Enforcing Wealth Taxes in the Developing World: Quasi-Experimental Evidence from Colombia
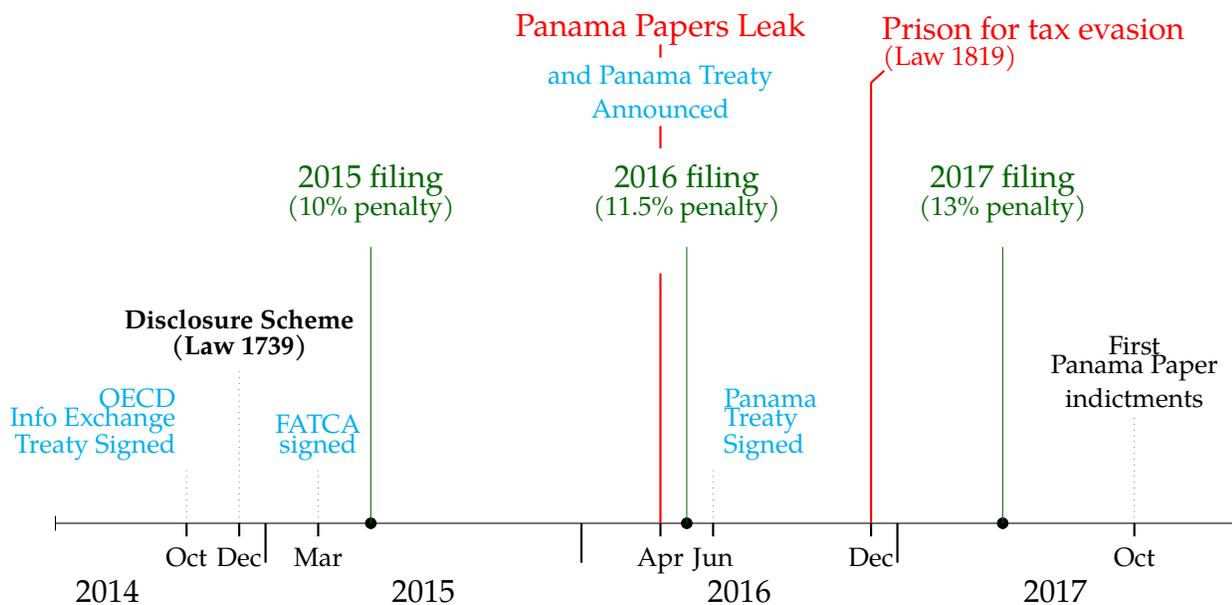
Juliana Londoño-Vélez and Javier Ávila-Mahecha

Online Appendix

Appendices
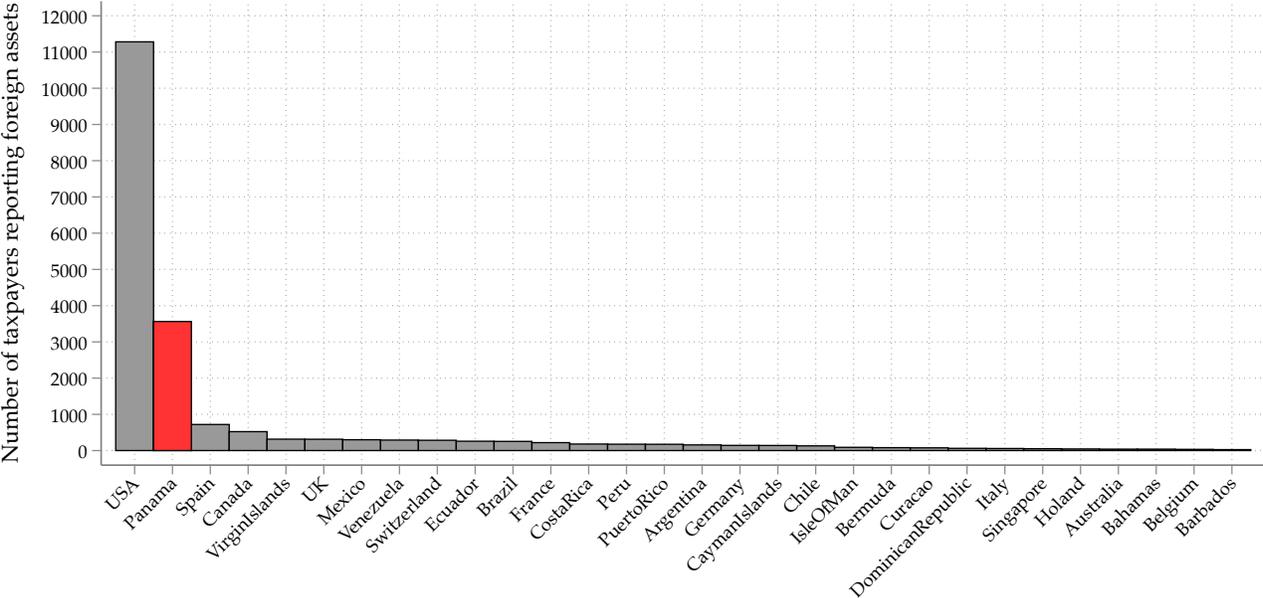
A   Online Figures and Tables

Figure A.1: Timeline of Events

Notes: This figure plots a timeline of the events taking place around Colombia's voluntary disclosure scheme between 2014 and 2018.

Figure A.2: Number of Taxpayers Reporting to Own Foreign Assets in Each Country in 2015


Notes: This figure plots the number of taxpayers reporting to own a foreign asset located in the 30 most common jurisdictions (form #160) in FY 2015, i.e., before the Panama Papers leak. 17,426 taxpayers filed a foreign asset information return in 2015. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Figure A.3: Comparison on Google Trends of Search Terms "Panama Papers" and "Impuesto Riqueza" (Wealth Tax) in Colombia
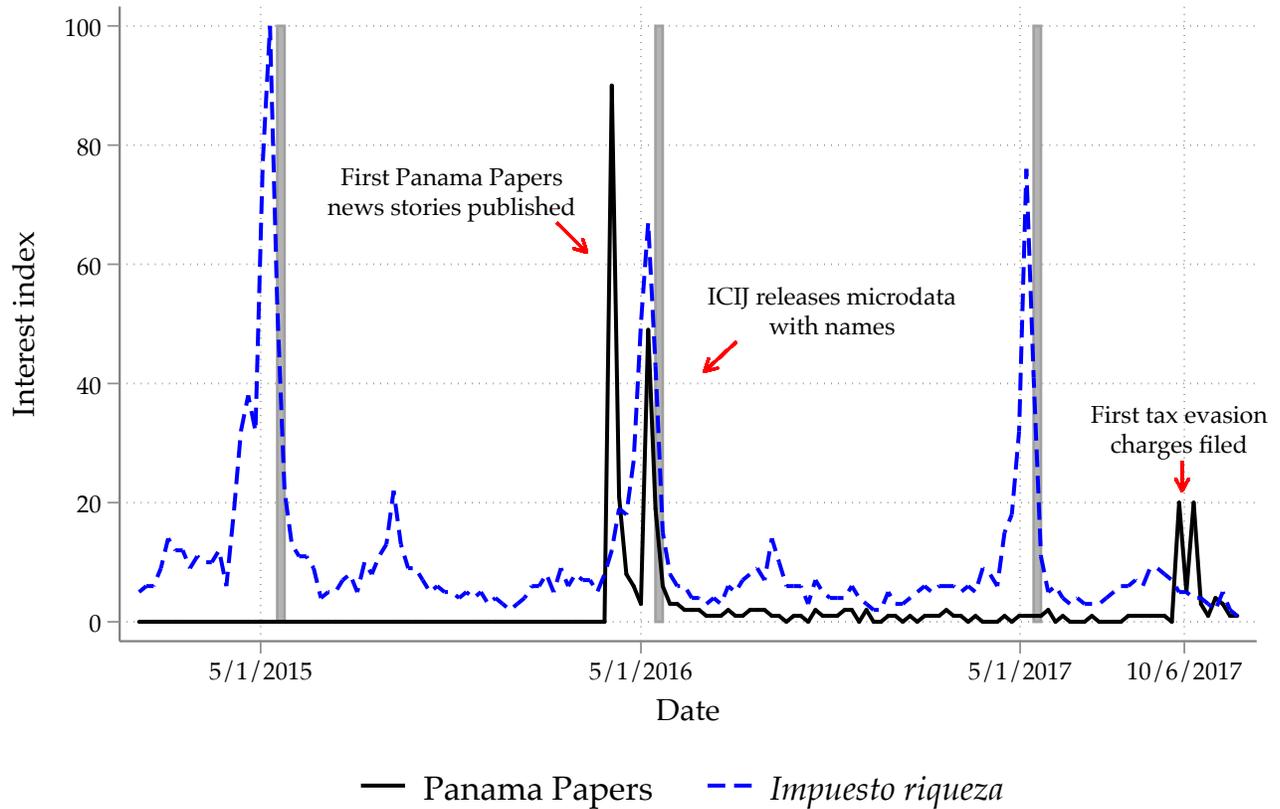

Notes: This figure plots the relative search interest in Colombia of the terms "Panama Papers" in black and "Impuesto riqueza" (wealth tax, in Spanish) in blue. The number in the y-axis represents search interest relative to the highest point on the chart for Colombia during the plotted period of time. 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% 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 Office of the Attorney General—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, and placed under house arrest later that month. *Source:* Google. 2017. "Google Trends." Accessed November 30, 2017. https://trends.google.com/trends/explore?date=2015-01-01%202017-11-30&geo=CO&q=Panama%20Papers,Impuesto%20riqueza.

