Voting with One's Neighbors: Evidence from Migration within Mexico*

Frederico Finan† Enrique Seira‡ Alberto Simpser§

November 6, 2020

Abstract

We study how proximate neighbors affect one's propensity to vote using data on 12 million registered voters in Mexico. To identify this effect, we exploit idiosyncratic variation at the neighborhood block level resulting from approximately one million relocation decisions. We find that when individuals move to blocks where people vote more (less) they themselves start voting more (less). We show that this finding is not the result of selection into neighborhoods or of place-based factors that determine turnout, but rather peer effects. Consistent with this claim, we find a contagion effect for non-movers and show that neighbors from the same block are much more likely to perform an electoral procedure on the same exact day as neighbors who live on different blocks within a neighborhood.

Keywords: micro-neighborhoods, peer effects, voting registration, voting in developing countries, turnout.

---

*We thank Cristian Challu and Andres Santos for excellent research assistance, and INE's Alejandro Andrade, Raul Caballero, Roberto Heycher Cardiel, and Giovanni Ybarias for their work in providing us with the raw data. We also thank John Marshall and seminar participants at the American Political Science Association Annual Meeting, the Society for Institutional and Organizational Economics Annual Meeting, and ITAM for valuable comments and suggestions.

†UC-Berkeley, ffinan@berkeley.edu

‡Seira: ITAM, enrique.seira@gmail.com.

§Simpser: ITAM, alberto.simpser@itam.mx.

1 Introduction

The San Diego neighborhoods of Kensington and Teralta are located right across the street from each other. In Kensington, 62% of its residents turned out to vote during the 2014 general election, while in Teralta, turnout was only 27%.[1] Such differences across neighborhoods in political behavior are observable wherever one looks. But why they exist remains unclear. One possibility is that like-minded people, with similar voting propensity, want to live in the same neighborhoods. After all, Teralta's residents are less white, less wealthy, and more likely to be foreign-born than Kensington's. But an alternative explanation is that the neighborhood itself affects turnout choices directly either through peer effects, its physical attributes, and/or its other economic and political forces.

This study is one of the first to systematically document neighborhood effects in voting behavior at the block level, and the first one to do so outside of the United States. Our analysis relies on uniquely granular data that describe individual-level turnout choices and place of residence across two recent elections for over 12 million Mexican citizens. Because of the fineness of these data, we are able to isolate the influence of neighbors who reside on the same *block*. This stands in contrast to a voluminous literature that has had to define a neighborhood using substantially larger geographical designations, such as the county, the commuting zone (Chetty and Hendren, 2018a), or the census tract (Kling et al., 2007). By focusing on the block, we not only obviate the potential identification concerns associated with sorting, but we also provide valuable insights into one of the potential mechanisms for why neighborhoods matter.

To estimate the effect of an individual's block neighborhood on her turnout, we focus on citizens who move from one place of residence to another. We find that moving to a block with higher (/lower) turnout is associated with a significant increase (/decrease) in the probability of turning out to vote. Specifically, a mover's probability of turning out to vote at destination changes by about half a percentage point for every 10 percentage point difference in average turnout between the origin and destination blocks.

By zeroing in on block-level variation within an electoral precinct, we can control for the confounds associated with geographical sorting. Precincts in Mexico encompass a median of only 19 blocks and are much smaller geographical units than the commonly-used definition of a neighborhood.[2] Our identification strategy assumes, we believe quite reasonably, that while people are generally able to pick a neighborhood to move to, the precise block within

[1] ["San Diego Neighborhoods Close in Distance, Miles Apart in Voter Turnout," Feb 10 2016, KPBS, www.kpbs.org, Claire Trageser and Megan Burks] Trageser and Burks: https://www.kpbs.org/news/2016/feb/10/san-diego-neighborhoods-close-distance-miles-apart/; other examples: https://www.wbez.org/shows/wbez-news/chicago-elections-mapped-voter-turnout-high-but-low-in-minority-neighborhoods/38728dc7-85a8-41b6-849a-bf8888a715b2

[2] Geographical units commonly used in studies of neighborhood effects in the U.S. include the census tract and the block group. A census tract is a set of block groups. On average, block groups contain about 39 blocks.

that neighborhood where they end up living is essentially the result of idiosyncratic factors such as the availability of housing units at a particular point in time.[3] Consistent with this identification assumption, we document that a mover's past voting behavior is not correlated with average turnout at her destination block. We also show that the correlation of an individual's traits and the average traits of their neighbors is practically zero once we account for origin-destination precinct pair fixed effects.

To investigate the channels underlying our estimated effect, we classify potential mechanisms into three categories: physical and infrastructural factors that affect the costs of voting, such as distance from home to polling station; political campaigns and clientelistic mobilization; and interpersonal influence (peer effects). Because our measure of exposure to one's neighbors behaviors varies at the block level, our empirical strategy in effect rules out any potential forces such as political campaigning, turnout buying, and clientelism, which by all accounts operate at geographical levels no smaller—and often much larger—than the precinct (Levitsky, 2003; Larreguy et al., 2016, 2017). In terms of a neighborhood's physical layout, we are able to control for among other things, distance from one's home to polling station. We find that while distance to the precinct—presumably an important driver of the cost of turning out to vote—significantly influences a mover's turnout probability, its inclusion does not affect our estimated effect.

Our results point to interpersonal influence as an important driver of neighborhood effects in voting. To further test for evidence of peer effects, we implement two additional identification strategies that differ from our main analysis. First, we study what happens to the turnout of residents at the destination block when others move into that block. In contrast with our main analysis, this strategy estimates the effect of movers' past turnout on the future turnout of non-movers at destination. We find that the arrivers' propensity to vote in the past increases the probability the their new neighbors at the destination block turn out, which is consistent with peer effects.

Second, we test for interpersonal influence relating to elections prior the act of voting itself. Specifically, we ask whether two citizens who live on the same block are more likely to update or renew their voter ID *on the same day* than two citizens who live on different blocks but in the same neighborhood. Using data on the universe of ID renewals and updates in the Mexico, we find that this is indeed the case. We take this as a strong indication that people interact with their neighbors about election-related matters in daily life.

This paper's main contributions are, first, to furnish credible evidence of block-level neighborhood effects in political participation in the developing world, where social norms,

[3] Bayer et al. (2008) adopt a similar identification strategy to estimate effect of social interactions among neighbors on labor market outcomes.

social networks, formal institutions, infrastructure, and political campaigns—ostensibly the main candidate factors underpinning neighborhood effects—often differ importantly in comparison with the United States and other wealthy Western nations. In the 2018 campaign season in Mexico, for example, close to one in three Mexicans talked to their neighbors about the elections sometimes or often, while only one in ten Brittons and fewer than one in five Americans, did.[4]

More generally, we document a negative relationship between a country's level of development and the fraction of its citizens who report speaking to their neighbors about elections or electoral campaigns (Figure 4). One hypothesis consistent with this relationship is that when formal institutions are weaker people have incentives to rely on, and invest in, social ties more than in societies with stronger and more reliable formal institutions. Whatever the reason, the aforementioned empirical pattern suggests that peer effects may matter more in the global South than in the global North, underscoring the importance of extending research on neighborhood effects to less-developed settings.

Second, our analysis speaks to the mechanisms responsible for neighborhood effects in political participation, by disentangling peer effects from the effects of local infrastructure and services, all the while ruling out campaign advertising, clientelism and other factors relevant to political participation that largely operate at levels of aggregation much greater than our geographical unit of analysis, the block.

Existing literature on neighborhood effects has focused almost exclusively on the United States context. A lot of this literature studies income, education, and health outcomes of the poor (Sampson et al., 2002). The study of neighborhood effects on political behavior has largely concerned itself with independent variables such as neighborhood levels of education, the distribution of partisan preferences, and the distribution of racial and ethnic traits (Gimpel et al., 2004; Huckfeldt and Sprague, 1987; Johnston et al., 2004; Baybeck and McClurg, 2005; Beck et al., 2002; Cohen and Dawson, 1993; Kenny, 1992; Huckfeldt et al., 1993; Rolfe, 2012). This literature has argued that large, politicized networks foster participation (McClurg, 2003; Kenny, 1992), that the presence of "political experts" in a network increases participation (McClurg, 2006), and that political disagreement depresses participation (Mutz, 2002). Both descriptive and injunctive social norms have been shown to influence individual political participation (Shulman and Levine, 2012; Gerber and Rogers, 2009). Barber and Imai (2014) study block-level influences on individual turnout in three US states, but their substantive focus is on racial and partisan neighborhood composition as potential drivers of participation. They aim to discern between competing psychologi-

[4]Authors' calculations based on an identical question asked in the Comparative National Elections Project (CNEP) surveys for Mexico (2018), Great Britain (2017), and the United States (2012).

cal theories about the relationship between turnout and neighborhood composition, finding support for the theory of voter empowerment but not for threat theory. Cantoni and Pons (2020) likewise investigate the influence of geographical context on individual turnout. In contrast with our focus on the very local, their geographical unit of analysis is the US state. Sinclair (2012) sets out to distinguish between two mechanisms of interpersonal influence: information vs. social pressure to conform. Her findings are complementary to ours. Collectively, the literature suggests that interpersonal influence is likely to account, at least partly, for neighborhood effects.

Randomization-based research on peer effects in turnout, meanwhile, is once- or twice-removed from the real-world experiences of citizens in their neighborhoods. Nickerson (2008), for example, randomly encouraged one of the adults in a two-adult household to turn out on election day, finding a spillover effect on the turnout of the other adult in treated households. Gerber et al. (2008) used randomly assigned mailings threatening social shaming upon failure to turn out to show that the threat increases turnout. Klofstad (2015) studied the influence of political discussion with randomly-assigned college roommates on turnout. While fascinating in their own right, these treatments were not designed to directly probe neighborhood effects. Our data, in contrast, reflect the experiences of large numbers of citizens "in the wild."

2 Context and Data

Mexico has had transparent and competitive national elections at least since 1997, following an important set of legal reforms that created an electoral authority independent of the executive branch (the *Instituto Federal Electoral* or IFE, now called *Instituto Nacional Electoral* or INE) and a transparent and reliable list of registered voters, among others improvements. INE—like IFE before it—is in charge of organizing elections for all national legislative and executive offices. INE's tasks include, among other things, continuously updating the voter registry, determining the locations of polling stations, assigning registered citizens to polling stations, and recruiting teams of citizen volunteers to staff the polling stations and count the votes. Elections for both houses of national congress take place every 3 years, and presidential elections take place every 6 years.

The basic unit of Mexico's electoral geography is the precinct. The number of precincts in the 2015 national elections was 66,224. Every precinct contains one or more polling stations, depending on the number of voters registered in the precinct. The median number of blocks per precinct in 2012 was 19 blocks. Citizens are assigned to vote at a specific precinct according to their place of residence.[5] All citizens 18 years of age and over are eligible to vote if they have registered to do so. In order to cast a vote on Election Day, citizens must show a current INE ID at the polling station. As in most democracies, failing to turn out to vote on Election Day in Mexico does not result in any sort of penalty.

In this paper, we use data from the presidential and legislative elections of 2012, and the legislative elections of 2015. For purposes of this study, we define a neighborhood as a precinct and a micro-neighborhood as a block. Our data set compiles information from the following sources:[6]

Voter list (*padrón electoral*). The voter registry is a cornerstone of Mexico's electoral system and INE devotes substantial resources to keeping it accurate and up to date. Every registered citizen is issued an INE ID card. Because the INE ID is, de facto, the main form of official identification in Mexico for both private-sector transactions and bureaucratic procedures, citizens have an incentive obtain it and to keep its information up to date. INE estimates that upwards of 97% of citizens 18 and over have a valid INE ID card. To obtain an INE ID, citizens have to present documentary evidence at an INE office, including proof of address, proof of birth, and photo ID. Citizens also report their educational attainment and occupation. This data source contains an up-to-date list of all citizens legally registered to vote in a particular election, recording, for every citizen, their gender, age, occupation, educational attainment, home address, and assigned precinct. The 2012 voter list contains information on 84,464,713 citizens, and the 2015 list contains 87,243,321 citizens.

Place of residence. We do not observe the exact home address of citizens. However, we do have data on the coordinates for the centroid of the block of their home residence, as well as the coordinates for the location of their assigned polling station. On the basis of these data, we are able to construct the linear distance between home and polling station. The mean distance is 0.4 km with a standard deviation of 0.56. In addition, we have a household identifier variable that takes the same value for all citizens who reside at the same address. We use this variable as a proxy for "family." The average "family" size is 4.7 adults, compared to 3.7 adults in the 2015 mid-census survey.

Turnout census. INE compiles a list of all registered voters who turned out to vote in each federal election. These data come from the precinct-specific booklet listing those citizens registered to vote at that precinct. When a citizen shows up at the polling station,

[5] There are some exceptions (for example, citizens who are traveling can vote in special precincts), but in practice these only apply to a small fraction of citizens.

[6] We never had access to personally identifiable information. The data belong to INE and we are not at liberty to post or share them.

poll workers mark that citizen in the booklet. INE then digitizes the booklets for all polling stations. We had access to the full turnout census for the 2015 elections, as well as to a random sample of about 13,554,266 citizens for the 2012 elections.[7]

ID registration, replacement, renewal, and updating procedures. Citizens can replace a lost card, renew an expired one, or update their home address, at any INE office at any time. In the year leading to the 2015 national elections (June 2014 to May 2015), citizens performed more than 11,895,125 procedures relating to their INE IDs at INE offices. We have transaction-level data describing the date of a procedure, the type of procedure, and the (anonymized) identity of the citizen who performed it.

