The Macroeconomics of Neighborhood Effects

Ricardo Marto*
University of Pennsylvania

Felipe Mazin†
University of Pennsylvania

July 15, 2020
Preliminary Version

Abstract

We use micro-level data on high school students from Brazil to estimate the importance of neighborhood effects. We rank neighborhoods by their income level and find that neighborhood effects are stronger in richer districts than in poor ones. Poor neighborhoods, on the other hand, tend to have stronger classroom peer effects. We then build an overlapping generations model with distinct neighborhoods to examine the dynamics of segregation, wealth accumulation, and intergenerational mobility. Neighborhoods shape the skills a child is born with and her labor earnings once she becomes an adult. Hence, parents have to tradeoff higher rental prices today for a higher probability of their children being skilled. We show that the presence of borrowing constraints traps some individuals in the inferior neighborhood, while wealthier households compensate low earnings with more bequests.

Keywords: Neighborhood Effects, Intergenerational Mobility, Brazil
JEL Classifications: E21, E24, R23, J62

*Department of Economics, University of Pennsylvania. Email address: rmarto@sas.upenn.edu.
†Department of Economics, University of Pennsylvania. Email address: feliperm@sas.upenn.edu.

1 Introduction

The neighborhood in which a child grows up shapes her opportunities later in life. For instance, a child who grows up in a poor neighborhood is more likely to do worse than her peers from affluent origins on a wide range of indicators even when controlling for family background[1]. The neighborhood choice has thus important implications for aggregate human capital, social mobility, and policy. In this paper, we quantify the extent of neighborhood effects—defined as the collection of the local factors not directly related to an individual's family—in children's human capital formation and their impact on wealth inequality. In particular, we examine the vicious circle where inequality in skills and income begets segregation, which in turn implies lower social mobility and perpetuates inequality.

This paper has two parts. First, we use microdata covering nearly the universe of students in the last year of elementary school in the city of Sao Paulo, Brazil, to quantify neighborhood effects in the accumulation of skills. Our strategy is to carry out a bottom-up estimation of these effects, which we define as the combination of within-classroom peer effects and school-level effects. In particular, we estimate the linear-in-means model of peer effects of student test scores on the average socioeconomic background of their classroom peers, controlling for school fixed effects as well as their own socioeconomic background. We then interpret the estimated school fixed effects as capturing selection bias and other relevant neighborhood characteristics. Adding this effect on top of the classroom peer effects provides us with an upper bound of the neighborhood effects, while the estimates of classroom peer effects serve as a lower bound of the neighborhood effects. We then aggregate the estimated ranges at the city's district level, rank the 96 districts by their annual income per capita and find that (i) classroom peer effects are significantly different from zero, (ii) Classroom peer effects together with school effects are increasing in districts' average income, (iii) there appears to be diminishing returns in terms of neighborhood effects from moving from a poorer district to a richer one, and (iv) classroom peer effects make up most of the neighborhood effects in poorest areas.

In the second part of the paper, we develop a dynastic overlapping generations model with heterogeneity in wealth and skills to study the dynamics of segregation and wealth accumulation across neighborhoods. We consider two neighborhoods. One of the neighborhoods is assumed to have strict zoning laws, where the supply of housing is limited and the rental price is determined in equilibrium. The second neighborhood has a rental price of zero and has capacity to accommodate all households who wish to live there. When parents choose a neighborhood to live in, that choice has long lasting effects on their children. In particular, children draw their skills from a distribution that depends on the *endogenous* average skill level in the neighborhood. Hence, moving to a worse neighborhood to save on rent affects their children's earnings once they become adults.

The presence of borrowing constraints, on the other hand, could further dampen segregation. If adults have limited resources or face tight borrowing constraints, they are unable to move to better

[1] Chetty and Hendren (2018a), Chetty and Hendren (2018b)

neighborhoods where rent is prohibitively high. Wealthier families, on the other hand, will be willing to pay higher housing prices to segregate themselves and have their children grow up with peers that have positive spillovers on them. Growing up with smarter classmates enhances the human capital of your own children. Once they become adults, their labor earnings—proportional to their skills—will be higher and hence will be able to accumulate more wealth. The children of poorer individuals instead are more likely to have skills that are less valued in the labor market. Their lower earnings constrains them live in cheaper neighborhoods unless they are able to save their way out of slums. Hence, location can be a trap.

In our numerical exercise, we find that the neighborhood with limited supply of housing ends up having a higher average skill level. However, and surprisingly as well, the neighborhood with unlimited housing space is more unequal. Indeed, rich and poor people live side by side in that neighborhood. The reason is that wealthy individuals can insure their children by endowing them with a large amount of bequests that compensates the probability of being low skilled. This effect could be a rationale for why so many slums in Brazil are side by side with wealthy housing.

Related literature. This paper relates to the strand of literature that studies the macroeconomic implications of neighborhood effects. It is closely related to Ferreira et al. (2018), who also focus on Brazil and similar peer effects to study the emergence of slums as a stepping stone in the urbanization process. Our paper adds wealth heterogeneity, borrowing constraints, and parents who are altruistic towards their children. Our paper is also closely related to Bilal and Rossi-Hansberg (2020), who view the choice of where to live as an an investment in a "location asset" that can be used as a way of smoothing consumption when the borrowing constraint on their financial assets is binding. They address the question of why there is so little mobility even though there exists positive returns from moving to better neighborhoods. Their answer is that people that cannot borrow choose to live in inferior neighborhoods by borrowing from their own and their children's future human capital. However, their paper completely abstracts from considerations about intergenerational mobility. Our paper adds this into the picture and shows that there is an implicit force guiding the choice of a location given by the continuation value of an agent (in our case through children).

The impact of neighborhood effects on intergenerational mobility have been recently documented and identified for the United States in a series of connected papers (Chetty et al. (2014),Chetty et al. (2016),Chetty and Hendren (2018a), Chetty and Hendren (2018b)) that use tax records and location data covering the U.S. population. Durlauf and Seshadri (2018) and Fogli and Guerrieri (2019) are two other papers that bear close relation to ours. They investigate the tradeoff between inequality and social mobility in the United States with a mechanism akin to ours, where residential segregation implies different opportunities across neighborhoods. Our focus on education also relates to the work of Eckert and Kleineberg (2019), who study the implications of neighborhood effects for educational policies. Their focus is on school financing in the U.S. and examine the general equilibrium effects of different policy interventions on the disparity of educational outcomes and welfare between children from low and highly skilled families, which is also attempt to do with our framework.

Our empirical work follows the empirical literature that estimates the role of peer effects in the education production function, building off the seminal work by Manski (1993). Recent papers that relate to ours is Ammermueller and Pischke (2009) and Agostinelli (2018).

Outline. The remainder of the paper is organized as follows. Section 2 presents our data on Brazilian neighborhoods and estimation of neighborhood effects, Section 3 introduces the dynamic model, Section 4 presents our calibration and quantitative results, and Section 5 concludes the paper.

2 Empirical Analysis

In this section we combine micro-data on pupils' test scores at the neighborhood level to assess the importance of peer effects on the accumulation of human capital. We start by describing our data and our empirical strategy, and then discuss our results on peer effects.

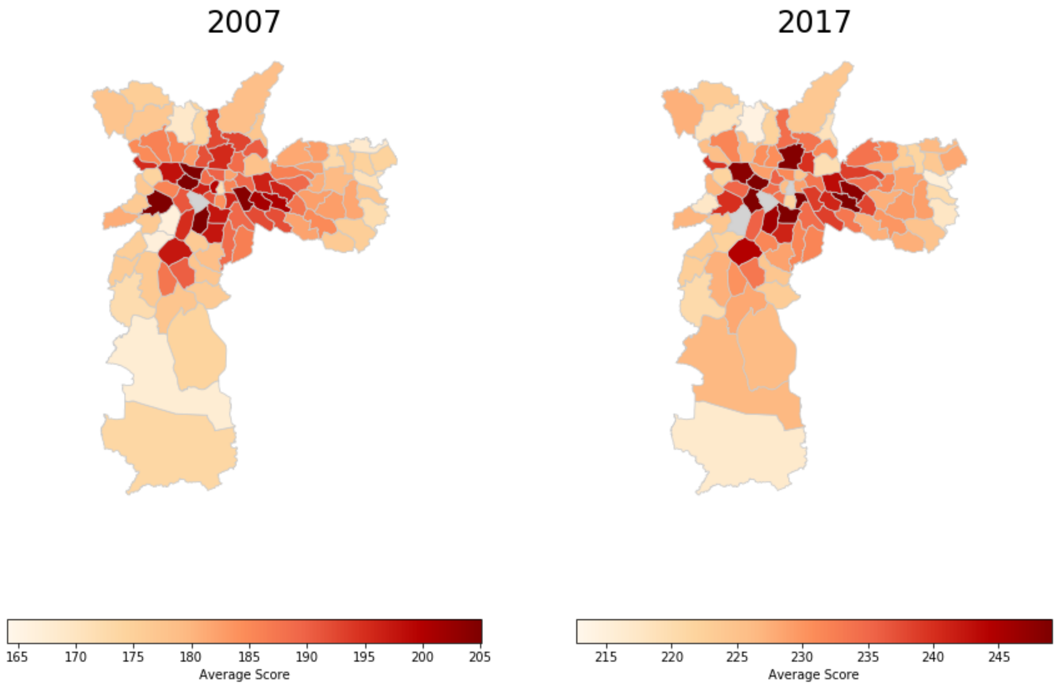
2.1 Data

We use detailed data on students and schools across the 96 districts of the city of Sao Paulo in Brazil. We focus on Sao Paulo because it has huge disparities in wealth and neighborhoods. To understand how the relevance of the location while growing up, we delve into the accumulation of human capital early in life. Our data consists of test scores for the standardized test "Prova Brasil," administered every two years to students in their final year of elementary, middle, and high school across the universe of public schools in Brazil. Note that although private schools have proliferated over the last decades, the majority of students are enrolled in public schools according to the 2019 school census[2]. The dataset is a repeated cross-section and covers the ten year period that goes from 2007 to 2017 and includes six instances of the test. The test is standardized (across time *and* school years) and measures the proficiency of students in Mathematics and Portuguese. Figure 1 shows the average proficiency in Portuguese and mathematics per each district of the city of Sao Paulo for the final year of elementary school[3]. The average score is an average in each of these subjects, with darker colors indicating better test scores. The neighborhoods with better test scores are also those with wealthier neighborhoods. The dataset also includes a socioeconomic supplement with questions concerning students' characteristics, family background, and their experience in school. Appendix A provides more details on the test.