Figure A.4: Probability of Participating in the 2015–17 Voluntary Disclosure Scheme, by Pre- and Post-Disclosure Wealth Group
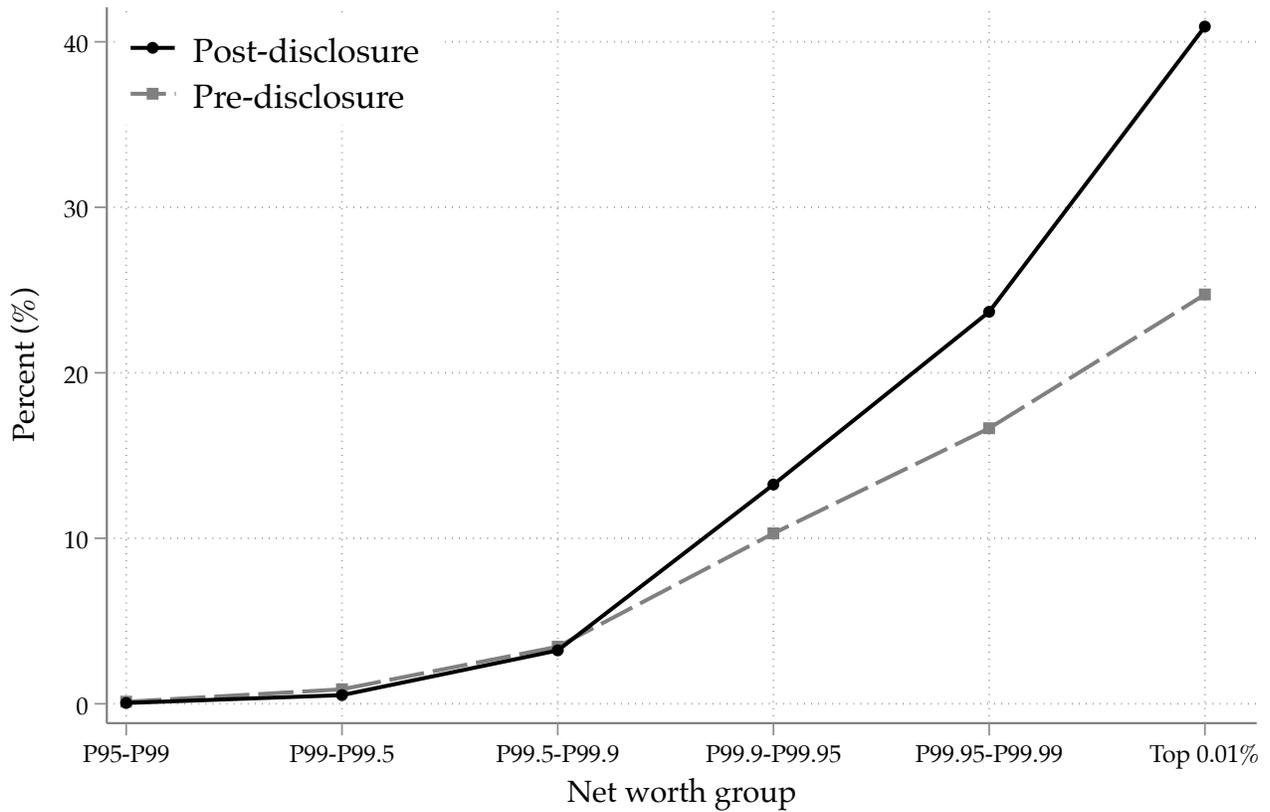
Notes: This figure plots the fraction of tax units in Colombia that participate in the 2015–17 voluntary disclosure program by bins of reported net worth. The figure ranks individuals by pre- and post-disclosure net worth, and shows that participation in the program is increasing in both measures of wealth. Ranking by pre-disclosure wealth, 24.7% of individuals in the wealthiest 0.01% disclosed hidden wealth (dashed gray line). Ranking by wealth including disclosures, 40.9% of individuals in the wealthiest 0.01% 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.5: Probability of Disclosing under Scheme versus Being Named in the Panama Papers
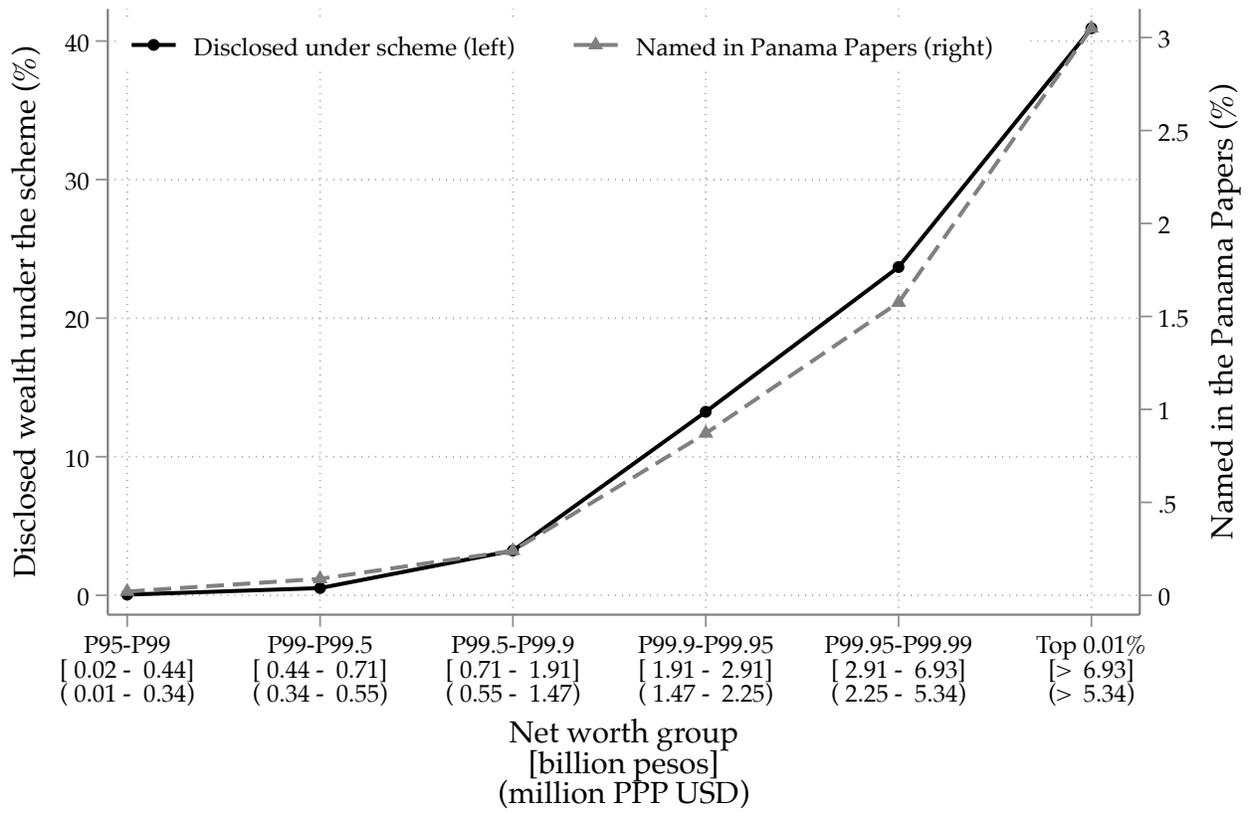
Notes: This figure compares the fraction of tax units in Colombia that (i) participate in the voluntary disclosure program or (ii) were named in the Panama Papers by bins of reported net worth. The wealth gradients in two series are remarkably similar: for the wealthiest 0.01%, these probabilities are roughly twice as large than for the second wealthiest group (P99.95-P99.99), three times as large than for the third wealthiest group (P99.9-P99.95), and thirteen times as large than for the fourth wealthiest group (P99.5-P99.9). In all, being named in the leak is nearly 43 times more likely for the wealthiest 0.01% than the top 5% overall. This odds ratio is smaller than the equivalent 55:1 ratio for willingness to disclose because the "poorest" evaders (i.e., those owning wealth worth PPP USD 0.01–0.55 million) also disclose domestic assets and fake debts. *Source:* Table A.1.