Socio-demographic information. The national statistical agency, INEGI, together with INE provide a version of the 2010 Population Census where the data are presented at the precinct level. These data cover close to 67 thousand precincts. We use the following variables to check for balance and as control variables: (i) percentage of the population constituted by: men, individuals residing in the same state where they were born, indigenous, catholic, without social security, employed; (ii) average years of education; and (iii) percentage of households with: water, sewage, dirt floor, electricity, radio, TV, refrigerator, car, computer, telephone, cell phone, and internet service.

2.1 Analysis sample

Our main analysis studies the voting behavior of registered citizens who move across blocks between the 2012 and the 2015 elections. We refer to such citizens as "movers." To create our analysis sample, we first merge the different data sources, and then drop observations with missing data for any of the regressors in the main analysis. Our merged dataset consists of all observations for citizens that meet the following criteria: (i) they are present in both the 2012 and 2015 voter lists, (ii) the data contain their block of residence in 2012 and 2015, (iii) there is information on their turnout behavior from the turnout census of 2015 and the 2012 turnout census sample, and (iv) the 2010 Population Census covers both their origin and destination precincts.[8] These data encompass 11,540,922 registered voters in 66,119 precincts. Of these, approximately 1.2 million citizens are movers as previously defined. By using an anonymous individual identifier, we are able to construct a two-period panel of citizens.

[7] This is before merging with the other data sources.

[8] For simplicity, we exclude from our sample the small subset of precincts that were reshaped or partitioned as part of INE's ongoing process of *reseccionamiento* between the 2012 and 2015 elections.

For the main analysis, we work with a subsample of the analysis sample that contains information for all the variables in the regression with the largest subset of controls (subsequently the "estimation sample"). Thus, the estimation sample is consistent across our different econometric specifications. In the end, our estimation sample consists of 515,362 movers, each of whom we observe both in 2012 and 2015.[9] In the Appendix, we present results with maximal samples for each of the specifications in our main analysis.

2.2 Descriptive statistics

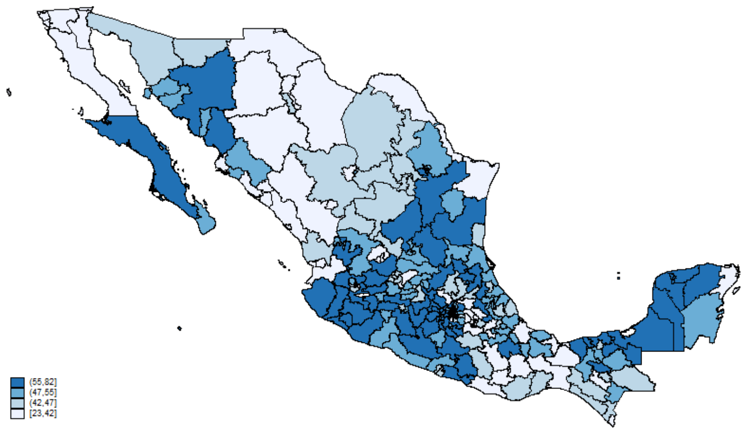
Variation in turnout. Turnout rates in Mexico exhibit substantial geographical variation. Figure 1(a) illustrates variation in turnout across electoral districts in the 2015 national legislative elections, with shading denoting quartiles of district-level turnout. Figure 1(b) zooms in on the state of Veracruz as an example, now displaying turnout rates at the precinct level. The shading indicates quartiles of precinct-level turnout rates. Note that turnout can vary substantially even across adjacent precincts. The standard deviation of turnout across all precincts in the 2015 election was 13.4 and the inter-quartile range was 17.2.

Turnout rates also vary across blocks within a precinct: the within-precinct standard deviation in block-level turnout is 0.30, while the equivalent between-precinct figure is 0.22. In cross-national perspective, the geographical variation in Mexican turnout appears to be quite typical. Table 1 provides information on cross-sectional variation in turnout across localities in all the countries for which we were able to obtain such information.

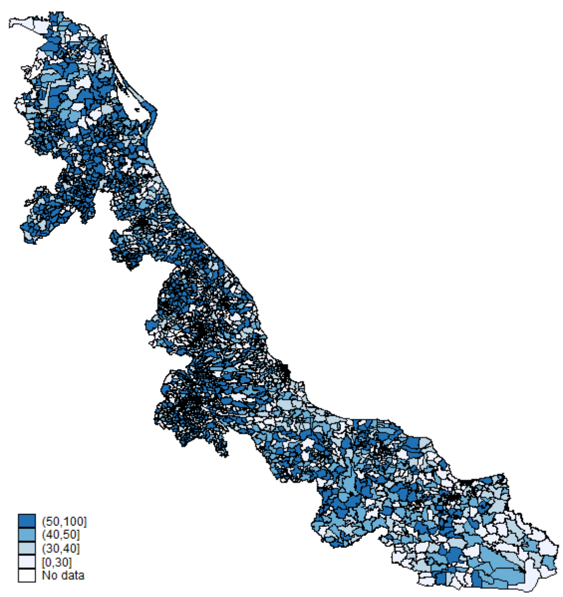
Our main explanatory variable is the difference in average turnout between the destination block in 2015 and origin block in 2012, for an individual mover. This is a measure of individual exposure to differences in the turnout environment of their micro-neighborhoods at origin vs. destination. Figure 2 displays the distribution of this variable. There is considerable variation in exposure across individual movers.

[9] We lose a total of 680,420 observations due to list-wise deletion of missing data. Of these, about 54 thousand are due to missing turnout either at origin or destination; about 191 thousand are due to missing data on educational attainment; about 181 thousand have missing data on household size, about 196 thousand lack distance to polling station, and about 67 thousand are missing block-level controls.

Figure 1: Variation in turnout across Mexico



(a) 2015 District turnout



(b) 2015 Precinct turnout

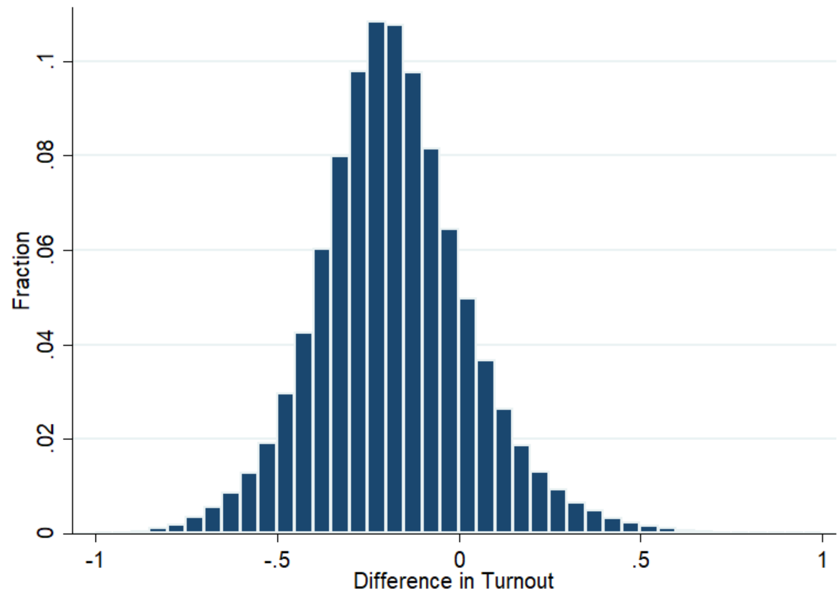
This figure shows the extent of variation in 2015 turnout. Panel (a) does this for the entire 300 Electoral Districts in Mexico. Panel (b) takes the state of Veracruz as an example and plots turnout at the precinct level. Each color represents a different quartile of the empirical distribution of turnout.

Table 1: **Variation in turnout across countries**

Country	Geographic Unit	Turnout	% of Elec./Unit	SD	IQR	N	Year	Type of Election
USA	Precinct	66.1	0.00%	12.6	14.3	84441	2012	Presidential
Chile	Comuna	41.1	0.29%	12.0	10.1	346	2013	Presidential
France	Departament	78.9	0.93%	10.0	3.8	107	2012	Presidential
Mexico	**Poll Station**	62.8	**0.00%**	9.6	12.1	**136128**	**2012**	**Presidential**
Costa Rica	Electoral District	66.5	0.21%	8.3	10.9	478	2014	Presidential
Panama	Electoral Circuit	77.9	2.56%	6.6	12.1	39	2014	Presidential
USA	State	61.3	2.38%	6.4	8.9	42	2012	Presidential
Mexico	**Electoral District**	62.7	**0.33%**	6.0	7.1	**300**	**2012**	**Presidential**
Canada	Federal Electoral Distrits	61.2	0.07%	5.7	7.2	302	2011	Parliamentary
South Africa	Province	71.7	0.43%	5.4	6.8	234	2014	Presidential
United Kingdom	Constituency	66.1	0.15%	0.5	7.8	650	2015	Parliamentary

This table shows various summary statistics of voter turnout. The statistics are computed across geographical units for different countries and elections and sorted by the standard deviation. Data for France obtained from The Guardian's France election results 2012 datablog <https://www.theguardian.com/news/datablog/2012/may/07/france-election-results-list#data>. Data for Panama obtained from the country's Electoral Tribune's 2014 elections results web page <https://www.tribunal-electoral.gob.pa/eventos-electorales/elecciones-generales-1994-2019/elecciones-2014/resultados-electorales-2014/>. Data for Chile obtained from Chile's Electoral Service's 2013 presidential elections results from their historical records website <https://historico.servel.cl/servel/app/index.php?r=EleccionesGenerico/Default/MesasElectores&id=2&n=2&v2=30&v3=0&v4=0&v5=0&v6=0&v7=0&v8=0>. Data for the United Kingdom obtained from the UK's Electoral Commissions electoral data <https://www.electoralcommission.org.uk/our-work/our-research/electoral-data/electoral-data-files-and-reports>. Data for Costa Rica obtained form CR's Supreme Electoral Tribune's statistics of election processes <http://www.tse.go.cr/estadisticas_elecciones.htm>. Data for Canada obtained from Elections Canada's 2011 general election official voting results raw data <http://www.elections.ca/content.aspx?section=res&dir=rep/off/41gedata&document=byed&lang=e>. Data for the United States obtained from United States Elections Project's 2012 General Election turnout rates data at the state level <http://www.electproject.org/2012g>and Harvard Election Data Archive's 2014 election data at the precinct level <https://projects.iq.harvard.edu/eda/data>. Data for South Africa was obtained from the Electoral Comission of South Africa's national and provincial election results <http://www.elections.org.za/content/Elections/National-and-provincial-elections-results/>.

Figure 2: Distribution of exposure to destination-origin block turnout differences



This figure plots the differential exposure of movers, defined as the difference in block-average turnout for an individual mover at the destination block in 2015 minus that at the origin block in 2012. This variable is bounded in [-1,1]. Because turnout is greater in presidential elections (such as the 2012 one) than in legislative ones (such as the 2015 election), the differences variable is negative in most cases.

Internal Migration. As mentioned previously, we observe the block of residence for registered citizens in 2012 and in 2015. In the merged dataset, 10.4% of citizens changed their home address between 2012 and 2015. Of these, 21.2% moved out of state, 21.7% moved within a state across municipalities, 30.8% moved within municipalities across precincts, and 26.3% moved across blocks within their precinct.[10] These movers originated in 20,151 precincts and arrived in 61,950 precincts. The origin precincts with the most movers saw 403 individuals move out. The 2000th-ranked precinct saw about 80 individuals move out. In terms of origin-destination precincts *pairs*, the pair with the highest flow contains 125 movers, while the 2000th-ranked pair contains 8.

In Table 2, we compare the characteristics of non-movers (column 1) to those of movers in the estimation sample (column 2). While both movers and non-movers are about 47% male, movers are more schooled (8.3% have a bachelor's degree vs 5.8% of non-movers, and 17.5% claim to be studying vs 12.7% of non-movers). Movers are also younger on average (34 years of age vs. 41 for non-movers). Movers are also less likely to have voted in 2012.

[10] Figures for the main estimation sample are similar.

Table 2: **Summary statistics for non-movers and movers**

	Non movers	Movers in estimation sample		
		All	Q1 to Q4	Q4 to Q1
	(1)	(2)	(3)	(4)
Voted in 2012 at origin (%)	65.88	54.40	45.25	52.57
Male (%)	47.91	47.45	50.17	49.91
Age (years)	40.82	34.37	33.82	33.04
High school (%)	16.49	22.53	28.35	26.87
Bachelor (%)	5.83	8.27	11.68	10.10
Domestic work (%)	31.26	24.51	19.98	21.36
Employee (%)	28.48	35.91	39.36	33.04
Student (%)	12.67	17.53	19.67	24.04
Household size at origin	4.58	4.07	3.98	4.47
Distance to polling station at origin (km)	0.40	0.38	0.45	0.37
Number of observations	10,333,006	515,362	6,285	10,130

Figures are means of 2012 variables for different subsamples of the analysis sample. Column 1 refers to all non-movers in the analysis sample. Column 2 contains all movers in the estimation sample for the main analysis (note that listwise deletion of missing data drops 629,567 movers from the analysis sample). Column 3 describes those movers in the estimation sample who moved from blocks located in precincts in the lowest precinct-turnout quartile in 2012 to blocks located in precincts in the highest precinct-turnout quartile in 2015, while column 4 describes those movers in the estimation sample who moved from blocks located in precincts in the highest precinct-turnout quartile in 2012 to blocks located in precincts in the lowest precinct-turnout quartile in 2015.

