92	4.24	0.44	1.75	0.12
SD	5.85	7.82	3.71	2.9	2.20	0.20
Min	1.19	0	0	0	0	0
Max	35.72	46.88	22.5	19.84	13.82	1.32
$\mathbb{1}$[Share $\geq$ 5%]	99.52%	94.23%	39.07%	5.30%	10.95%	0%

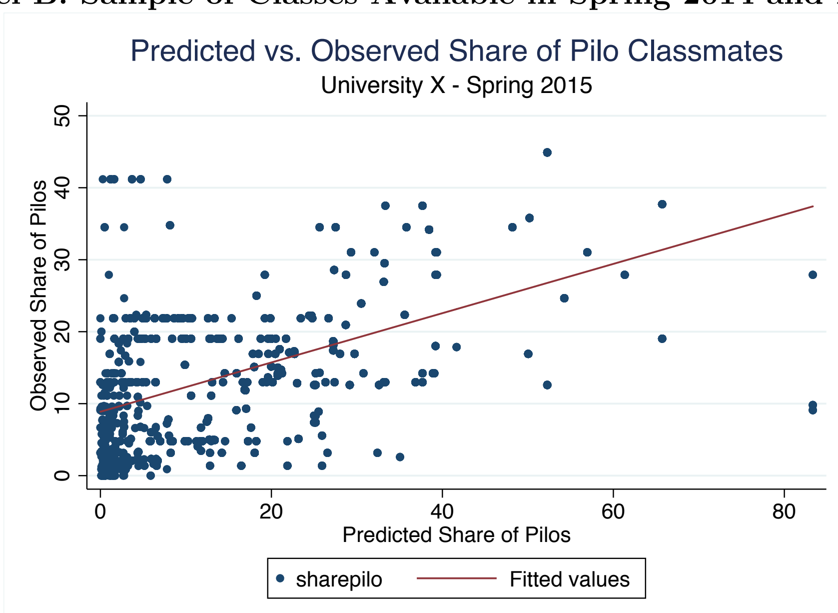
Note: Differences between Observed and Predicted columns reflect some combination of new classes opening in Spring 2015, new majors (Government, Art History), and sorting in response to SPP.

Figure B.9: Correlation of Predicted and Observed Share of Pilo Classmates

Panel A: Raw



Panel B: Sample of Classes Available in Spring 2014 and 2015:



Note: In Panel A, differences between Observed and Predicted reflect some combination of new classes opening in Spring 2015, new majors (Government, Art History), and sorting in response to SPP. Panel B corrects this by including only classes that were available in both Spring 2014 and 2015.