Figure A.6: Foreign Assets Reported in 2017

(a) Probability of Filing a Foreign Asset Information Return
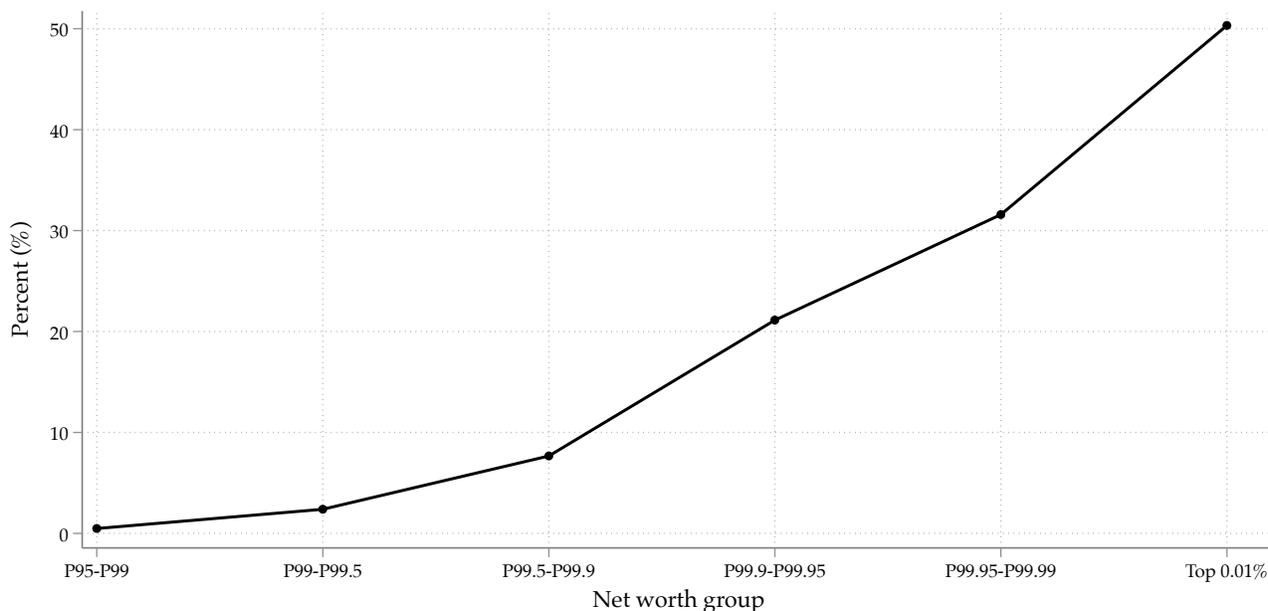
(b) Two Most Popular Destinations
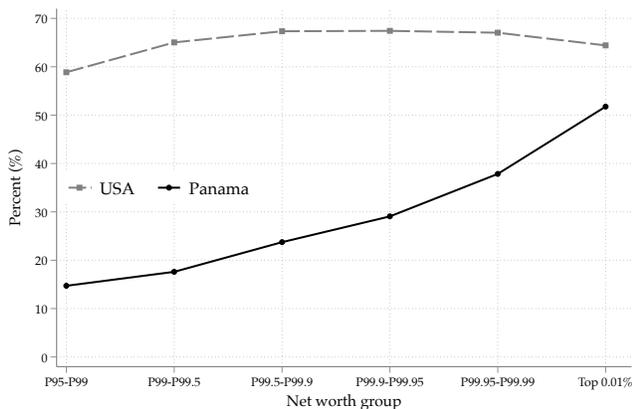
(c) Next Three Most Popular Destinations
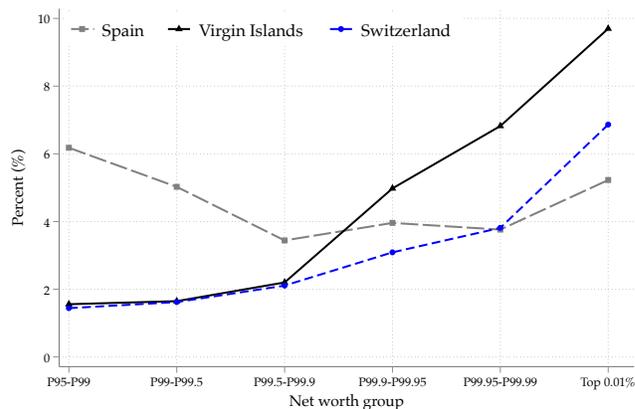
Notes: These figures show foreign asset ownership in FY 2017 by bins of net wealth. Panel (a) plots the likelihood of filing a foreign asset information return (form #160), and shows the probability of reporting to own a foreign asset increases to more than 50% for the wealthiest 0.01%. Panels (b) and (c) plot the likelihood of owning a foreign asset in a given location, conditional on owning any foreign asset. 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. Individuals are ranked based on either wealth reported either in the FY 2017 wealth tax return or the FY 2016 income tax return. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Figure A.7: The Impact of a Voluntary Disclosure Program on Reported Wealth and Income: 2016 Disclosers

(a) Wealth (Assets minus Debt) and Assets
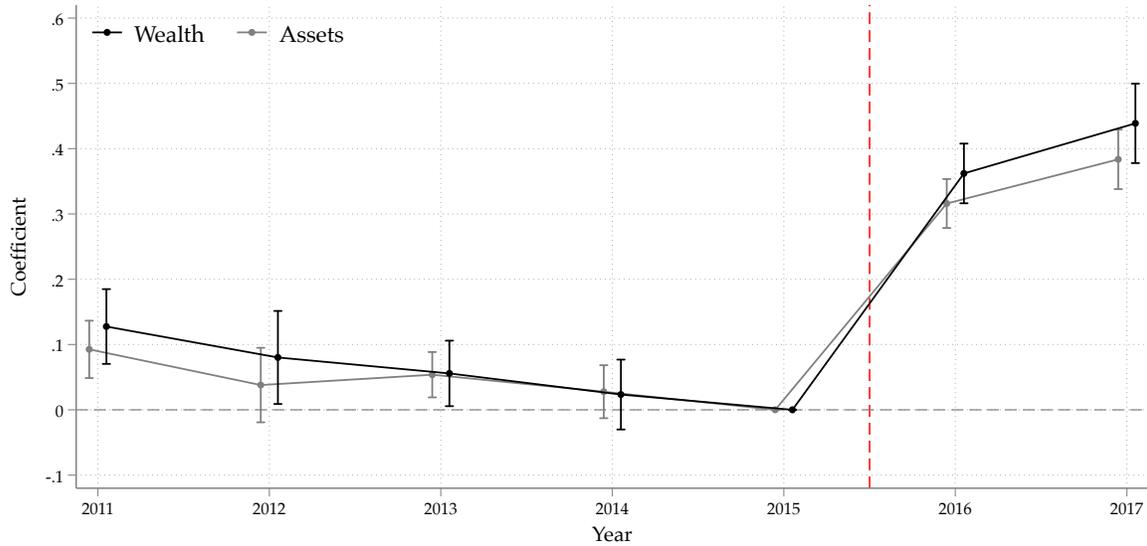
(b) Income: Interest, Foreign, Total Gross, and Tax
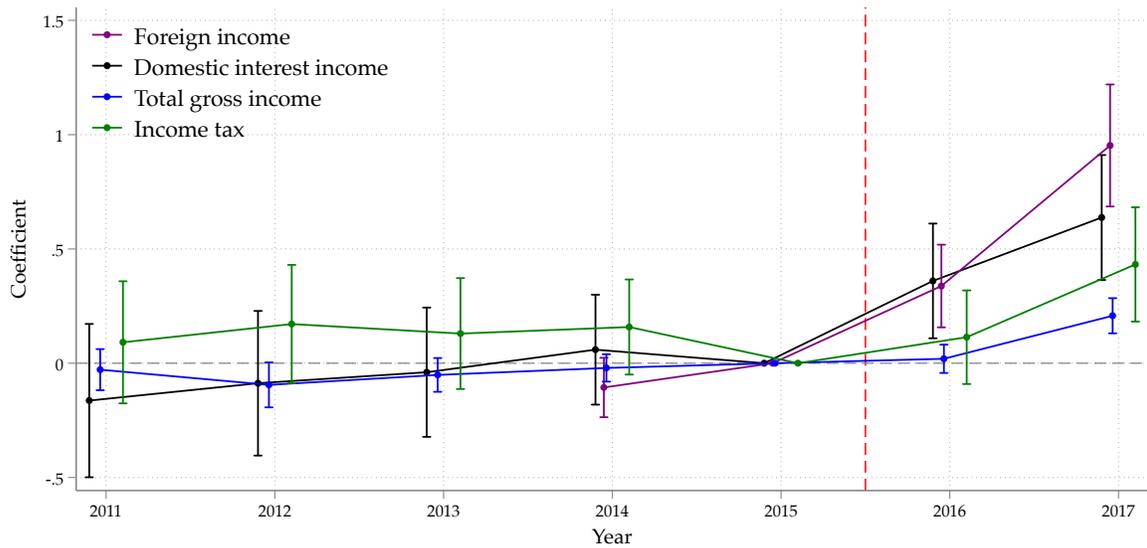
Notes: These figures compare outcomes between 2,074 taxpayers that first disclosed hidden assets or inexistent liabilities in 2016 and 43,181 that never disclosed between 2015 and 2017. The inverse hyperbolic sine transformation of a given outcome is regressed on individual fixed effects and a voluntary 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% 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 (i) filed income taxes annually between 2011 and 2017; and (ii) 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.5. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Figure A.8: The Impact of a Voluntary Disclosure Program on the Probability of Reporting Strictly Positive Values of Capital Income