[2] In the country, the numbers are 80.8%, 84.6% and 87.5% for elementary, middle and high school, respectively. In the city of Sao Paulo the overall share is smaller, 64.4%.

[3] We chose to analyze the performance of students in elementary school given that these students, generally aged 10 or 11, tend to study close to where they grow up.

Figure 1: The evolution of average proficiency in the city of Sao Paulo.



Note: Average proficiency in Portuguese and mathematics measured by district. The middle point of the scale is defined as the mean average score across districts for that year. The minimum and maximum points are two standard deviations away from the middle point. Standard deviations for 2007 and 2017 are similar (10.3 and 9.1, respectively).

We complement this dataset with the geographic location of schools reported by the Department of Education of the state of Sao Paulo to identify the neighborhoods where schools are located.[4] We then match each school to a one of the 96 districts. Table 1 presents descriptive statistics about our dataset.

Table 1: Descriptive Statistics

	2007		2011				2017		
	SP	BRA	SP		BRA		SP	BRA	
	Public	Public	Public	Private	Public	Private	Public	Public	Private
Mean	179.70	181.96	195.41	246.32	196.83	239.81	226.45	214.76	245.60
Std. Dev.	41.79	39.01	42.60	37.83	42.66	41.49	43.18	44.88	40.48
# Classes	5136	92,324	5009	39	96,667	1,345	3973	98,752	1,145
# Schools	1,044	37,471	1,153	27	41,418	1,033	1,148	48,189	975
# Students	153,783	2,309,702	138,964	999	2,284,211	31,384	100,188	2,171,042	22,154

[4] We verified the accuracy of these locations using Google Maps.

2.2 Measuring Neighborhood Effects

The objective of this section is to use the data on test scores to measure neighborhood effects. In particular, we will consider social effects that come from the interaction with other students (usually known and henceforth referred to as peer effects) and a combination of other effects defined at the class, school, and district levels that are also part of the broadly defined neighborhood effects.

To understand these neighborhood effects, we follow Manski (1993) and divide these along three different dimensions: (i) endogenous effects, (ii) exogenous effects, and (iii) correlated effects. To understand each of these effects separately, we focus on the linear relationship established in equation (1) as in Ammermueller and Pischke (2009). The dependent variable, s_{icsdt}, refers to the score of student i in class c, school s, and district d at time t. On the right-hand side of the equation, we have our three dimensions of neighborhood effects: the endogenous peer effects captured by $\bar{s}_{(-i)csdt}$, the exogenous peer effects given by $\bar{x}_{(-i)csdt}$, and the correlated effects pertaining to the student's class ν_{csdt}, school u_{sdt}, and district θ_{dt}. We also include all the exogenous variables related to test scores, summarized by x_{icsdt} (such as gender, race, family background, and effort in class), a time fixed effect μ_t. The constant term is c and the error term is ε_{icsdt}. We thus have the relationship between test scores and neighborhood effects given by

$$s_{icsdt} = c + \underbrace{\varphi\ \bar{s}_{(-i)csdt}}_{\text{Endogenous peer effects}} + \underbrace{\bar{x}'_{(-i)csdt}\ \delta}_{\text{Exogenous peer effects}} + \underbrace{\theta_{dt} + u_{sdt} + \nu_{csdt}}_{\text{Correlated peer effects}} + \underbrace{x'_{icsdt}\ \beta}_{\text{Controls}} + \mu_t + \varepsilon_{icsdt} \tag{1}$$

Note that the endogenous peer effects are defined by the average traits of a student's classmates on her own performance (excluding herself). These effects are said to be endogenous because the knowledge of the remaining students affects an individual student's learning, which is measured by the average of their test scores $\bar{s}_{(-i)csdt}$ and its marginal effect is given by φ. The peer effects are exogenous when her peers' attributes affect the student's performance (such as their race, gender, or family background). The peers' average characteristics are denoted by $\bar{x}_{(-i)csdt}$ and its marginal effect is given by δ. Finally, correlated effects arise when a group of students is subject to common factors. We consider the possibility of factors at the district, school, and class levels, where θ_{dt}, u_{sdt}, and ν_{csdt} capture the observed and unobserved effects at each of these levels of aggregation.

Note that a simultaneity problem emerge once we allow for $\varphi \neq 0$. The nature of the problem is as follows: a student's test scores s_{icsdt} is affected by the unobservable errors ε_{icsdt}. Since these are present for each student, the errors affect $\bar{s}_{(-i)csdt}$. This will lead to a bias if equation (1) is estimated by least squares. This is particularly relevant when comparing the effects of different targeted interventions on education. Our goal, however, is the measure the average neighborhood effect, which encompasses both the endogenous and exogenous peer effects. For that reason, we resort to estimating the reduced-form equation below

$$
\begin{aligned}
s_{icsdt} &= \frac{c}{1-\varphi} + \bar{x}'_{(-i)csdt}\frac{\delta+\varphi\beta}{1-\varphi} + \theta_{dt} + u_{sdt} + \nu_{csdt} + x'_{icsdt}\,\beta + \mu_t + \varepsilon_{icsdt} \\
&= \tilde{c} + \underbrace{\bar{x}'_{(-i)csdt}\,\lambda}_{\text{(Adjusted) peer effects}} + \underbrace{\theta_{dt} + u_{sdt} + \nu_{csdt}}_{\text{Correlated peer effects}} + \underbrace{x'_{icsdt}\,\beta}_{\text{Controls}} + \mu_t + \varepsilon_{icsdt},
\end{aligned}
\tag{2}
$$

where the parameter vector λ captures the marginal peer effects, which is our object of interest. The estimation of λ is not as trivial as it seems. Consider the possibility of some of the correlated peer effects (θ_{dt}, u_{sdt} or ν_{csdt}) being correlated with the predetermined variables ($\bar{x}_{(-i)csdt}$). This could arise, for instance, if wealthier families lived in certain districts of the city (i.e. $\bar{x}_{(-i)csdt}$ being correlated with an unobserved part of θ_{dt}) or if weaker students were sorted into a particular set of weaker schools (i.e. $\bar{x}_{(-i)csdt}$ being correlated with an unobserved part of u_{sdt}). If such conditions hold, then the estimation of λ by least squares will be biased.

Indeed, the task of capturing the collection of local factors that shapes the accumulation of human capital is made complicated by the existence of sorting and selection biases. As discussed by Heckman and Vytlacil (2007), disentangling these biases econometrically is extremely demanding either in terms of data requirements or restrictions needed. For instance, although the public education system for primary schools in Sao Paulo is centrally funded by the state or municipal governments, there is still a nonnegligible amount of selection among schools as documented by Alves et al. (2015). On the demand side, parents try to avoid schools known in the neighborhood to be bad. They also circumvent the restriction on distance from home by using other addresses, such as their work's. On the supply side, schools try to avoid some students known to be problematic by informally blocking their enrollment or transferring them to other schools. These factors thus imply the existence of some correlation between $\bar{x}_{(-i)csdt}$ and u_{sdt}. To address these issues, we control for school fixed effects and assume the assignment of students to classes to be random. This allows us to use within-school variation to estimate the peer effects λ according to the following two education production functions

$$
s_{icsdt} = \tilde{c} + \bar{x}'_{(-i)csdt}\,\lambda + x'_{icsdt}\,\beta + \alpha_{sd} + \mu_t + \varepsilon_{icsdt} \tag{3}
$$

and

$$
s_{icsdt} = \tilde{c} + \bar{x}'_{(-i)csdt}\,\lambda + x'_{icsdt}\,\beta + \tilde{\alpha}_{sdt} + \varepsilon_{icsdt}, \tag{4}
$$

where α_{sd} is a school (or district) fixed effect and $\tilde{\alpha}_{sdt}$ is a school (or district) time fixed effect.[5]

The regression results are presented in Table 2. Only coefficients for the peer effects are shown and these should be interpreted in terms of standard deviations of the entire sample. All variables that

[5] Note that the identification of peer effects already implies the identification of more general neighborhood effects. The estimate thus serves as a lower bound for them. Moreover, the estimated fixed effects, which capture both neighborhood effects at the school and district levels together with the selection bias, can be added to the estimated peer effects to give us an upper bound for what the neighborhood effects actually are.

represent shares of a classroom are peer effects calculated from dummy variables. Standard errors are clustered at the classroom level. The controls included are the individual values for the peer effect variables as well as gender and race (race only as control for individual effects since it is not significant as a peer effect) and other inputs such as class size and student and parent efforts. A full description of variables can be found in Table A.1 in the Appendix.

