Appendix

Let the Rebels Rule?
Evidence on the Economic Effects of Rebel Governance in Colombia

Author: Santiago Pérez-Cardona

Table of Contents

A Data Appendix

This appendix provides detailed information on data sources, sample availability and other details for the different variables used in the paper.

Agricultural Product Prices

Information on agricultural product prices was provided by the Center for Production and Sectorial Trade Studies (*Centro de Estudios sobre Producción y Comercio Sectorial* - CEPCO) at the Central Bank of Colombia. CEPCO gathered data on crops prices from two primary sources: i) the Agricultural Price Information System (*Sistema de Información de Precios de Sector Agropecuario* - SIPSA) from the National Department of Statistics (*Departamento Adminsitrativo Nacional de Estadística* - DANE); ii) official record from producers associations (*agremiaciones*). To compute each price per kilogram, CEPCO takes the average price for each crop across 2013 and wholesale markets (*mercados mayoristas*).

Three aspects of the prices data are worth-noting. First, price information is only available for 86 crop types from 122 in the CNA. Nonetheless, these 86 crops represent 79.7% of the total sown area in the CNA. Second, the prices available include transportation costs and intermediation margins, which might exaggerate agricultural revenues for the local producer. Third, for sugar cane, crude palm oil and cottonseed production, CEPCO applies conversion factors to the final product price (i.e., sugar, oil, cotton) to retrieve a better estimate of the producers income. See de Roux (2020); de Roux et al. (2019) for previous work using this data.

Rural Census (2014)

Most of the information I use comes from Colombia's National Rural Census. The CNA covers the entire rural population in Colombia, and its information corresponds to the year 2013. Moreover, I access restrictive information of each estate's exact GPS location to compute the distance to the DMZ. I now explain in detail each variable relevant to my analysis.

Individuals and Dwelling Characteristics

- *Distance to the DMZ's Border*: Define as the geodesic distance from the farms GPS location to the nearest point of the DMZ's Border.
- *Literacy (=1)*: Indicator that equals one if the individual reports knowing how to read and write.
- *Years of Education*: I construct this variable using each individual's highest grade achieved. I then multiply each grade by the average number of years one has to study to complete it.
- *Health Insurance (=1)*: Indicator that equals one if the individual reports having health insurance.

- *Concrete Walls (=1)*: Indicator that equals one if the dwelling's walls materials are in concrete or better. Better materials include bricks, stones, or polished wood. Worse materials include, for example, cartons, raw lumber, *bahareque*, canned waste, or no walls.

- *Concrete Floor (=1)*: Indicator that equals one if the dwelling's floors materials are in concrete or better. Better materials include tile, bricks, marble, or polished wood. Worse materials include raw lumber, ground, or sand.

- *Electricity (=1)*: Indicator that equals one if the dwelling owner reports having electricity.

- *Sewerage (=1)*: Indicator that equals one if the dwelling owner reports having access to a sewerage system.

- *Aqueduct (=1)*: Indicator that equals one if the dwelling owner reports having access to a aqueduct system.

- *Forced Displacement (=1)*: Indicator that equals one if the household head reports that at least one individual within the family was a victim of forced displacement. Forced displacement refers to an involuntary or coerced migration of a person or people away from their home as a result of violence.

- *Land Dispossession (=1)*: Indicator that equals one if the household head reports that at least one individual within the family was a victim of land dispossession. Forced displacement refers to an involuntary transfer of right over a estate as a result of violence.

- *Land Abandonment (=1)*: Indicator that equals one if the household head reports that at least one individual within the family was a victim of land abandonment. Land abandonment refers to an involuntary or coerced migration of a person or people away from their estate.

Agricultural Production

- *Total Revenue*: Since CNA do not provide information on producers revenues, I have to compute them using information on the total production quantity and average prices per crop. To do so, I compute the total amount produced of each crop by farm and multiply it by price provided by CEPCO. This leaves me with the total revenue per crop at the crop-farm level. I further clean this variable winsorizing at the 1% and 99% levels.

- *Revenue Per Hectare (Yield)*: I then divide the revenue per crop at the crop-farm level by the total number of hectares each producer devotes to each crop. This leaves me with the revenue per hectare at the crop-farm level. I further clean this variable winsorizing at the 1% and 99% levels.

- *Subsistence (=1)*: Indicator that equals one if the farms agricultural production is used only for self-consumption.
- *Sale (=1)*: Indicator that equals one if the farms agricultural production is used only for market sale.
- *Cash Crops*: Indicator that equals one for crops that require centralized processing to be valuable, and it cannot be directly consumed by an individual worker. I categorize the following crops as cash crops: African oil palm, rice, yellow corn, white corn, cocoa, sugar cane, rubber tree, and coffee. I then compute the area share devoted to cash crops and the revenue share they yield to each producer.
- *Perennial Crops*: Indicator that equals one for crops that don't need to be replanted each year, as they automatically grow back after each harvest. I follow DANE's official classification of perennial crops. I then compute the area share devoted to perennial crops and the revenue share they yield to each producer.
- *Transitory Crops*: Indicator that equals one for crops that need to be replanted each year, as they do not grow back after each harvest. I follow DANE's official classification of transitory crops. I then compute the area share devoted to transitory crops and the revenue share they yield to each producer.
- *Private Property (=1)*: Indicator that equals one if the farm is privately own.
- *Collective Property (=1)*: Indicator that equals one if the farm is collectively own.
- *Irrigation System (=1)*: Indicator that equals one if the farm has an irrigation system for agricultural production.
- *Fertilizer Use (=1)*: Indicator that equals one if the producer uses fertilizer in the agricultural production process.
- *Pest Control (=1)*: Indicator that equals one if the producer uses pest control in the agricultural production process.
- *Machinery (=1)*: Indicator that equals one if the farm has machinery for agricultural production.
- *Buildings (=1)*: Indicator that equals one if the farm has buildings and infrastructure for agricultural production.
- *Producers Association (=1)*: Indicator that equals one if the producer belongs to a producers association.
- *Cooperativa (=1)*: Indicator that equals one if the producer belongs to a *cooperativa*. *Cooperativa* typically refers to a association of small producers.
- *Colective Work (=1)*: Indicator that equals one if the farms uses collective work within the production process.

Geographic Information

To study geographic characteristics at each side of the DMZ's boundary, I use grid of $10km$ by $10km$ cells with the average value of each characteristic within the cell. I use this size of cells, because the precision of most information varies at this level.

- *Elevation*: Average elevation comes from WorldClim, a online database of high spatial resolution global weather and climate data. I use elevation data with a resolution of 30 second resolution (about 1 km cells at the equator).

- *Rainfall*: Average rainfall comes from WorldClim, a online database of high spatial resolution global weather and climate data. I use rainfall data with a resolution of 2.5 minutes resolution (about 4.5 km cells at the equator). Specifically, I usec compute the rainfall's mean and standard deviation between 1980 and 1989.

- *Land Suitability*: This information comes from The Atlas of the Biosphere at the University of Wisconsin's Nelson Institute Center for Sustainability and the Global Environment (SAGE). The dataset represents the fraction of each grid cell that is suitable to be used for agriculture. It is based on the temperature and soil conditions of each grid cell, in a resolution of $10km$ by $10km$.

- *Cropland (1992)*: This information comes from The Atlas of the Biosphere at the University of Wisconsin's Nelson Institute Center for Sustainability and the Global Environment (SAGE). The dataset represents the fraction of each gridcell's surface that is covered by crops. It is based on the temperature and soil conditions of each grid cell, in a resolution of $10km$ by $10km$.

- *River (=1)*: This information comes from official information of Colombia's National Mapping Institute - IGAC. I construct a indicator variables that takes the value of one of there is a river within each cell.

Fiscal Outcomes

To study differential investment by the government after the DMZ as an alternative explanation, I use fiscal information from Colombia's National Planning Department (*Departamento Nacional de Planeación* - DNP):

- *Total SGP*: Measured in millions of 2010 COP. This variable measures each municipality-year total transfers from the central government to the municipal government (i.e., *Sistema General de Participaciones*- SGP). Information is available for all municipalities between 1994 and 2008, based on the data in the provided yearly by DNP.

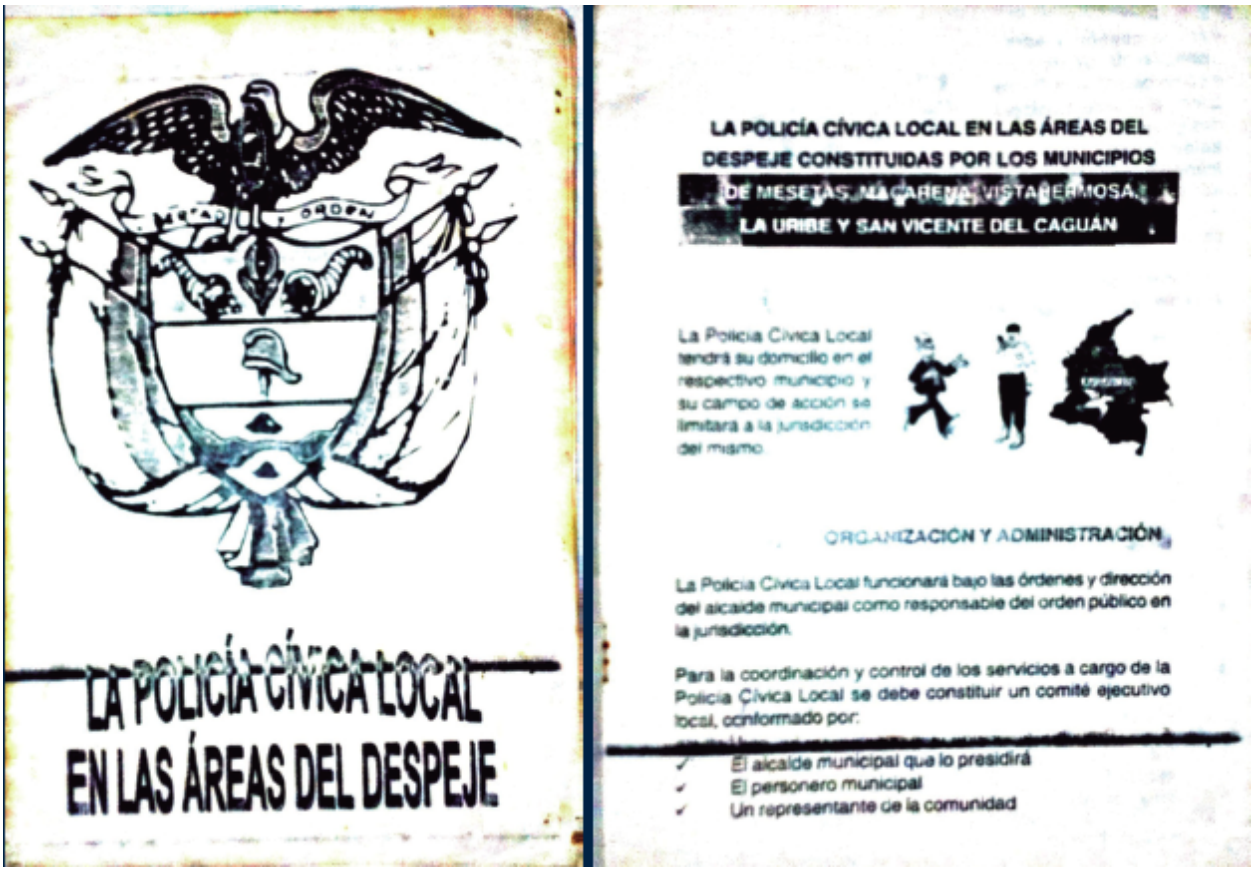
- *Municipal Investment*: Measured in millions of 2010 COP. This variable measures each municipality-year capital expenses (i.e., investment). Information is available for all municipalities between 1994 and 2008, based on the data in the municipal balance sheets provided yearly by DNP.

B Additional Figures and Tables

Figure B1: President-elect Pastrana meets Manuel Marulanda in 1998

Notes: Obtained from El Tiempo (2020). From left to right: Victor Julio Suárez (a high-ranking member of the FARC), Andrés Pastrana (president-elect), Manuel Marulanda (FARC leader), and Víctor Ricardo (high-ranking aide to the president-elect).