(a) Any Positive Income: Interest, Foreign, Total Gross, and Tax
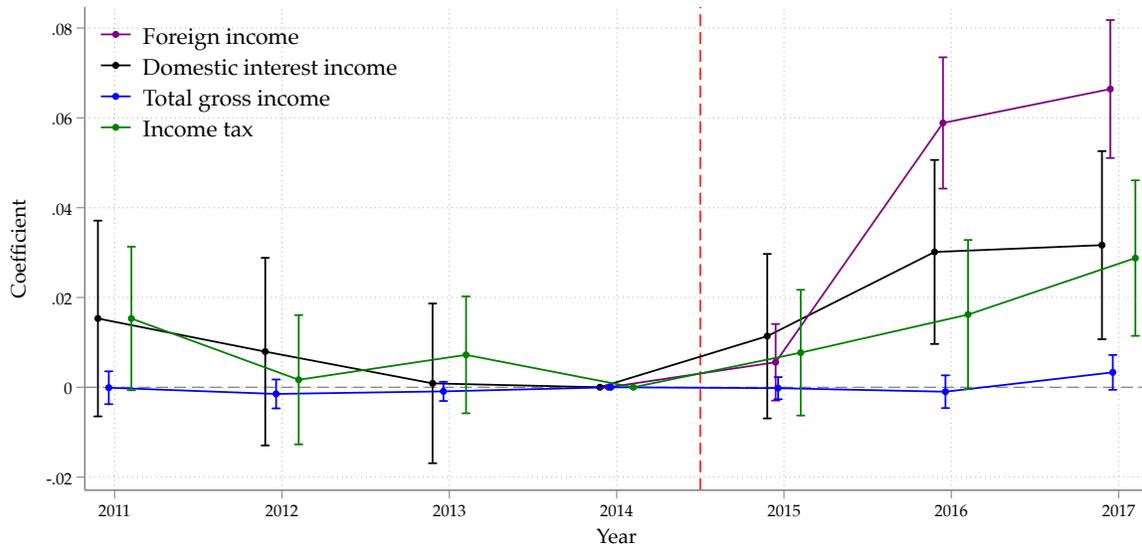
(b) Any Positive Irregular Income: Taxable and Tax
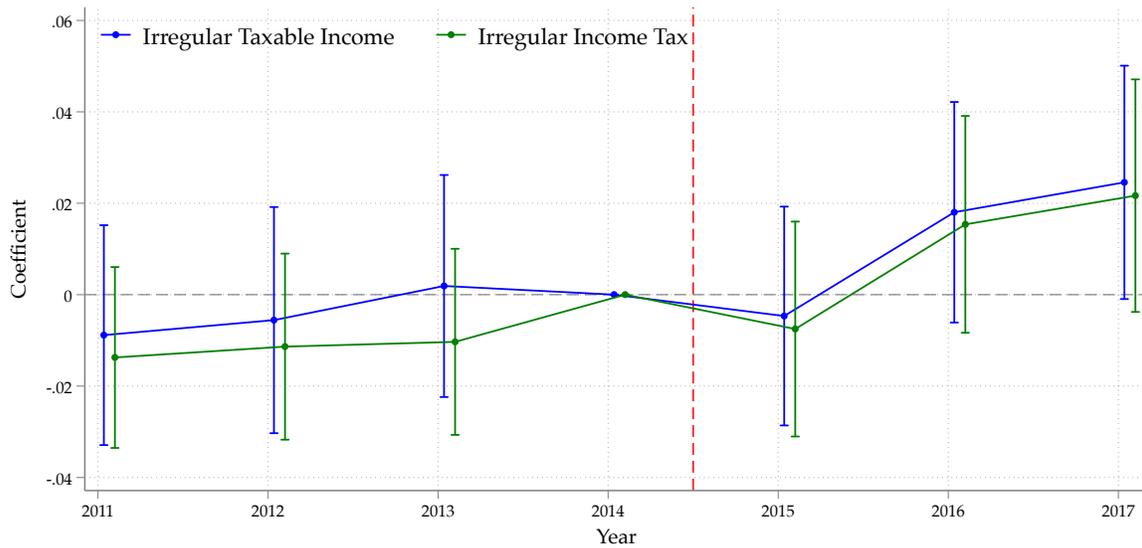
Notes: These figures present the effect of the tax incentives provided by the voluntary disclosure scheme on the probability of reporting positive values of selected regular income categories in Panel (a) and irregular income categories in Panel (b). The markers plot the event study coefficients from specification (1), in which the outcome is dichotomized and regressed on individual fixed effects and a voluntary discloser dummy interacted with year fixed effects (2014 is the omitted category). The vertical lines represent the associated 95% confidence intervals. Standard errors are clustered at the taxpayer level. The red dashed vertical line marks the period individuals disclosed hidden wealth. The corresponding difference-in-differences coefficients are presented in Table A.4, Panel B. The sample is a balanced panel of 44,958 individuals that (i) filed income taxes annually between 2011 and 2017; and (ii) filed the wealth tax in 2015, 2016, or 2017. Among these 44,958 individuals, 1,777 voluntarily disclosed in 2015, while 43,181 never disclosed under the scheme; those who first disclosed *after* 2015 are excluded from the estimation sample. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Figure A.9: Rise in the Effective Tax Rate is Especially Large at the Top
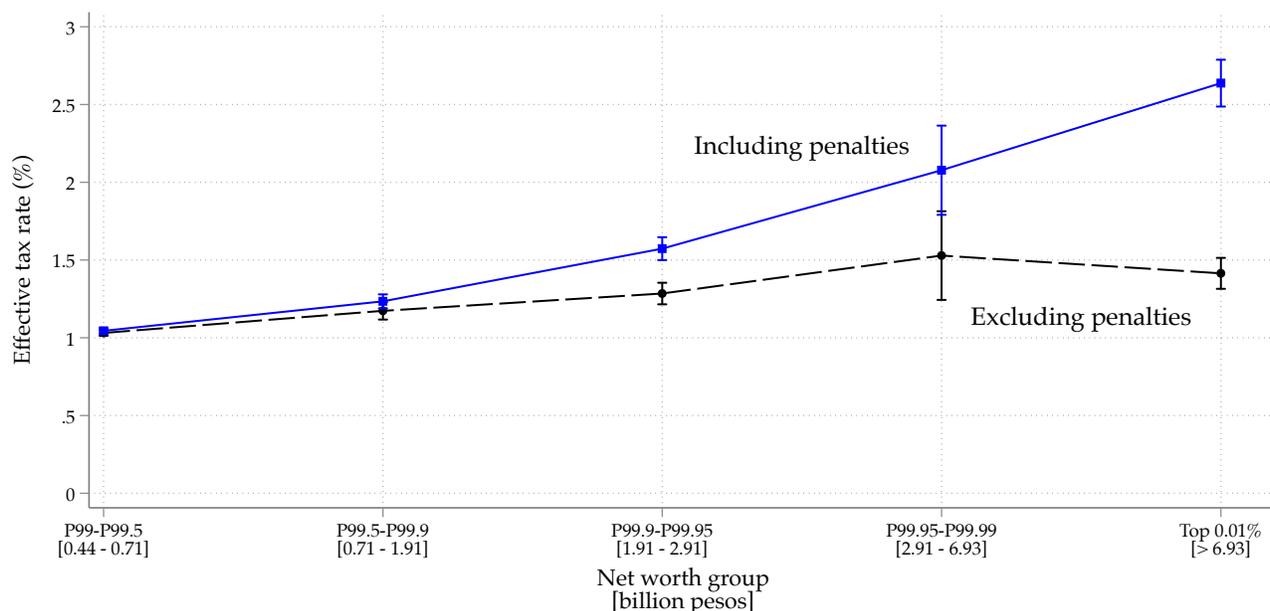
Notes: This figure illustrates how the voluntary disclosure scheme raised overall tax progressivity. The figure plots average personal income and wealth taxes paid on income and wealth in 2017, expressed as a share of net wealth for subgroups of individuals in the wealthiest 1% of the distribution. Individuals are ranked by their net wealth reported before the voluntary disclosure scheme (FY 2013) including any disclosures made under the scheme. The dashed black curve plots ipersonal ncome *and* wealth taxes in 2017 (FY 2016), while the solid blue curve adds the penalties associated with the disclosure program that year. 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 A.10: Probability of Participating in the Voluntary Disclosure Program, by Post-Disclosure Wealth and Year of First Disclosure
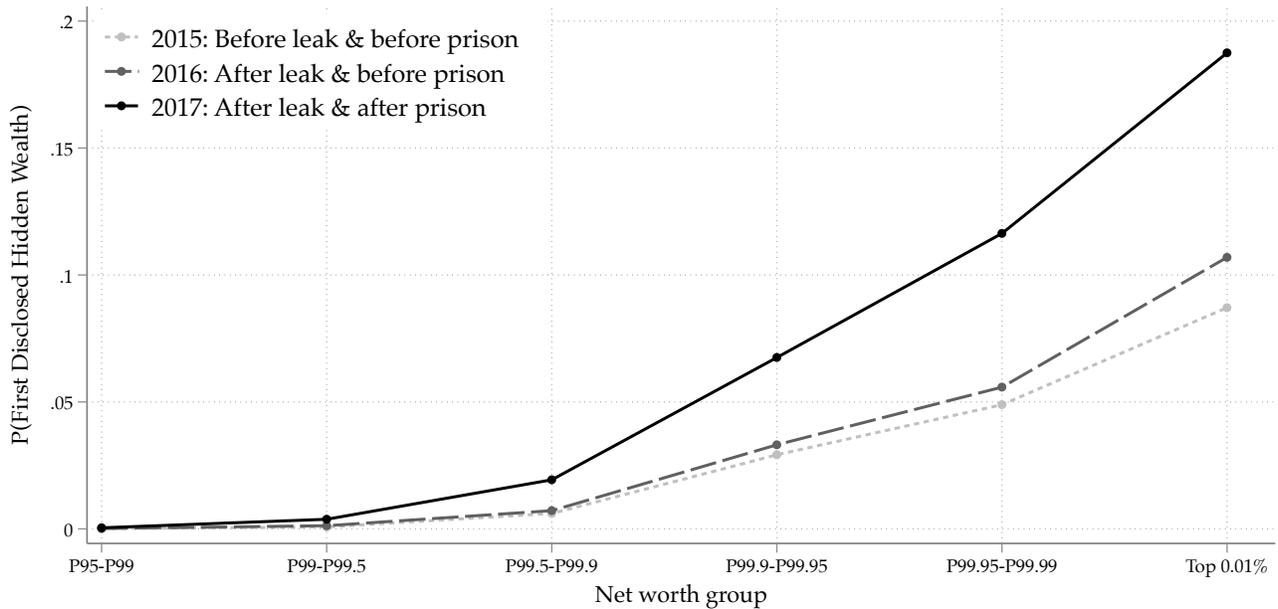
Notes: This figure plots the fraction of tax units in Colombia that file taxes and participate in the voluntary disclosure program, by bins of net worth (including disclosures) and the year in which they first disclosed hidden wealth. The figure shows that there is a spike in disclosures in 2017 relative to previous years. *Sources:* Authors' calculations using administrative tax microdata from DIAN.

