In search of better opportunities: Sorting and agglomeration effects among young college graduates in Colombia

Marigee Bacolod[*‡]
Naval Postgraduate School

Jorge De la Roca[*§]
University of Southern California

María Marta Ferreyra[*¶]
The World Bank

This version, January 2021

Abstract: We study the sorting of workers that leads to differences in skills across cities of different sizes upon labor market entry. Using administrative data of young college graduates in Colombia, we show such sorting affects the estimation of agglomeration effects. We find a substantial effect of college city size on wages, a much larger one than that of high school or work city size. We analyze how sorting concentrates population and skill in big cities. The most talented individuals sort into big cities, primarily because they move for college and remain for work. Individuals who move for work after college to smaller cities are relatively less able than others in their college cities but become the highest earners in their destination cities. Meanwhile, those who move after college to bigger cities, though relatively talented in their college cities, are not the highest earners in their destination cities.

Key words: agglomeration effects, regional migration, spatial inequality
JEL classification: J15, R23

[*]We thank Andrea Franco for her exceptional research assistance. Thanks to the Editor, two referees, Juan Pablo Chauvin, David Cuberes, Carlos Hurtado, Luis E. Quintero, Harris Selod, and participants of seminars at Clark University, Johns Hopkins Carey Business School, and the Urban Economics Association annual meeting for helpful comments and discussions. Jorge De la Roca thankfully acknowledges funding from Spain's State Research Agency MDM-2016-0684 under the María de Maeztu Unit of Excellence Programme. Maria Marta Ferreyra thanks The World Bank for funding.

[‡]Corresponding author. Graduate School of Defense Management, Naval Postgraduate School, 555 Dyer Road, Monterey, CA 93943, (e-mail: mbacolod@nps.edu; website: https://sites.google.com/site/marigeebacolod/).

[§]Sol Price School of Public Policy, University of Southern California, 650 Childs Way RGL 326, Los Angeles, CA 90089, USA (e-mail: jdelaroc@usc.edu; website: http://jorgedelaroca.name).

[¶]Office of the Chief Economist for Latin America and the Caribbean, The World Bank, 1818 H Street, NW, MSN I8-808, Washington, DC 20433 (e-mail: mferreyra@worldbank.org; website: https://sites.google.com/site/mariamartaferreyraphd/).

1. Introduction

Workers earn higher wages in bigger cities. One reason for the higher observed wages is the learning advantages that bigger cities provide and the broader array of workers' opportunities to match with a firm. An alternative explanation is the sorting of more skilled workers into bigger cities. We know workers in bigger cities exhibit higher educational attainment and work in occupations that demand more skills. Nevertheless, we continue to know little about the contribution of internal migration within countries to the sorting of more talented workers into bigger cities.

Recent studies that examine differences in skills between workers in big and small cities rely on rich administrative panel data that follow workers throughout their work lives (Combes, Duranton, Gobillon, and Roux, 2012*b*, Carlsen, Rattsø, and Stokke, 2016, Matano and Naticchioni, 2016, De la Roca and Puga, 2017). However, sorting may occur before entering the labor force as more skilled individuals may leave their small towns upon college graduation and move to a big city to commence their job search and career. Furthermore, sorting may occur even earlier as high-ability high school graduates may move to a big city to pursue a college degree and remain there after graduation to work. Therefore, the sorting of skilled individuals into big cities could occur prior to labor market entry and might be a dynamic process.

In this paper, we study the dynamic sorting that leads to differences in skills across cities of different sizes and its consequences on estimated agglomeration effects for work and college city.[1] Using rich administrative data for young college graduates in Colombia, we find a significant urban wage premium in line with the previous literature, an elasticity of wages with respect to city population of 0.052 (Duranton, 2016).[2] More importantly, our estimates reflect a substantial effect on wages for the size of the city where individuals obtain their college degree. This effect is relatively more significant and larger than for the size of the city where individuals work or complete high school. We find this to be the case even when controlling for various mediating factors such as individual ability, college selectivity, parental income and education, and study field in college.

We obtain a much higher elasticity with respect to college (0.066) than work city size (0.020) when we estimate them jointly. To understand the much larger effect of college city size, we focus on individuals who move for work after college (henceforth, work movers). These movers separately identify the college and work city size agglomeration effects. We further examine how college and work city agglomeration effects vary between those who move to smaller or bigger cities than those where they attend college. Those who move to smaller cities for work have lower standardized test scores, come from less advantaged households, and attend less selective institutions than other workers in the medium and big cities where they attend college. However, they are positively selected in their smaller work cities: they earn considerably higher wages, have much higher standardized test scores, come from more advantaged households, and attend selective institutions at a much higher rate. Hence, they become the big fish in the small pond.