Column (1) refers to equation (3) without the school fixed effects. Most coefficients show the expected signs already documented in the literature: class size and peers' retention, abandonment, time spent working (either at home or outside) are negative, whereas the effort of peers and their parents either directly reported or indirectly measured by the share of exam takers in the classroom are positive. There are, however, two regressors whose estimated coefficients show statistically significant and sizeable negative effects when we would undoubtedly expect them to be positive: the share of peers' parents with a college degree (the effect is also substantially negative for a student's own parents). This is puzzling but we offer a possible explanation.[6] The dataset only comprises students from public schools and public schools in Brazil are on average much worse than private schools. Parents with college degrees are, on average, more able to afford private primary school for their children. This could signal that those who send their children to study at a public school do not probably value education as much. Therefore, the negative return on parents' college education probably captures a strong negative selection bias.[7]

Columns (2) and (3) are the relevant results for equation (3), while Columns (4) and (5) are the relevant ones for equation (4). Columns (2) and (4) control for district fixed effects, implying that within a district there is nearly random assignment of families to homes and their children to schools and classrooms. Column (3) and (5), on the other hand, adopt a more conservative assumption and assumes that the random assignment happens within each school, but probably not within a district.[8] Our preferred specifications are the ones presented in Columns (3) and (5) for being more realistic. Overall, the results of these regressions show that the estimated coefficients change significantly when we control for school effects and use within-school variation for estimation relative to Column (1).[9]

We note two effects of particular interest. A negative one resulting from peers who have been retained in a school year in the past, who work outside (for a wage or not), and who report low effort

[6] The average test score for students whose parents have a college degree is lower than the average score for students whose parents have at most a high school degree for every single year.

[7] This peculiarity is not found for test results for students in middle school however. This strengthens the argument, since private middle school education is more expensive than private elementary education. The selection bias thus gets attenuated and average test scores are the highest for parents with college degrees. This is the reason why the parents' college variable is explicitly included in the regression (a dummy for a student's own parents as well, but not shown in the table).

[8] Regression (5) does not include some variables of family background due to some numerical restrictions (estimation algorithm does not converge).

[9] Most regressors are in terms of shares of a total thus belonging to the $[0, 1]$ interval. That makes it straightforward to compare them and see which ones are more relevant.

Table 2: Neighborhood Effects

	(1)	(2)	(3)	(4)	(5)
Class size	-0.0059***	-0.0013**	0.0022***	-0.0017**	0.0099***
	(0.0006)	(0.0006)	(0.0007)	(0.0007)	(0.0011)
Share of exam takers in class	0.4185***	0.3638***	0.2963***	0.3692***	0.3408***
	(0.0326)	(0.0321)	(0.0316)	(0.0326)	(0.0352)
Share of male classmates	-0.0808***	-0.0920***	-0.1323***	-0.0914***	-0.1073***
	(0.0251)	(0.0245)	(0.0237)	(0.0245)	(0.0245)
Share of no female adult in family	-0.1585***	-0.1487***	-0.0884*	-0.1349**	
	(0.0550)	(0.0542)	(0.0525)	(0.0543)	
Share of female adult but not mom	-0.0140	-0.0030	-0.0154	0.0229	
	(0.0721)	(0.0706)	(0.0679)	(0.0706)	
Share of male adult but not dad	-0.0834*	-0.1499***	-0.0372	-0.1603***	
	(0.0454)	(0.0439)	(0.0422)	(0.0440)	
Share of no male adult in family	-0.0924***	-0.1531***	-0.0936***	-0.1606***	
	(0.0258)	(0.0254)	(0.0250)	(0.0253)	
Share of low effort in math	-0.0110	-0.0246	0.0091	-0.0240	-0.1421**
	(0.0805)	(0.0764)	(0.0702)	(0.0746)	(0.0671)
Share of middle effort in math	-0.1161***	-0.1161***	-0.0766***	-0.1147***	-0.2054***
	(0.0287)	(0.0280)	(0.0271)	(0.0280)	(0.0267)
Share of low effort in Portuguese	-0.3560***	-0.3352***	-0.2651***	-0.3359***	-0.4761***
	(0.0614)	(0.0597)	(0.0570)	(0.0593)	(0.0536)
Share of middle effort in Portuguese	-0.0476*	-0.0396	-0.0087	-0.0447*	-0.1900***
	(0.0267)	(0.0260)	(0.0252)	(0.0260)	(0.0246)
Dad's education in years (PE)	0.0177***	0.0188***	0.0094***	0.0193***	0.0073***
	(0.0027)	(0.0027)	(0.0026)	(0.0027)	(0.0026)
Mom's education in years (PE)	0.0313***	0.0345***	0.0151***	0.0344***	0.0121***
	(0.0030)	(0.0029)	(0.0029)	(0.0029)	(0.0029)
Share of dads with college degree	-0.1050*	-0.1549***	-0.1358***	-0.1722***	-0.1248**
	(0.0558)	(0.0544)	(0.0513)	(0.0546)	(0.0515)
Share of moms with college degree	-0.2441***	-0.3432***	-0.2801***	-0.3434***	-0.2214***
	(0.0520)	(0.0509)	(0.0485)	(0.0508)	(0.0482)
Share of retention	-0.2962***	-0.3015***	-0.2554***	-0.3049***	-0.1860***
	(0.0225)	(0.0219)	(0.0219)	(0.0220)	(0.0233)
Share of abandonment	-0.1829***	-0.1635***	-0.0983**	-0.1615***	-0.0801**
	(0.0414)	(0.0402)	(0.0384)	(0.0401)	(0.0388)
Share of parent incentive	0.2634***	0.1853***	0.0908*	0.1925***	0.0947*
	(0.0532)	(0.0516)	(0.0496)	(0.0523)	(0.0534)
Share of domestic work	-0.1267***	-0.0532**	-0.0230	-0.0554***	-0.0379*
	(0.0215)	(0.0212)	(0.0204)	(0.0211)	(0.0204)
Share of working classmates	-0.5385***	-0.4863***	-0.3919***	-0.4926***	-0.3777***
	(0.0323)	(0.0316)	(0.0306)	(0.0316)	(0.0301)
Fixed effects	No	Yes	Yes	Yes	Yes
		(District)	(School)	(District × Time)	(School × Time)
Time effects	Yes	Yes	Yes	No	No
Adjusted R2	0.3014	0.3072	0.3195	0.3086	0.3162
Number of observations	208,940	208,940	208,940	208,940	219,427

SEs are clustered at the classroom level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

in doing homework. These peers have strong negative effects on a student's learning. On the other hand, the positive effects relate to the share of peers taking the exam and parental effort. The former may be the result of a competition effect: the more peers take the exam, the more pressure they may impose on their classmates. The latter shows the relevance of parents being supportive of their children studies, which spills over to classmates. Beyond the intricacies of which factors are the most important peer effects, we want to show that these exist and are sizable. Next, we want to compute an aggregate measure of the neighborhood effects.

We measure the aggregate peer effect as $\bar{x}'_{(-i)csdt}\lambda$ in equation (2) at a certain location for our benchmark specification in Column (3). We use the estimated coefficients that are significant at least at the 5% significance level and the aggregation is be done at the district level. Figure 2 plots the estimated peer and neighborhood effects aggregated for nearly all districts[10] in the city against these districts' household income per capita measured by the 2010 Census, and each subplot is for a different year. The y-axis corresponds to the difference between each district's aggregate peer effect and the lowest aggregate peer effect in that year. The x-axis ranks districts by their annual household income per capita in 2010. Each blue dot represents a district and thus shows the expected gain in test scores of moving a student from the worst district to that particular district. The orange dots, on the other hand, add to the blue dots the change in school fixed effects from moving from the worst district. These dots measure not only the gains from having better peers, but also the gains from better schools—hence the aggregate neighborhood effects.[11]

We can draw many conclusions from the results depicted in Figure 2. First, neighborhood effects, measured by classroom peer effects, exist. This is shown by their strictly positive value. Second, the vertical difference between the orange dots and the blue dots capture a mixture of the remaining components of neighborhood effects and the selection bias. If we assume selection bias to be positively correlated with the effects of the other components, then we have an upper bound (in absolute value) on how sizable these remaining effects are. Note that for some districts, the orange dots lie below the blue dots. That means that even though there is a gain coming from peer effects in that district, the combination of selection bias and the remaining neighborhood effects is worse there as compared to the benchmark. Third, peer effects are very weakly increasing in household income per capita, but the upper bound on neighborhood effects increases more strongly with income. Fourth, there appears to be diminishing returns in neighborhood effects of moving from a poorer district to a richer one. Finally, the results show that peer effects make up most of the neighborhood effects in poorest areas. This is consistent with the idea that poor districts have little to offer in terms of public goods that

[10]Two outliers are excluded: the districts of Morumbi and Vila Andrade. Both districts are contiguous and present great social disparities: upscale neighborhoods and slums exist alongside each other. That implies low values for neighborhood effects, since test score data comes exclusively from schools in the poor areas, but high average income per capita. However, a more granular approach to the geolocation data must correct for this issue.