The rightmost two columns of Table 2 compare two subsets of movers in our main estimation sample: those who moved from the lowest quartile to the highest quartile of precinct-level turnout (column 3) vs. those who moved from the highest to the lowest quartile (column 4). These two groups are quite similar in age and gender, but those moving to higher-turnout precincts are more likely to be employees and less likely to be students. They also voted at somewhat lower rates in 2012 in comparison with those who moved from high- to low-turnout.

3 Empirical Approach

Our main analysis aims to identify the influence of one's close neighbors (i.e., those residing on the same block) on one's decision to turn out to vote. To estimate the causal effect of neighbors on voting, we exploit the granularity of our data to isolate *block-level* variation in turnout within precincts. Specifically, we estimate the following econometric model:

$$Vote_i^{2015} = \alpha + \beta \Delta Turnout_{o(b),d(b)} + \phi Vote_i^{2012} + \delta X_i + \eta Z_{d(b)} + \theta_{o(p),d(p)} + e_{iod}. \quad (1)$$

The dependent variable, $Vote_i^{2015}$, is an indicator for whether person i living in destination block b voted in 2015. Our key independent variable, $\Delta Turnout_{o(b),d(b)}$, is the difference in average turnout between the destination block $d(b)$ in 2015 minus the origin block $o(b)$ average turnout in 2012.[11] Equation 1 also includes an indicator, $Vote_i^{2012}$, for whether the mover i voted in 2012. This is an important control because one's propensity to vote tends to be stable across time. Our model also accounts for a set of individual controls, X_i – including the person's age, gender, household size, education level and occupation both in 2012 and 2015, distance to the polling station in 2012 and 2015, moving distance, the number of voter IDs the person has held, and an indicator for whether the move included crossing a state boundary. It also controls for 11 destination block variables $-Z_d(b)-$ described in Table 3 containing the average characteristics of block-level neighbors. The model also includes precinct fixed effects—in our most demanding specification, *precinct-pair* fixed effects $\theta_{o(p),d(p)}$. The fixed effects imply that we compare individuals who moved from and to the same precincts, but happen to live in different blocks. Our key identifying assumption is that, conditional on the precinct (and our rich set of controls), movers do not sort across blocks on the basis of their block neighbors' potential turnout.

3.1 Identification strategy

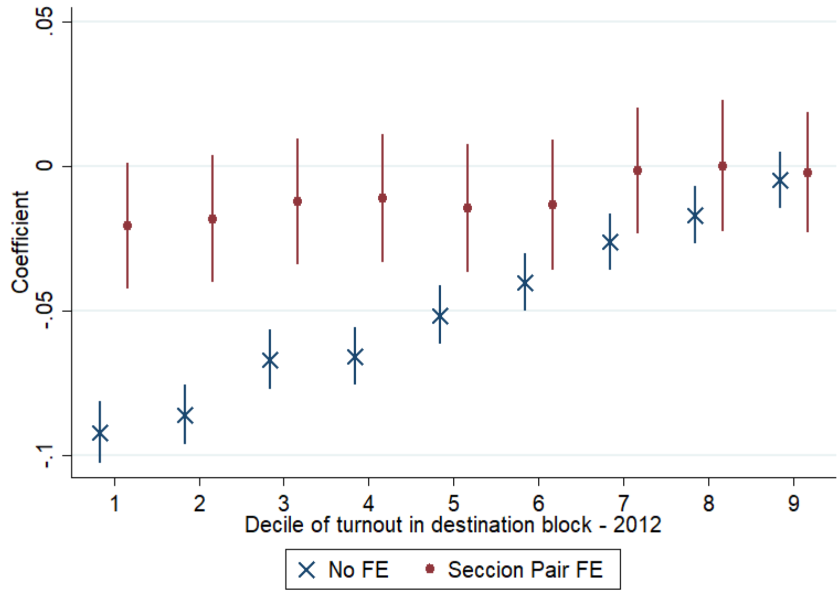
The main concern with using movers as a source of identifying variation is that people do not choose where to live at random. It is possible, for example, that individuals who are more likely to vote also tend to move to places where turnout rates are higher. In order to deal with this concern, we only make use of variation in turnout across micro-neighborhoods (blocks) within a given neighborhood (precinct), discarding all variation in turnout across neighborhoods (precincts). Accordingly, our identifying assumption is that, within a given precinct, movers do not sort across blocks on the basis of their neighbors' turnout or factors correlated with turnout.

We believe this is a plausible assumption for at least two reasons. First, the housing market within a precinct is very thin—the median precinct is comprised of only 19 blocks — and thus limits an individual's ability to choose a specific block. During a 3-month period, on average only about 0.5% of houses/apartments in a block become available in Mexico.[12] This is less than half than what Bayer et al. (2008), which employs a similar research design, finds for Boston. Second, before they move, individuals are unlikely to know about turnout rates for specific blocks or about the characteristics of their future neighbors on a given block, as compared to other blocks *within* a neighborhood.

[11] The fact that average turnout at the national level is different across these two elections is not an issue for our empirical strategy since identification relies on comparison across individuals, not across elections. Nevertheless, we construct an alternative measure of exposure based on relative turnout in each of the elections, which we describe later in the paper.

[12] Authors' calculations based on INE data.

Figure 3: Selection of movers on turnout at the destination block



We estimate a regression where the dependent variable is an indicator for whether an individual mover voted in 2012 (before moving) and the explanatory variables are indicators for deciles of block-level 2012 turnout at the destination block (with the 10^{th} decile as the omitted reference category), plus the same controls as the model described in column 3 of Table 4. Standard errors are clustered at the origin precinct level. The analysis is run on the estimation sample (i.e., the same used in Table 4 further below). The figure plots the estimated coefficients on the destination-block turnout indicators, together with 95% confidence intervals. We estimate one specification without precinct fixed effects (cross markers) and one specification with origin-destination precinct pair fixed effects (dot markers). We test the hypothesis of equality of the coefficients on the decile indicator variables, and while we reject equality for the specification without precinct fixed effects (p-value = 0.00), we cannot reject equality for the specifications with precinct pair fixed effects (p-value=0.60).

The data support our identification assumption. Figure 3 shows that, while people who voted in 2012 moved to blocks with higher average turnout than those who did not vote (star markers), that correlation disappears once precinct fixed effects are controlled for (dot markers). To construct the graph, we first calculated deciles of block-level turnout for all destination blocks in 2012 and created dummy variables for each decile. We then regressed an indicator for whether a mover voted in 2012 on these dummies. Figure 3 plots the coefficients on each of the dummy variables (except for the reference category, which is the 10^{th} turnout decile), both with and without precinct fixed effects. We cannot reject the hypothesis that the coefficients on the dummies are equal to each other in the specification with precinct fixed effects. In other words, movers appear to sort according to neighborhood (precinct) turnout, but they *do not* sort by turnout across micro-neighborhoods (blocks) within a neighborhood.

We further test our identification assumption by studying whether individuals select into residential blocks populated by people with similar socio-demographic traits. Even if individuals do not care about turnout, they could potentially sort on factors correlated with turnout. Specifically, we compute correlations between the characteristics of an individual and the characteristics of his or her neighbors living in the destination block (Table 3). For each individual in our sample, we identify all registered voters on her destination block, excluding the individual's household members, and calculate their average characteristics: age, family size, gender, education level, propensity to vote, and self-reported occupation—student, employee, homemaker, or worker. We report the mean of these characteristics in column 1. We next compute the correlation between the individual's trait X and the average trait of her non-family block neighbors, $\overline{X}$, by running the following regression: $X_{i,j} = \alpha + \beta \overline{X}_{i's\, block\, neighbors,j} + \epsilon_j$, where i is a randomly-picked citizen living in block j.[13] Column 2 reports β estimates from this regression equation. In column 3, we report the results of the same analysis adding destination-precinct fixed effects to the regression. Comparing column 3 with column 2, we see that the inclusion of precinct fixed effects causes the β estimates to decrease quite substantially (between 40 and 150 percent).

Table 3: **Selection on block-level socio-demographic characteristics**

	Individual mean (1)	Unconditional (2)	Section FE (3)
Age 45-49	0.160	0.019	0.007
Age 35-44	0.239	0.029	0.017
Age 25-34	0.338	0.035	0.017
Male	0.448	-0.002	-0.006
Household size	4.45	0.088	-0.043
High-school graduate	0.204	0.329	0.131
College graduate	0.067	0.361	0.171
Housework	0.411	0.114	0.014
Employee	0.218	0.308	0.035
Student	0.053	0.053	0.013
Self-employed	0.066	0.113	0.025
Predicted probability of voting	0.431	0.064	0.016

The table (specifically, columns 2 and 3) displays conditional correlations between the demographic characteristics of a randomly picked individual (one per block) and those of her non-family block neighbors. The socio-demographic variables are listed in the rows of the table. They include indicators for age group, an indicator for being male, the number of household family members (proxied by the number of registered voters with the same address), indicators for educational attainment, and indicators for the four occupations most commonly reported to INE (housework, employee, student, and self-employed). The last line displays the probability of voting as predicted by these demographic variables (details of the prediction method are provided in the text). Column 1 displays the mean of each variable. Column 2 reports the β estimates from the regression equation: $X_{i,j} = \alpha + \beta \overline{X}_{i's\, block\, neighbors,j} + \epsilon_j$, as described in the text. Column 3 adds precinct fixed effects to this regression.

[13] The random pick is for computational simplicity.

Although the coefficients reported in column 3 are very small, they are not exactly zero. Thus, one might be concerned that these correlations, although small, might be sufficiently large to induce spurious (i.e., non-causal) correlation in voting behavior between an individual and her block neighbors. To test whether this is the case, we isolate the part of individual voting behavior that is due to the aforementioned characteristics. We then compute the correlation between the individual and her neighbors in voting as explained by these characteristics. Operationally, we regress an indicator variable for having voted in 2012 (Y_k) on the set of characteristics X_k (i.e., those listed on the first eleven lines of the table) and calculate predicted values $\widehat{Y}_k$, for all individuals k in the data.[14] We then regress the predicted values $\widehat{Y}_{ij}$ for individuals i who reside on block j on their non-household block neighbors' average predicted values: $\widehat{Y}_{ij} = \gamma + \delta \overline{\widehat{Y}}_{i's\ block\ neighbors,j} + \nu_j$, with and without precinct fixed effects. A large coefficient δ would indicate that the residual correlations in the first 11 lines of column 3 are large enough that selection is likely to remain a problem for identification.

The estimated δ coefficients are shown in the last line of columns 2 and 3. Loosely speaking, these estimates represent the degree to which the traits on the first 11 lines of the table indirectly induce residential selection on turnout. The estimate in column 3, 0.017, implies that a one standard deviation increase in an individual's block neighbors' average predicted vote is associated with a 0.02 standard deviation increase in the individual's own likelihood of voting—a very small correlation. That is, once precinct fixed effects are accounted for, block-level sorting based on socio-demographic characteristics at most can explain a tiny fraction of individual voting behavior.

3.2 Local nature of interactions

By discarding all variation at the precinct level through the inclusion of precinct-pair fixed effects, we are implicitly focusing on social interactions at a very local level. We believe we are on firm ground in supposing that those interactions matter for political behavior. Many studies have found that neighbors close by are important sources of information and constitute important contacts in an individual's social network. Lee and Campbell (1999) find that 31% of neighbors in the closest 10 housing units are judged as personally close or very close by Nashville survey respondents. Otani (1999) uses the US General Social Survey and finds that neighbors comprise about 1/5 of people listed as part of an individual's social network in the US and Japan. Huckfeldt (1980) argues that "neighbors are a potential powerful source of contextual influence on political behavior" (see also Rolfe and Chan (2017)). Survey evidence indicates that Mexican citizens speak quite often with their neighbors about

[14] We use a logit regression to maintain predicted values in the (0,1) range.

electoral campaigns. In 2018, 31% of Mexicans reported doing so either sometimes or often. Mexicans also reported speaking about electoral campaigns sometimes or often to friends (44%) and family (56%).[15]

One of the most successful interventions to date to increase turnout is by Gerber et al. (2008), and it involved sending a letter about the turnout record of neighbors, showing that people do care about opinions of their neighbors regarding voting. Relatedly, DellaVigna et al. (2017) show that people vote in part to be able to tell others that they did so, since "it is common for neighbors, friends, and family to ask whether we voted." Overall, this literature suggests that localized interactions matter for voting. This is precisely what we test in the paper. To the extent that people on blocks other than one's block of residence influence voting, our design may provide a lower bound of the effect of immediate neighbors.

4 Main Results

Our main analysis estimates the effect of exposure to different micro-neighborhoods for citizens who move from one such micro-neighborhood to another in the period 2012-2015. As mentioned previously, the exposure variable consists of the difference in average block turnout at destination minus at origin. The dependent variable is an indicator for whether the mover turned out to vote in the 2015 elections. In all specifications we control for whether the mover had voted in 2012.[16] Standard errors are clustered at the precinct level.

In column 1, we find that exposure to a higher turnout micro-neighborhood is positively correlated with one's propensity to vote. The coefficient of 0.048 implies that a one standard deviation increase in destination-origin exposure is associated with 0.1 percentage point increase in the likelihood of voting. We also find the common result that past turnout predicts future turnout (e.g., Green and Shachar (2000)). In column 2, we add individual controls. These controls are predictive of an individual's decision to turn out in the expected direction. Distance from the home block to the precinct, a driver of voting costs, is negatively associated with propensity to vote, while age is positively correlated with voting. The inclusion of these additional controls leaves our main coefficient of interest virtually unchanged at 0.043.