[1]From here on out, the terms 'agglomeration effects' or 'agglomeration returns' refer to the estimated elasticity of wages with respect to city size, which could be due to true agglomeration economies or workers' sorting.

[2]Our data capture the universe of young college graduate workers in Colombia employed in the formal sector, described in more detail below.

In contrast, individuals who move to bigger cities for work are slightly positively selected compared to other individuals in the small and medium cities where they attend college. Nonetheless, they are negatively selected in their big work cities: they have substantially lower standardized scores, come from less advantaged households, and attend selective institutions at a much lower rate than other workers in big cities. Hence, work movers to bigger cities are relatively talented in their small college cities but become the small fish in the big pond. Taken together, the observed wages for the two types of after-college moves help explain our finding of a much higher elasticity of wages with respect to college than work city size.

Potential reasons for the relative importance of college city size come directly from the micro-foundations of agglomeration: matching, learning, and the importance of networks (Duranton and Puga, 2004, Puga, 2010). Thicker markets in bigger cities may better match college students' skills and preferences with study fields or urban amenities. Bigger cities may also offer more settings for learning, experimentation, and knowledge spillovers, such as those arising from increased summer internships or professional interactions. Further, they may provide larger and denser networks that facilitate, for instance, a job search.

Sorting may also be behind the observed college city size premium as more talented high school students may move into bigger cities to pursue post-secondary education (Ahlin, Andersson, and Thulin, 2017, Stenholt Sørensen and Holm, 2020). We investigate this earlier round of sorting by exploring the determinants of moving for college to a bigger city.[3] Not surprisingly, we find that individuals who attend high school in cities with little college availability are more likely to move for college. Individuals with more educated parents and more affluent backgrounds are also more likely to move for college, especially to attend a private institution. Surprisingly, ability—as proxied by Saber 11, a college academic readiness test—has a modest effect on the probability of moving for college to a bigger city.

Through this pre-labor market sorting, population quantity and quality concentrate in big cities as individuals move in search of better opportunities. Before any move, the spatial distribution of ability among high school graduates in our data is already more favorable in big than in small and medium cities. Migration only intensifies this advantage, as big cities attract migrants (population quantity) who are highly talented (population quality). Meanwhile, the relatively less able either remain in small and medium cities or move to such smaller locations from bigger ones.

Our findings highlight the importance of pre-labor market locational data. Administrative data sets usually follow an individual from the beginning of her work life. Researchers use repeated observations per worker to handle the inherent self-selection into the work city, identifying (work) agglomeration effects from migrants (Combes, Duranton, and Gobillon, 2008, D'Costa and Overman, 2014, De la Roca and Puga, 2017). In contrast, our pre-labor market locational data enable us to examine the spatial trajectory leading individuals to their first work city, thereby underscoring at what point of this trajectory a big city pays off the most. Moreover, our data let us identify mobility patterns consistent with a comparative advantage narrative in the spirit of Roy (1951): for instance, the most talented high school students from smaller cities move for college to big cities

[3] Since very few individuals move to a smaller city for college, we do not study this type of move in our main analysis. However, our results are robust to the inclusion of these individuals.

and succeed there, whereas the relatively less able, who attend college in the big city, move to a smaller city for work, where they become the highest earners.

Our study thus contributes to the urban literature on sorting and agglomeration. Existing evidence for developed countries shows a central role in the pre-labor market sorting of more educated and skilled individuals into big cities. College graduates with higher grades in high school and more affluent parental backgrounds sort into big cities in Sweden and Finland, mainly due to college enrollment (Ahlin, Andersson, and Thulin, 2017, Eliasson, Haapanen, and Westerlund, 2020).[4] Winters (2011) finds that US cities with a high share of college graduates tend to attract more undergraduates who end up developing local human capital and remaining there after graduation. In a related study to ours, Suhonen (2013) shows that students' pre-university characteristics such as high school grades and parental education can fully explain the positive earnings gap between Helsinki and smaller cities in Finland among young college graduates.[5]

Meanwhile, Bosquet and Overman (2019) show the significant role of birthplace on adult wages in the UK, which accounts for 0.042 or two-thirds of the 0.068 raw elasticity of wages with respect to city size. They conclude that lifetime immobility explains much of the correlation between birthplace and current city size: in their sample of Britons, 44 percent of workers never leave their home town. Our data reveal even lower mobility rates, as 67 percent of individuals in our sample do not move at all, 82 percent do not move for college, and 55 percent of those who move for college do not move for work. These rates are more in line with the low college mobility rates across macro-regions in Italy provided by Brunello and Cappellari (2008).

To our knowledge, our study is the first that uses data outside of a relatively homogeneous developed-country setting to explore pre-labor market sorting and its impact on the estimation of agglomeration effects. Pre-labor market sorting is more consequential in a developing country than in a developed country due to wider college access and quality disparities between big and small cities and tighter credit constraints that restrict young people's ability to move for college. Indeed, we document that college availability and household income are primary drivers of young workers' decision to move for college.