Figure B2: *El Caguan's civic* police pamphlet

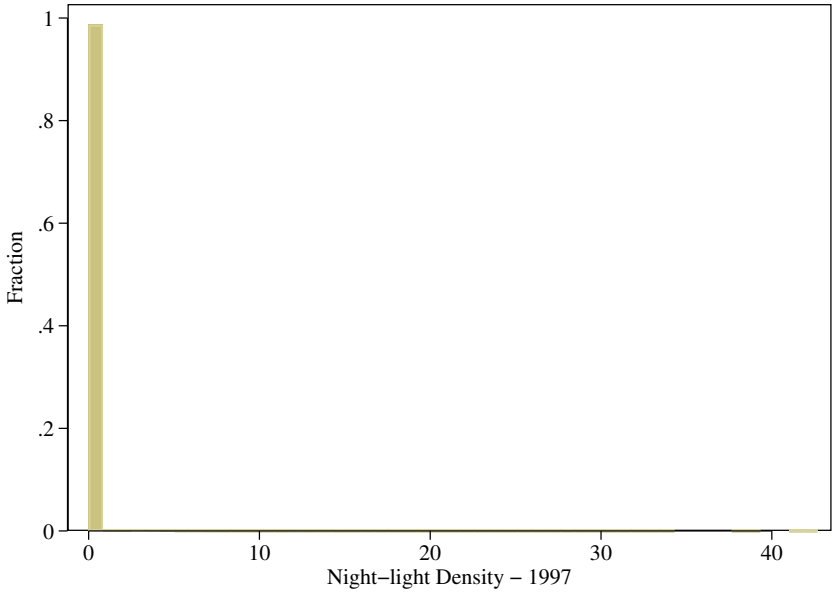


Notes: XXXXX Source Reyes (2012).

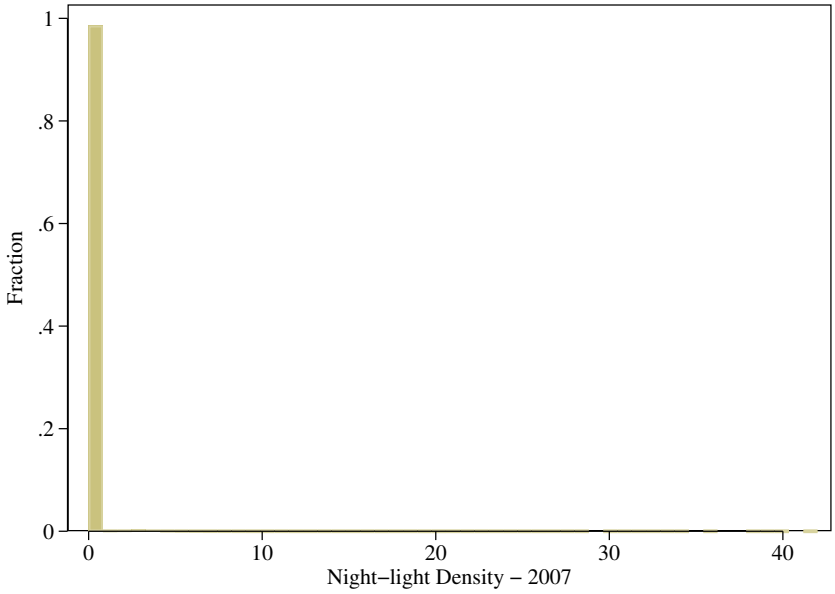
Figure B3: San Vicente's Police station in 2012

Notes: Source Reyes (2012). Sandbags surround San Vicente's police station in 2012 as a defense mechanism from FARC attacks.

Figure B4: Night-light Density



(a) 1997



(b) 2007

Notes: These figures show the nigh-light density distribution of 1km 1 1km cells within 100 km of the DMZ's border.

Table B1: Armed Actors Presence in Vistahermosa - ELCA

	Presence of Armed Actors (= 1)									
	2001	2002	2003	2004	2005	2006	2007	2009	2009	2010
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Vistahermosa (=1)	0.61**	0.44*	0.19	0.24	0.29	0.29	0.30	0.10	0.18	0.41**
$\Pr(\lvert\beta^{RI:Vistahermosa}\rvert > \lvert\beta^{Non-Affected}\rvert)$	[0.014]	[0.066]	[0.562]	[0.594]	[0.142]	[0.156]	[0.142]	[0.990]	[0.218]	[0.044]
DV. Mean Non-Exposed	0.38	0.30	0.30	0.25	0.21	0.21	0.19	0.14	0.065	0.81

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01. Unit of observations are ELCA's communities (i.e., most likely neighborhoods). I compare theecha variable's mean of communities in Vistahermosa, which was part of the DMZ, with other communities in the oriental region of Colombia. In squared brackets, I present randomization inference p-values with 500 replications.

Table B2: Multiple Armed Actors Presence in Vistahermosa - ELCA

	More than One Armed Actors (= 1)									
	2001	2002	2003	2004	2005	2006	2007	2009	2009	2010
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Vistahermosa (=1)	-0.16	-0.24	-0.63**	-0.56*	0.03	0.03	0.00	-0.55	-0.50	-0.10
$\Pr(\lvert\beta^{RI:Vistahermosa}\rvert > \lvert\beta^{Non-Affected}\rvert)$	[0.636]	[0.562]	[0.032]	[0.088]	[0.576]	[0.546]	[0.990]	[0.364]	[0.206]	[0.294]
DV. Mean Non-Exposed	0.66	0.57	0.63	0.56	0.46	0.46	0.60	0.55	0.050	0.40

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01. Unit of observations are ELCA's communities (i.e., most likely neighborhoods). I compare theecha variable's mean of communities in Vistahermosa, which was part of the DMZ, with other communities in the oriental region of Colombia. In squared brackets, I present randomization inference p-values with 500 replications.

Table B3: Armed Conflict Events in Vistahermosa - ELCA

	Migration (=1)	Forced Recruitment (=1)	Armed Actors Overpower (=1)	Safe (=1)
	(1)	(2)	(3)	(4)
Vistahermosa (=1)	0.48	0.52	0.66***	-0.08
$\Pr(\lvert\beta^{RI:Vistahermosa}\rvert > \lvert\beta^{Non-Affected}\rvert)$	[0.228]	[0.152]	[0.000]	[0.990]
DV. Mean Non-Exposed	0.52	0.22	0.33	0.83

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01. Unit of observations are ELCA's communities (i.e., most likely neighborhoods). I compare eacha variable's mean of communities in Vistahermosa, which was part of the DMZ, with other communities in the oriental region of Colombia. In squared brackets, I present randomization inference p-values with 500 replications.

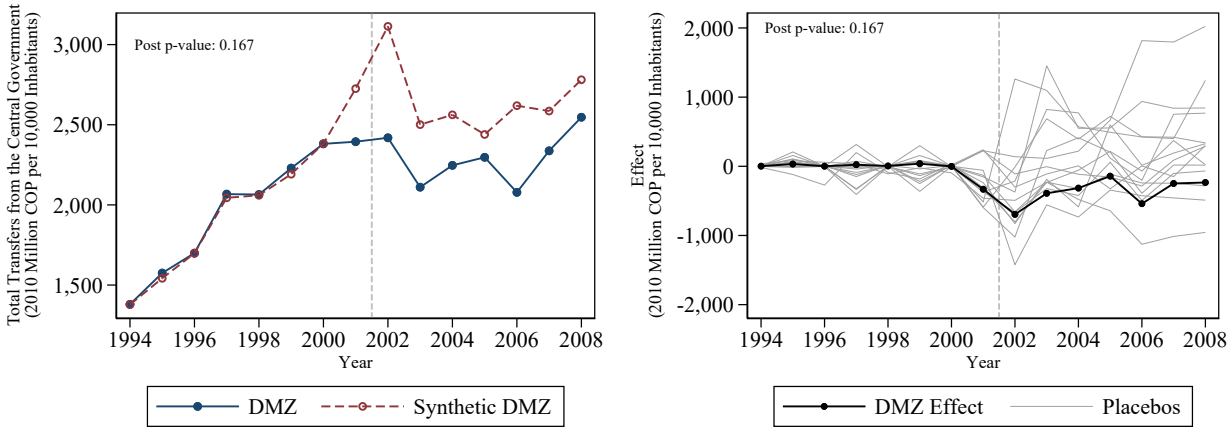
Table B4: Armed Actors Rule in Vistahermosa - ELCA

	Violent Attacks (=1)	Forced Displacement (=1)	Forced Rule of Law (=1)	Taxation (=1)
	(1)	(2)	(3)	(4)
Vistahermosa (=1)	-0.30	0.13	-0.20	-0.34
$\Pr(\lvert\beta^{RI:Vistahermosa}\rvert > \lvert\beta^{Non-Affected}\rvert)$	[0.564]	[0.624]	[0.600]	[0.578]
DV. Mean Non-Exposed	0.55	0.37	0.70	0.40

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01. Unit of observations are ELCA's communities (i.e., most likely neighborhoods). I compare theecha variable's mean of communities in Vistahermosa, which was part of the DMZ, with other communities in the oriental region of Colombia. In squared brackets, I present randomization inference p-values with 500 replications.

C Synthetic Control Analysis

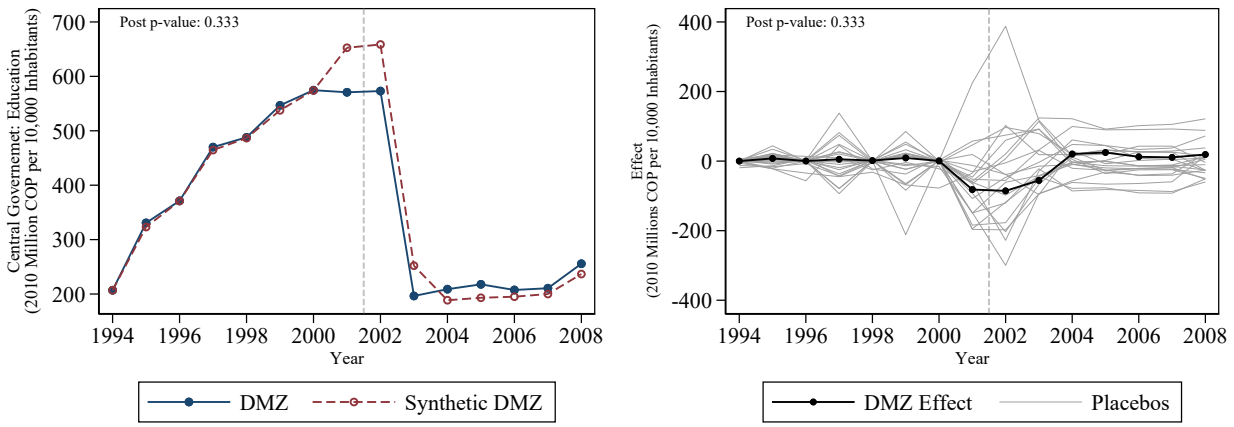
Figure C1: Synthetic Control Results: Total SGP



(a) Synthetic Control **(b)** Placebos

Notes: Panels (a) shows observed values of fiscal expenditure between 1995 and 2008 (solid line) and counterfactuals from the synthetic control (dashed line). The outcome variable is the total value of 2010 COP transferred from the central government to the municipalities as part of the *Sistema General de Participaciones* (SGP). To feet the synthetic control I use the dependent variable in 1994, 1996, 1998 and 2000. I collapsed DMZ municipalities into one single entity, restricted sample to municpalities in the departments of Meta, Caqueta and Guaviare, and dropped each department's capital. The "Post p-value" follows Abadie (2021) and refers to the proportions of placebo standardized effects that are at least as large as the main standardized effect for the treated unit.

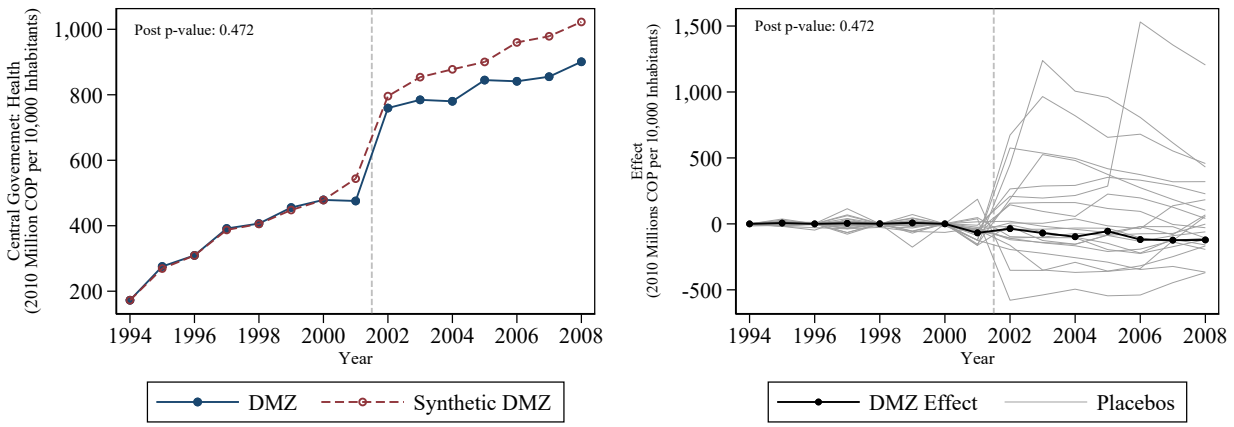
Figure C2: Synthetic Control Results: SGP Education



(a) Synthetic Control **(b)** Placebos

Notes: Panels (a) shows observed values of fiscal expenditure between 1995 and 2008 (solid line) and counterfactuals from the synthetic control (dashed line). The outcome variable is the total value of 2010 COP transferred from the central government to the municipalities for the educational system as part of the *Sistema General de Participaciones* (SGP). To feet the synthetic control I use the dependent variable in 1994, 1996, 1998 and 2000. I collapsed DMZ municipalities into one single entity, restricted sample to municpalities in the departments of Meta, Caqueta and Guaviare, and dropped each department's capital. The "Post p-value" follows Abadie (2021) and refers to the proportions of placebo standardized effects that are at least as large as the main standardized effect for the treated unit.