Column 3 augments the model with block level controls. These controls consist of the block average values of each of the 11 traits presented in Table 3. Consistent with the lack of sorting by block that we documented earlier, our coefficient of interest is unaffected by the inclusion of these block-level controls. We then add precinct-pair fixed effects (Column 4). In this specification, our most stringent, we are comparing individuals who both used to live

[15] All figures in this paragraph based on the 2018 CNEP Mexico survey.

[16] Thus, we are effectively estimating the dependent variable in changes.

in precinct A and both moved into precinct B, but happen to live in different blocks. The estimated effect of the exposure variable is even stronger, with a point estimate of 0.055, implying that a one standard deviation increase in destination-origin exposure is associated with a 1.2 percentage point increase in the mover's likelihood of voting.

Table 4: **Effect of destination-origin exposure on individual turnout of movers**

	Dependent variable: individual mover's turnout in 2015				
	(1)	(2)	(3)	(4)	(5)
Destination-origin block turnout difference	0.048***	0.043***	0.043***	0.055***	
	(0.004)	(0.004)	(0.004)	(0.011)	
Destination-origin block turnout percentile diff.					0.002***
					(0.000)
Individual mover's turnout in 2012	0.236***	0.209***	0.209***	0.253***	0.265***
	(0.001)	(0.001)	(0.001)	(0.005)	(0.005)
Male		-0.028***	-0.028***	-0.034***	-0.034***
		(0.002)	(0.002)	(0.005)	(0.005)
Age		0.004***	0.004***	0.004***	0.003***
		(0.000)	(0.000)	(0.000)	(0.000)
Household size 2015		0.004***	0.004***	0.003***	0.003***
		(0.000)	(0.000)	(0.001)	(0.001)
Distance to polling station in 2015 (km)		-0.015***	-0.014***	-0.015***	-0.013**
		(0.001)	(0.001)	(0.004)	(0.004)
N	515362	515362	515362	515362	515362
R^2	0.206	0.244	0.244	0.735	0.737
Mean of dependent variable	0.351	0.351	0.351	0.351	0.351
Mean of destination-origin block turnout diff.	-0.182	-0.182	-0.182	-0.182	-3.097
Effect per 1 SD chg. in block turnout diff.	0.00997	0.00905	0.00900	0.0115	0.0505
Individuallevel controls		✓	✓	✓	✓
Block-level controls			✓	✓	✓
Destination precinct fixed effects	✓	✓	✓		
Origin-destination precinct pair fixed effects				✓	✓

This table reports the estimated relationship between individual movers' turnout in 2015 (dependent variable) and the difference in average block turnout between the mover's destination block in 2015 and her origin block in 2012 (explanatory variable). Coefficients are estimates of: $Vote_i^{2015} = \alpha + \beta \Delta Turnout_{o(b),d(b)} + \phi Vote_i^{2012} + \delta X_i + \eta Z_{d(b)} + \theta_{o(p),d(p)} + e_{iod}$., further described in the text. Each column corresponds to a different regression model. All regressions are run on the same estimation sample. Standard errors, clustered at the level of origin precinct, are provided in parentheses below the regression coefficients. Column 1 only includes as controls the mover's past turnout (in 2012) and destination-precinct fixed effects. Column 2 adds controls for mover's age in 2012, gender, household size in 2015, distance from home block to polling station at origin in 2012 and at destination in 2015, the number of voter IDs the citizen has historically held, an indicator for whether the citizen performed any transactions related with her voter ID in the period covered by the data, an indicator for whether the citizen crossed a state boundary when moving, distance between previous (2012) and actual (2015) households, and education level and occupation in both 2012 and 2015. Column 3 adds as control variables the eleven destination block-level average neighbor characteristics described in the first eleven lines of Table 3, as well as the average block neighbors' distance to the polling station. Columns 4 and 5 add origin-destination precinct *pair* fixed effects. Column 5 defines the exposure variable (destination-origin block turnout difference) in percentiles of block-level turnout. To construct this variable, we calculate the 2015 percentile of block-level turnout for the destination block, and subtract the 2012 percentile of block-level turnout for the origin block (thus this variable potentially ranges in -100 to 100).
Significance: * p<0.05; ** p<0.01; *** p<0.001

Relative differences in turnout In addition to measuring exposure as an "absolute" difference in destination (2015) minus origin (2012) turnout, we also create a "relative" measure of exposure, based on turnout percentiles in each of the two elections in our data. Specifically, we compute for each block its percentile in the turnout distribution in a given election. We then calculate the difference in percentiles between movers' destination and original blocks. This measure washes out level differences in national turnout across election years. The results, presented in column 5, are qualitatively similar. A standard deviation (31 point) increase in the destination-origin turnout percentile is associated with around 5 percentage points increase in the likelihood that the mover votes in the 2015 elections.

Positive vs. negative exposures The estimated coefficient on the destination-origin turnout difference variable in Equation 1 corresponds to an average effect. It is possible, however, that the effect of moving to a higher-turnout block could differ from the effect of moving to a lower-turnout block. We examine the linearity of the effect in Figure OA-1, where we plot a bin scatter of the relationship between our explanatory and dependent variables after partialling out our controls (a specification similar to the one presented in column 4). We do not find evidence that the relationship is non-linear. Moreover, moving to higher turnout blocks increases the mover's likelihood of voting, while moving to lower turnout blocks lowers it.

Intensity of exposure. In Figure OA-2, we investigate whether the estimated effect varies according to how long the mover has resided in their destination block before the 2015 elections. To estimate these differential effects, we classify voters by the number of quarters that they have lived in their destination block. We then re-estimate Equation 1 separately for each group of voters. The figure displays these coefficient estimates, along with confidence intervals, by time of exposure. The effect becomes stronger with the time that the mover has spent at the destination block up until the 4^{th} quarter. After that, the effect remains fairly constant.

Homophily. We now check whether socio-demographic similarity with one's neighbors at destination enhances the effect of exposure. Specifically, we allow the effect of destination-origin turnout exposure to vary according to the similarity of the mover and her block neighbors at destination. We implement this by adding to Equation 1 an interaction term between the explanatory variable, $\Delta Turnout_{o(b),d(b)}$, and the absolute value of the difference between the mover's trait and the average of the same trait for the block neighbor's (excluding members of the mover's household). We do this for one trait at a time, for gender, age, years of education, occupation, and household size. The results, presented in Table OA-2, show that even though the differential effects all go in the expected direction (i.e., similarity strengthens the effect of the exposure), although the effect sizes are small in magnitude. For example, a one standard deviation in the difference between the age of the mover and that of her neighbors results in an increase of 0.2 percentage points in the effect of the main explanatory variable.

Clientelism. Could clientelism explain the micro-neighborhood effects on turnout that we have documented? There is, in fact, little question that clientelistic exchanges are common in present-day Mexico. Nevertheless, all the available evidence suggests that clientelism in Mexico (as virtually everywhere else) operates on geographical units much larger than the block—our geographical unit of analysis. Shefner (2001), for example, documents a squatter community near Mexico City where clientelism operates at the level of "colonia," which is much larger than precinct (and in some cases larger than a zip code). Hagene and González-Fuente (2016) similarly find clientelism at the level of community in Xico, Veracruz (population over 35,000) and in San Lorenzo Acopilco in Mexico City (population 24,000). In Xico, brokers "went to all the houses, including the remote settlements, and those of opposition" and gave "live animals, goats and chickens" (p.14). These ethnographies suggest that clientelism does not vary at the block level because blocks are only small subsets of any relevant electorate and because people share a community-wide sense of belonging, which could even make it counterproductive to treat blocks differently within one same community. Larreguy et al. (2016) and Larreguy et al. (2017) argue that brokers employed by political parties in Mexico are responsible for at least a full precinct, not a block. In this case, the argument is about the possibility of monitoring the brokers' effort: precinct level electoral results are readily available (not so at the block level). Organizational brokers, as Holland and Palmer-Rubin (2015) refer to leaders who can deliver the full set of votes of their organization's members, operate on an even larger scale than run-of-the mill party brokers. In cross-national perspective, Mexico appears quite typical in this sense.[17]

5 Additional Analyses

In this section, we present additional analyses aimed at probing the mechanism at work. Specifically, we test for two kinds of empirical implications of peer effects. First, if peer

[17] Levitsky (2003), for example, relates that clientelism in Argentina operates at the level of a "base unit" (*Unidad Básica*), which on average encompasses 1750 citizens in La Matanza and 2400 citizens in San Miguel Tucuman (p.60), and Stokes et al. (2013) exemplify the lowest-level broker in Venezuela (*jefe zonal*) as someone who was responsible for 13 blocks.

effects are at play, then block-level turnout should not only affect the voting behavior of movers, but also of the people in the blocks that received movers—although we do not expect the effect to be as large as for a mover, whose whole environment changes. Second, peer effects require that neighbors interact with each other regarding matters related to the election. We test whether block neighbors are more likely to attend an INE office to obtain or replace their voter ID on the same day than people who live in different blocks within the same precinct.

Peer effects of movers on non-movers

In this analysis, the mover's past voting behavior is the treatment variable. The dependent variable is the voting behavior of residents at the mover's destination block (who themselves are not movers). We henceforth refer to movers as "arrivers." This provides a separate test for peer effects that complements our main analysis. In this case, the only variable that changes for non-movers (whose voting behavior is the dependent variable) is the past voting behavior of the arrivers. The notional comparison is between two blocks that both received a mover, but one of the movers had voted in the past while the other had not. Also in contrast with our main analysis, this test uses data for both movers and non-movers.

We estimate the following econometric specification for "stayer" (i.e., non-mover) individuals in our sample:

$$V_{ib(d),2015} = \alpha + \beta T_{b(d),2012} + \theta X_i + \mu A_{b(o)} + \delta_d + e_{ib(d)}, \tag{2}$$

where $V_{ib(d),2015}$ is an indicator for whether individual non-mover i in block $b(d)$ voted in the 2015 elections. Our main independent variable, $T_{b(d),2012}$, measures the average 2012 turnout of all those movers who arrived in block $b(d)$.[18] The regression includes the same set of individual controls X_i that we used in the main analysis, as well as a set of controls $A_{b(o)}$ based on arrivers' traits, including census information about the arrivers' precinct of origin as well as average individual characteristics of arrivers. Finally, the regression includes destination-precinct fixed effect δ_d.

[18] For example, if 3 movers arrived in block $b(d)$, of which 2 had voted in 2012 and 1 had not, then $T_{b(d),2012} = 2/3$.

Table 5: **Effect of arrivers' past turnout on the turnout of non-movers at destination**

	Dependent variable Individual non-mover's turnout in 2015		
	(1)	(2)	(3)
Average past (2012) turnout of arrivers to the block	0.005**	0.005**	0.005**
	(0.002)	(0.002)	(0.002)
Individual non-mover's turnout in 2012	0.311***	0.311***	0.310***
	(0.001)	(0.001)	(0.001)
Male	-0.036***	-0.036***	-0.036***
	(0.002)	(0.002)	(0.002)
Age	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)
Household size 2015	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)
Distance to polling station in 2015 (km)	-0.024***	-0.024***	-0.024***
	(0.003)	(0.003)	(0.003)
N	562241	562241	529028
R^2	0.229	0.229	0.231
Mean of dependent variable	0.410	0.410	0.410
Mean of avg. past (2012) turnout of arrivers to the block	0.510	0.510	0.510
Effect per one SD chg. in avg. past arriver turnout	0.00201	0.00193	0.00192
Destination precinct fixed effects	✓	✓	✓
Non-mover individual and block level controls	✓	✓	✓
Arriver individual-level controls		✓	✓
Dropping citizens in arrivers' households			✓

This table reports regression estimates of the association between the average past (2012) turnout of movers arriving at a given block ("arrivers") and the individual 2015 turnout of non-movers at that block. We estimate OLS regressions as described by equation 2 in the text. The main explanatory variable is the average 2012 turnout of the set of arrivers to block $b(d)$ where non-mover i lives. Thus, the dataset on which these regressions are estimated contains one row per non-mover. The dependent variable is individual-level turnout for the non-mover in 2015. The equation includes destination precinct fixed effects. The set of control variables in X_i is the same as the individual and block-level controls in column 3 of Table 4, excluding distance between address in 2012 and address in 2015, and the indicator for whether a mover crossed state lines, since the analysis here focuses on non-movers. We exclude from the sample blocks where the proportion of arrivers constitutes either less than 10% or more than 90% of the population of the block. The specification in column 2 additionally includes controls for the average individual characteristics of arrivers (age and gender), the number of movers arriving to the block, and indices for the average economic, demographic, household, and education census variables corresponding to the arrivers' precincts of origin (these indices were constructed via principal-components analysis of census variables; see Appendix for details). The specification in column 3 excludes from the sample citizens residing in the same households as the arrivers. Standard errors, displayed in parentheses below the coefficient estimates, are clustered at the block level.
Significance: * $p<0.05$; ** $p<0.01$; *** $p<0.001$

We present the estimation results of Equation 2 in Table 5. In column 1, we see that the average 2012 turnout of arrivers positively correlates with the turnout decisions of non-movers in 2015. The point estimate of 0.005 (robust standard error = 0.002) implies that a one standard deviation increase in the average 2012 turnout of the arrivers is associated with an increase in a non-mover's propensity to vote of 0.2 percentage points.[19] As in our main analysis, the inclusion of precinct fixed effects strongly controls for selection of movers into destinations. Nevertheless, as additional checks we control for the number of arrivers and their average characteristics (column 2). The inclusion of these controls does not affect the point estimate. Column 3 further excludes all individuals who reside in the same household as the arriver at destination from the analysis. Again, the point estimate remains unchanged. Overall, we take these results as further evidence for peer effects.