Figure A.11: Wealth Inequality in Colombia Including Hidden Offshore Wealth

(a) Top 1 Percent

(b) Top 0.1 Percent
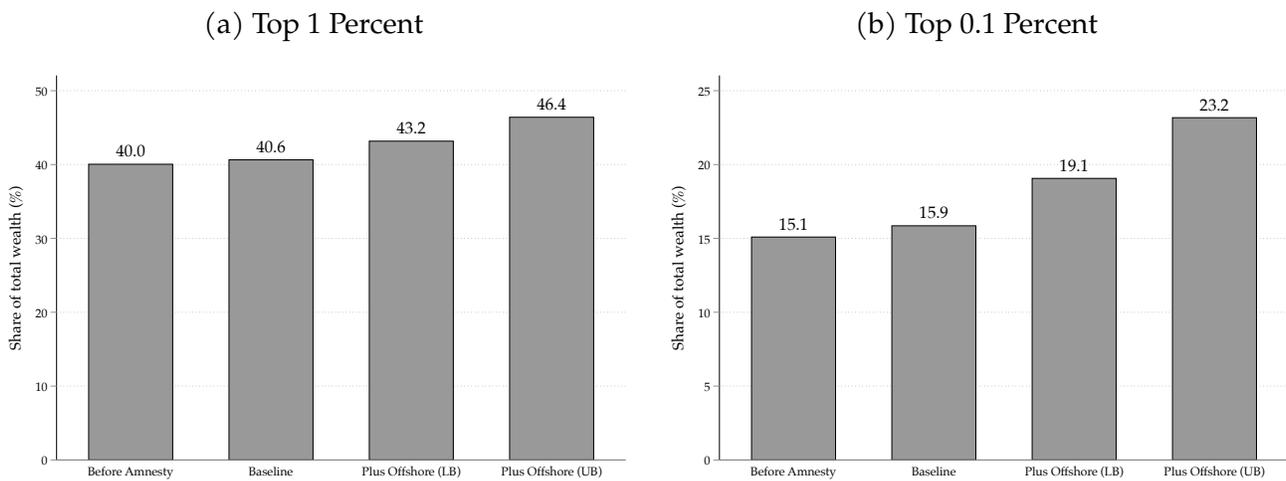
Notes: These figures presents estimates of top wealth shares in Colombia for 2017. The first bars report wealth inequality excluding offshore assets disclosed under the voluntary disclosure scheme. In the baseline estimates, the top 1% owns 40.6% of total wealth in Panel (a), and the top 0.1% owns 15.9% of total wealth in Panel (b). The third and fourth bars present estimates for top shares corrected by including unreported offshore wealth: using data from Alstadsater et al. (2018b) and the Panama Papers leak, the lower bound assumes unreported offshore wealth today represents 6.2% of GDP, while the upper bound assumes it represents 15%. See Appendix C. *Sources:* Table C.2.