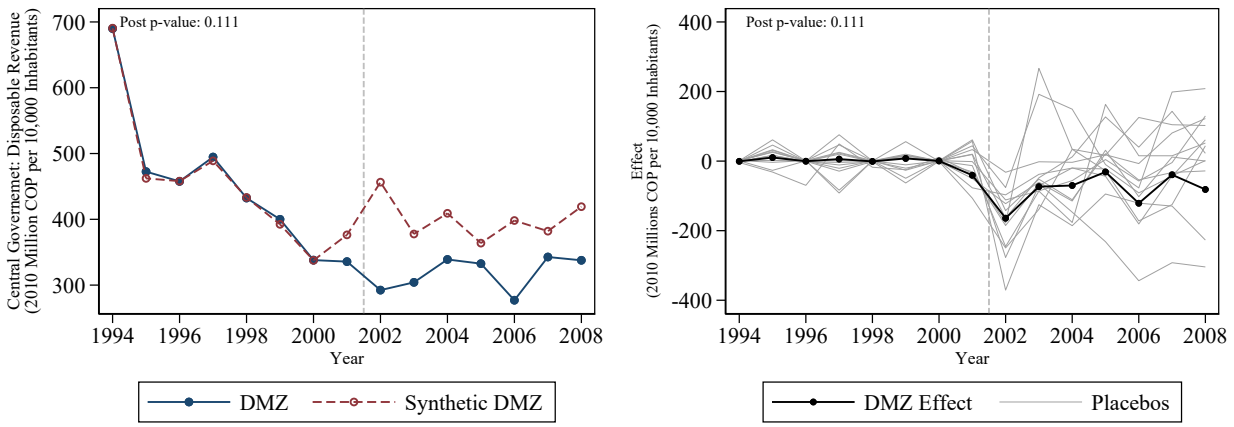
Figure C3: Synthetic Control Results: SGP Health



(a) Synthetic Control **(b)** Placebos

Notes: Panels (a) shows observed values of fiscal expenditure between 1995 and 2008 (solid line) and counterfactuals from the synthetic control (dashed line). The outcome variable is the total value of 2010 COP transferred from the central government to the municipalities for the health system as part of the *Sistema General de Participaciones* (SGP). To feet the synthetic control I use the dependent variable in 1994, 1996, 1998 and 2000. I collapsed DMZ municipalities into one single entity, restricted sample to municpalities in the departments of Meta, Caqueta and Guaviare, and dropped each department's capital. The "Post p-value" follows Abadie (2021) and refers to the proportions of placebo standardized effects that are at least as large as the main standardized effect for the treated unit.

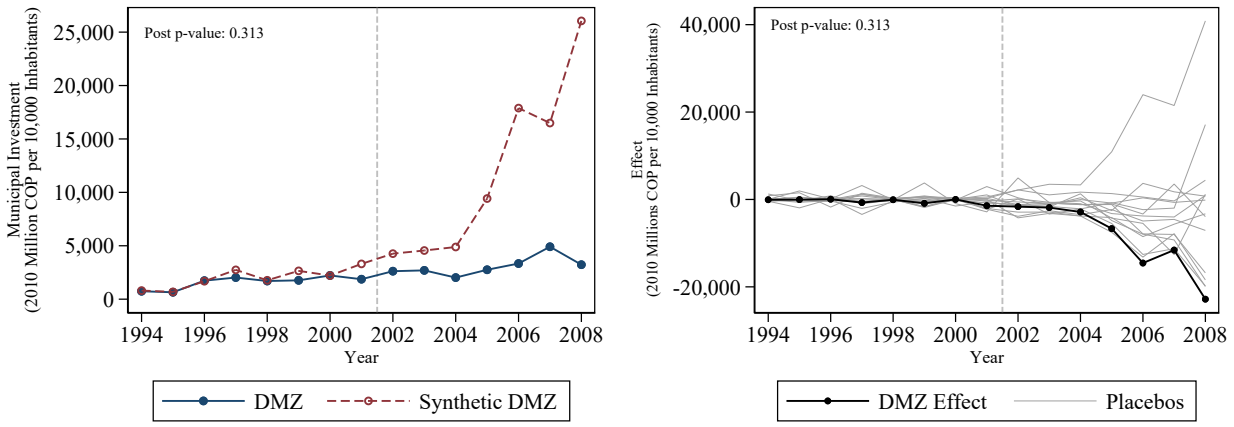
Figure C4: Synthetic Control Results: SGP Disposable Revenues



(a) Synthetic Control **(b)** Placebos

Notes: Panels (a) shows observed values of fiscal expenditure between 1995 and 2008 (solid line) and counterfactuals from the synthetic control (dashed line). The outcome variable is the total value of 2010 COP transferred from the central government to the municipalities as disposable income (i.e., *libre destinación)* as part of the *Sistema General de Participaciones* (SGP). To feet the synthetic control I use the dependent variable in 1994, 1996, 1998 and 2000. I collapsed DMZ municipalities into one single entity, restricted sample to municpalities in the departments of Meta, Caqueta and Guaviare, and dropped each department's capital. The "Post p-value" follows Abadie (2021) and refers to the proportions of placebo standardized effects that are at least as large as the main standardized effect for the treated unit.

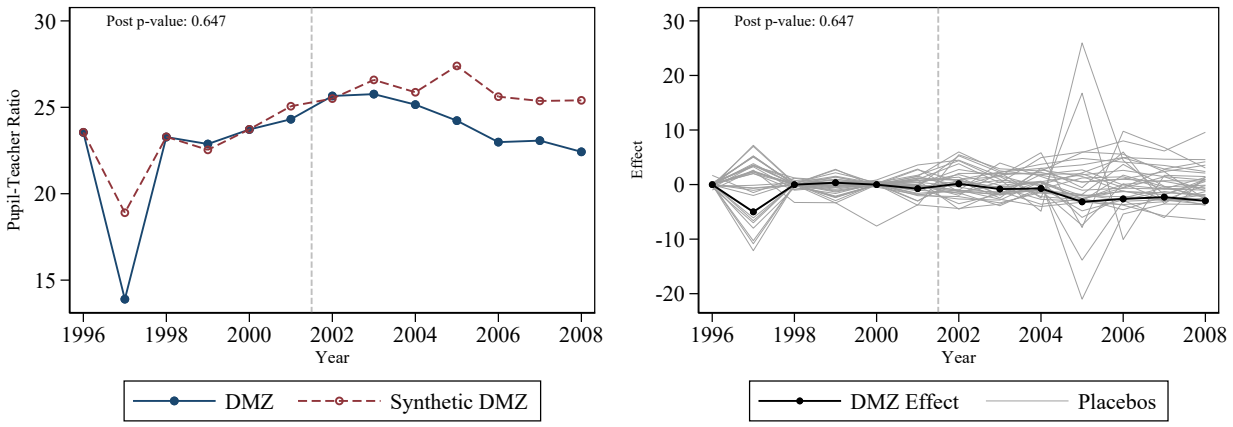
Figure C5: Synthetic Control Results: Municipal Investment



(a) Synthetic Control **(b)** Placebos

Notes: Panels (a) shows observed values of fiscal expenditure between 1995 and 2008 (solid line) and counterfactuals from the synthetic control (dashed line). The outcome variable is the total value of 2010 COP of municipal investment (i.e. capital expenditure). To feet the synthetic control I use the dependent variable in 1994, 1996, 1998 and 2000. I collapsed DMZ municipalities into one single entity, restricted sample to municpalities in the departments of Meta, Caqueta and Guaviare, and dropped each department's capital. The "Post p-value" follows Abadie (2021) and refers to the proportions of placebo standardized effects that are at least as large as the main standardized effect for the treated unit.

Figure C6: Synthetic Control Results: Pupil-Teacher Ratio

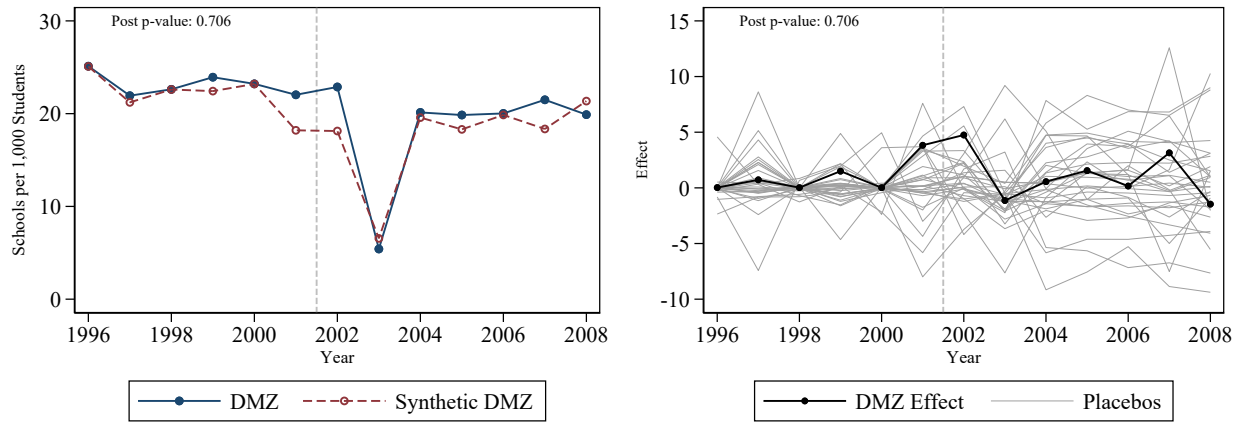


(a) Municipal Investment - Synthetic Control

(b) Municipal Investment - Placebos

Notes: Panels (a) shows observed values of fiscal expenditure between 1995 and 2008 (solid line) and counterfactuals from the synthetic control (dashed line). The outcome variable is the total value of 2010 COP of municipal investment (i.e. capital expenditure). To feet the synthetic control I use the dependent variable in 1994, 1996, 1998 and 2000. I collapsed DMZ municipalities into one single entity, restricted sample to municpalities in the departments of Meta, Caqueta and Guaviare, and dropped each department's capital. The "Post p-value" follows Abadie (2021) and refers to the proportions of placebo standardized effects that are at least as large as the main standardized effect for the treated unit.

Figure C7: Synthetic Control Results: Schools per 1,000 Students



(a) Municipal Investment - Synthetic Control

(b) Municipal Investment - Placebos

Notes: Panels (a) shows observed values of fiscal expenditure between 1995 and 2008 (solid line) and counterfactuals from the synthetic control (dashed line). The outcome variable is the total value of 2010 COP of municipal investment (i.e. capital expenditure). To feet the synthetic control I use the dependent variable in 1994, 1996, 1998 and 2000. I collapsed DMZ municipalities into one single entity, restricted sample to municpalities in the departments of Meta, Caqueta and Guaviare, and dropped each department's capital. The "Post p-value" follows Abadie (2021) and refers to the proportions of placebo standardized effects that are at least as large as the main standardized effect for the treated unit.