Beyond voting: Peer effects in bureaucratic electoral procedures

We now use our data on procedures relating to voter registration and the voter ID in order to further test for peer effects. Insofar as people interact with their neighbors about voting-related matters in daily life, peer effects should be apparent not just in turnout behavior. They should also influence a person's decision of when to register to vote, renew one's voter ID, or any other related bureaucratic procedures that precede an election and are required to turn out and cast a vote.

Specifically, we test whether citizens who live on the same block are more likely to perform such procedures on the same day than citizens who live on different blocks (within the same precinct). As explained in Section 2, to conduct any of these procedures, citizens attend an INE office in person. INE offices are open throughout the year, and both appointments and walk-ins are accepted. If citizens decided independently when to go to an INE office to update or renew their voter IDs, one would not expect to observe any difference between micro-neighbors (i.e. block neighbors) and neighbors who live on different blocks within the same precinct in the likelihood of performing a procedure on the same day. Conversely, if citizens were more likely to discuss election-related matters with their micro-neighbors, then one would expect micro-neighbors to be more likely to conduct electoral procedures on the same day than neighbors on different blocks. We view this as a strong test for peer effects, as we are hard-pressed to find an alternative explanation for why, within a precinct, block neighbors are more likely to go to an INE office *on the same day* than neighbors on different blocks, other than peer effects.

We construct our estimation sample on the basis of the universe of electoral procedures in the year-or-so preceding the 2015 elections (June 2014 to May 2015).[20] We restrict the sample to registered voters who performed a procedure in this time window. We further restrict the

[19] One way to think about this estimate is as a reduced form effect, where the first stage would be the effect of having voted in 2012 on voting in 2015 for the arriver (which is approximately 0.2-0.25 on the basis of Table 4). Under this interpretation, the treatment on the treated would be approximately five times as large as the reduced form, that is, about 1 to 1.25 pp for a one standard deviation change in the 2012 turnout of arrivers.

[20] Not all electoral procedures can be performed right up to Election Day; INE sets specific deadlines for registration, renewal, and correction of voter IDs.

sample to precincts that have at least 40 such voters, living on at least two blocks, with at least two voters on each block. Within every precinct in this sample, we randomly draw 10 pairs of citizens who live on the same block and an additional 10 pairs of citizens who live on different blocks within the same precinct.[21] As constructed, our estimation sample contains approximately 970,000 pairs of citizens. We estimate the following regression:

$$Procedure_{ij} = \theta_p + \beta I(Same\ Block)_{ij} + \beta' \mathbf{X}_{ij} + e_{ijp}, \tag{3}$$

where $Procedure_{i,j}$ is an indicator that takes the value of 1 when citizens i and j, $i \neq j$, (who by construction live in the same precinct) performed an electoral procedure on the same day, and 0 otherwise. The indicator $I(Same\,Block)_{ij}$ equals 1 if citizens i and j lives on the same block according to the 2015 voter list. We include precinct fixed effects θ_p, which absorb all cross-precinct variation.

Table 6: **Neighbor influence in electoral procedures**

	All procedures	Enrollment	Replacement
	(1)	(2)	(3)
Same block	0.012***	0.007***	0.013***
	(0.000)	(0.001)	(0.001)
Precinct fixed effects	✓	✓	✓
With controls	✓	✓	✓
Observations	972619	214072	530010
R^2	0.109	0.094	0.144
Mean number of same-day procedures for non-block neighbors	0.018	0.016	0.020

This table reports estimates of equation 3. We cluster standard errors at the precinct level. Column 1 reports results using all procedures performed, while columns 2 and 3 focus on enrollment and INE ID replacement procedures, respectively. The number of observations is highest in column 1 as the sampling of pairs draws from the larger population of those doing any procedure. The last row of the table displays the mean probability that a pair of randomly chosen registered voters living in different blocks within the same precinct perform a procedure at the INE office on exactly the same day.
Significance: * $p<0.05$; ** $p<0.01$; *** $p<0.001$

The evidence that we presented earlier in the paper shows that there is virtually no selection into blocks once precinct fixed effects are considered. Nevertheless, one might hypothesize that people of similar age, education, or occupation, or of the same gender, might be both more likely to live on the same block and to attend an INE office on the same day. To guard against this possibility, we control for a set of indicators $\mathbf{X}_{ij}$, each taking the value of 1 when citizens i and j are of the same gender or are in the same category of age, gender, occupation, schooling level, or quartile of distance from home to polling booth.

[21] A given citizen may appear in more than one pair, but we dropped identical pairs.

Note also that the control group (precinct neighbors living on different blocks) is extremely similar to the treatment group (precinct neighbors living on the same block) in terms of socio-demographic traits, local infrastructure, exposure to political information, and exposure to mobilization efforts, importantly limiting the scope for selection in the first place.

The first column in Table 6 presents the results for a dependent variable that pools across all different types of procedures.[22] The coefficient on the indicator for living on the same block implies that block neighbors are 1 percentage point more likely to conduct a procedure on the same day vis-a-vis a pair of precinct neighbors living on different blocks. Given a sample mean of 0.018, this implies a 66 percent increase in the probability of conducting a procedure on the same day. In columns 2 and 3, we respectively focus on registering to vote and on replacing one's voter ID—the two most common procedures. In both cases, block neighbors are significantly more likely to perform these procedures on the same day compared to neighbors who live on different blocks. The marginal effects are similar in magnitude as that in column 1. We take these findings as evidence that micro-neighbors are more likely to interact with each other over voting matters than non-block neighbors, consistent with a peer effects mechanism for our main findings.

6 Conclusions

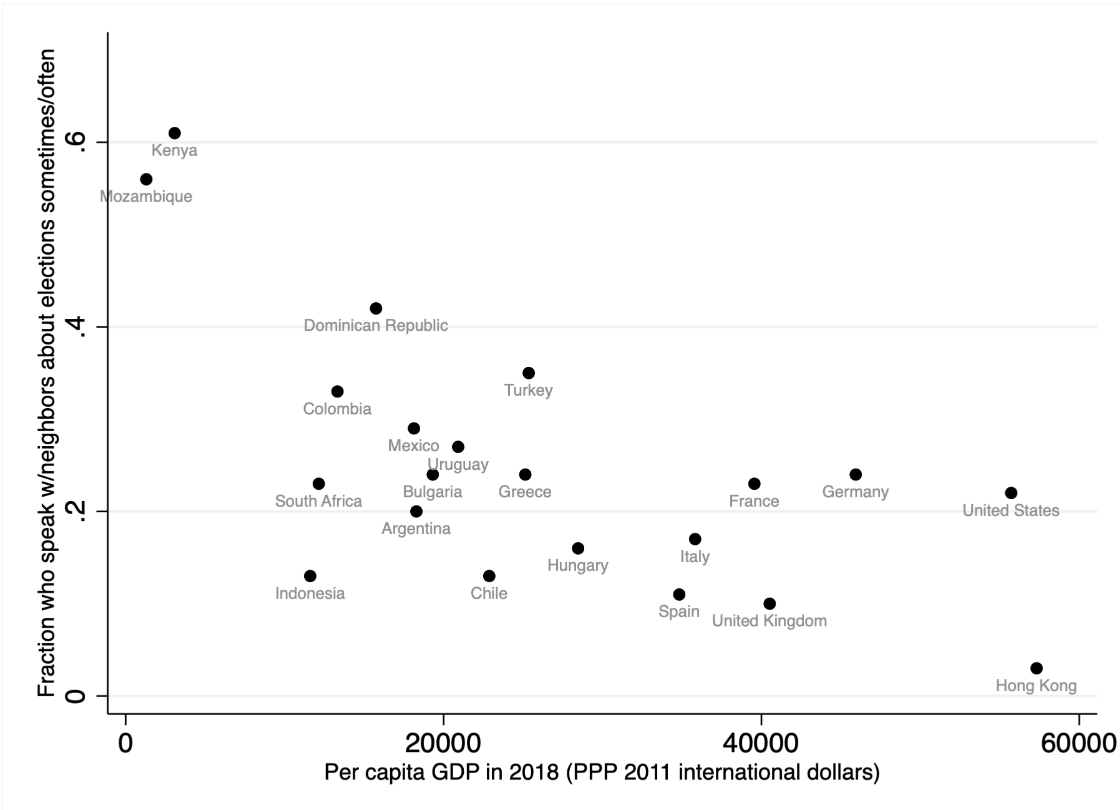
Explaining variation in turnout is a challenge that has long vexed political scientists. In this study, we suggest that neighborhoods may play an important role in explaining part of this puzzle. In particular, we show that the actions of one's neighbors affect not only one's propensity to vote, but also to register and perform other electoral procedures. Our emphasis on proximate neighbors—which we operationalize as neighbors on one's same block—connects with recent work on intergenerational mobility that underscores the importance of context at the very local, sub-neighborhood level (Chetty and Hendren, 2018b; Pinsker, 2019). Our findings show that what is true for intergenerational mobility is also true for voter turnout. While our results point to peer effects as a key channel of influence, it remains unclear what the sources of these effects might be. Whether they come from social conformity or from neighbors learning from one another remains an open question for future research.

We believe our general findings are likely to extend beyond the Mexican case, for two reasons. First, cross-locality variation in voter turnout is very large in many countries, as Table 1 shows. Second, cross-national survey evidence suggests that the fraction of citizens who talk to their neighbors frequently about elections is substantial virtually everywhere

[22] That is, the dependent variable takes the value of 1 if citizens i and j attended an INE office on the same day to perform a procedure *of any type*. The set of procedures includes: change of address, data corrections, enrollment, and replacement of voter ID.

(Figure 4).

Figure 4: Fraction of people who talk to neighbors about elections by country



Source: Authors' calculations based on Comparative National Election Project data for the question: "How frequently did you speak to your neighbors about electoral campaigns: very frequently, sometimes, rarely, never." Income is per-capita GDP in 2018, expressed in 2011 purchasing power parity international dollars, taken from the World Development Indicators.

Finally, we believe this is the first systematic evidence on peer effects in turnout outside the global North. Most electoral systems today, and most democracies, are in the global South. As such, the evidence from Mexico is perhaps more representative of the modal case than most of the existing literature, which has focused largely on the United States. Figure 4, for example, shows that the fraction of voters who speak to their neighbors about elections is strongly and inversely associated with a country's per-capita GDP. This suggests that peer effects could matter even more in middle- and low-income democracies than in high-income countries.

References

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 2005, *113* (1), 151–184.

Barber, Michael and Kosuke Imai, "Estimating neighborhood effects on turnout from geocoded voter registration records," 2014. Unpublished manuscript.

Baybeck, Brady and Scott D McClurg, "What do they know and how do they know it? An examination of citizen awareness of context," *American Politics Research*, 2005, *33* (4), 492–520.

Bayer, Patrick, Stephen L Ross, and Giorgio Topa, "Place of work and place of residence: Informal hiring networks and labor market outcomes," *Journal of Political Economy*, 2008, *116* (6), 1150–1196.

Beck, Paul Allen, Russell J Dalton, Steven Greene, and Robert Huckfeldt, "The social calculus of voting: Interpersonal, media, and organizational influences on presidential choices," *American Political Science Review*, 2002, pp. 57–73.

Cantoni, Enrico and Vincent Pons, "Does Context Trump Individual Drivers of Voting Behavior?: Evidence from US Movers," 2020. Unpublished manuscript, Harvard Business School.

Card, David and Jesse Rothstein, "Racial segregation and the black–white test score gap," *Journal of Public Economics*, 2007, *91* (11), 2158 – 2184.

Chetty, Raj and Nathaniel Hendren, "The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects," *The Quarterly Journal of Economics*, 2018, *133* (3), 1107–1162.

_ and _ , "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates*," *The Quarterly Journal of Economics*, 02 2018, *133* (3), 1163–1228.

_ , _ , Patrick Kline, and Emmanuel Saez, " Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States *," *The Quarterly Journal of Economics*, 09 2014, *129* (4), 1553–1623.

Cohen, Cathy J and Michael C Dawson, "Neighborhood poverty and African American politics," *American Political Science Review*, 1993, pp. 286–302.

DellaVigna, Stephano, Ulrike Malmendier, John List, and Gautham Rao, "Voting to Tell Others," *Review of Economic Studies*, 2017, *84* (1), 143–181.

Evans, William N., Wallace E. Oates, and Robert M. Schwab, "Measuring Peer Group Effects: A Study of Teenage Behavior," *Journal of Political Economy*, 1992, *100* (5), 966–991.

Gerber, Alan S and Todd Rogers, "Descriptive social norms and motivation to vote: Everybody's voting and so should you," *Journal of Politics*, 2009, *71* (1), 178–191.

_ , Donald P Green, and Christopher W Larimer, "Social pressure and voter turnout: Evidence from a large-scale field experiment," *American political Science review*, 2008, *102* (1), 33–48.

Gimpel, James G, Joshua J Dyck, and Daron R Shaw, "Registrants, voters, and turnout variability across neighborhoods," *Political Behavior*, 2004, *26* (4), 343–375.