Our study further highlights the role of parental location decisions, or to some extent, luck. Individuals born in big cities or who grew up there do not need to move to attend college, in contrast with those who grew up in small cities with limited college access. Given the spatial disparities in college access and quality in Colombia, policies that address poverty and inequality should consider lowering college mobility costs for talented high school graduates from small towns and cities.

The rest of this paper is organized as follows. Section 2 describes the data and institutional setting. Section 3 presents our estimation of agglomeration effects and analyzes the identification of work and college agglomeration effects. Section 4 presents our findings on pre-labor market

[4] Ahlin, Andersson, and Thulin (2017) focus on urban and rural differences in Sweden and do not investigate the consequences of such sorting on wages. Eliasson, Haapanen, and Westerlund (2020) document the regional migration flows of college graduates in both Finland and Sweden, showing the self-selection into larger local labor markets of the most skilled in terms of high school grades and of the most affluent in terms of parental education and income.

[5] Other studies have also examined the internal migration of college graduates from smaller to big cities, e.g. in Finland (Haapanen and Tervo, 2012), in Germany (Busch and Weigert, 2010), in The Netherlands (Venhorst and Cörvers, 2018), among others.

sorting. We discuss the implications of our findings on the spatial distribution of skills in section 5, and conclude in section 6.

2. Data

The higher education system in Colombia

In their final year, high-school students in Colombia must take Saber 11, a standardized test covering multiple academic fields similar to high school exit exams in the US. Qualified high school graduates can enroll in either of two program types offered in Colombia's higher education system: four- and five-year programs similar to US bachelor's programs or shorter two- or three-year programs. Higher education institutions (HEI) that offer bachelor's programs include universities (henceforth, colleges), technological schools, and technical and technological institutes (hereafter, T&T). In what follows, 'moving for college' stands for moving to pursue a bachelor's degree, regardless of the HEI type attended. Similarly, 'college graduate' stands for having attained a bachelor's degree, regardless of the HEI type attended.

Most departments in Colombia have a large public college, with the largest ones usually being the most selective.[6] Average tuition at public colleges is significantly lower than in private colleges. Calculations by Carranza and Ferreyra (2019) show that for a student with annual family income equal to 12 monthly minimum wages, annual tuition for a bachelor's program represents 19 percent of income at a public college and 100 percent of income at a private college. In addition to costs, college average student ability varies substantially within and across the public and private sectors (Carranza and Ferreyra, 2019). We describe the types of colleges and other HEI types available to the individuals in our data when presenting our sample below.

Data sources and sample

Our sample draws from two source data sets. The first one is the universe of high school students who took Saber 11 between 2000 and 2009. For these students, we observe test scores (standardized by semester) and a rich set of characteristics such as gender, ethnic origin, and parental education and income. Importantly, this data set records the municipality of the high school attended by the student. The second data set is administrative data from the Labor Observatory of Education on college graduates. For these individuals, we know the type of HEI, degree attained, and study field. We also observe wages, actual years of work experience, and work municipality for those who work in the formal sector between 2007 and 2013.

Combining these two sources yields a data set of young HEI graduates ($n = 729{,}726$). For this set of individuals, we observe the type of HEI, degree attained, study field, Saber 11 standardized score, background characteristics, and crucially, the municipalities when completing high school, attending college, and beginning to work.[7] We first restrict the data to those who obtained a

[6] A department is similar to a US state and is a relatively autonomous country subdivision.

[7] For convenience, throughout the text and unless expressly noted, the term 'attend college' stands for having attended any HEI type that provides a bachelor's degree. Note that we cannot follow high school students who did not pursue post-secondary education. Further, we do not observe labor market outcomes for individuals who dropped out of a HEI, or college graduates who did not work from 2007 to 2013 or worked only in the informal sector.

bachelor's degree ($n = 473{,}775$) to eliminate vocational degree holders and focus on a more homogeneous set of graduates. We further limit the sample to individuals who had graduated from college by 2012 (since the last year we observe wages is 2013), were aged between 20 and 30 at college graduation, took the Saber 11 test between the ages of 14 and 24, and were 35 years of age or younger at the time we observe their wages. We also drop individuals who completed their bachelor's degree less than four years after taking Saber 11. These exclusion criteria leave us with 389,749 individuals. These age restrictions ensure we exclude non-traditional college students who attend college in their 30s or older to focus on individuals at a similar stage in their lifecycle—an early career stage when they are more likely to migrate.