D Robustness Checks

D.1 RDD Specification

Table D1: RDD Specification Robustness - Human Capital

	Main	Bandwidth				Polynomial		Kernel	
	Estimate	0.50	0.75	1.25	1.50	Quadratic	Cubic	Epanechnikov	Uniform
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Literacy (=1)								
RD Estimate	-0.01	-0.00	-0.00	-0.02**	-0.02**	0.00	-0.00	-0.01	-0.02***
	(0.009)	(0.013)	(0.010)	(0.009)	(0.009)	(0.011)	(0.013)	(0.008)	(0.008)
BW	15.3	7.67	11.5	19.1	23.0	15.5	20.7	15.3	15.3
Obs.	18250	9171	13933	22119	26460	18388	23469	18250	18250
Dep. Var. Mean	0.84	0.84	0.84	0.83	0.83	0.84	0.83	0.84	0.84
Dep. Var. Std.	0.37	0.36	0.37	0.37	0.37	0.37	0.37	0.37	0.37
	Panel B: Years of Education								
RD Estimate	0.40***	-0.45**	0.18	0.34***	0.28**	0.27*	0.20	0.49***	0.46***
	(0.126)	(0.204)	(0.154)	(0.124)	(0.128)	(0.145)	(0.154)	(0.128)	(0.124)
BW	11.3	5.66	8.48	14.1	16.9	15.8	22.4	11.3	11.3
Obs.	12338	6014	9253	15026	17836	16807	23228	12338	12338
Dep. Var. Mean	7.53	7.37	7.46	7.52	7.51	7.52	7.48	7.53	7.53
Dep. Var. Std.	4.06	4.02	4.01	4.06	4.03	4.03	4.04	4.06	4.06
	Panel C: Health Insurance (=1)								
RD Estimate	0.01	0.00	0.01	0.01*	0.01	0.01	0.02	0.02*	0.02**
	(0.008)	(0.011)	(0.009)	(0.007)	(0.006)	(0.007)	(0.010)	(0.008)	(0.007)
BW	7.58	3.79	5.68	9.47	11.3	18.8	17.0	7.58	7.58
Obs.	8750	4499	6604	11199	13478	21166	19420	8750	8750
Dep. Var. Mean	0.98	0.98	0.98	0.97	0.97	0.97	0.97	0.98	0.98
Dep. Var. Std.	0.16	0.14	0.16	0.16	0.16	0.16	0.16	0.16	0.16

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is an indicator variable that takes the value of one for individuals that know how to read and write, in Panel B it is the total years of formal education, and in Panel C it is an indicator that takes the value of one for individuals with health insurance. The unit of observation is at the individual level in all columns. Estimates in column (1) include regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects, and control for age, age squared and sex. Columns (2)-(5) vary the optimal bandwidthl columns (6) and (7) vary the local linear polynomial function at each side of the boundary, and columns (8) and (9) vary the kernel weighting. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. The unit of observation is at the individual level. Standard errors clustered at the rural districts level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table D2: RDD Specification Robustness - Dwelling Characteristics and Public Conveniences

	Main	Bandwidth				Polynomial		Kernel	
	Estimate	0.50	0.75	1.25	1.50	Quadratic	Cubic	Epanechnikov	Uniform
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Concrete Walls (=1)								
RD Estimate	-0.02	-0.02	-0.04	0.00	0.02	-0.03	-0.05	-0.02	0.00
	(0.040)	(0.056)	(0.050)	(0.050)	(0.047)	(0.044)	(0.046)	(0.041)	(0.042)
BW	8.84	4.42	6.63	11.0	13.2	14.6	21.0	8.84	8.84
Obs.	3392	1751	2542	4268	4993	5535	7579	3392	3392
Dep. Var. Mean	0.23	0.21	0.21	0.25	0.28	0.28	0.31	0.23	0.23
Dep. Var. Std.	0.42	0.41	0.41	0.44	0.45	0.45	0.46	0.42	0.42
	Panel B: Concrete Floor (=1)								
RD Estimate	-0.02	-0.02	-0.03	-0.01	0.00	-0.03	-0.04	-0.02	-0.01
	(0.025)	(0.036)	(0.028)	(0.021)	(0.020)	(0.027)	(0.030)	(0.024)	(0.020)
BW	7.82	3.91	5.87	9.78	11.7	13.6	21.1	7.82	7.82
Obs.	2976	1571	2279	3728	4469	5131	7597	2976	2976
Dep. Var. Mean	0.04	0.04	0.04	0.04	0.05	0.05	0.07	0.04	0.04
Dep. Var. Std.	0.19	0.19	0.19	0.20	0.22	0.22	0.25	0.19	0.19
	Panel C: Electricity (=1)								
RD Estimate	-0.01	-0.02	-0.01	-0.02	-0.02	-0.04	-0.02	-0.02	-0.04
	(0.066)	(0.044)	(0.062)	(0.075)	(0.073)	(0.069)	(0.084)	(0.065)	(0.062)
BW	10.8	5.41	8.12	13.5	16.2	19.6	21.0	10.8	10.8
Obs.	4224	2158	3201	5188	6155	7314	7737	4224	4224
Dep. Var. Mean	0.43	0.35	0.40	0.47	0.49	0.50	0.50	0.43	0.43
Dep. Var. Std.	0.49	0.48	0.49	0.50	0.50	0.50	0.50	0.49	0.49
	Panel D: Sewerage (=1)								
RD Estimate	0.02**	0.01	0.02	0.02*	0.01	0.03**	0.00	0.02**	0.02**
	(0.010)	(0.012)	(0.011)	(0.009)	(0.008)	(0.012)	(0.015)	(0.009)	(0.009)
BW	6.78	3.39	5.08	8.47	10.1	9.40	9.43	6.78	6.78
Obs.	2614	1380	2008	3287	3902	3624	3635	2614	2614
Dep. Var. Mean	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Dep. Var. Std.	0.10	0.10	0.10	0.11	0.11	0.11	0.11	0.10	0.10
	Panel E: Aqueduct (=1)								
RD Estimate	0.11**	0.04	0.05*	0.13**	0.11**	0.08*	0.07	0.12***	0.13***
	(0.046)	(0.026)	(0.028)	(0.049)	(0.054)	(0.048)	(0.064)	(0.046)	(0.048)
BW	7.98	3.99	5.98	9.97	11.9	19.8	16.0	7.98	7.98
Obs.	3054	1609	2330	3826	4548	7268	5988	3054	3054
Dep. Var. Mean	0.11	0.10	0.10	0.14	0.16	0.19	0.20	0.11	0.11
Dep. Var. Std.	0.31	0.30	0.30	0.35	0.37	0.39	0.40	0.31	0.31

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is an indicator that equals one if the dwelling's walls materials are in concrete or better, while in columns Panel B it is an indicator that equals one if the dwelling's floor materials are in concrete or better. Dependent variable in Panel C is an indicator that equals one if the dwelling has access to electricity, in Panel D it is an indicator that equals one if the dwelling has access to sewerage system, and in Panel E it is an indicator that equals one if the dwelling has access to aqueduct system. The unit of observation is at the dwelling level in all columns. Estimates in column (1) include regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects, and control for age, age squared and sex. Columns (2)-(5) vary the optimal bandwidth columns (6) and (7) vary the local linear polynomial function at each side of the boundary, and columns (8) and (9) vary the kernel weighting. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. The unit of observation is at the individual level. Standard errors clustered at the rural districts level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D3: RDD Specification Robustness - Agricultural Yield

	Main	Bandwidth				Polynomial		Kernel	
	Estimate	0.50	0.75	1.25	1.50	Quadratic	Cubic	Epanechnikov	Uniform
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Total Per Hectare								
RD Estimate	0.16***	0.17***	0.17***	0.15***	0.16***	0.20***	0.20***	0.16***	0.12***
	(0.020)	(0.035)	(0.032)	(0.034)	(0.032)	(0.030)	(0.034)	(0.020)	(0.019)
BW	12.1	6.06	9.09	15.1	18.1	12.4	15.7	12.1	12.1
Obs.	8466	4157	6307	10500	12295	8665	10831	8466	8466
Dep. Var. Mean	4.40	4.09	4.19	4.52	4.59	4.42	4.57	4.40	4.40
Dep. Var. Std.	4.20	3.79	4.04	4.25	4.36	4.21	4.31	4.20	4.20
	Panel B: Total Revenue								
RD Estimate	0.35***	0.32***	0.41***	0.42***	0.40***	0.49***	0.53***	0.35***	0.42***
	(0.084)	(0.123)	(0.091)	(0.096)	(0.103)	(0.093)	(0.098)	(0.082)	(0.079)
BW	9.17	4.59	6.88	11.4	13.7	16.0	24.9	9.17	9.17
Obs.	6364	3274	4705	8083	9405	11066	17541	6364	6364
Dep. Var. Mean	15.29	15.19	15.73	15.69	15.81	15.97	18.85	15.29	15.29
Dep. Var. Std.	37.81	38.48	39.39	37.01	37.21	37.17	46.02	37.81	37.81
	Panel C: Subsistence (=1)								
RD Estimate	0.12**	0.15*	0.16**	0.09	0.05	0.13**	0.13*	0.12**	0.08
	(0.060)	(0.080)	(0.079)	(0.079)	(0.082)	(0.065)	(0.072)	(0.059)	(0.056)
BW	7.47	3.74	5.60	9.34	11.2	14.4	21.2	7.47	7.47
Obs.	4287	2320	3312	5400	6499	8087	11699	4287	4287
Dep. Var. Mean	0.42	0.42	0.41	0.41	0.40	0.38	0.35	0.42	0.42
Dep. Var. Std.	0.49	0.49	0.49	0.49	0.49	0.49	0.48	0.49	0.49
	Panel D: Sale (=1)								
RD Estimate	-0.02	-0.03	-0.02	-0.02	-0.01	-0.03*	-0.03*	-0.02	-0.02**
	(0.013)	(0.028)	(0.020)	(0.014)	(0.012)	(0.014)	(0.016)	(0.012)	(0.011)
BW	9.74	4.87	7.31	12.1	14.6	17.2	26.8	9.74	9.74
Obs.	5627	2937	4216	6908	8216	9484	16777	5627	5627
Dep. Var. Mean	0.02	0.02	0.02	0.02	0.03	0.03	0.05	0.02	0.02
Dep. Var. Std.	0.15	0.14	0.14	0.16	0.16	0.17	0.22	0.15	0.15

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is the inverse hyperbolic sine of the revenue per hectare in Colombian 2013 millions of COP, while in Panel B it is the inverse hyperbolic sine of the total revenue in Colombian 2013 millions of COP. Dependent variable in Panel C is an indicator that equals one if the farms agricultural production is used only for self-consumption, and in Panel D it is an indicator that equals one if the farms agricultural production is used only for market sale. The unit of observation is at the farm-crop level in columns Panel A and B, while it is at the farm level in panel C and D. Estimates in column (1) include regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects, and include boundary segment fixed effects, and control for farm extension. Columns (2)-(5) vary the optimal bandwidth columns (6) and (7) vary the local linear polynomial function at each side of the boundary, and columns (8) and (9) vary the kernel weighting. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. The unit of observation is at the individual level. Standard errors clustered at the rural districts level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D4: RDD Specification Robustness - Agricultural Choices

	Main	Bandwidth				Polynomial		Kernel	
	Estimate	0.50	0.75	1.25	1.50	Quadratic	Cubic	Epanechnikov	Uniform
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Cash Crops - Area Share								
RD Estimate	-0.07*	-0.10*	-0.09	-0.06	-0.05	-0.08*	-0.09	-0.06	-0.05
	(0.041)	(0.057)	(0.063)	(0.064)	(0.063)	(0.046)	(0.054)	(0.041)	(0.040)
BW	14.3	7.16	10.7	17.8	21.4	22.8	24.2	14.3	14.3
Obs.	3476	1743	2667	4343	5278	5731	6205	3476	3476
Dep. Var. Mean	0.33	0.32	0.33	0.31	0.31	0.31	0.30	0.33	0.33
Dep. Var. Std.	0.34	0.35	0.34	0.33	0.34	0.34	0.34	0.34	0.34
	Panel B: Cash Crops - Revenue Share								
RD Estimate	-0.11**	-0.13*	-0.13*	-0.09	-0.08	-0.11**	-0.10*	-0.10**	-0.09**
	(0.046)	(0.070)	(0.073)	(0.075)	(0.075)	(0.049)	(0.056)	(0.045)	(0.044)
BW	11.6	5.83	8.74	14.5	17.4	21.2	27.2	11.6	11.6
Obs.	2907	1469	2166	3577	4246	5225	7920	2907	2907
Dep. Var. Mean	0.31	0.31	0.32	0.31	0.29	0.29	0.27	0.31	0.31
Dep. Var. Std.	0.37	0.37	0.38	0.37	0.36	0.36	0.35	0.37	0.37
	Panel C: Perennial Crops - Area Share								
RD Estimate	0.07**	0.07	0.07	0.07	0.08	0.07	0.05	0.08**	0.09***
	(0.035)	(0.063)	(0.069)	(0.060)	(0.055)	(0.042)	(0.057)	(0.034)	(0.031)
BW	21.8	10.9	16.3	27.2	32.7	32.9	29.0	21.8	21.8
Obs.	5391	2707	3989	7935	9653	9710	8603	5391	5391
Dep. Var. Mean	0.63	0.59	0.63	0.63	0.64	0.64	0.64	0.63	0.63
Dep. Var. Std.	0.32	0.30	0.31	0.33	0.33	0.33	0.33	0.32	0.32
	Panel D: Perennial Crops - Revenue Share								
RD Estimate	0.13***	0.06	0.11**	0.14***	0.17***	0.10***	0.08*	0.14***	0.14***
	(0.033)	(0.048)	(0.045)	(0.042)	(0.039)	(0.035)	(0.042)	(0.033)	(0.032)
BW	18.2	9.10	13.6	22.7	27.3	26.5	25.0	18.2	18.2
Obs.	4437	2249	3304	5680	7947	7616	6515	4437	4437
Dep. Var. Mean	0.63	0.57	0.60	0.64	0.64	0.64	0.64	0.63	0.63
Dep. Var. Std.	0.35	0.35	0.35	0.36	0.35	0.35	0.35	0.35	0.35
	Panel E: Transitory Crops - Area Share								
RD Estimate	-0.07**	-0.07	-0.07	-0.07	-0.08	-0.07	-0.05	-0.08**	-0.09***
	(0.035)	(0.063)	(0.069)	(0.060)	(0.055)	(0.042)	(0.057)	(0.034)	(0.031)
BW	21.8	10.9	16.3	27.2	32.7	32.9	29.0	21.8	21.8
Obs.	5391	2707	3989	7935	9653	9710	8603	5391	5391
Dep. Var. Mean	0.37	0.41	0.37	0.37	0.36	0.36	0.36	0.37	0.37
Dep. Var. Std.	0.32	0.30	0.31	0.33	0.33	0.33	0.33	0.32	0.32
	Panel F: Transitory Crops - Revenue Share								
RD Estimate	-0.10***	-0.02	-0.07	-0.11**	-0.10**	-0.05	0.03	-0.12***	-0.13***
	(0.032)	(0.073)	(0.070)	(0.053)	(0.048)	(0.041)	(0.058)	(0.031)	(0.028)
BW	25.6	12.8	19.2	32.0	38.4	32.9	24.1	25.6	25.6
Obs.	7271	3146	4743	9486	10994	9710	6177	7271	7271
Dep. Var. Mean	0.34	0.37	0.34	0.33	0.34	0.33	0.33	0.34	0.34
Dep. Var. Std.	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34