Glaeser, Edward L, Bruce I Sacerdote, and Jose A Scheinkman, "The social multiplier," *Journal of the European Economic Association*, 2003, *1* (2-3), 345–353.

Green, Donald P and Ron Shachar, "Habit formation and political behaviour: Evidence of consuetude in voter turnout," *British Journal of Political Science*, 2000, pp. 561–573.

Hagene, Turid and Íñigo González-Fuente, "Deep politics: Community adaptations to political clientelism in twenty-first-century Mexico," *Latin American Research Review*, 2016, *51* (2), 3–23.

Holland, Alisha C and Brian Palmer-Rubin, "Beyond the machine: clientelist brokers and interest organizations in Latin America," *Comparative Political Studies*, 2015, *48* (9), 1186–1223.

Huckfeldt, R., "Variable responses to neighborhood social contexts: Assimilation, conflict, and tipping points," *Political Behavior*, 1980, (2), 231–257.

Huckfeldt, Robert and John Sprague, "Networks in context: The social flow of political information," *American Political Science Review*, 1987, pp. 1197–1216.

_ and _ , "Discussant effects on vote choice: Intimacy, structure, and interdependence," *Journal of Politics*, 1991, *53* (1), 122–158.

_ , Eric Plutzer, and John Sprague, "Alternative contexts of political behavior: Churches, neighborhoods, and individuals," *Journal of Politics*, 1993, *55* (2), 365–381.

Johnston, Ron, Kelvyn Jones, Rebecca Sarker, Carol Propper, Simon Burgess, and Anne Bolster, "Party support and the neighbourhood effect: spatial polarisation of the British electorate, 1991–2001," *Political Geography*, 2004, *23* (4), 367–402.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman, "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment*," *The Quarterly Journal of Economics*, 05 2001, *116* (2), 607–654.

Kenny, Christopher B, "Political participation and effects from the social environment," *American Journal of Political Science*, 1992, pp. 259–267.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental analysis of neighborhood effects," *Econometrica*, 2007, *75* (1), 83–119.

Klofstad, Casey A, "Exposure to political discussion in college is associated with higher rates of political participation over time," *Political Communication*, 2015, *32* (2), 292–309.

Larreguy, Horacio, Cesar E Montiel Olea, and Pablo Querubin, "Political Brokers: Partisans or Agents? Evidence from the Mexican Teachers' Union," *American Journal of Political Science*, 2017, *61* (4), 877–891.

_ , John Marshall, and Pablo Querubin, "Parties, brokers, and voter mobilization: How turnout buying depends upon the party's capacity to monitor brokers," *American Political Science Review*, 2016, *110* (1), 160–179.

Lee, Barrett A. and Karen E. Campbell, *Neighbor Networks of Black and White Americans*, Networks in the Global Village: Life in Contemporary Communities. By Barry Wellman. Boulder, CO: Westview, 1999.

Levitsky, Steven, *Transforming labor-based parties in Latin America: Argentine Peronism in comparative perspective*, Cambridge University Press, 2003.

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 1993, *60* (3), 531–542.

McClurg, Scott D, "Social networks and political participation: The role of social interaction in explaining political participation," *Political research quarterly*, 2003, *56* (4), 449–464.

_ , "The electoral relevance of political talk: Examining disagreement and expertise effects in social networks on political participation," *American Journal of Political Science*, 2006, *50* (3), 737–754.

Mutz, Diana C, "The consequences of cross-cutting networks for political participation," *American Journal of Political Science*, 2002, pp. 838–855.

Nickerson, David W, "Is voting contagious? Evidence from two field experiments," *American political Science review*, 2008, *102* (1), 49–57.

Oster, Emily, "Unobservable selection and coefficient stability: Theory and evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

Otani, Shinsuke, *Personal Community Networks in Contemporary Japan*, Networks in the Global Village: Life in Contemporary Communities. By Barry Wellman. Boulder, CO: Westview, 1999.

Pinsker, Joe, "How Do Rich Neighborhoods Exist So Close to Poor Ones?," *The Atlantic*, June 2019, pp. 561–573.

Rolfe, Meredith, *Voter turnout: A social theory of political participation*, Cambridge University Press, 2012.

_ **and Stephanie Chan**, "Voting and Political Participation," *The Oxford Handbook of Political Networks*, 2017.

Sampson, Robert J., Jeffrey D. Morenoff, and Thomas Gannon-Rowley, "Assessing "neighborhood effects": Social processes and new directions in research," *Annual Review of Sociology*, 2002, *28* (1), 443–478.

Schelling, Thomas C., "Models of Segregation," *The American Economic Review*, 1969, *59* (2), 488–493.

Shefner, Jon, "Coalitions and clientelism in Mexico," *Theory and Society*, 2001, *30* (5), 593–628.

Shulman, Hillary C and Timothy R Levine, "Exploring social norms as a group-level phenomenon: Do political participation norms exist and influence political participation on college campuses?," *Journal of Communication*, 2012, *62* (3), 532–552.

Sinclair, Betsy, *The social citizen: Peer networks and political behavior*, University of Chicago Press, 2012.

Stokes, Susan C, Thad Dunning, Marcelo Nazareno, and Valeria Brusco, *Brokers, voters, and clientelism: The puzzle of distributive politics*, Cambridge University Press, 2013.

Young, Peyton, *Social Dynamics*, MIT Press, 2001.

Learning to Vote

Appendix – For Online Publication

Frederico Finan, Enrique Seira, and Alberto Simpser

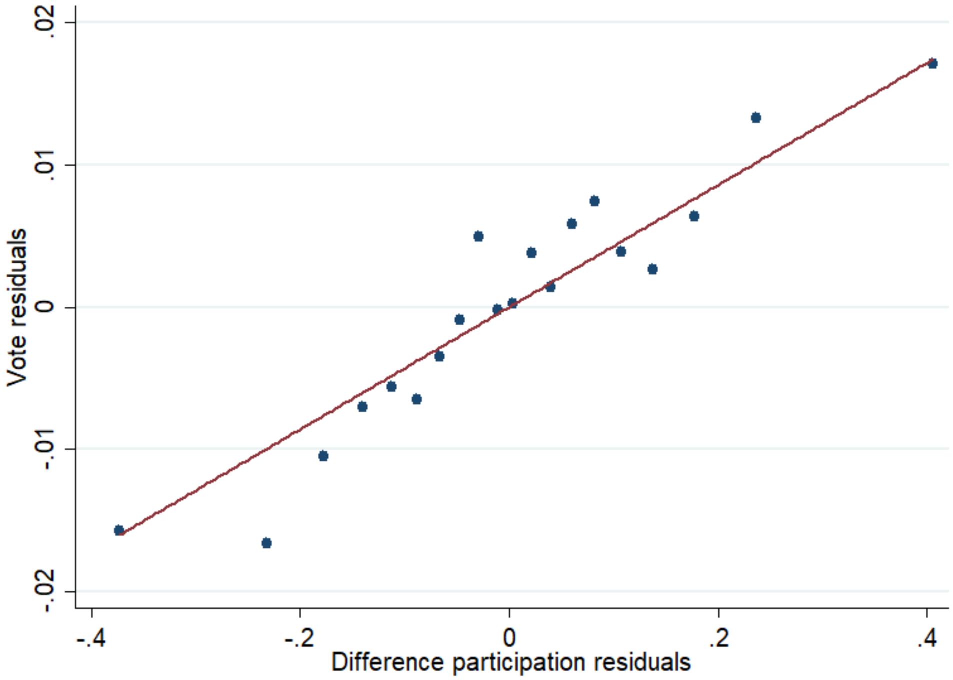
Appendix A. Further results

Table OA-1: **Main results with maximal sample**

	Dependent variable: individual mover's turnout in 2015					
	(1a)	(1b)	(2)	(3)	(4)	(5)
Destination-origin block turnout difference	0.0687***	0.0662***	0.0767***	0.0426***	0.0550***	
	(0.003)	(0.003)	(0.003)	(0.004)	(0.011)	
Destination-origin block turnout percentile diff.						0.00163***
						(0.000)
Individual mover's turnout in 2012	0.241***	0.244***	0.207***	0.209***	0.253***	0.271***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)	(0.005)
Male			-0.0293***	-0.0285***	-0.0342***	-0.0338***
			(0.002)	(0.002)	(0.005)	(0.005)
Age			0.00362***	0.00364***	0.00350***	0.00342***
			(0.000)	(0.000)	(0.000)	(0.000)
Household size 2015			0.00419***	0.00398***	0.00261***	0.00239**
			(0.000)	(0.000)	(0.001)	(0.001)
Distance to polling station in 2015 (km)			-0.0133***	-0.0134***	-0.0132**	-0.0130**
			(0.001)	(0.001)	(0.004)	(0.004)
N	1153091	1206229	583474	516335	516335	528698
R^2	0.199	0.200	0.243	0.244	0.735	0.727
Mean of dependent variable	0.379	0.380	0.350	0.351	0.351	0.353
Mean of destination-origin block turnout diff.	-0.166	-0.158	-0.179	-0.182	-0.182	-3.080
Effect per 1 SD chg. in block turnout diff.	0.0148	0.0141	0.0165	0.00895	0.0115	0.0484
Individual-level controls			✓	✓	✓	✓
Block-level controls				✓	✓	✓
Destination precinct fixed effects	✓	✓	✓	✓		
Origin-destination precinct pair fixed effects					✓	✓

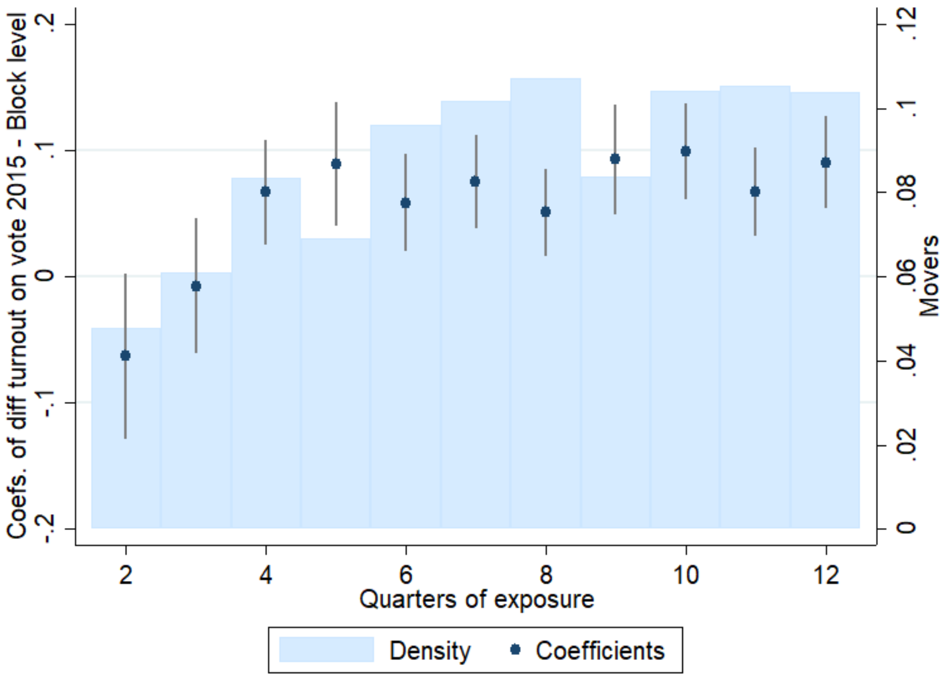
This table is analogous to Table 4 in the main paper, except that instead of keeping the same sample across columns, we try to maximize sample size in each columns. The different sample sizes arise because the control variables change across columns and these controls have missing values. Results are very similar nonetheless.
Significance: * p<0.05; ** p<0.01; *** p<0.001

Figure OA-1: Symmetry and linearity of effects



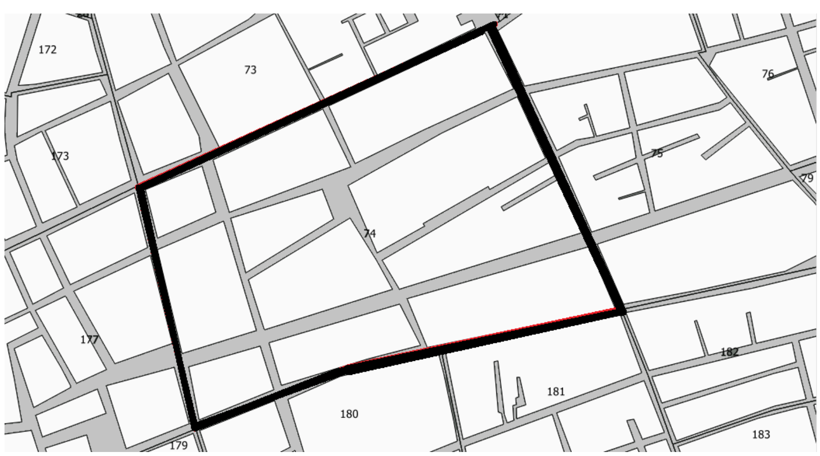
This figure explores whether the effect estimated in column 3 of Table 4 shows symmetry and linearity, i.e. whether increases in exposure to voting in destination blocks has effects that are similar but of the opposite sign as decreases in exposure to voting. To this end it shows a residuals-residuals partial-out plot for the main estimating equation 1 (using the Frisch–Waugh–Lovell theorem). The first regression estimates equation 1 without the main explanatory variable $\Delta Turnout_{o(b),d(b)}$ and predicts residuals $\hat{e}_{1i}$; the second regression uses $\Delta Turnout_{o(b),d(b)}$ as the dependent variable on the same covariates and predicts residuals $\hat{e}_{2i}$. It then plots a bin scatter of $\hat{e}_{1i}$ against 20-percentile bins of $\hat{e}_{2i}$, along with a linear regression line.