Table A.1: Participation in the Panama Papers and Colombia's Voluntary Disclosure Scheme and Magnitude of Disclosures

	Percentile Cutoff		Probability of...		Magnitude of Disclosures under the Voluntary Disclosure Scheme											
					Share of all wealth				Share of disclosers' wealth				Share of total disclosed under scheme			
	COP (2017 billion)	PPP USD (2017 million)	Disclosing under scheme	Appearing in Panama Papers	Total wealth	Foreign assets	Domestic assets	Debts	Total wealth	Foreign assets	Domestic assets	Debts	Total wealth	Foreign assets	Domestic assets	Debts
	(1)	(2)	(%) (3)	(%) (4)	(%) (5)	(%) (6)	(%) (7)	(%) (8)	(%) (9)	(%) (10)	(%) (11)	(%) (12)	(%) (13)	(%) (14)	(%) (15)	(%) (16)
P95-P99	0.02	0.01	0.05	0.02	0.01	0.01	0.00	0.00	27.94	19.27	5.44	3.23	0.29	0.24	0.49	0.59
P99-P99.5	0.44	0.34	0.52	0.09	0.16	0.11	0.03	0.02	30.40	20.80	5.67	3.94	0.90	0.76	1.35	1.80
P99.5-P99.9	0.71	0.55	3.22	0.24	0.93	0.66	0.16	0.11	28.90	20.34	5.09	3.47	9.34	8.13	13.06	17.47
P99.9-P99.95	1.91	1.47	13.24	0.87	3.98	2.96	0.59	0.42	30.02	22.39	4.44	3.20	9.27	8.52	11.13	15.19
P99.95-P99.99	2.91	2.25	23.68	1.58	7.26	5.61	1.00	0.65	30.66	23.70	4.21	2.75	25.11	23.96	27.77	34.52
P99.99-P100	6.93	5.34	40.92	3.05	15.34	13.07	1.50	0.76	37.48	31.95	3.68	1.86	55.08	58.38	46.20	30.43

Notes: To rank individuals in the wealth distribution, wealth is computed at the micro level using wealth reported in the income tax statement in FY 2013—the last year before the disclosure program—but include any wealth disclosed under the scheme. Column 1 presents the minimum wealth thresholds for top groups in 2017 billion pesos. The market exchange rate for 1 million pesos is USD 335.13 in 12/31/2017. Column 2 translates this to PPP USD using the World Bank's PPP conversion factor for GDP (1,297.107 in 2017). Column 3 plots the likelihood of participating in Colombia's voluntary disclosure program in 2015, 2016, or 2017. Column 4 plots the likelihood of being named in the Panama Papers. Columns 5 through 16 present the size of disclosures, expressed as a share of total wealth (i.e., including zeros for non-disclosers), as a share of disclosers' wealth, and as a share of total disclosed under the scheme. Total wealth is decomposed into foreign assets, domestic assets, and debt. 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") and 1,085 were named in the Panama Papers. *Sources*: Authors' calculations using administrative tax microdata from DIAN and ICIJ.

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

| | Tax Haven | Type of Asset | | | | |
| | | Bank Deposits | Portfolio Securities | Trusts | Real Estate | Cars |
	(1)	(2)	(3)	(4)	(5)	(6)
Discloser	0.192***	0.018**	0.171***	0.029***	-0.081***	-0.016***
	(0.008)	(0.008)	(0.008)	(0.003)	(0.006)	(0.002)
Constant	0.337***	0.388***	0.415***	0.017***	0.183***	0.022***
	(0.005)	(0.005)	(0.006)	(0.001)	(0.004)	(0.002)
N	13740	13740	13740	13740	13740	13740
R^2	0.037	0	0.029	0.007	0.013	0.004

Notes: This table compares the location and type of foreign assets reported in 2017 between wealth tax filers who disclosed any foreign asset under Colombia's voluntary disclosure scheme versus not. Each column represents a separate regression with a different dependent variable obtained from the foreign asset information return (information tax return #160). 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 2017 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 any 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 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: Comparing Disclosers and Non-disclosers' Baseline Covariates

	Non-disclosers (1)	2015 Disclosers (2)	p-value (3)
Male	0.614	0.614	0.998
Born after 1985	0.000	0.000	
Capital rentier	0.275	0.325	0.000
Wage-earner	0.245	0.248	0.771
Other	0.479	0.427	0.000
In Panama Papers	0.004	0.019	0.000
Gross wealth (in million pesos)	2,865.517	4,736.594	0.059
Gross wealth (winsorized 0-99.9)	2,801.433	3,552.229	0.000
Gross wealth (winsorized 0-99)	2,704.481	3,209.146	0.000
Gross wealth (arcsinh)	22.192	22.171	0.377
Net wealth (in million pesos)	2,365.410	3,725.093	0.153
Net wealth (winsorized 0-99.9)	2,313.418	2,686.526	0.000
Net wealth (winsorized 0-99)	2,236.890	2,445.120	0.001
Net wealth (arcsinh)	22.013	21.834	0.000
Foreign income (in million pesos)	5.755	1.700	0.000
Foreign income (arcsinh)	0.330	0.457	0.090
Dividend income (in million pesos)	69.751	110.139	0.027
Dividend income (arcsinh)	7.514	9.080	0.000
Interest income (in million pesos)	19.197	47.410	0.000
Interest income (arcsinh)	11.869	13.691	0.000
Total gross income (in million pesos)	876.666	1,211.621	0.182
Total gross income (arcsinh)	20.109	20.146	0.369
Taxable income (in million pesos)	134.042	205.124	0.013
Taxable income (arcsinh)	18.782	18.996	0.000
Income tax (in million pesos)	29.784	52.981	0.014
Income tax (arcsinh)	15.319	15.331	0.933
Gross K gains and other income (in million pesos)	93.685	104.278	0.405
Gross K gains and other income (arcsinh)	4.591	5.226	0.002
Net K gains and other income (in million pesos)	30.276	36.078	0.357
Net K gains and other income (arcsinh)	2.665	3.083	0.012
Taxable K gains and other income (in million pesos)	22.603	30.106	0.212
Taxable K gains and other income (arcsinh)	2.323	2.567	0.112
K gains and other income taxes (in million pesos)	2.258	3.000	0.218
K gains and other income taxes (arcsinh)	1.815	2.127	0.018
N	43,181	1,777	

Notes: This table compares baseline covariates among 44,958 taxpayers, 1,777 of whom voluntarily disclosed hidden assets or fake debts in 2015 and 43,181 of whom never disclosed under the scheme. Columns 1 and 2 report the 2014 means for non-disclosers and 2015 disclosers, respectively. Column 3 reports the *p*-value of a *t*-test for the difference of means. The sample of taxpayers is the same as in Table 1 and Figure 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. Values are expressed in 2017 pesos, and 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 A.4: Robustness Checks: The Effect of a Voluntary Wealth Disclosure Program on Reported Income and Wealth

	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)
					Panel A. ArcSinh: no winsorizing							
δ	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)
$\bar{y}_c$	22.018	21.823	0.324	7.622	11.297	20.089	18.741	15.205	4.252	2.695	2.421	1.031
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
					Panel B. Dummy for strictly positive values							
δ	0	0.001	0.044***	-0.002	0.018***	0.001	0.001	0.012**	0.003	0.015**	0.016**	0.019***
	(0.000)	(0.001)	(0.005)	(0.008)	(0.007)	(0.001)	(0.002)	(0.005)	(0.008)	(0.007)	(0.007)	(0.006)
$\bar{y}_c$	1.000	0.999	0.0138	0.346	0.721	0.996	0.992	0.906	0.222	0.153	0.139	0.066
N	314,706	314,706	179,832	179,832	314,706	314,706	314,706	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	0.263	0.245	0.241	0.24
					Panel C. Levels (in million pesos)							
δ	2021.490***	1586.210***	8.298***	80.673***	5.844*	-60.53	24.642**	7.748**	-28.171*	0.986	2.324	-0.299
	(681.692)	(180.578)	(1.500)	(28.006)	(3.116)	(193.269)	(11.324)	(3.733)	(15.463)	(7.054)	(6.613)	(1.263)
$\bar{y}_c$	2489.714	2038.660	5.409	72.012	20.842	890.369	123.418	26.409	72.462	18.938	11.385	1.466
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	0.791	0.815	0.506	0.609	0.248	0.685	0.709	0.706	0.148	0.164	0.163	0.161
					Panel D. Levels: winsorizing at top 0.1% each year (in million pesos)							
δ	1326.350***	1413.511***	7.808***	42.251***	7.858***	43.151	15.431***	4.707***	0.777	8.043**	8.040**	0.727**
	(81.334)	(75.080)	(1.282)	(11.549)	(1.595)	(45.563)	(3.602)	(1.175)	(8.052)	(3.839)	(3.218)	(0.329)
$\bar{y}_c$	2432.978	1993.944	4.485	64.967	19.037	831.345	121.943	25.921	65.399	15.088	8.525	0.900
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	0.874	0.863	0.586	0.648	0.581	0.843	0.741	0.738	0.21	0.187	0.185	0.189
					Panel E. Levels: winsorizing at top 1% each year (in million pesos)							
δ	1116.518***	1176.559***	2.131***	21.910***	5.226***	41.387*	15.269***	4.642***	3.397	4.886***	4.592***	0.449***
	(54.471)	(49.964)	(0.235)	(5.390)	(0.940)	(22.634)	(2.613)	(0.843)	(4.675)	(1.468)	(1.101)	(0.109)
$\bar{y}_c$	2347.821	1925.880	0.543	53.249	16.750	732.548	116.873	24.245	53.210	8.978	5.599	0.438
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	0.874	0.866	0.592	0.667	0.63	0.845	0.772	0.768	0.217	0.202	0.202	0.208

Notes: This table reports how the difference-in-differences δ coefficient from Table 1—reproduced in Panel A—changes under different functional forms for the outcome variables. The functional form of the outcome variables is binary (equalling 1 for strictly positive values) in Panel B and linear 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. $\bar{y}_c$ represents the mean of the outcome variable for non-disclosers across all pre-event years. $*p < 0.1, **p < 0.05, ***p < 0.01$. *Source:* Authors' calculations using administrative tax microdata from DIAN.