Notes: This table shows estimates of τ in equation 1. The dependent variable in columns Panel A is the share of area with cash crops, while in columns Panel B it is the share of revenue from cash crops. The dependent variable in columns Panel C is the share of area with perennial crops, while in columns Panel D it is the share of revenue from perennial crops.The dependent variable in columns Panel E is the share of area with transitory crops, while in columns Panel F it is the share of revenue from transitory crops. The unit of observation is at the farm level in all columns. Estimates in column (1) use a triangular kernel, local linear polynomial at each side of the boundary, and include boundary segment fixed effects, and control for farm extension. Columns (2)-(5) vary the optimal bandwidth columns (6) and (7) vary the local linear polynomial function at each side of the boundary, and columns (8) and (9) vary the kernel weighting. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. Standard errors clustered at the rural districts level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D.2 Donut Regression Discontinuity Designs

Table D5: Donut RDD - Human Capital

	Main	Donut Size				
	Estimate	0.5 km	1.0 km	1.5 km	2.0 km	2.5 km
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Literacy (=1)					
RD Estimate	-0.01	-0.01	-0.01	0.01	0.00	-0.01
	(0.009)	(0.009)	(0.011)	(0.013)	(0.013)	(0.014)
BW	15.3	15.3	15.3	15.3	15.3	15.3
Obs.	18250	17924	17186	16454	15740	15176
Dep. Var. Mean	0.84	0.84	0.84	0.84	0.84	0.84
Dep. Var. Std.	0.37	0.37	0.37	0.37	0.37	0.37
	Panel B: Years of Education					
RD Estimate	0.40***	0.71***	1.10***	1.25***	1.69***	1.40***
	(0.126)	(0.149)	(0.186)	(0.251)	(0.301)	(0.279)
BW	11.3	11.3	11.3	11.3	11.3	11.3
Obs.	12338	12050	11432	10824	10180	9674
Dep. Var. Mean	7.53	7.54	7.54	7.55	7.57	7.58
Dep. Var. Std.	4.06	4.06	4.07	4.07	4.08	4.07
	Panel C: Health Insurance (=1)					
RD Estimate	0.01	0.01	0.02	0.02**	0.05***	0.05**
	(0.008)	(0.010)	(0.012)	(0.009)	(0.014)	(0.023)
BW	7.58	7.58	7.58	7.58	7.58	7.58
Obs.	8750	8444	7764	7046	6355	5802
Dep. Var. Mean	0.98	0.98	0.98	0.98	0.97	0.97
Dep. Var. Std.	0.16	0.16	0.15	0.16	0.16	0.17

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is an indicator variable that takes the value of one for individuals that know how to read and write, in Panel B it is the total years of formal education, and in Panel C it is an indicator that takes the value of one for individuals with health insurance. The unit of observation is at the individual level in all columns. All regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects, and control for age, age squared and sex. In columns (2) to (6) I implement a RD Donut design by dropping observations close to the boundary. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. The unit of observation is at the individual level. Standard errors clustered at the rural districts level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table D6: Donut RDD - Dwelling Characteristics and Public Conveniences

	Main	Donut Size				
	Estimate	0.5 km	1.0 km	1.5 km	2.0 km	2.5 km
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Concrete Walls (=1)					
RD Estimate	-0.02	-0.06	-0.09	-0.02	0.12	0.10
	(0.040)	(0.045)	(0.054)	(0.066)	(0.089)	(0.110)
BW	8.84	8.84	8.84	8.84	8.84	8.84
Obs.	3392	3280	3045	2794	2576	2384
Dep. Var. Mean	0.23	0.23	0.23	0.23	0.24	0.24
Dep. Var. Std.	0.42	0.42	0.42	0.42	0.43	0.43
	Panel B: Concrete Floor (=1)					
RD Estimate	-0.02	-0.04	-0.05**	-0.01	0.03	0.06
	(0.025)	(0.028)	(0.024)	(0.017)	(0.029)	(0.040)
BW	7.82	7.82	7.82	7.82	7.82	7.82
Obs.	2976	2865	2630	2379	2162	1970
Dep. Var. Mean	0.04	0.04	0.04	0.04	0.04	0.04
Dep. Var. Std.	0.19	0.19	0.19	0.19	0.19	0.20
	Panel C: Electricity (=1)					
RD Estimate	-0.01	0.02	-0.03	-0.08	0.05	-0.04
	(0.066)	(0.068)	(0.071)	(0.076)	(0.091)	(0.103)
BW	10.8	10.8	10.8	10.8	10.8	10.8
Obs.	4224	4111	3870	3616	3392	3198
Dep. Var. Mean	0.43	0.43	0.44	0.45	0.45	0.46
Dep. Var. Std.	0.49	0.50	0.50	0.50	0.50	0.50
	Panel D: Sewerage (=1)					
RD Estimate	0.02**	0.02**	0.05***	0.03***	0.04*	0.04*
	(0.010)	(0.012)	(0.012)	(0.012)	(0.021)	(0.020)
BW	6.78	6.78	6.78	6.78	6.78	6.78
Obs.	2614	2502	2264	2010	1788	1595
Dep. Var. Mean	0.01	0.01	0.01	0.01	0.01	0.01
Dep. Var. Std.	0.10	0.10	0.10	0.09	0.09	0.09
	Panel E: Aqueduct (=1)					
RD Estimate	0.11**	0.14***	0.16***	0.15**	0.35***	0.45***
	(0.046)	(0.047)	(0.058)	(0.061)	(0.078)	(0.086)
BW	7.98	7.98	7.98	7.98	7.98	7.98
Obs.	3054	2942	2704	2450	2228	2035
Dep. Var. Mean	0.11	0.11	0.12	0.11	0.12	0.12
Dep. Var. Std.	0.31	0.32	0.32	0.32	0.33	0.32

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is an indicator that equals one if the dwelling's walls materials are in concrete or better, while in columns Panel B it is an indicator that equals one if the dwelling's floor materials are in concrete or better. Dependent variable in Panel C is an indicator that equals one if the dwelling has access to electricity, in Panel D it is an indicator that equals one if the dwelling has access to sewerage system, and in Panel E it is an indicator that equals one if the dwelling has access to aqueduct system. The unit of observation is at the dwelling level in all columns. All regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects. In columns (2) to (6) I implement a RD Donut design by dropping observations close to the boundary. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. The unit of observation is at the individual level. Standard errors clustered at the rural districts level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table D7: Donut RDD - Agricultural Yield

	Main	Donut Size				
	Estimate	0.5 km	1.0 km	1.5 km	2.0 km	2.5 km
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Revenue Per Hectare					
RD Estimate	0.16***	0.16***	0.20***	0.21***	0.19***	0.13***
	(0.020)	(0.021)	(0.024)	(0.026)	(0.032)	(0.039)
BW	12.1	12.1	12.1	12.1	12.1	12.1
Obs.	8466	8288	7805	7376	7013	6690
Dep. Var. Mean	4.40	4.41	4.46	4.48	4.51	4.53
Dep. Var. Std.	4.20	4.19	4.27	4.27	4.31	4.34
	Panel B: Total Revenue					
RD Estimate	0.35***	0.40***	0.64***	0.91***	0.56***	0.16
	(0.084)	(0.093)	(0.091)	(0.111)	(0.156)	(0.178)
BW	9.17	9.17	9.17	9.17	9.17	9.17
Obs.	6364	6186	5703	5274	4911	4588
Dep. Var. Mean	15.29	15.46	15.71	15.79	15.99	15.55
Dep. Var. Std.	37.81	38.04	38.54	38.52	39.01	36.93
	Panel C: Subsistence (=1)					
RD Estimate	0.12**	0.12*	0.16**	0.14*	-0.02	-0.15
	(0.060)	(0.065)	(0.070)	(0.072)	(0.093)	(0.107)
BW	7.47	7.47	7.47	7.47	7.47	7.47
Obs.	4287	4071	3725	3345	3030	2734
Dep. Var. Mean	0.42	0.42	0.42	0.43	0.43	0.43
Dep. Var. Std.	0.49	0.49	0.49	0.49	0.50	0.49
	Panel D: Sale (=1)					
RD Estimate	-0.02	-0.01	-0.01	-0.01	-0.02*	-0.02
	(0.013)	(0.007)	(0.007)	(0.009)	(0.013)	(0.018)
BW	9.74	9.74	9.74	9.74	9.74	9.74
Obs.	5627	5411	5065	4685	4370	4074
Dep. Var. Mean	0.02	0.02	0.02	0.02	0.02	0.02
Dep. Var. Std.	0.15	0.14	0.14	0.14	0.14	0.14

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is the inverse hyperbolic sine of the revenue per hectare in Colombian 2013 millions of COP, while in Panel B it is the inverse hyperbolic sine of the total revenue in Colombian 2013 millions of COP. Dependent variable in Panel C is an indicator that equals one if the farms agricultural production is used only for self-consumption, and in Panel D it is an indicator that equals one if the farms agricultural production is used only for market sale. The unit of observation is at the farm-crop level in columns Panel A and B, while it is at the farm level in panel C and D. All regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects, and control for age, age squared and sex. In columns (2) to (6) I implement a RD Donut design by dropping observations close to the boundary. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. The unit of observation is at the individual level. Standard errors clustered at the rural districts level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D8: Donut RDD - Agricultural Choices

	Main	Donut Size				
	Estimate	0.5 km	1.0 km	1.5 km	2.0 km	2.5 km
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Cash Crops - Area Share					
RD Estimate	-0.07*	-0.05	-0.06	-0.05	-0.06	-0.04
	(0.041)	(0.044)	(0.052)	(0.059)	(0.059)	(0.065)
BW	14.3	14.3	14.3	14.3	14.3	14.3
Obs.	3476	3398	3227	3065	2939	2811
Dep. Var. Mean	0.33	0.33	0.33	0.33	0.33	0.33
Dep. Var. Std.	0.34	0.34	0.34	0.33	0.34	0.33
	Panel B: Cash Crops - Revenue Share					
RD Estimate	-0.11**	-0.08*	-0.12*	-0.12*	-0.14**	-0.11
	(0.046)	(0.049)	(0.060)	(0.068)	(0.069)	(0.079)
BW	11.6	11.6	11.6	11.6	11.6	11.6
Obs.	2907	2829	2658	2496	2370	2242
Dep. Var. Mean	0.31	0.31	0.31	0.31	0.32	0.32
Dep. Var. Std.	0.37	0.37	0.37	0.37	0.37	0.37
	Panel C: Perennial Crops - Area Share					
RD Estimate	0.07**	0.09**	0.12***	0.11***	0.09**	0.07*
	(0.035)	(0.036)	(0.036)	(0.036)	(0.038)	(0.041)
BW	21.8	21.8	21.8	21.8	21.8	21.8
Obs.	5391	5313	5142	4980	4854	4726
Dep. Var. Mean	0.63	0.63	0.63	0.63	0.64	0.64
Dep. Var. Std.	0.32	0.32	0.32	0.32	0.32	0.32
	Panel D: Perennial Crops - Revenue Share					
RD Estimate	0.13***	0.15***	0.21***	0.18***	0.15***	0.15***
	(0.033)	(0.035)	(0.038)	(0.040)	(0.047)	(0.055)
BW	18.2	18.2	18.2	18.2	18.2	18.2
Obs.	4437	4359	4188	4026	3900	3772
Dep. Var. Mean	0.63	0.63	0.63	0.63	0.63	0.63
Dep. Var. Std.	0.35	0.35	0.35	0.35	0.35	0.35
	Panel E: Transitory Crops - Area Share					
RD Estimate	-0.07**	-0.09**	-0.12***	-0.11***	-0.09**	-0.07*
	(0.035)	(0.036)	(0.036)	(0.036)	(0.038)	(0.041)
BW	21.8	21.8	21.8	21.8	21.8	21.8
Obs.	5391	5313	5142	4980	4854	4726
Dep. Var. Mean	0.37	0.37	0.37	0.37	0.36	0.36
Dep. Var. Std.	0.32	0.32	0.32	0.32	0.32	0.32
	Panel F: Transitory Crops - Revenue Share					
RD Estimate	-0.10***	-0.11***	-0.15***	-0.15***	-0.14***	-0.15***
	(0.032)	(0.033)	(0.032)	(0.032)	(0.033)	(0.036)
BW	25.6	25.6	25.6	25.6	25.6	25.6
Obs.	7271	7193	7022	6860	6734	6606
Dep. Var. Mean	0.34	0.34	0.34	0.34	0.34	0.34
Dep. Var. Std.	0.34	0.34	0.34	0.34	0.34	0.34