Figure OA-2: Exposure time at destination block



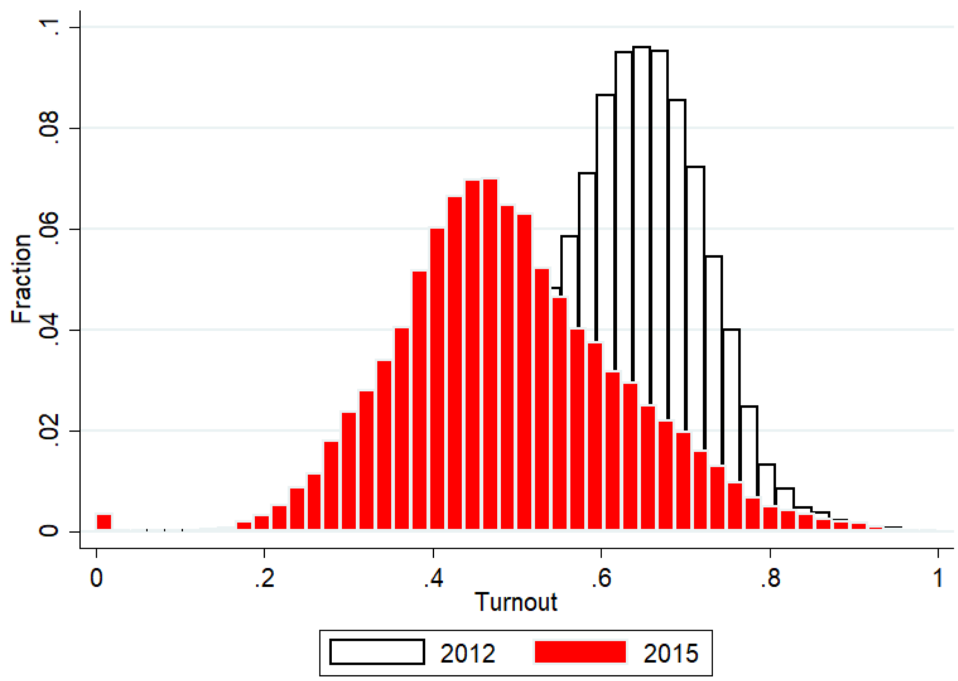
This figure estimates regression of column 3 of Table 4, splitting the sample according to the number of quarters that movers have resided at the destination block at the time of the 2015 election. Regressions are estimated separately for each subsample. The dots correspond to the regression coefficients for the destination minus origin turnout variable. We also plot in the background a histogram that displays the fraction of movers in each subsample.

Figure OA-3: Blocks in a precinct



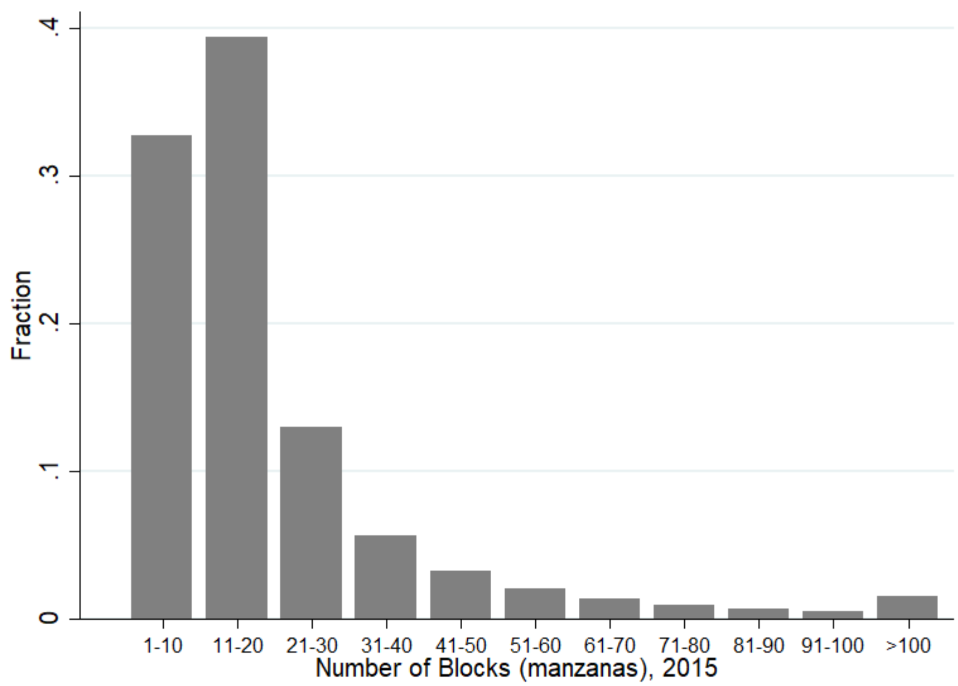
Panel A shows an example of blocks within a precinct in Mexico City. The black lines delimit one precinct and the grey lines delimit blocks.

Figure OA-4: Precinct level turnout



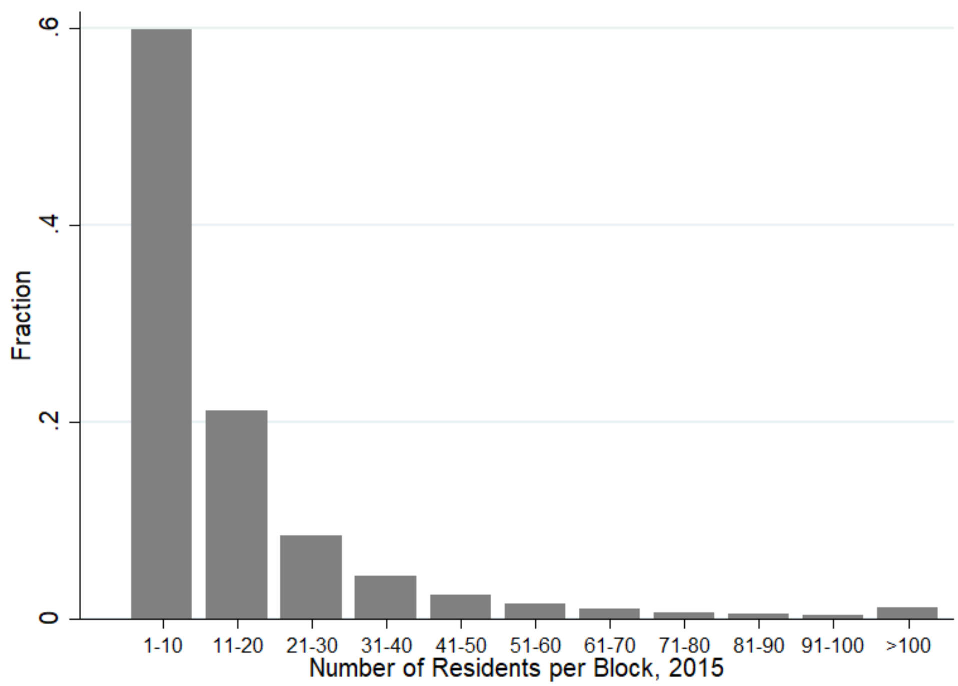
This figure shows the distribution of voter turnout at the precinct level for the 2012 and 2015 elections in the analysis sample.

Figure OA-5: Number of blocks per precinct



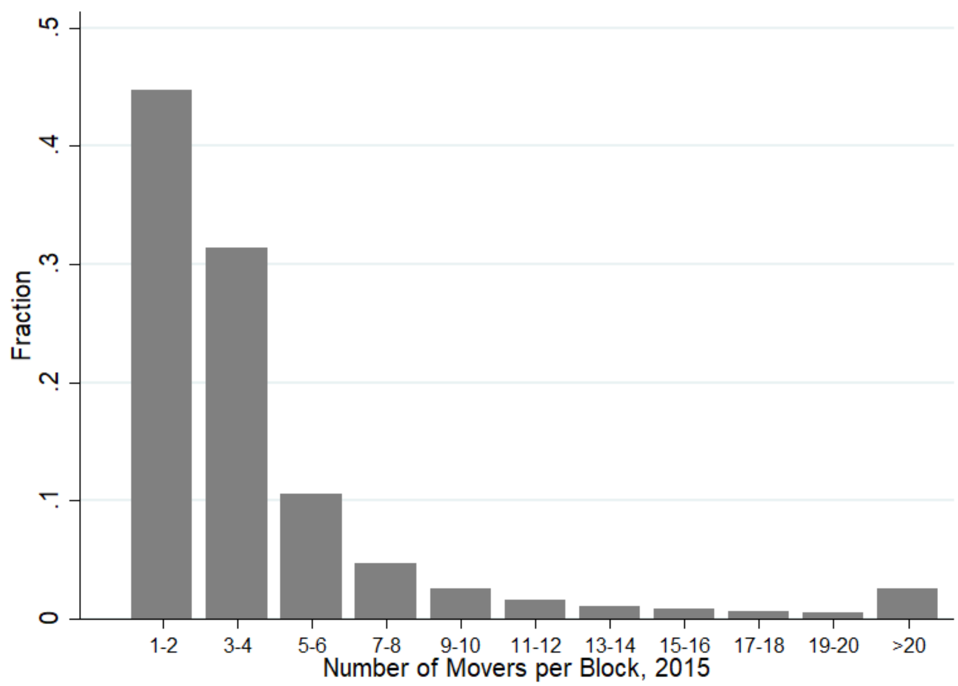
Histogram of the fraction of blocks per precinct in 2015 where an observation is a precinct.

Figure OA-6: Number of residents per block



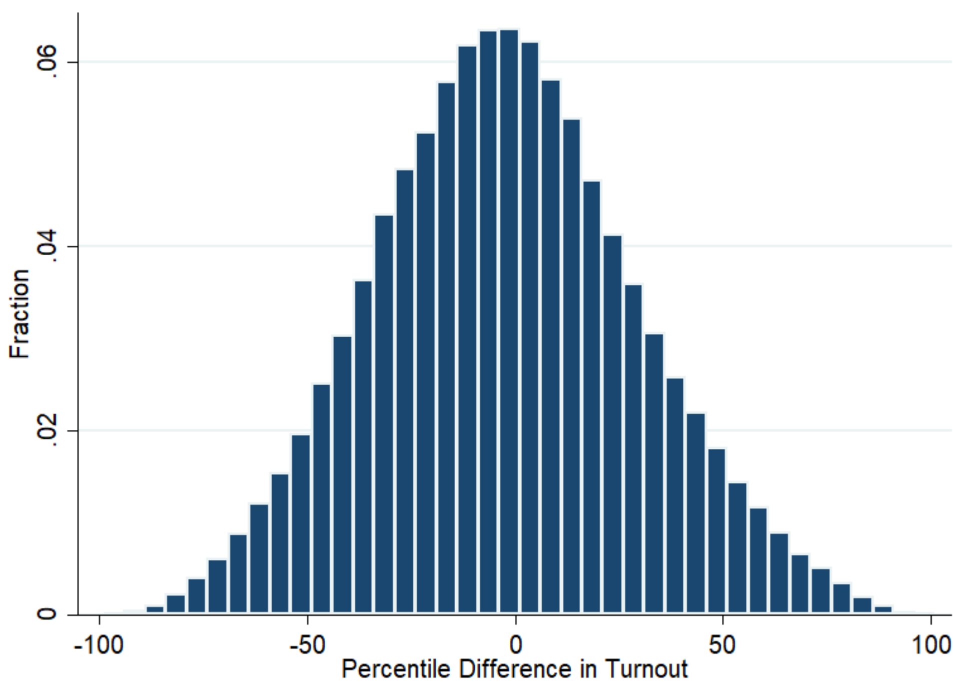
Histogram of the fraction of residents per block in 2015 where an observation is a block.

Figure OA-7: Number of movers per block



Histogram of the number of movers per destination block where an observation is a block.

Figure OA-8: Percentile turnout differential exposure (block level)



This figure plots the differential exposure of movers to percentile differences in their block turnouts. An observation is a mover. For each mover we first assign a percentile to average turnout of registered voters in the destination block in 2015 as well as a percentile for the average turnout of registered voters in the origin block in 2012, and take the difference of percentiles (2015 destination block percentile minus 2012 origin block percentile). The variable is bounded in [-100,100].