Table A.5: 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 tax breaks 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. δ represents the difference-in-differences coefficient from specification (2), in which the inverse hyperbolic sine transformation of the outcome is regressed on individual fixed effects, year fixed effects, and an interaction of the voluntary discloser dummy and post-reform years (2015 is the omitted category). The standard errors in parentheses are clustered at the taxpayer level. The sample is a balanced panel of 45,255 individuals that (i) filed income taxes annually between 2011 and 2017; and (ii) filed the wealth tax in 2015, 2016, or 2017. Among these 45,255 individuals, 2,074 voluntarily disclosed in 2016 for the first time, while 43,181 never disclosed under the scheme; those who first disclosed 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 variables 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.

B   Merging the leaked Panama Papers microdata with individual tax returns

The Panana Papers microdata was last downloaded on January 4, 2017. It included information from 1,751 shareholders of offshore entities with a personal or entity contact address in Colombia. Using personal names, we are able to match 1,208 individuals to their tax records, i.e., a match rate of almost 70%. This high match rate is partly thanks to the naming custom of Hispanic America—which involves two given name, plus a paternal surname, followed by a maternal surname—often practiced in Colombia. For instance, in the (imaginary) case of a person named Juliana María Londoño Vélez, Juliana would be the first given name, María the second given name, Londoño the paternal surname, and Vélez the maternal surname.

There are four main reasons why 543 (=1,751-1,208) individuals named in the Panama Papers did not match with a personal tax return:

First, incomplete name information, common names, and homonyms affected the tax authority's ability to identify an individual in the tax returns. 68 cases were either homonyms or exact duplicates, so we drop 34 cases, leaving 509 unmatched cases. Of these, only 30% had two given names and two surnames (which is half the share among matched cases), 30% had only one given name and one surname (e.g., Juliana Londoño), 15% had two first names but only one surname (e.g., Juliana María Londoño), and 25% had two surnames but only one given name (e.g., Juliana Londoño Vélez). Moreover, even when there was a unique full name in the Panama Papers, i.e., two given names and two surnames, it may match with more than one taxpayer if that name is common among the population of tax filers. In some cases, the tax authority succeeded to identify the correct individual among the homonyms. Yet, in most cases, it did not match with any of the alternatives, considering the cost of identifying and contacting the wrong individual too burdensome.

Figure B.1 plots the likelihood of being identified in the tax records as a function of the names available in the Panama Papers. While people appearing with two given names and two surnames have an 82% chance of being identified in the tax returns, this likelihood drops to 24% for those with one given name and one surname or two given names but only one surname.

Second, an individual in the Panama Papers may be legally exempt from filing taxes in Colombia. Such is the case of minors who appear in the Panama Papers as the ultimate beneficiaries of trusts but who, by virtue of being aged below 14, are not required to file taxes in Colombia (their parents file for them). Indeed, there is a considerable number of families named in the Panama Papers as a result of turning to Mossack Fonseca to create a private offshore family trust.[13] Further, some individuals linked to Colombia by ICIJ may not file taxes in Colombia if they are not residents for tax purposes. For instance, a Panamanian intermediary—a middleman that asked Mossack Fonseca to create an offshore firm for a Colombian client—will not be required to file taxes in Colombia. We observe 142 of 1,717 individuals (=1,751-34 duplicates) with a foreign name (e.g., John Smith), and the likelihood of being matched to the Colombian tax returns is 28.3 percentage points lower for these individuals (the difference is highly significant).

Third, an extreme tax evader may have not filed taxes in Colombia between 1993 and 2015, thus passing completely undetected to the tax authority for more than two decades. However, the requirement for filing income taxes is based not only on gross income received, but also on assets owned, credit card purchases, bank deposits, and financial investments, making the probability of extreme evasion due to non-filing less likely.

Fourth, it is possible that some cases may have been incorrectly linked to Colombia by ICIJ. ICIJ performed automatic identification of country names within all the addresses using geocoding tools and manual reviews. Once a country was determined, the connected record—an entity, an officer or

---

[13]27% of 1,717 (=1,751-34 duplicates) individuals uniquely named in the Panama Papers exactly share their two surnames with others, as occurs with siblings.

an intermediary—was linked to that country. Admittedly, however, this identification process has some limitations, including potential data-entry and country-matching errors. Lastly, errors with the Colombian tax authority's information system might also explain why it could not match some of the individuals named in the Panama Papers.

Figure B.1: Match rate by name available in the Panama Papers
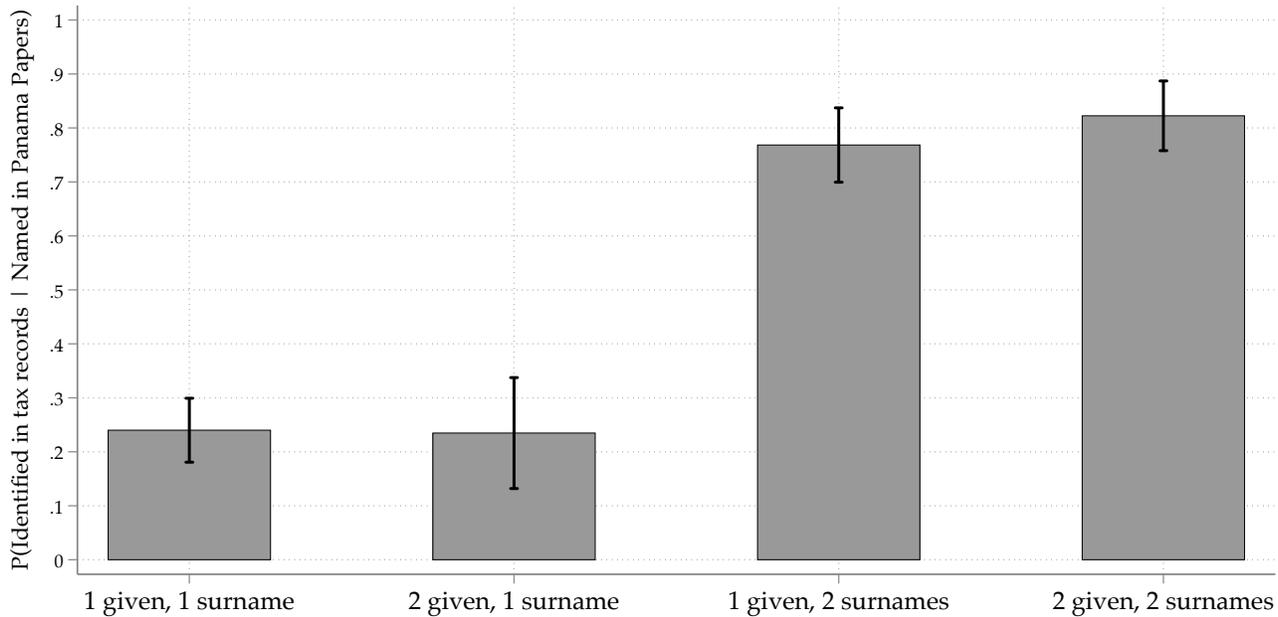
Notes: This figure plots the likelihood of being identified in the tax records as a function of the information about names available in the leaked Panama Papers microdata. While people who have two given names and two surnames in the Panama Papers have an 82% chance of being identified in the tax returns, this likelihood is only 24% for individuals with two given names but only one surname. Overall, 70% of the 1,717 unique individuals named in the Panama Papers linked to Colombia were identified in the tax returns. *Source:* Authors' calculations using administrative data from DIAN and ICIJ.

C Measuring Wealth Inequality in Colombia

Measuring top wealth shares (e.g., the fraction of total wealth held by the top 1%) 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.