Notes: This table shows estimates of τ in equation 1. The dependent variable in columns Panel A is the share of area with cash crops, while in columns Panel B it is the share of revenue from cash crops. The dependent variable in columns Panel C is the share of area with perennial crops, while in columns Panel D it is the share of revenue from perennial crops.The dependent variable in columns Panel E is the share of area with transitory crops, while in columns Panel F it is the share of revenue from transitory crops. The unit of observation is at the farm level in all columns. All regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects, and control for age, age squared and sex. In columns (2) to (6) I implement a RD Donut design by dropping observations close to the boundary. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. Standard errors clustered at the rural districts level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01 .

D.3 Robustness to Variance-covariance Matrix Assumptions

Table D9: Variance-covariance Matrix Assumptions - Human Capital

	Standard Errors Structure					
	NN-Cluster		Cluster			
	Vereda	Municipality	*Vereda*	Municipality	250m Bins	500m Bins
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Literacy (=1)					
RD Estimate	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.009)	(0.009)	(0.018)	(0.028)	(0.015)	(0.015)
BW	15.3	15.3	16.9	16.8	16.3	15.8
Obs.	18250	18250	19818	19713	19244	18771
Dep. Var. Mean	0.84	0.84	0.84	0.84	0.84	0.84
Dep. Var. Std.	0.37	0.37	0.37	0.37	0.37	0.37
	Panel B: Years of Education					
RD Estimate	0.40***	0.40***	0.36	0.35	0.40*	0.39*
	(0.126)	(0.126)	(0.363)	(0.358)	(0.220)	(0.233)
BW	11.3	11.9	14.5	13.7	12.7	12.9
Obs.	12338	12847	15640	14712	13660	13936
Dep. Var. Mean	7.53	7.53	7.50	7.54	7.54	7.55
Dep. Var. Std.	4.06	4.06	4.04	4.06	4.06	4.06
	Panel C: Health Insurance (=1)					
RD Estimate	0.01	0.01*	0.01	0.01	0.01	0.01
	(0.008)	(0.007)	(0.018)	(0.013)	(0.013)	(0.012)
BW	7.58	9.83	9.95	9.37	10.7	10.2
Obs.	8750	11500	11687	11045	12659	11971
Dep. Var. Mean	0.98	0.97	0.97	0.97	0.97	0.97
Dep. Var. Std.	0.16	0.16	0.16	0.16	0.16	0.16

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is an indicator variable that takes the value of one for individuals that know how to read and write, in Panel B it is the total years of formal education, and in Panel C it is an indicator that takes the value of one for individuals with health insurance. The unit of observation is at the individual level in all columns. All regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects, and control for age, age squared and sex. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. The unit of observation is at the individual level. While column (1) presents baseline specification for the standard errors, columns (2)-(6) show robustness of the results to alternative specifications. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table D10: Variance-covariance Matrix Assumptions - Dwelling Characteristics and Public Conveniences

	Standard Errors Structure					
	NN-Cluster		Cluster			
	Vereda	Municipality	*Vereda*	Municipality	250m Bins	500m Bins
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Concrete Walls (=1)					
RD Estimate	-0.02	-0.02	-0.01	-0.02	-0.02	-0.02
	(0.040)	(0.051)	(0.052)	(0.054)	(0.045)	(0.047)
BW	8.84	9.99	9.78	10.1	8.76	9.01
Obs.	3392	3816	3738	3871	3377	3472
Dep. Var. Mean	0.23	0.24	0.24	0.24	0.23	0.23
Dep. Var. Std.	0.42	0.43	0.43	0.43	0.42	0.42
	Panel B: Concrete Floor (=1)					
RD Estimate	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
	(0.025)	(0.024)	(0.026)	(0.024)	(0.022)	(0.022)
BW	7.82	7.30	8.24	7.92	8.28	8.21
Obs.	2976	2776	3149	3018	3165	3148
Dep. Var. Mean	0.04	0.04	0.04	0.04	0.04	0.04
Dep. Var. Std.	0.19	0.19	0.20	0.19	0.20	0.20
	Panel C: Electricity (=1)					
RD Estimate	-0.01	-0.02	-0.02	-0.02	-0.01	-0.01
	(0.066)	(0.075)	(0.108)	(0.123)	(0.057)	(0.062)
BW	10.8	12.2	11.6	12.0	9.71	10.0
Obs.	4224	4678	4551	4642	3797	3912
Dep. Var. Mean	0.43	0.45	0.44	0.44	0.41	0.41
Dep. Var. Std.	0.49	0.50	0.50	0.50	0.49	0.49
	Panel D: Sewerage (=1)					
RD Estimate	0.02**	0.02*	0.02	0.02	0.02	0.02
	(0.010)	(0.011)	(0.016)	(0.014)	(0.017)	(0.017)
BW	6.78	4.96	5.32	7.49	5.25	5.19
Obs.	2614	1973	2093	2839	2069	2038
Dep. Var. Mean	0.01	0.01	0.01	0.01	0.01	0.01
Dep. Var. Std.	0.10	0.10	0.10	0.10	0.10	0.10
	Panel E: Aqueduct (=1)					
RD Estimate	0.11**	0.13***	0.13*	0.13	0.12**	0.12**
	(0.046)	(0.049)	(0.077)	(0.079)	(0.052)	(0.056)
BW	7.98	9.90	10.0	10.1	9.11	9.50
Obs.	3054	3791	3846	3872	3518	3666
Dep. Var. Mean	0.11	0.14	0.14	0.14	0.13	0.14
Dep. Var. Std.	0.31	0.35	0.35	0.35	0.34	0.34

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is an indicator that equals one if the dwelling's walls materials are in concrete or better, while in columns Panel B it is an indicator that equals one if the dwelling's floor materials are in concrete or better. Dependent variable in Panel C is an indicator that equals one if the dwelling has access to electricity, in Panel D it is an indicator that equals one if the dwelling has access to sewerage system, and in Panel E it is an indicator that equals one if the dwelling has access to aqueduct system. The unit of observation is at the dwelling level in all columns. All regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. The unit of observation is at the individual level. While column (1) presents baseline specification for the standard errors, columns (2)-(6) show robustness of the results to alternative specifications. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table D11: Variance-covariance Matrix Assumptions - Agricultural Yield

	Standard Errors Structure					
	NN-Cluster		Cluster			
	Vereda	Municipality	*Vereda*	Municipality	250m Bins	500m Bins
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Revenue Per Hectare					
RD Estimate	0.16***	0.17***	0.15***	0.17**	0.15***	0.15***
	(0.020)	(0.034)	(0.037)	(0.076)	(0.038)	(0.043)
BW	12.1	12.0	12.8	10.9	12.8	12.5
Obs.	8466	8434	8886	7704	8877	8706
Dep. Var. Mean	4.40	4.40	4.45	4.33	4.44	4.43
Dep. Var. Std.	4.20	4.20	4.23	4.13	4.23	4.22
	Panel B: Total Revenue					
RD Estimate	0.35***	0.36***	0.39**	0.39	0.42***	0.42***
	(0.084)	(0.087)	(0.180)	(0.243)	(0.150)	(0.157)
BW	9.17	8.23	12.2	11.0	12.2	11.5
Obs.	6364	5728	8557	7847	8506	8139
Dep. Var. Mean	15.29	15.13	15.70	15.72	15.68	15.65
Dep. Var. Std.	37.81	37.98	37.18	37.11	36.99	36.90
	Panel C: Subsistence (=1)					
RD Estimate	0.12**	0.15**	0.09	0.12	0.12**	0.12**
	(0.060)	(0.077)	(0.084)	(0.109)	(0.052)	(0.056)
BW	7.47	6.80	9.82	9.02	7.37	7.42
Obs.	4287	3946	5654	5219	4227	4239
Dep. Var. Mean	0.42	0.42	0.41	0.41	0.42	0.42
Dep. Var. Std.	0.49	0.49	0.49	0.49	0.49	0.49
	Panel D: Sale (=1)					
RD Estimate	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
	(0.013)	(0.016)	(0.016)	(0.021)	(0.013)	(0.016)
BW	9.74	9.87	10.1	11.0	9.74	10.6
Obs.	5627	5669	5826	6333	5619	6140
Dep. Var. Mean	0.02	0.02	0.02	0.02	0.02	0.02
Dep. Var. Std.	0.15	0.15	0.15	0.15	0.15	0.15

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is the inverse hyperbolic sine of the revenue per hectare in Colombian 2013 millions of COP, while in Panel B it is the inverse hyperbolic sine of the total revenue in Colombian 2013 millions of COP. Dependent variable in Panel C is an indicator that equals one if the farms agricultural production is used only for self-consumption, and in Panel D it is an indicator that equals one if the farms agricultural production is used only for market sale. The unit of observation is at the farm-crop level in columns Panel A and B, while it is at the farm level in panel C and D. All regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects, and form's size. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. The unit of observation is at the individual level. While column (1) presents baseline specification for the standard errors, columns (2)-(6) show robustness of the results to alternative specifications. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table D12: Variance-covariance Matrix Assumptions - Agricultural Choices

	Standard Errors Structure					
	NN-Cluster		Cluster			
	Vereda	Municipality	*Vereda*	Municipality	250m Bins	500m Bins
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Cash Crops - Area Share					
RD Estimate	-0.07*	-0.07	-0.07	-0.07	-0.08**	-0.08**
	(0.041)	(0.065)	(0.044)	(0.079)	(0.031)	(0.034)
BW	14.3	17.2	14.7	17.0	12.0	12.6
Obs.	3476	4183	3634	4140	2988	3095
Dep. Var. Mean	0.33	0.32	0.33	0.32	0.33	0.33
Dep. Var. Std.	0.34	0.33	0.34	0.33	0.34	0.34
	Panel B: Cash Crops - Revenue Share					
RD Estimate	-0.11**	-0.10	-0.10**	-0.10	-0.12***	-0.12***
	(0.046)	(0.077)	(0.052)	(0.098)	(0.038)	(0.042)
BW	11.6	14.3	12.3	14.8	10.0	9.80
Obs.	2907	3481	3036	3660	2490	2437
Dep. Var. Mean	0.31	0.31	0.32	0.31	0.32	0.32
Dep. Var. Std.	0.37	0.37	0.37	0.36	0.37	0.37
	Panel C: Perennial Crops - Area Share					
RD Estimate	0.07**	0.07	0.08*	0.08	0.07***	0.07**
	(0.035)	(0.070)	(0.047)	(0.102)	(0.028)	(0.031)
BW	21.8	24.5	23.9	24.7	18.7	18.7
Obs.	5391	6314	6104	6418	4638	4640
Dep. Var. Mean	0.63	0.64	0.64	0.64	0.63	0.63
Dep. Var. Std.	0.32	0.33	0.32	0.33	0.32	0.32
	Panel D: Perennial Crops - Revenue Share					
RD Estimate	0.13***	0.13***	0.13***	0.13*	0.12***	0.12***
	(0.033)	(0.043)	(0.042)	(0.067)	(0.031)	(0.037)
BW	18.2	19.6	17.0	17.0	15.9	15.9
Obs.	4437	4817	4147	4147	3889	3889
Dep. Var. Mean	0.63	0.63	0.62	0.62	0.62	0.62
Dep. Var. Std.	0.35	0.35	0.35	0.35	0.35	0.35
	Panel E: Transitory Crops - Area Share					
RD Estimate	-0.07**	-0.07	-0.08*	-0.08	-0.07***	-0.07**
	(0.035)	(0.070)	(0.047)	(0.102)	(0.028)	(0.031)
BW	21.8	24.5	23.9	24.7	18.7	18.7
Obs.	5391	6314	6104	6418	4638	4640
Dep. Var. Mean	0.37	0.36	0.36	0.36	0.37	0.37
Dep. Var. Std.	0.32	0.33	0.32	0.33	0.32	0.32
	Panel F: Transitory Crops - Revenue Share					
RD Estimate	-0.10***	-0.11*	-0.10**	-0.11	-0.08***	-0.08**
	(0.032)	(0.063)	(0.045)	(0.104)	(0.028)	(0.034)
BW	25.6	27.5	25.0	26.8	20.0	20.9
Obs.	7271	8078	6522	7707	4902	5128
Dep. Var. Mean	0.34	0.33	0.33	0.34	0.34	0.34
Dep. Var. Std.	0.34	0.34	0.34	0.34	0.34	0.34