Table OA-2: **Heterogeneity of exposure to neighbors**

Variable	Block (baseline)		Block pair FE	
	Coefficient (1)	1 sd change (2)	Coefficient (3)	1 sd change (4)
$\Delta Turnout_{o(b),d(b)}$*blockdiff_gender	-0.0584***	-0.0010	-0.0616	-0.0010
	(0.017)		(0.053)	
$\Delta Turnout_{o(b),d(b)}$*blockdiff_age	-0.0018***	-0.0009	-0.0024	-0.0012
	(0.000)		(0.001)	
$\Delta Turnout_{o(b),d(b)}$*blockdiff_years_education	0.0007	0.0001	0.0039	0.0006
	(0.002)		(0.005)	
$\Delta Turnout_{o(b),d(b)}$*blockdiff_housekeeper	0.0146	0.0002	0.0597	0.0009
	(0.016)		(0.047)	
$\Delta Turnout_{o(b),d(b)}$*blockdiff_employee	-0.0358*	-0.0006	-0.0051	-0.0001
	(0.015)		(0.047)	
$\Delta Turnout_{o(b),d(b)}$*blockdiff_student	-0.0092	-0.0001	0.0423	0.0005
	(0.018)		(0.055)	
$\Delta Turnout_{o(b),d(b)}$*blockdiff_household_size	-0.0029	-0.0004	-0.0028	'0.0004
	(0.002)		(0.005)	
N	512394		512394	
R^2	0.244		0.735	
$\lvert Avg.\ \Delta Turnout_{o(b),d(b)}\rvert$	0.183		0.183	

To explore heterogeneity of the effect of exposure to neighbors, this table uses the main specification in equation 1 and adds covariates (z^k) of "similarity" between the mover and her neighbors, and their interactions with our main explanatory variable $\Delta Turnout_{o(b),d(b)}$. These covariates are the absolute value of the difference between the mover's k=gender, age, years of education in 2015, household size in 2015, and dummies of the three main occupations in the dataset (housekeeper, employee, student) and the same variables averaged across her (non-family $f(i)$ member) neighbors: $z^k = |x_i^k - Avg(x_j^k)|_{j \neq f(i)}$. The table reports only the interaction coefficients on $z^k \times \Delta Turnout_{o(b),d(b)}$ along with their standard errors (columns 1 and 3). Column 1 uses the baseline specification corresponding to column 2 in Table 4, including destination precinct fixed effects, while column 3 corresponds to the same specification as in column 2 with added pair precinct fixed effects. A negative value indicates that the more different they are the smaller the effect of living in the same block on turnout. Even columns correspond to the implied magnitude of the effect on 2015 vote of the mover from a 1 standard deviation change in z^k, multiplied by the absolute value of the average of $\Delta Turnout_{o(b),d(b)}$ (0.182). For instance, a 1 standard deviation in the difference between the age of the mover and that of their neighbors results in a decrease of 0.1 percentage points of the effect of $\Delta Turnout_{o(b),d(b)}$. Significance: * p<0.05; ** p<0.01; *** p<0.001

Table OA-3: **Neighbor influence in electoral procedures - Restricted time period**

	All procedures	Enrollment	Replacement
	(1)	(2)	(3)
Same block	0.012***	0.009***	0.014***
	(0.000)	(0.001)	(0.001)
Precinct fixed effects	✓	✓	✓
With controls	✓	✓	✓
Observations	970846	213084	516229
R^2	0.111	0.091	0.144
Mean same-day procedure for non-block neighbors	0.018	0.017	0.020

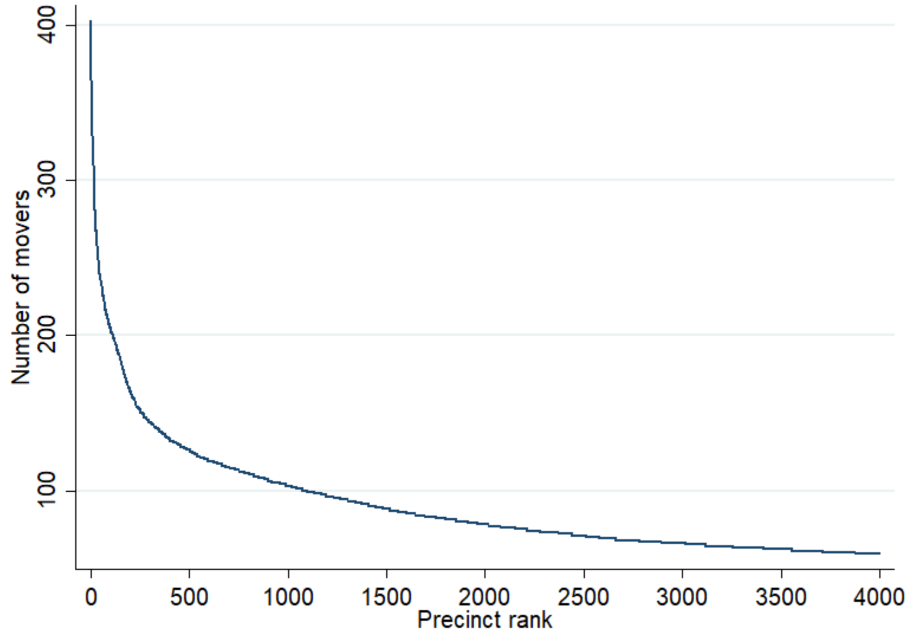
This table is analogous to Table 6, but we restrict the sample period from June 2014 to February 2015.
Significance: * p<0.05; ** p<0.01; *** p<0.001

Table OA-4: **Variable list**

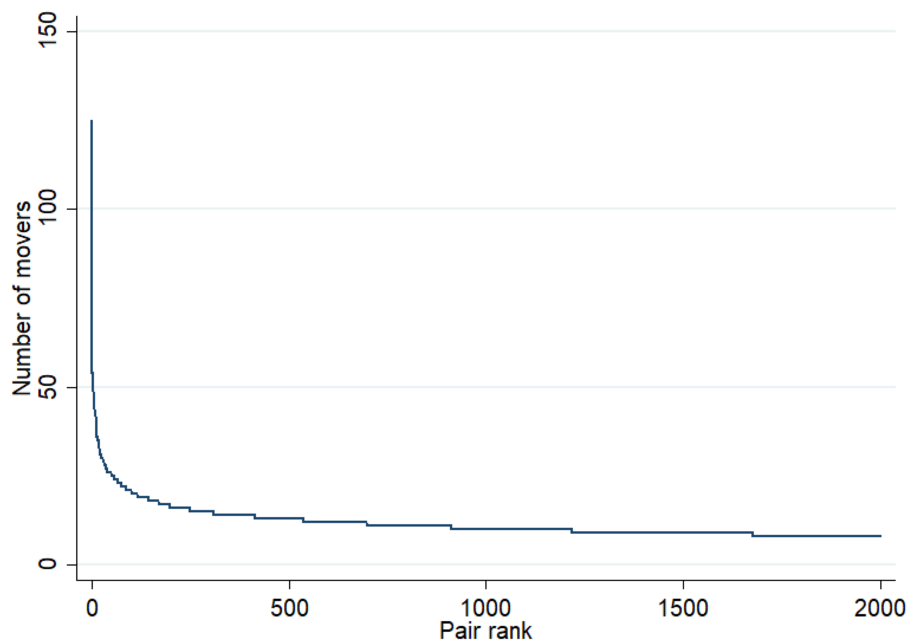
Variable	Source	Notes
Precinct code/identifier	Census 2010 at precinct level	
PCA of demographics variables	Census 2010 at precinct level	Total population, male population, female population, population by gender and age intervals, number of houses, population by religious beliefs, population by marital status.
PCA of education variables	Census 2010 at precinct level	Children who do not attend school by age interval and gender, children who attend school by age interval and gender, population with illiteracy by age interval and gender, population by educational attainment by age interval and gender.
PCA of economic variables	Census 2010 at precinct level	Employed population, unemployed population, economically activate population, inactive population.
PCA of household variables	Census 2010 at precinct level	Number of occupied houses, number of unoccupied houses, average number of occupants per house, number of houses by type of floor, number of houses by number of rooms, number of houses with various goods (refrigerator, radio, television, car, etc.), number of houses by access to public services (water, electricity, sewage, etc.).
Unique citizen code/identifier	Padron Electoral 2012, 2015, 2018	
Socio-demographic variables	Padron Electoral 2012, 2015, 2018	Gender, Age, Occupation (housework, student, employee, etc.) and education level attainment (no education, high school, undergraduate, etc.)
Citizen location	Padron Electoral 2012, 2015, 2018	State, Locality, Precinct and Block of residence
Polling station	Padron Electoral 2012, 2015, 2018	
Unique citizen code/identifier	Individual vote 2012, 2015	
Dummy citizen voted	Individual vote 2012, 2015	
Unique citizen code/identifier	Voter ID Procedures 2015, 2018	
Type of procedure	Voter ID Procedures 2015, 2018	Registration, change of address, etc.
Date of procedure	Voter ID Procedures 2015, 2018	
Numbers of voter id had	Voter ID Procedures 2015, 2018	
Unique citizen code/identifier	Households 2012, 2015	
Household code/identifier	Households 2012, 2015	
Number of household members	Households 2012, 2015	
Block code/identifier	Blocks shapefile	
Block centroid coordinates (latitude and longitude)	Blocks shapefile	
Polling station code/identifier	Polling Station Coordinates	
Coordinates (latitude and longitude)	Polling Station Coordinates	
Precinct code/identifier	Electoral results 2012, 2015	
Total votes at the precinct level	Electoral results 2012, 2015	

Number of registered voters at the precinct level	Electoral results 2012, 2015	
Turnout at the precinct level	Electoral results 2012, 2015	

Figure OA-9: Flow of movers



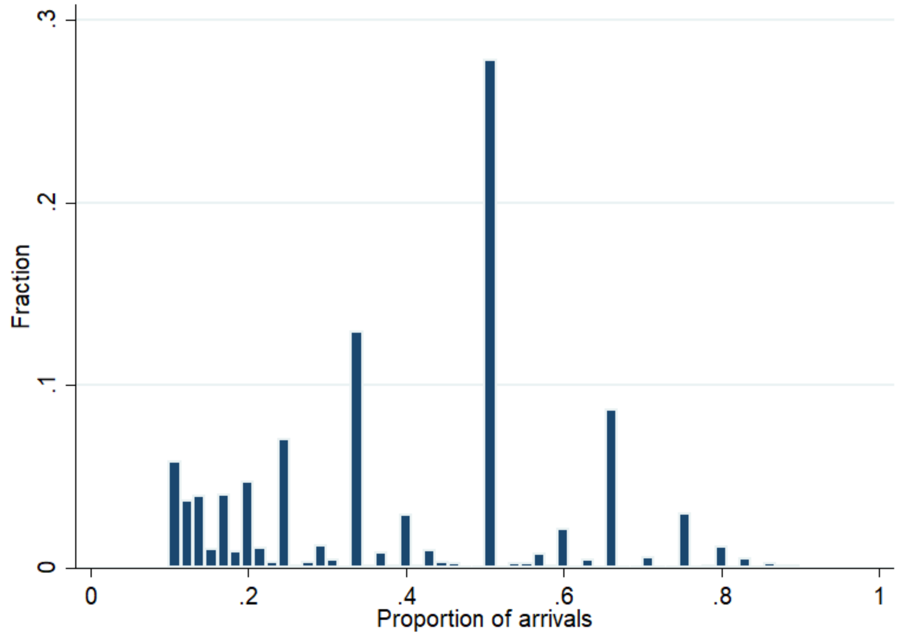
(a) Movers by origin precinct



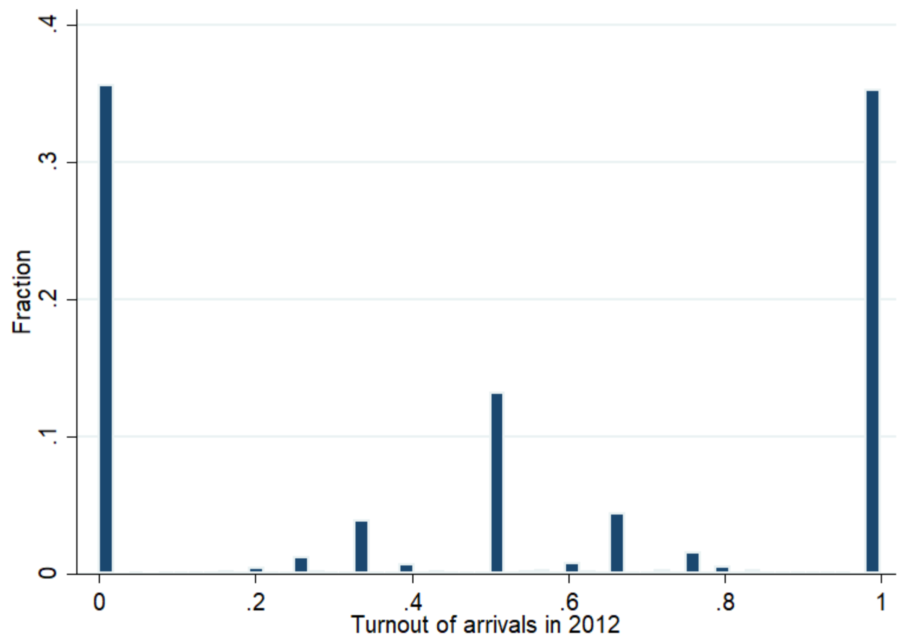
(b) Movers by origin-destination precinct pair

Panel (a) shows the number of movers by precinct of origin. For each of the precincts we counted the number of voters that were registered there in 2012 and moved to another precinct by 2015. Panel (a) plots this count ordered by precincts with the largest the number of such movers. We truncate the figure at the top 10,000 origin precincts. Panel (b) shows the number of movers across precincts between 2012 and 2015 for all origin-destination precinct *pairs*. The figure plots the 2,000 precincts with the highest flow, ordered from highest (left) to lowest (right).

Figure OA-10: Arrivers to a block



(a) Arrivers as fraction of block's registered voter



(b) Average 2012 turnout of block' arrivers

Panel (a) shows the proportion of arrivers in 2015 in a block aS a proportion of that block's registered voters in 2015. Panel (b) shows the distribution of turnout in 2012 of the arrivers at the block level, which is our measure of "influence". Only blocks with at least 10% and no more than 90% of its residents being arrivers are included in the figures.