[11]The red line represents the estimated median for a district's expected gain in peer effects conditional on its income. The conditional median is used to correct for outliers. The black line represents the estimated median of a district's total gain (peer plus fixed effects) conditional on the district's income.

are relevant to children's education (e.g. good public schools, parks and recreational areas, security, public healthcare, access to safe water and basic sanitation and so on). This is also consistent with good parents being able to afford and sorting into better neighborhoods. These results will guide our modeling assumptions and be central to the experiments we conduct in the next sections.

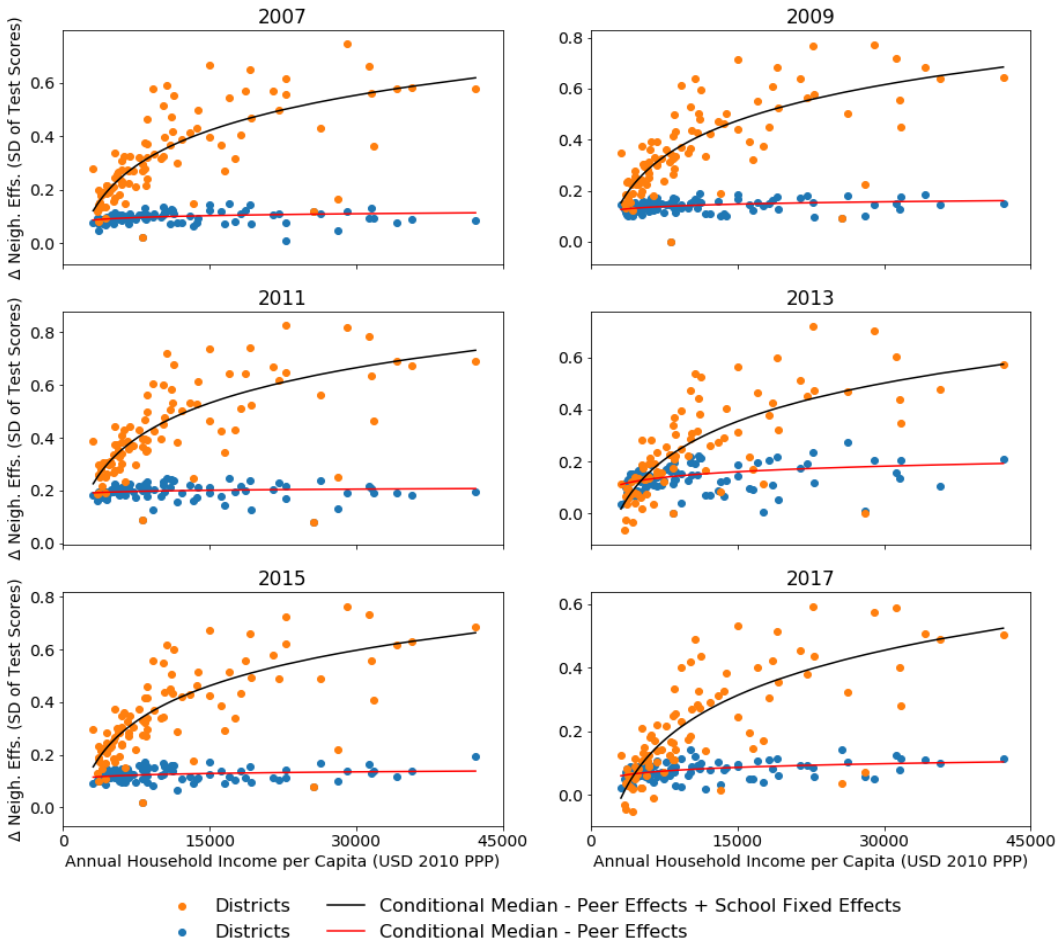


Figure 2: Estimated peer effects and school fixed effects.

3 Model

In this section we build a dynastic model with household heterogeneity in wealth and skills to study the dynamics of segregation and wealth accumulation across neighborhoods. The economy is populated by a continuum of households who live for two periods, childhood and adulthood (y and o for young and old respectively). Adults choose a neighborhood $\eta \in \{\eta_A, ..., \eta_Z\}$ to live in in exchange for rent p^η. In turn, this decision has a detrimental effect on the level of skills s their children are born with. In particualr, there are peer effects from living in a particular neighborhood η. To capture the idea that skills are affected by the peers one grows up with, s is assumed to be drawn from a distribution that depends on the average skill level of the neighborhood S^η. Therefore, when parents choose where to live, they tradeoff neighborhood rent for their children's skills. Parents are linked to their children altruistically à la Becker (1974) by caring about the expected value of their children prior to knowing their skills. Note that this differs from the paternalistic view of Fogli and Guerrieri (2019) and Ferreira et al. (2018), who assume that parents only care about their children's expected wage or human capital. Parents being altruistic elicits the important relationship between parental bequests and children's skills. Parents who expect their children to be less able will want to insure them by transferring them a significant amount of resources. There is also a unit measure of representative firms hiring labor L at wage w and renting capital K at rate R across neighborhoods to produce output. In what follows, we focus on a stationary economy.

3.1 Households

Old. The old generation derives utility from their own consumption c, represented by $u(c)$, as well as the discounted expected utility of their children. The latter is represented by V^y and the expectation is taken over the children's skills, with $\beta\gamma$ denoting the parents discount rate. Children skills are drawn from the conditional distribution $G_s(\cdot|S^\eta)$ and depend on the average skill level of the inhabitants in the neighborhood S^η given by

$$S^\eta \quad = \quad \left[\frac{\int s^\rho \mu(da, ds; \eta)}{\int \mu(da, ds; \eta)}\right]^{1/\rho} \qquad \forall\, \eta, \tag{5}$$

with ρ capturing the degree of skills substitution in the neighborhood and $\mu(a, s; \eta)$ denoting the measure of (old and young) individuals with wealth a and skills s living in neighborhood η. The neighborhood a child is born into has a lasting effect on her future labor earnings and thus parents have to tradeoff a cheaper neighborhood for a more likely skill level for their children. In addition to choosing their own consumption, parents choose the amount of wealth to transfer to their children $\tilde{a}$. They are endowed with one unit of time that is supplied inelastically in the labor market in exchange for wage w and receive the proceeds from the savings made when young, which pay interest R. Living in neighborhood η has rent denoted by p^η. The value of an adult is then given by

$$V^o(a, s; \eta) \quad = \quad \max_{c,\tilde{a}} \quad u(c) + \nu^{o,\eta} + \beta\gamma \int V^y(\tilde{a}, \tilde{s}; \eta)\, dG_{\tilde{s}}(S^\eta) \tag{6}$$

$$\text{s.t. } c + \tilde{a} + (1 + \tau^\eta)p^\eta \quad = \quad R\, a + w\, s$$

$$\tilde{a} \quad \geq \quad \underline{a}^{o,\eta} \;\geq\; 0 \quad \text{ and } \;(5)$$

where $\nu^{o,\eta}$ is an additional value derived from living in neighborhood η and $\underline{a}^{o,\eta}$ is the borrowing constraint faced by the old in neighborhood η. Note that the old has to repay any outstanding loans incurred while young and must therefore leave non-negative bequests to their children.

Young. A young individual born in neighborhood η with skill s and parental transfer a chooses how much to consume today c and how much to save for the old age a'. She ages deterministically into a parent who then chooses where to live and raise her children. The continuation value of the children is denoted by $\widehat{V}$, which corresponds to the value from choosing the best neighborhood to live in as given by

$$\widehat{V}(a,s) \quad = \quad \max_{\eta}\ V^o(a,s;\eta) \tag{7}$$

The young individual then solves the following problem

$$V^y(a,s;\eta) \quad = \quad \max_{c,a'} \quad u(c) + \nu^{y,\eta} + \beta\,\widehat{V}(a',s) \tag{8}$$

$$\text{s.t. } c + a' \quad = \quad R\,a + Tr^{\eta}$$

$$a' \quad \geq \quad \underline{a}^{y,\eta}$$

where $\nu^{y,\eta}$ is an additional value derived from living in neighborhood η and $\underline{a}^{y,\eta}$ is the borrowing constraint faced by the young in neighborhood η, which is allowed to be negative.

3.2 Remaining Agents

Firms. Firms are representative and perfectly competitive. They hire the human capital of adults in the economy and rent capital across the different neighborhoods to produce output according to technology F. The constant returns to scale technology is given by $zF(K,L)$ and capital depreciates at rate δ.

Housing market. There is an exogenous stock of houses in each neighborhood denoted by M^{η}. We let the first neighborhood be available to all households who want to move in in exchange for a rent of 0, i.e. $p^{\eta_1} = 0$.