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is the share of area with cash crops, while in Panel B it is the share of revenue from cash crops. The dependent variable in Panel C is the share of area with perennial crops, while in Panel D it is the share of revenue from perennial crops.The dependent variable in Panel E is the share of area with transitory crops, while in Panel F it is the share of revenue from transitory crops. The unit of observation is at the farm level in all columns. All regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects, and farm's size. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. While column (1) presents baseline specification for the standard errors, columns (2)-(6) show robustness of the results to alternative specifications. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D.4 Selective Migration: Trimming

Table D13: Trimming for Selective Migration - Human Capital

	% Trimmed					
	0%	2.5%	5%	10%	15%	20%
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Literacy (=1)					
RD Estimate	-0.01	-0.01	-0.01	-0.02*	-0.02*	-0.02*
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.011)
BW	15.3	15.3	15.3	15.3	15.3	15.3
Obs.	18250	18064	17789	17470	17155	16767
Dep. Var. Mean	0.84	0.84	0.83	0.83	0.83	0.83
Dep. Var. Std.	0.37	0.37	0.37	0.37	0.37	0.38
	Years of Education					
RD Estimate	0.40***	0.35***	0.31**	0.22	0.23*	0.30**
	(0.126)	(0.126)	(0.130)	(0.137)	(0.136)	(0.145)
BW	11.3	11.3	11.3	11.3	11.3	11.3
Obs.	12338	12277	12109	11890	11630	11378
Dep. Var. Mean	7.53	7.51	7.49	7.44	7.47	7.49
Dep. Var. Std.	4.06	4.05	4.06	4.03	4.04	4.02
	Panel C: Health Insurance (=1)					
RD Estimate	0.01	0.01	0.01	0.01	0.01	0.02**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)
BW	7.58	7.58	7.58	7.58	7.58	7.58
Obs.	8750	8703	8602	8452	8278	8138
Dep. Var. Mean	0.98	0.97	0.98	0.98	0.97	0.97
Dep. Var. Std.	0.16	0.16	0.15	0.16	0.16	0.16

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is an indicator variable that takes the value of one for individuals that know how to read and write, in Panel B it is the total years of formal education, and in Panel C it is an indicator that takes the value of one for individuals with health insurance. The unit of observation is at the individual level in all columns. All regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects, and control for age, age squared and sex. In columns (2) to (6), I trim the top 2.5, 5, 10, 15, and 20% of the most well-off individuals and farms. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. The unit of observation is at the individual level. Standard errors clustered at the rural districts level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table D14: Trimming for Selective Migration - Dwelling Characteristics and Public Conveniences

	% Trimmed					
	0%	2.5%	5%	10%	15%	20%
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Concrete Walls (=1)					
RD Estimate	-0.02	-0.02	-0.05	-0.08**	-0.11***	-0.11***
	(0.040)	(0.040)	(0.039)	(0.038)	(0.039)	(0.039)
BW	8.84	8.84	8.84	8.84	8.84	8.84
Obs.	3392	3373	3328	3270	3203	3148
Dep. Var. Mean	0.23	0.22	0.21	0.20	0.19	0.19
Dep. Var. Std.	0.42	0.42	0.41	0.40	0.39	0.39
	Panel B: Concrete Floor (=1)					
RD Estimate	-0.02	-0.02	-0.03	-0.04*	-0.04**	-0.05**
	(0.025)	(0.024)	(0.024)	(0.022)	(0.023)	(0.022)
BW	7.82	7.82	7.82	7.82	7.82	7.82
Obs.	2976	2958	2918	2864	2798	2747
Dep. Var. Mean	0.04	0.03	0.03	0.03	0.03	0.02
Dep. Var. Std.	0.19	0.18	0.17	0.16	0.16	0.15
	Panel C: Electricity (=1)					
RD Estimate	-0.01	-0.02	-0.04	-0.06	-0.09	-0.11*
	(0.066)	(0.066)	(0.066)	(0.065)	(0.065)	(0.066)
BW	10.8	10.8	10.8	10.8	10.8	10.8
Obs.	4224	4202	4151	4083	4005	3928
Dep. Var. Mean	0.43	0.42	0.42	0.41	0.40	0.39
Dep. Var. Std.	0.49	0.49	0.49	0.49	0.49	0.49
	Panel D: Sewerage (=1)					
RD Estimate	0.02**	0.01	0.01	0.00	0.00	0.00
	(0.010)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)
BW	6.78	6.78	6.78	6.78	6.78	6.78
Obs.	2614	2596	2558	2506	2447	2401
Dep. Var. Mean	0.01	0.01	0.01	0.01	0.01	0.01
Dep. Var. Std.	0.10	0.09	0.08	0.08	0.08	0.08
	Panel E: Aqueduct (=1)					
RD Estimate	0.11**	0.10**	0.08*	0.05	0.04	0.00
	(0.046)	(0.045)	(0.043)	(0.042)	(0.040)	(0.038)
BW	7.98	7.98	7.98	7.98	7.98	7.98
Obs.	3054	3035	2994	2940	2873	2821
Dep. Var. Mean	0.11	0.11	0.10	0.09	0.09	0.08
Dep. Var. Std.	0.31	0.31	0.30	0.29	0.29	0.27

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is an indicator that equals one if the dwelling's walls materials are in concrete or better, while in columns Panel B it is an indicator that equals one if the dwelling's floor materials are in concrete or better. Dependent variable in Panel C is an indicator that equals one if the dwelling has access to electricity, in Panel D it is an indicator that equals one if the dwelling has access to sewerage system, and in Panel E it is an indicator that equals one if the dwelling has access to aqueduct system. The unit of observation is at the dwelling level in all columns. All regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects. In columns (2) to (6), I trim the top 2.5, 5, 10, 15, and 20% of the most well-off individuals and farms. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. The unit of observation is at the individual level. Standard errors clustered at the rural districts level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table D15: Trimming for Selective Migration - Agricultural Yield

	% Trimmed					
	0%	2.5%	5%	10%	15%	20%
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Revenue Per Hectare					
RD Estimate	0.16***	0.15***	0.15***	0.15***	0.14***	0.13***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.021)	(0.021)
BW	12.1	12.1	12.1	12.1	12.1	12.1
Obs.	8466	8409	8368	8284	8189	8043
Dep. Var. Mean	4.40	4.28	4.23	4.17	4.11	4.04
Dep. Var. Std.	4.20	3.90	3.85	3.83	3.81	3.80
	Panel B: Total Revenue					
RD Estimate	0.35***	0.33***	0.29***	0.25***	0.22***	0.18**
	(0.084)	(0.084)	(0.085)	(0.086)	(0.084)	(0.085)
BW	9.17	9.17	9.17	9.17	9.17	9.17
Obs.	6364	6342	6287	6209	6092	5966
Dep. Var. Mean	15.29	14.00	12.75	12.05	11.44	10.94
Dep. Var. Std.	37.81	30.70	27.27	26.68	26.56	26.62

Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is the inverse hyperbolic sine of the revenue per hectare in Colombian 2013 millions of COP, while in Panel B it is the inverse hyperbolic sine of the total revenue in Colombian 2013 millions of COP. The unit of observation is at the farm-crop level in columns Panel A and B, while it is at the farm level in panel C and D. All regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects, and form's size. In columns (2) to (6), I trim the top 2.5, 5, 10, 15, and 20% of the most well-off individuals and farms. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. The unit of observation is at the individual level. Standard errors clustered at the rural districts level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

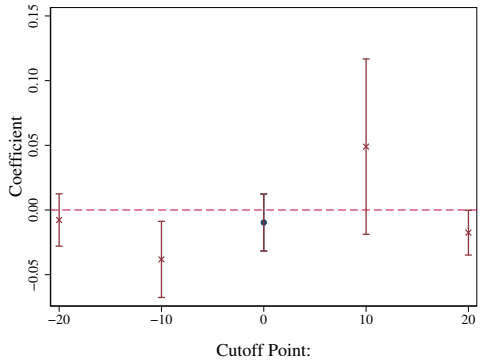
Table D16: Trimming for Selective Migration - Agricultural Choices

	% Trimmed					
	0%	2.5%	5%	10%	15%	20%
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Cash Crops - Area Share					
RD Estimate	-0.07*	-0.07*	-0.07*	0.07*	0.10***	0.15***
	(0.041)	(0.041)	(0.041)	(0.037)	(0.035)	(0.034)
BW	14.3	14.3	14.3	14.3	14.3	14.3
Obs.	3476	3476	3476	3184	3076	2972
Dep. Var. Mean	0.33	0.33	0.33	0.27	0.25	0.24
Dep. Var. Std.	0.34	0.34	0.34	0.28	0.27	0.26
	Panel B: Cash Crops - Revenue Share					
RD Estimate	-0.11**	-0.11**	-0.11**	0.13***	0.17***	0.19***
	(0.046)	(0.046)	(0.046)	(0.041)	(0.034)	(0.032)
BW	11.6	11.6	11.6	11.6	11.6	11.6
Obs.	2907	2907	2907	2537	2487	2392
Dep. Var. Mean	0.31	0.31	0.31	0.21	0.20	0.18
Dep. Var. Std.	0.37	0.37	0.37	0.28	0.27	0.25
	Panel C: Perennial Crops - Area Share					
RD Estimate	0.07**	0.07**	0.07**	0.07**	0.07**	0.07**
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
BW	21.8	21.8	21.8	21.8	21.8	21.8
Obs.	5391	5391	5391	5391	5391	5391
Dep. Var. Mean	0.63	0.63	0.63	0.63	0.63	0.63
Dep. Var. Std.	0.32	0.32	0.32	0.32	0.32	0.32
	Panel D: Perennial Crops - Revenue Share					
RD Estimate	0.13***	0.13***	0.13***	0.13***	0.13***	0.13***
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
BW	18.2	18.2	18.2	18.2	18.2	18.2
Obs.	4437	4437	4437	4437	4437	4437
Dep. Var. Mean	0.63	0.63	0.63	0.63	0.63	0.63
Dep. Var. Std.	0.35	0.35	0.35	0.35	0.35	0.35
	Panel E: Transitory Crops - Area Share					
RD Estimate	-0.07**	-0.07**	-0.07**	-0.02	0.01	0.05*
	(0.035)	(0.035)	(0.035)	(0.035)	(0.033)	(0.030)
BW	21.8	21.8	21.8	21.8	21.8	21.8
Obs.	5391	5391	5391	4950	4817	4595
Dep. Var. Mean	0.37	0.37	0.37	0.31	0.30	0.28
Dep. Var. Std.	0.32	0.32	0.32	0.28	0.27	0.26
	Panel F: Transitory Crops - Revenue Share					
RD Estimate	-0.10***	-0.10***	-0.10***	-0.02	0.01	0.05*
	(0.032)	(0.032)	(0.032)	(0.032)	(0.030)	(0.030)
BW	25.6	25.6	25.6	25.6	25.6	25.6
Obs.	7271	7271	7271	6552	6367	6151
Dep. Var. Mean	0.34	0.34	0.34	0.27	0.25	0.24
Dep. Var. Std.	0.34	0.34	0.34	0.28	0.27	0.26

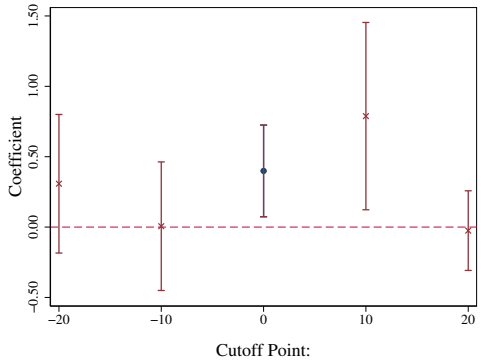
Notes: This table shows estimates of τ in equation 1. The dependent variable in Panel A is the share of area with cash crops, while in Panel B it is the share of revenue from cash crops. The dependent variable in Panel C is the share of area with perennial crops, while in Panel D it is the share of revenue from perennial crops.The dependent variable in Panel E is the share of area with transitory crops, while in Panel F it is the share of revenue from transitory crops. The unit of observation is at the farm level in all columns. All regressions use a triangular kernel, local linear polynomial at each side of the boundary, include boundary segment fixed effects, and farm's size. In columns (2) to (6), I trim the top 2.5, 5, 10, 15, and 20% of the most well-off individuals and farms. The RDD MSE optimal bandwidths are determined using the procedure suggested by Cattaneo et al. (2020a). Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. Standard errors clustered at the rural districts level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D.5 Spatial Placebo Analysis