C.1 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. 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.[14] 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 our 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% of individuals own 71% of all wealth reported in the survey. This is very close to the equivalent top share of 75.3% in the U.S., based on data from the Survey of Consumer Finances from 2013 (Saez and Zucman, 2016). Moreover, the wealthiest 1% in Colombia own 25.8% of total wealth according to IEFIC survey, which is significantly less than the U.S.' 35.8% estimate for households using the Survey of Consumer Finances. Figure C.1 plots wealth decomposition by net wealth groups. The figure shows that 50% 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% of wealth. Finally, for individuals in the top 1%, the share of wealth belonging to real estate falls to 80.7% while the shares of financial and business assets increase. However, it is clear from Figure C.1 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 (see https://www.forbes.com/forbes-400/), 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).

---

[14]In fact, the sum of survey weights add up to 6,719,291, i.e., 21% of all tax units.

Figure C.1: Wealth Decomposition in Survey Data
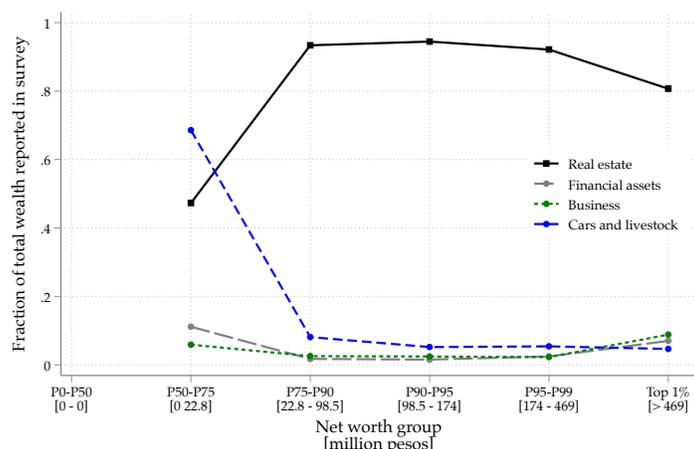
Notes: This figure plots wealth decomposition by asset types across wealth groups using household survey data from 2017. Individuals aged 20 and above are included. For assets reported at the family level (real estate, business assets, vehicles, and livestock—and the outstanding debt of each asset), 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 DANE, "Encuesta de Carga Financiera y Educación Financiera de los Hogares, IEFIC," 2017. Accessed 2018-07-29. http://microdatos.dane.gov.co/index.php/catalog/626/get_microdata.

Table C.1: Net wealth groups in survey data

Fractile	Min (million pesos)	Mean (million pesos)	Mean P_0–P_1 (million pesos)	Share of total wealth (%)	Sum of survey weights
(1)	(2)	(3)	(4)	(5)	(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 DANE, "Encuesta de Carga Financiera y Educación Financiera de los Hogares, IEFIC," 2017. Accessed 2018-07-29. http://microdatos.dane.gov.co/index.php/catalog/626/get_microdata.

XX

C.2 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 can be 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^*:

$$W^* = [K \cdot (1 - \alpha) + \alpha \cdot K \cdot \delta] - L \tag{4}$$

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 (4) 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. 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% of market values in Bogota, and 60% 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) 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.

C.3 Unreported Offshore Wealth

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. (2018b). 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. (2018b) estimate that total offshore wealth by Colombians is 9.0% of GDP. This places Colombia just below the world average of 9.8% 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. This means 6.2% of GDP remains concealed offshore.

Who holds this offshore wealth? Following Alstadsater et al. (2019), 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 C.2 shows the total amount of offshore assets disclosed by wealth group, ranking individuals by their pre- and post-disclosure wealth. The figure shows that offshore wealth is extremely concentrated at the top: 99% of disclosed offshore wealth is owned by the wealthiest 1% and 58% is owned by the wealthiest 0.01%. 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.

Figure C.2: Distribution of Hidden Offshore Assets in 2015–2017, by Pre/Post-Disclosure Top Wealth Group
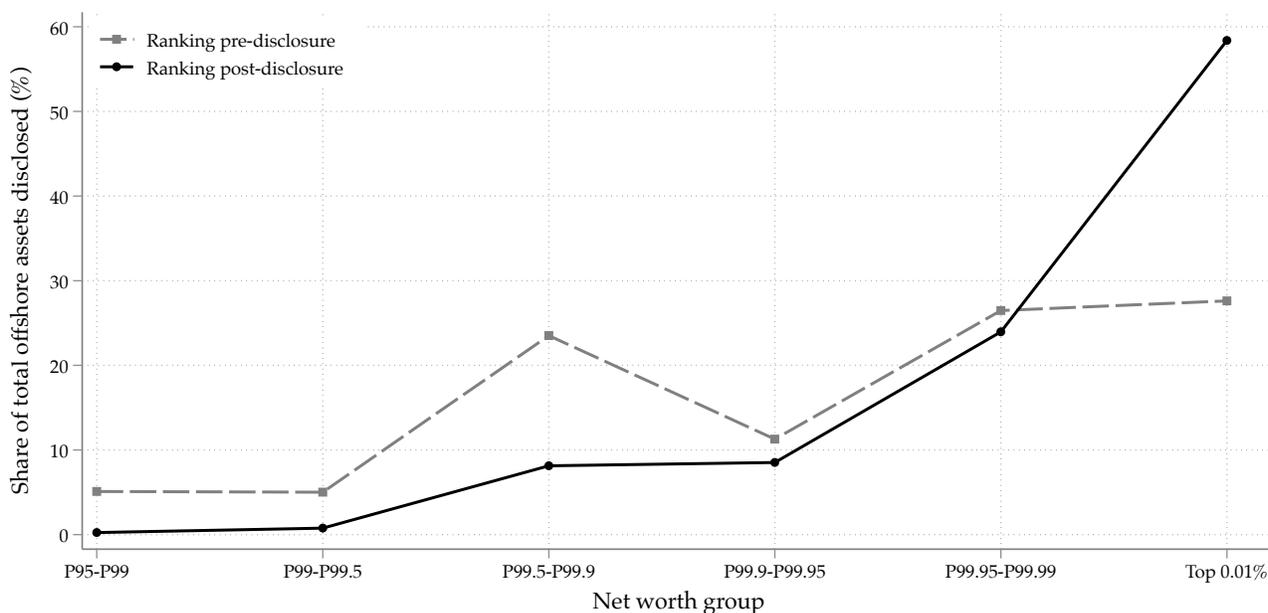
Notes: This figure shows the fraction of total disclosures of hidden offshore assets during the 2015–2017 voluntary disclosure program for each wealth group, ranking by pre- and post-disclosure net worth. The figure shows that the volume of offshore assets disclosed in 2015–2017 is increasing in net worth. Tax filers in the wealthiest 0.01% post-disclosure disclosed 58% 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:* Table A.1 for the series ranking post-disclosure; otherwise, authors' calculations using administrative tax microdata from DIAN.

Note, however, that the estimates from Alstadsater et al. (2018b) 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, 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% increase. If the increase in

the use of offshore structures also reflects rises in assets held offshore, then offshore wealth today could reach 40% of GDP ($40\% = (1 + 3.45) \times 9\%$) and 37.25% of GDP would be unreported to the tax agency.

Instead, we assume that the increase in unreported offshore wealth is roughly one-half the rise in the stock of offshore entities created by Mossack Fonseca between 2007 and 2015, such that unreported offshore wealth represents 15% of GDP. This would place offshore wealth owned by Colombians—both reported and unreported, expressed as share of GDP—above the equivalent shares owned by Americans, Frenchmen, and Germans.

The results of this exercise on wealth inequality measures are presented in Table C.2.

Table C.2: Top Wealth Shares in Colombia Using Tax and Survey Data, Including and Excluding Hidden Offshore Wealth

	Top 5%	Top 1%	Top 0.5%	Top 0.1%	Top 0.05%	Top 0.01%	Top 5% to 1%	Top 1% to 0.5%	Top 0.5% to 0.1%	Top 0.1% to 0.05%	Top 0.05% to 0.01%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(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 Disclosure Scheme	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. (2018b) and the Panama Papers leak, the lower bound assumes unreported offshore wealth today represents 6.2% of GDP, while the upper bound assumes it represents 15%. The last row augments unreported wealth by the observed increased in the stock of entities ever incorporated through Mossack Fonseca between 2007 and 2015. *Sources:* Authors' calculations using administrative tax microdata from DIAN and IEFIC from DANE.