Government. The government taxes rent across neighborhoods and redistributes the proceeds to the young individuals to balance its budget according to

$$\sum_{\eta} Tr^{\eta} \int d\mu^y(a,s;\eta) \quad = \quad \sum_{\eta} \tau^{\eta} p^{\eta} \int d\mu^o(a,s;\eta) \tag{9}$$

3.3 Equilibrium

Given an interest rate R, a stationary recursive equilibrium is a value function for the young V^y and the old V^o, a value from choosing the optimal neighborhood $\hat{V}$, policy functions $c^y, a^{y\prime}, c^o, \tilde{a}^o, \eta^{\star}$, a wage w, neighborhood rents p^{η}, government transfers Tr^{η}, and a measure μ, with μ^y and μ^o corresponding to the measure of young and old individuals respectively, such that

1. Given prices, V^y solves the problem of the young, respectively, and $c^y, a^{y\prime}$ are the associated policy functions.

2. Given prices and the density of skills with S^η satisfying (5), V^o solves the problem of the old and $c^o, \tilde{a}^o$ are the associated policy functions.

3. V is the value derived from optimally choosing a neighborhood and $\eta^\star$ is the associated policy function (7).

4. Factors prices are given by their marginal productivities

$$R = zF_K(K,L) + (1-\delta) \tag{10}$$

$$w = zF_L(K,L) \tag{11}$$

5. The government balances its budget as in (9).

6. Markets for labor and housing clear

$$L = \sum_\eta \int s \, d\mu^o(a,s;\eta) \tag{12}$$

$$M^\eta = \int d\mu^y(a,s;\eta) + \int d\mu^o(a,s;\eta) \qquad \forall\, \eta \tag{13}$$

7. The measure satisfies $\forall\, \mathcal{B}^a, \mathcal{B}^s, \mathcal{B}^\eta$

$$\mu^o(\mathcal{B}^a, s, \mathcal{B}^\eta) = \sum_\eta \int Q^{yo}\left((a,s,\eta),(\mathcal{B}^a \times s \times \mathcal{B}^\eta)\right) \mu^y(da, s, \eta) \qquad \forall\, s \tag{14}$$

$$\mu^y(\mathcal{B}^a, \mathcal{B}^s, \eta) = \int Q^{oy}\left((a,s,\eta),(\mathcal{B}^a \times \mathcal{B}^s \times \eta)\right) d\mu^o(a,s,\eta) \qquad \forall\, \eta \tag{15}$$

where Q^{yo} is the transition function from young to old and Q^{oy} the transition function from old to young defined respectively as

$$Q^{yo}\left((a,s,\eta),(\mathcal{B}^a \times s \times \mathcal{B}^\eta)\right) = \mathbf{1}_{\{a^{y\prime}(a,s,\eta)\,\in\,\mathcal{B}^a\}} \mathbf{1}_{\{\eta^\star(a^{y\prime}(a,s,\eta),s)\,\in\,\mathcal{B}^\eta\}} \tag{16}$$

$$Q^{oy}\left((a,s,\eta),(\mathcal{B}^a \times \mathcal{B}^s \times \eta)\right) = \mathbf{1}_{\{\tilde{a}^o(a,s,\eta)\,\in\,\mathcal{B}^a\}} \int \mathbf{1}_{\{\tilde{s}\,\in\,\mathcal{B}^s\}} dG_s(\tilde{s}|S^\eta), \tag{17}$$

where S^η satisfies equation (5) for each neighborhood η. Since the countable sum of measures is a measure, we have for a Borel σ-algebra $\mathcal{B} = \mathcal{B}^a \times \mathcal{B}^s \times \mathcal{B}^\eta$ that total population satisfies

$$\mu(\mathcal{B}) = \mu^y(\mathcal{B}) + \mu^o(\mathcal{B}) \tag{18}$$

3.4 Characterizing the Trade-offs

Choosing a neighborhood. The decision of where to live and how much resources to pass on to the next generation poses a tradeoff to the adult. She can either spend more today in rent to ensure her child draws a high skill level that can pay off when the child becomes an adult or, alternatively, the parent can choose to live in a cheaper neighborhood and transfer more bequests to her child to ensure she has enough resources to compensate her for the possibility of being low-skilled.

To understand the dynamics at play, we assume for simplicity that there are two neighborhoods (A and B) with the housing supply in A being restricted—hence $p^A > 0$—while rent in B is free (i.e. $p^B = 0$). We abstract from property taxes, transfers, as well as the instantaneous value from living in a particular neighborhood. Let x denote the resources an individual of wealth a and skills s has, such that $x = R\,a + w\,s$. We are interested in the defining the indifference curve that makes an adult being indifferent between living in neighborhood A and neighborhood B. For that purpose, we fix the average skill in neighborhood B to be 1 and take the rental price in neighborhood A as give. Then, we can easily draw the optimal choice of a neighborhood on a $S^A x$-plane, which we do below. First, we analyze the indifference set implicitly determined by the relationship $V^o(a, s, \eta_A) = V^o(a, s, \eta_B)$, given by

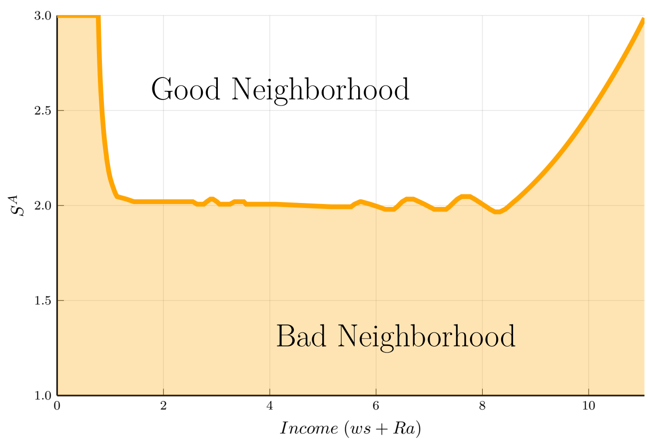
$$
\begin{aligned}
F(x, S^A) \equiv u\left(x - \tilde{a}(x; S^A) - p^A\right) + \beta\gamma \int V^y\left(\tilde{a}(x; S^A), \tilde{s}\right) dG_{\tilde{s}}\left(S^A\right) & \\
- u\left(x - \tilde{a}(x; S^B)\right) - \beta\gamma \int V^y\left(\tilde{a}(x; S^B), \tilde{s}\right) dG_{\tilde{s}}(S^B) &= 0
\end{aligned}
$$

and inspect how the average skill level in neighborhood A changes for different values of an individual's resources using the implicit function theorem as follows

$$
\begin{aligned}
\frac{\partial S^A}{\partial x} &= -\frac{F_x}{F_{S^A}} \\
&= \frac{u_c\left(x - \tilde{a}(x; S^A) - p^A\right) - u_c\left(x - \tilde{a}(x; S^B)\right)}{-\beta\gamma \int \frac{\partial}{\partial S^A} V^y\left(\tilde{a}(x; S^A), \tilde{s}\right) g_{\tilde{s}}(S^A) d\tilde{s}},
\end{aligned}
$$

where the Euler equation and Leibniz rule were used in the second line. The numerator is positive for $\tilde{a}(x; S^A) + p^A > \tilde{a}(x; S^B)$ and negative otherwise. Suppose the denominator is positive. Then, the region where living in A is more favorable than in B depends on the difference in the optimal bequests the parent has chosen. The indifference curve is increasing in resources for wealthy adults (i.e. $\frac{\partial S^A}{\partial x} > 0$). The intuition is as follows. For wealthy parents, living in A is relatively less attractive since they are able to rely more on savings to increase the utility of their children by leaving them more bequests. Alternatively, adults who face a tight borrowing constraint may never have a chance to afford the rent in A. Yet, once their constraint is relaxed and they have more resources available to pass on to their next generation, they find it more profitable to move to B. In that region of the state space, the indifference curve is downward-sloping. Finally, there is a an interval of individual's resources for whom having an additional dollar does not affect the choice of where to live. The reason is that they are not wealthy enough to overcome the possibility of their children growing with low skills in B, nor poor enough that makes living in A too expensive. Figure 3 below illustrates these indifference regions for different values of the average skill in (S^A), where both S^B and p^A have been fixed at 1.

Figure 3: Neighborhood choice for different average skills in A



The optimal choice of a neighborhood hinges on the differences in neighborhoods' rental prices and in their child's expected values from the different alternatives. We have explored the bequest channel for the choice of neighborhood. Note that there is also a maximum price the old is willing to pay for rent in neighborhood A, denoted by $\overline{p}^A$, which depends on the optimal choice of bequests she would have given to her child if she had chosen to live in B and her own discounted consumption if she were living in B. The following proposition explores the particular case of log utility, for which there is a closed for solution for the maximum price the adult is willing to pay to live in a particular neighborhood.