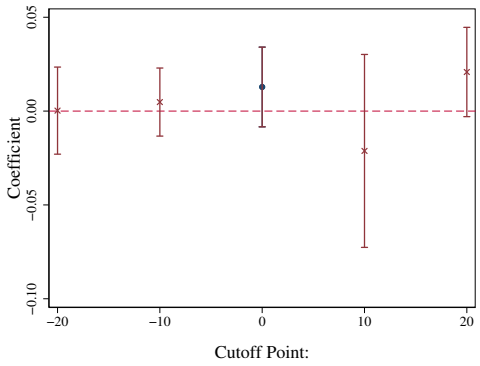
Figure D1: Spatial Placebo: Human Capital



(a) Literacy (=1)



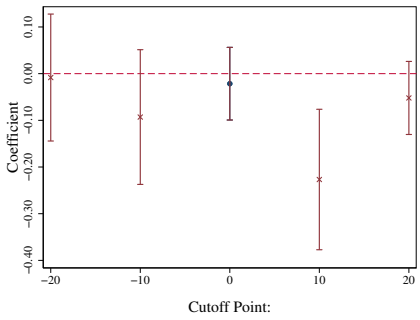
(b) Years of Education



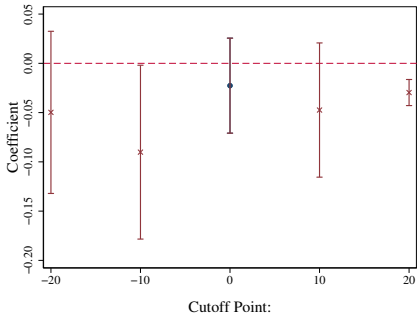
(c) Health Insurance (1)

Notes: This figures plots the discontinuity at the boundary. The dependent variable in panel (a) is an indicator variable that takes the value of one for individuals that know how to read and write, in panel (b) it is the total years of formal education, and panel (c) it is an indicator that takes the value of one for individuals with health insurance. The points represents the average value of the outcome variable in bins of width of 1.4 km. The regressions are estimated using local linear polynomials in the outcome of interest estimated separately on each side of the border within a fixed bandwidth of 10 km. Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. I present the corresponding estimate of τ in equation 1 in Table 5.

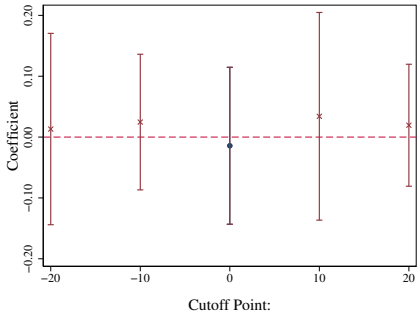
Figure D2: Spatial Placebo: Dwelling Characteristics



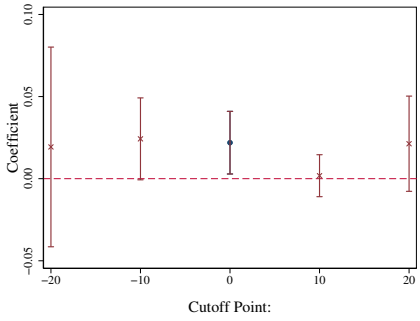
(a) Concrete Walls (=1)



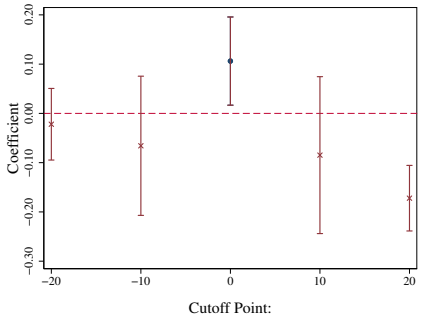
(b) Concrete Floor (=1)



(c) Electricity (=1)



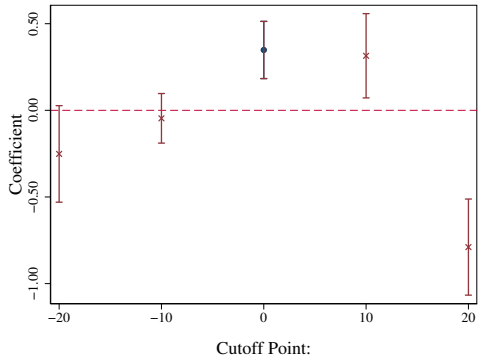
(d) Sewerage (=1)



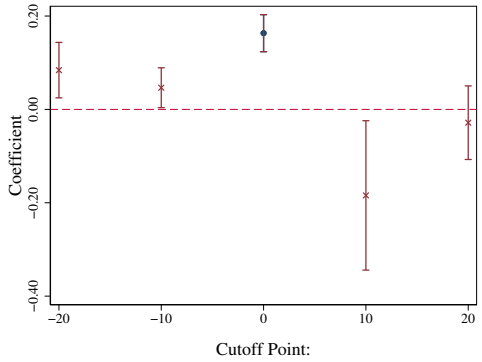
(e) Aqueduct (=1)

Notes: This figures plots the discontinuity at the boundary. The dependent variable in panel (a) is an indicator that equals one if the dwelling's walls materials are in concrete or better, while in panel (b) it is an indicator that equals one if the dwelling's floor materials are in concrete or better. Dependent variable in panel (c) is an indicator that equals one if the dwelling has access to electricity, in panel (d) it is an indicator that equals one if the dwelling has access to sewerage system, and in panel (e) it is an indicator that equals one if the dwelling has access to aqueduct system. The points represents the average value of the outcome variable in bins of width of 1.4 km. The regressions are estimated using local linear polynomials in the outcome of interest estimated separately on each side of the border within a fixed bandwidth of 10 km. Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. I present the corresponding estimate of τ in equation 1 in Table 6.

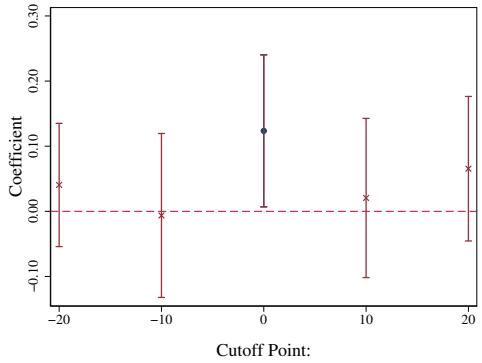
Figure D3: Spatial Placebo: Yield & Agricultural Production



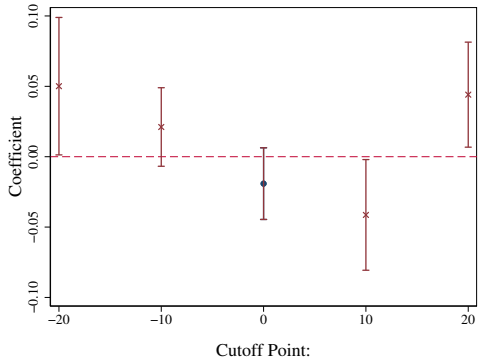
(a) Total Revenue (IHS)



(b) Revenue per Hectare (IHS)



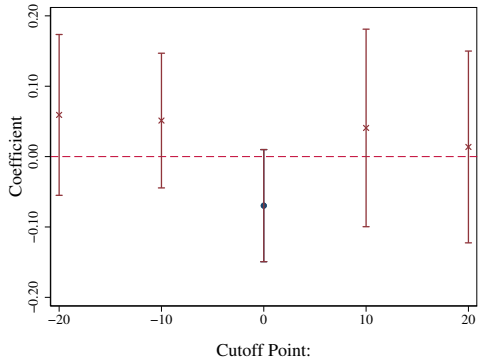
(c) Subsistence (=1)



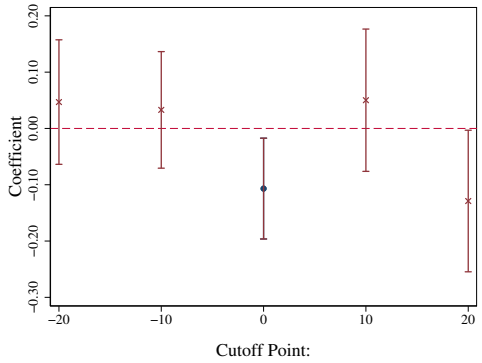
(d) Sale (=1)

Notes: This figures plots the discontinuity at the boundary. The dependent variable in panel (a) is the inverse hyperbolic sine of the total revenue in Colombian 2013 millions of COP, while in panel (b) it is the inverse hyperbolic sine of the revenue per hectare in Colombian 2013 millions of COP. Dependent variable in (c) is an indicator that equals one if the farms agricultural production is used only for self-consumption, and in panel (d) it is an indicator that equals one if the farms agricultural production is used only for market sale. The points represents the average value of the outcome variable in bins of width of 1.4 km. The regressions are estimated using local linear polynomials in the outcome of interest estimated separately on each side of the border within a fixed bandwidth of 10 km. Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas.I present the corresponding estimate of τ in equation 1 in Table 7.

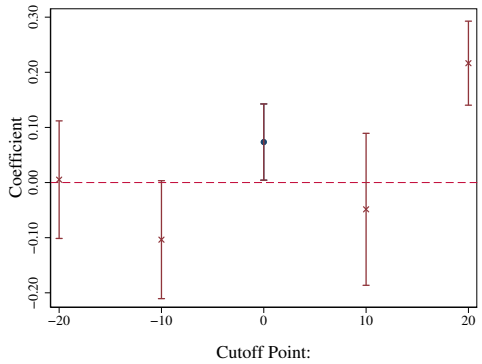
Figure D4: Spatial Placebo: Crop Specialization



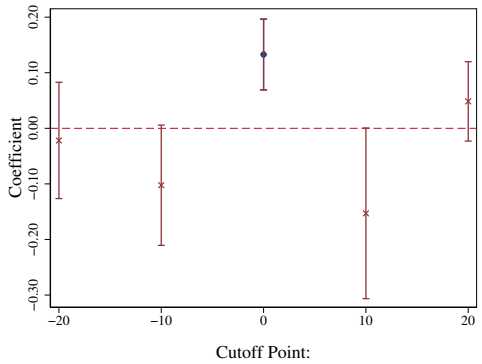
(a) Cash Crops - Area Share



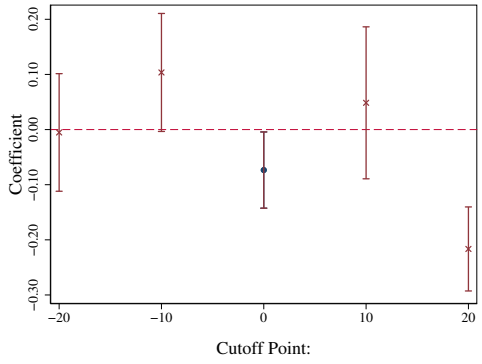
(b) Cash Crops - Revenue Share



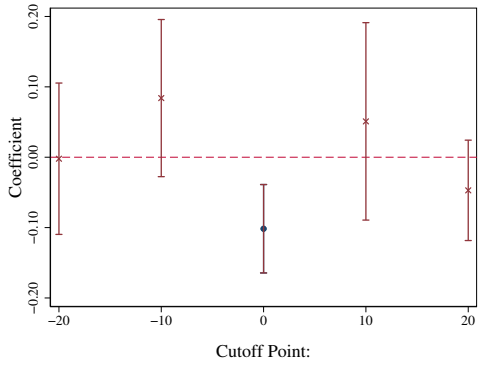
(c) Perennial Crops - Area Share



(d) Perennial Crops - Revenue Share



(e) Transitory Crops - Area Share



(f) Transitory Crops - Revenue Share

Notes: This figures plots the discontinuity at the boundary. The dependent variable in panel (a) is the share of area with cash crops, while in panel (b) it is the share of revenue from cash crops. The dependent variable in (c) is the share of area with perennial crops, while in panel (d) it is the share of revenue from perennial crops.The dependent variable in panel (e) is the share of area with transitory crops, while in columns panel (f) it is the share of revenue from transitory crops. The points represents the average value of the outcome variable in bins of width of 1.4 km. The regressions are estimated using local linear polynomials in the outcome of interest estimated separately on each side of the border within a fixed bandwidth of 10 km. Sample is restricted to border segments of San Vicente, La Macarena, Vistahermosa, and Mesetas. I present the corresponding estimate of τ in equation 1 in Table 8.