Proposition 1. *Let preferences be logarithmic. Then an individual with wealth a and skills s is willing to pay at most price $\overline{p}^\eta$ for the best neighborhood. Let the best neighborhood be denoted by A and the second best neighborhood by η, then that maximum price an adult is willing to pay to live in A is given by*

$$\begin{aligned}
\overline{p}^A &= \frac{1}{1+\tau^A}\left[\underbrace{\tilde{a}^\eta - \tilde{a}^A}_{\textit{Differences in bequests}} + \underbrace{c^\eta\left(1 - \hat{\nu}e^{\beta\gamma\left[\mathbb{E}[V^y(\tilde{a}^\eta,s;\eta)] - \mathbb{E}[V^y(\tilde{a}^A,s;A)]\right]}\right) + (1+\tau^\eta)p^\eta}_{\textit{Discounted consumption}}\right] \\
&< \frac{1}{1+\tau^A}\left[w\,s + R\,a - \tilde{a}^A\right]
\end{aligned}$$

Consumption, savings, and bequests. Next, we focus on how individuals make their consumption, savings, and bequest decisions. For an old individual with characteristics (a, s, η), the Euler equations relates the marginal utility of her own consumption with the expected marginal utility of her child's consumption in the following period, which reads as

$$u_c\left(c^o(a,s;\eta)\right) \quad \geq \quad \beta\gamma\, R \int u_c(c^y(\tilde{a}^o(a,s;\eta),\tilde{s};\eta)\, dG_{\tilde{s}}(S^\eta), \tag{19}$$

and holds with equality if $\tilde{a}^o(a,s;\eta) > \underline{a}^{o,\eta}$. Similarly, a young individual born with $(\tilde{a},\tilde{s};\eta)$ decides optimally how much to consume today by trading off consumption when old given the given optimal neighborhood choice $\eta^\star$. We thus have that

$$u_c\left(c^y(\tilde{a},\tilde{s};\eta)\right) \quad \geq \quad \beta R\, u_c\left(c^o(a^{y\prime}(\tilde{a},\tilde{s};\eta),\tilde{s};\eta^\star(a^{y\prime}(\tilde{a},\tilde{s};\eta),\tilde{s})\right), \tag{20}$$

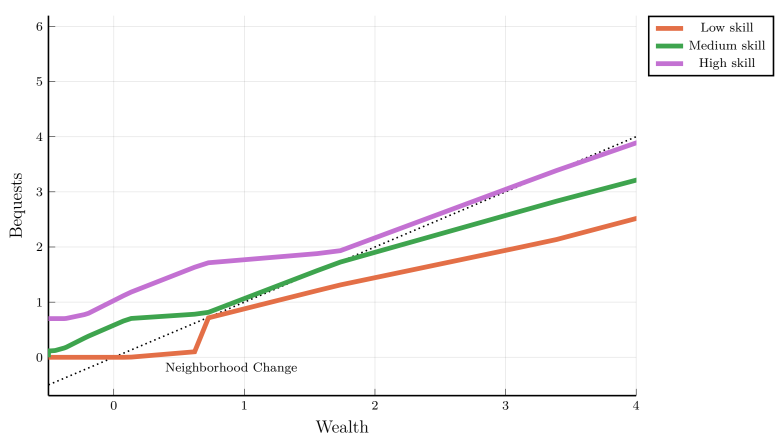
which holds with equality if $a^{y\prime}(\tilde{a},\tilde{s};\eta) > \underline{a}^{y,\eta}$.

Figure 4 below shows the bequest policy $\tilde{a}^o(a,s)$ prior to the old's neighbhorhood choice for different values of assets a and skills s, given by

$$\tilde{a}^o(a,s) \quad = \quad \begin{cases} \tilde{a}^o(a,s,A) & \text{if} \quad V^o(a,s;A) \geq V^o(a,s;B) \\ \tilde{a}^o(a,s,B) & \text{if} \quad V^o(a,s;A) < V^o(a,s;B). \end{cases}$$

The jumps in the policy function indicate changes in the optimal choice of neighborhood. For instance, a low-skilled adult born in A (in red) who attains a level of wealth above 0.7 prefers to move from neighborhood A to B, where she can transfer more bequests to her child to overcome the possibility of her becoming low-skilled herself. Note that the children who are born in A will tend to receive less bequests than their peers in B as will become clear in our quantitative exercise. The reason is that parents who choose to live in A face a trade off between paying a high rental price and leaving more bequests to their children.

Figure 4: Bequest policy prior to neighborhood choice



4 (A Somewhat Quantitative) Analysis

We use the model to elicit the tradeoffs we have described before and use it as a laboratory to understand both the distributional and spatial effects of different fiscal policies. We start by discussing our definition of neighborhood, our calibration strategy for the remaining parameters, and then present our quantitative results.

4.1 Neighborhoods

Our analysis focuses on two distinct neighborhoods: one where land available for rent is limited and therefore the rental price is demand-determined (denoted by A); and a second neighborhood, where land is plenty and hence the price of rent is zero. This distinction is motivated by the empirical analysis. We group neighborhoods that have at most one "favela" (slum) in its district according to the 2010 Brazilian Census and consider this group as the neighborhood with zoning restrictions and hence limited supply of housing. This group represents 31 neighborhoods in Sao Paulo, which corresponds to 18.7% of the population. The average monthly income in this group is high at US$2,824 and the majority of its population had achieved some schooling (literacy rate of 95.6%). We group the remaining 65 neighborhoods, which have at least two slums in their district, in B. This group represents more than 9 million inhabitants (81.3% of the population), has a lower average income than the other neighborhood (US$1,152 per month), and has lower literacy rate as shown in Table 3. We use these population ratios to set the supply of housing in neighborhood A with $M^A = 0.187$ and set the price in B to $p^B = 0$. The parametric assumption for the distribution of skills drawn by the young in each neighborhood is assumed to be a Gamma distribution, as in Ferreira et al. (2018), with shape parameter given by (S^η/θ) and a common scale parameter θ. The mean of the distribution is thus the average skill level in the neighborhood, S^η, and the variance is θS^η. We set $\theta = 0.07$. The elasticity of substitution in skills within a neighborhood is set to $\rho = 1$, which implies that the average skill level S^η is a simple mean of the skills of the individuals in the neighborhood.

Table 3: Neighborhood types

	No Slums	**Slums**
	Neighborhood A	Neighborhood B
Population	2,101,238	9,108,435
Number of neighborhoods	31	65
Literacy rate (%)	95.6	89.3
Average monthly income (R$ / 2019 US$ PPP)	R$ 6,343	R$ 2,588
	US$ 2,824	US$ 1,152

4.2 Calibration

Preferences. Individuals' preferences are represented by the CRRA utility function

$$u(c) = \frac{c^{1-\sigma} - 1}{1 - \sigma}, \tag{21}$$

where σ is the intertemporal elasticity of substitution, set to $\sigma = 2$. For simplicity, we set the neighborhood-specific utility flow while old and young to be zero for both neighborhoods (i.e. $\nu^{o,\eta} = \nu^{y,\eta} = 0$). We think of a model period as being 20 years and hence have the individuals' discount factor is set to $\beta = 0.96^{20} = 0.442$. The adjustment in the intergenerational discount factor is chosen to be $\gamma = 0.8$ and thus the discount factor towards children is $\beta\gamma = 0.3536$.

Borrowing constraints. The borrowing constraint of the young is set to $\underline{a}^{y,\eta} = -0.5$ for both neighborhoods. Since the old cannot die with debt, their borrowing constraints is set to $\underline{a}^{o,\eta} = 0$ for both neighborhoods.

Firms. The representative firm produces output according to a Cobb-Douglas technology that combines capital and labor, with the capital share of output set to $\alpha = 0.33$. Total factor productivity is set to $z = 1$ and capital is assumed to fully depreciate within the period and thus $\delta = 1$. We consider neighborhoods to be like a small open economy, where the interest rate is exogenous and assume that outside investors can provide the needed capital demanded by firms and absorb household's savings so as to ensure the interest rate is annually at 4% and hence $R = 1.04^{20} - 1 = 1.191$.

Government policy. In our baseline scenario, we start with no fiscal intervention and thus set taxes on rent across neighborhoods to $\tau^A = \tau^B = 0$. The resultant endogenous government transfers to the young are hence zero. Table 4 summarizes the calibrated parameters.

Table 4: Calibrated parameters

Variable	Description	Value
	Individuals	
σ	Intertemporal elasticity of sub	2.0
β	Discount factor of young	0.442
γ	Intergenerational discount factor	0.8
$\nu^{o,\eta}$	Utility flow from living in η while old	(0.0, 0.0)
$\nu^{y,\eta}$	Utility flow from living in η while young	(0.0, 0.0)
	Neighborhoods	
M^{η}	Supply of housing in η	(0.187, 1.0)
ρ	Elasticity of substitution in skills	1.0
θ	Scale parameter in skills dist.	0.07
$\underline{a}^{o,\eta}$	Old credit constraint in η	(0.0, 0.0)
$\underline{a}^{y,\eta}$	Young credit constraint in η	(-0.5, -0.5)
	Firms	
α	Capital share of income	0.33
δ	Depreciation rate	1.0
R	Real interest rate	1.191
	Government	
τ^{η}	Tax on rents	(0.0, 0.0)

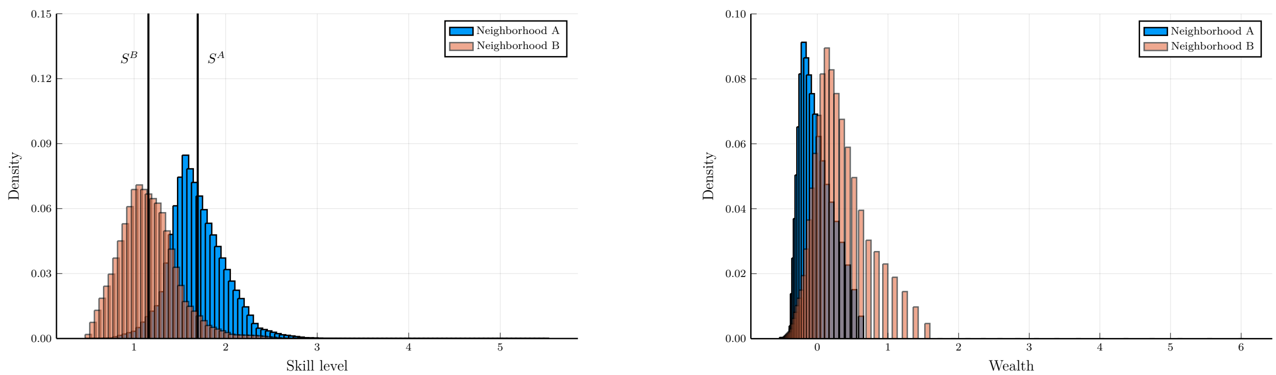
4.3 Baseline Results

We first use the model to understand how people sort into neighborhoods and how that choice shapes neighborhood-specific characteristics, such as the skill and wealth distributions. We then delve into the lasting effects sorting across distinct neighborhoods has on intergenerational mobility, in particular on whether children of low income parents can move to the best neighborhood or whether those whose parents were high skilled can live up to their parents' skills.

More highly skilled individuals live in neighborhood A than in B. Yet, not all wealthy individuals segregate in A. Figure 5a shows the stationary skills distribution in each neighborhood (blue for A and orange for B), with the vertical line denoting the average skill level S^{η}. The average skill in neighborhood A is almost 1.5 times larger than in neighborhood B, which reflects the concentration of highly-skilled population in A. Adults who choose to live in A do so to increase the probability of their children drawing better skill levels, which in turn will allow them to be more successful in the labor market in the future. However, wealthy parents have enough resources to ensure their children will survive in the future without relying too much on their labor earnings and are therefore willing to live in B and spend less on rent. To circumvent the possibility of their children drawing low skill levels, they leave a considerable amount of bequests to their children. The resulting distribution of wealth is thus more unequal in B than A as Figure 5b shows.

The total amount of savings accumulated by parents in B is much larger than what the adults accumulate in A as summarized Table 5. Although aggregate consumption is also higher in B, adults in A enjoy far more consumption in terms of income than their peers, with higher consumption to income ratios. In contrast, the youth who grow up in B are wealthier on average as a result of the high bequests they received and as a result consume far more their peers in A (both in absolute terms as well as out of their wealth).

Figure 5: Stationary density of skills and wealth in each neighborhood



(a) Skills density

(b) Wealth density

Note: The vertical line is the average skill level in each neighborhood (S^η)

In terms of intergenerational mobility, being born poor and low skilled is almost an absorbing state. Once an adult, the individual would have to have accumulated enough wealth to pay the rent in neighborhood A in order to ensure her child has better chances of drawing a high skill level and be willing to leave a significant amount of bequests. Yet, since there is still some probability of a child being born brighter than their parents—even in the worse neighborhood—there is still some degree of intergenerational mobility. Take a parent who has yet to decide which neighborhood to live in and is in the lowest quartile of the joint wealth distribution (by joint we mean the distribution of wealth across both neighborhoods pulled together). Her child has a 36.5% probability of being in the top wealth quartile of the joint distribution when adult. If instead that parent was in the top quartile of the wealth distribution, her child would almost be twice as likely to be in the top quartile of the wealth distribution. Table B.2 in the Appendix presents additional details on intergenerational mobility by wealth. Since there is no direct effect of parent's skills to their children's, there is a substantial amount of skill mobility across generations as shows Table B.3.

Delving into the distribution within neighborhoods gives us a broader picture of intergenerational mobility as Table B.4 shows. Although the adult child of a poor parent in neighborhood B faces the same probabilities of mobility as an adult whose parents were in neighborhood A, the adult child of a wealthy parent who was living in B is more likely to become wealthy in A than her peer in the same quartile of the wealth distribution who was born in A. This higher probability is the result of the wealth distribution in neighborhood being B more spread out than A. Wealthy children in B are wealthier than children born in A and those that are not highly skilled may prefer to move to A to ensure their children has higher chances of drawing higher skill levels.[12]

4.4 Subsidizing property

We now use the model to examine the effect of a policy that subsidizes property in A by imposing a lump-sum tax on the youth living in B. We chose this policy after seeing that there was a substantial amount of the population who are wealthy and prefer to live in B. The goal of the subsidy is to encourage people to move to the neighborhood that gives them more skills. We thus set the subsidy to 10% of the rent (i.e. $\tau^A = -0.1$). The resulting lump sum tax amounts to $Tr^B = -0.011$. With the introduction of the policy, the price of rent increases in A increases by 18% as a result of the rise in demand for housing in the neighborhood (and the price net of the subsidy is still 6% higher than before). The subsidy also has the effect of increasing the average skill level in both neighborhoods, but the increase is larger in B than in A (more than 2% versus 0.3%). As a result, parents transfer on average more bequests in B to compensate their children from growing up in bad neighborhood.[13]

In terms of intergenerational mobility, the housing subsidy greatly reduces the persistence of top wealth inequality. The child of a parent in the highest wealth quartile will only become a wealthy individual with 41% probability (a decline of 24 p.p.) as shows Table B.2 in the Appendix (see values in brackets).

Table 5: Results: Main aggregates

Variable	Description	Baseline	Subsidizing Property
S^η	Average skills in η	(1.695, 1.156)	(1.700, 1.181)
p^η	Rent in η	(0.387 ,0.0)	(0.457, 0.0)
w	Wage	0.353	0.353
L	Aggregate labor	0.628	0.639
L^η	Labor supplied by old in η (%)	(25.2, 74.2)	(24.9, 75.1)
$\int a\, d\mu / Y$	Aggregate savings/Output	0.430	0.468
$\int c\, d\mu / Y$	Aggregate consumption/Output	0.570	0.532
$\int a^y\, d\mu^y$	Young's savings in η	(-0.028, 0.007)	(-0.027, 0.08)
$\int a^o\, d\mu^o$	Old's bequests in η	(0.008, 0.166)	(0.017, 0.195)
$\int c^y\, d\mu^y$	Young's consumption in η	(0.005, 0.015)	(0.008, 0.014)
$\int c^o\, d\mu^o$	Old's consumption in η	(0.046, 0.137)	(0.055, 0.142)

Note: Values in parenthesis correspond to neighborhood A and B, respectively.

[12] Table B.5 presents the mobility statistics from a parent (prior to choosing a neighborhood) to her adult child for different quartiles of the skills distribution.

[13] Figure B.1 in the Appendix shows the stationary skills and wealth distribution in each neighborhood after the policy.

5 Conclusion

This paper studies the importance of location in determining social mobility by exploring the channel of neighborhood effects in the accumulation of skills. We use detailed micro-level data from 96 neighborhoods of the city of Sao Paulo to estimate neighborhood effects—measured as one's peers contribution to a student's performance in high school together with the school effect. We then rank neighborhoods by their annual income per capita and find that the combination of peer and school effects are larger in richer neighborhoods than in poor neighborhoods (the school effect strongly dominates though).

We argue that when parents choose a neighborhood to live in, that choice has long lasting effects on their children. In particular, when parents are borrowing constrained and decide to move to a different location to save on rent, their children's skills will be affected by their peers in the new neighborhood. These skills will have an impact on their earnings once they become adults. Hence, location could be more of a trap than an asset. To study this in further detail, we build an overlapping generations model that includes two neighborhoods. One of the neighborhoods is assumed to have strict zoning laws, where the supply of housing is limited and the rental price is determined in equilibrium. The second neighborhood has a rental price of 0 and has capacity to accommodate all households who wish to live there. Children draw their skills from a distribution that depends on the average skill level in the neighborhood. We find that the neighborhood with limited supply of housing ends up having a higher average skill level. However, and more surprisingly, the neighborhood with unlimited housing space is more unequal. Rich and poor people live side by side in that neighborhood. The reason is that wealthy individuals can insure their children by endowing them with a large amount of bequests that compensates the probability of being low skilled. This effect could be a rationale for why so many slums in Brazil are side by side with wealthy housing.

References

F. Agostinelli. Investing in Chldren's Skills: An Equilibrium Analysis of Social Interactions and Parental Investments. Unpublished manuscript, University of Pennsylvania, 2018.

L. Alves, A. A. G. Batista, V. M. Ribeiro, and M. Érnica. Seleção velada em escolas públicas: práticas, processos e princípios geradores. *Educação e Pesquisa*, 41(1):137–152, 2015.

A. Ammermueller and J.-S. Pischke. Peer Effects in European Primary Schools: Evidence from Progress in International Reading Literacy Study. *Journal of Labor Economics*, 27(3):315–348, 2009.

G. S. Becker. A Theory of Social Interactions. *Journal of Political Economy*, 82(6):1063–1093, 1974.

A. Bilal and E. Rossi-Hansberg. Location as an Asset. Working Paper, 2020.

R. Chetty and N. Hendren. The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. *The Quaterly Journal of Economics*, 133(3):1107–1162, 2018a.

R. Chetty and N. Hendren. The Impacts of Neighborhoods on Intergenerational Mobility II: County-level Estimates. *The Quaterly Journal of Economics*, 133(3):1163–1228, 2018b.

R. Chetty, N. Hendren, P. Kline, and E. Saez. Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quaterly Journal of Economics*, 129(4): 1553–1623, 2014.

R. Chetty, N. Hendren, and L. F. Katz. The Effects of Exposure of Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review*, 106(4): 855–902, 2016.

S. N. Durlauf and A. Seshadri. *Understanding the Great Gatsby Curve*, volume 32, pages 333–393. 2018.

F. Eckert and T. Kleineberg. Can We Save the American Dream? A Dynamic General Equilibrium Analysis of the Effects of School Financing on Local Opportunities. Unpublished manuscript, Yale University, 2019.

P. C. Ferreira, A. Monge-Naranjo, and L. T. de M. Pereira. Of Cities and Slums. Working Papers 2018-005. Human Capital and Economic Opportunity Working Group, 2018.

A. Fogli and V. Guerrieri. The End of the American Dream? Inequality and Segregation in US cities. Working Paper, 2019.

J. J. Heckman and E. Vytlacil. Extension of MTE to the analysis of more than two treatments and associated outcomes. In J. J. Heckman and E. E. Leamer, editors, *Handbook of Econometrics*, volume 6B, pages 4978–5026. North-Holland, Amsterdam, 2007.

C. Manski. Identification of endogenous social effects. *The Review of Economic Studies*, 60(3):531–542, 1993.

Appendix

A Data Appendix

Table A.1: Variable description

Variable name	Type	Individual effect in reg's	Peer effect in reg's	Description
Sex	Categorical (2)	All	All	Student's sex. 2 categories: (i) male or (ii) female
Race	Categorical (5)	All	None	Student's race. In Brazil this is generally understood as one's color of skin. 5 categories: (i) white, (ii) brown, (iii) black, (iv) yellow and (v) indigenous.
Female adult	Categorical (3)	All	All except (5)	If student lives with a female adult at home. 3 categories: (i) mom, (ii) woman (but not mom), and (iii) no female adult.
Male adult	Categorical (3)	All	All except (5)	Same as the case of a female adult, but male adult.
Effort in mathematics	Categorical (3)	All	All	Answer to whether student does math homework. Questionnaire presents 4 possible choices: (i) always or almost always, (ii) sometimes, (iii) never or almost never and (iv) 'teacher does not assign homework'. The last choice is discarded and considered NaN.
Effort in Portuguese	Categorical (3)	All	All	Same as the case of math effort, but for Portuguese.
Mom's education	Integer	All	All	Categorical variable turned integer because of dimensionality. Categories and respective integer values: (i) Never studied: 0 (ii) Did not finish elementary school: 2 (iii) Finished elementary, but not middle school: 4 (iv) Finished middle, but not high school: 8 (v) Finished high school, but not college: 11 (vi) Finished college: 15 (vii) Student does not know: NaN
Dad's education	Integer	All	All	Same as the case of mom's education, but dad's.
Mom w. coll. degree	Dummy	All	All	Value 1 if mom finished college and 0 otherwise.
Dad w. coll. degree	Dummy	All	All	Value 1 if dad finished college and 0 o.w.
Parent incentive	Dummy	All	All	Value 1 if parents incentivize to study and 0 o.w.
Retention	Dummy	All	All	Value 1 if student has ever repeated an academic year of school and 0 o.w.

Table A1 Continued: Variable description

Variable name	Type	Individual effect in reg's	Peer effect in reg's	Description
Abandonment	Dummy	All	All	Value 1 if student has ever abandoned school; 0 o.w.
Domestic work	Dummy	All	All	Value 1 if student does domestic work for more than one hour on weekdays and 0 o.w.
Work	Dummy	All	All	Value 1 if student works outside the home (for a salary or not) and 0 o.w.
Exam taker	Dummy	-	All	Value 1 if student took the exam and 0 otherwise. Used to calculate the share of peers taking the exam.
Class size	Integer	-	All	Number of students in classroom.

B Additional Model Results

Table B.2: Intergenerational mobility by wealth: From Parent to Adult child

Parent \ Adult Child	a_{Q1}	a_{Q2}	a_{Q3}	a_{Q4}
a_{Q1}	0.365	0.132	0.138	0.365
(with subsidy)	(0.380)	(0.227)	(0.175)	(0.219)
a_{Q2}	0.365	0.132	0.138	0.365
(with subsidy)	(0.380)	(0.227)	(0.175)	(0.219)
a_{Q3}	0.288	0.116	0.151	0.445
(with subsidy)	(0.295)	(0.203)	(0.223)	(0.279)
a_{Q4}	0.088	0.072	0.186	0.654
(with subsidy)	(0.108)	(0.151)	(0.329)	(0.411)

Table B.3: Intergenerational mobility by skills: From Parent to Adult child

Parent \ Adult Child	s_{Q1}	s_{Q2}	s_{Q3}	s_{Q4}
s_{Q1}	0.793	0.046	0.044	0.117
(with subsidy)	(0.808)	(0.065)	(0.052)	(0.075)
s_{Q2}	0.799	0.044	0.043	0.114
(with subsidy)	(0.809)	(0.064)	(0.052)	(0.075)
s_{Q3}	0.838	0.034	0.032	0.095
(with subsidy)	(0.814)	(0.063)	(0.051)	(0.073)
s_{Q4}	0.859	0.029	0.027	0.085
(with subsidy)	(0.818)	(0.062)	(0.050)	(0.071)

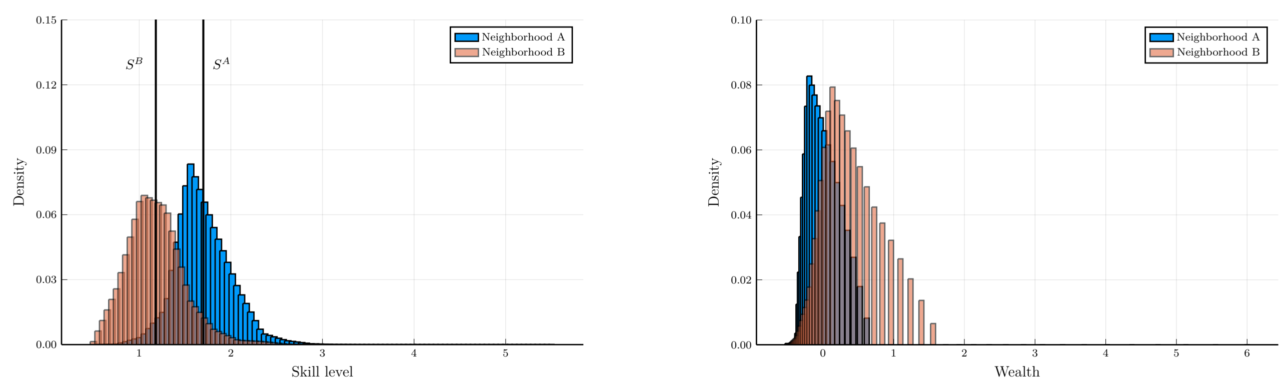
Table B.4: Intergenerational mobility by wealth and neighborhood: From Parent to Adult child by wealth

Parent	Adult child	η^A				η^B			
		a_{Q1}	a_{Q2}	a_{Q3}	a_{Q4}	a_{Q1}	a_{Q2}	a_{Q3}	a_{Q4}
η^A	a_{Q1}	0.054	0.202	0.121	0.320	0.119	0.062	0.044	0.079
	a_{Q2}	0.054	0.202	0.121	0.320	0.119	0.062	0.044	0.079
	a_{Q3}	0.054	0.202	0.121	0.320	0.119	0.062	0.044	0.079
	a_{Q4}	0.025	0.110	0.076	0.486	0.091	0.050	0.058	0.105
η^B	a_{Q1}	0.054	0.202	0.121	0.320	0.119	0.062	0.044	0.079
	a_{Q2}	0.050	0.187	0.114	0.345	0.115	0.060	0.046	0.083
	a_{Q3}	0.034	0.137	0.089	0.436	0.100	0.054	0.053	0.097
	a_{Q4}	0.010	0.061	0.052	0.573	0.076	0.044	0.065	0.119

Table B.5: Intergenerational mobility by skills and neighborhood: From Parent to Adult child by wealth

Parent	Adult child	η^A				η^B			
		s_{Q1}	s_{Q2}	s_{Q3}	s_{Q4}	s_{Q1}	s_{Q2}	s_{Q3}	s_{Q4}
η^A	s_{Q1}	0.334	0.181	0.038	0.143	0.257	0.016	0.011	0.021
	s_{Q2}	0.351	0.186	0.034	0.124	0.262	0.014	0.009	0.019
	s_{Q3}	0.357	0.188	0.033	0.118	0.264	0.014	0.008	0.019
	s_{Q4}	0.363	0.190	0.032	0.112	0.265	0.013	0.008	0.018
η^B	s_{Q1}	0.323	0.178	0.040	0.155	0.254	0.016	0.012	0.023
	s_{Q2}	0.324	0.178	0.040	0.153	0.254	0.016	0.011	0.023
	s_{Q3}	0.333	0.181	0.038	0.145	0.256	0.016	0.011	0.022
	s_{Q4}	0.353	0.187	0.034	0.122	0.263	0.014	0.009	0.019

Figure B.1: Stationary density of skills and wealth in each neighborhood after the policy



(a) Skills density

(b) Wealth density

Note: The vertical line is the average skill level in each neighborhood (S^η)