We further restrict our sample by dropping those with missing locations and all observations in 2008 because the geographic identifiers in that year were incorrectly recorded (resulting in $n = 310{,}581$). Excluding workers with zero observed wages—either because they are long-term unemployed, not in the labor force, or informally employed—we have $n = 252{,}788$ individuals.[8] Our next restriction excludes workers in the health sector, who are subject to a mandatory rural service requirement for doctors and college graduates in medical fields.[9]

For ease of interpretation in our results below, we also leave out individuals who move to a smaller city for college. Only 2.7 percent of the sample at this stage moves to a smaller city for college, whereas 17.6 percent of individuals move to a bigger city. Nonetheless, the inclusion of those who move to a smaller city for college does not alter our results. Our final sample comprises 211,948 individuals for whom we use their last recorded annual wage when observed over multiple years.[10] We deflate nominal annual wages using the Colombian CPI (Dec 2008 = 100).

Definition of cities

For the analyses that follow, we define a 'city' as the metropolitan area constructed from the aggregation of municipalities interconnected through commuting flows, as laid out in Duranton (2015). Duranton (2015) takes all pairs of origin and destination municipalities and designs an algorithm that flags those pairs where the share of commuters from the origin is above some chosen threshold. The algorithm then identifies the commuting shed as the set of municipalities with a commuter share above the threshold located around each urban core or central city. Both the urban core and the commuting shed thus constitute the metropolitan area. Duranton (2015) proposes a commuting threshold of 10 percent, a reasonable cutoff given that Colombian municipalities tend to be fairly large. From here on out, when we refer to cities, we allude to the metropolitan areas as defined by this algorithm.

Out of the 1,122 municipalities in Colombia, the algorithm identifies 21 cities that encompass 121 municipalities, while most cities have a single municipality. Data on the population size of cities come from the 2010 Colombian census, the same data used to define cities in Duranton (2015).

[8] Our data do not allow us to identify an individual's work status outside formal employee or formal self-employee.

[9] Thus, mobility after college for health care workers is not voluntary, and salaries are set at the national level.

[10] We observe only a small subset of individuals over multiple years: 38.1 percent work only for one year, 24.8 percent for two years, 16.5 percent for three years, and only 20.6 percent work between four and seven years. We use the last observed wage to avoid capturing an internship spell immediately upon graduating from college.

Table 1: Summary statistics

	mean	S.D.		mean	S.D.
Individual characteristics			*Educational background*		
Standardized Saber 11	0.94	1.10	Attended a college	0.83	0.37
Female	0.57	0.49	Attended a technological school	0.16	0.36
Age	26.59	2.41	Attended a public college	0.39	0.49
Ethnic origin	0.01	0.10	Attended a public high school	0.43	0.49
Parental characteristics			*Higher education access*		
Very high household income	0.21	0.41	# of public colleges	3.46	2.69
High household income	0.20	0.40	# of private colleges	11.63	10.78
Moderate household income	0.22	0.41	# of technological schools	26.30	27.93
Low household income	0.29	0.45	# of T&T institutes	15.29	16.60
Very low household income	0.09	0.28			
Mother, college degree	0.30	0.46			
Mother, vocational degree	0.21	0.41			
Mother, high school degree	0.28	0.45			
Mother, $\leqslant$ high school degree	0.21	0.40			

Notes: The number of observations is 211,948. T&T stands for technical and technological.

Next, we classify a city as big if it has a population of 2.5 million or more (in order of size, Bogota, Medellin, and Cali), medium if it has 400,000 to 2.5 million inhabitants (Barranquilla, Cartagena, Bucaramanga, Cucuta, Pereira, Ibague, Manizales, Armenia, and Santa Marta), and small if it has less than 400,000 inhabitants.[11]

To define a migratory move across cities, we first create a matrix of road distances for all municipalities. We then define an individual in our data as having moved for college if her high school city is different from her college city *and* the road distance between those two cities is greater than 50 kilometers. We define moves after college for work similarly.

Descriptive statistics

Table 1 reports summary statistics for our sample of college graduates working in the formal sector. Our sample includes young workers (average age is 26.6) who are majority female (57%). The average Saber 11 ability score is 0.94 standard deviations above the national average. While a fifth (21.2%) of individuals come from households with very high income (i.e., annual income is more than five times the annual minimum wage), only 9 percent come from households with very low income (i.e., with annual incomes below the annual minimum wage). A high share (30%) of individuals have mothers with a college degree. We would expect these characteristics, given that the individuals in our sample are all college graduates in a developing country setting.

On average, the individuals in our sample have the following HEI's available to them within 50 km of their high school city: 3.5 public colleges, 11.6 private colleges, 26.3 technological schools,

[11] Based on this classification, 32.3 percent of Colombia's population lives in a big city, 17.1 percent in a medium city, and 50.5 percent in a small city or rural area.

and 15.3 T&T institutes. The majority (83%) attended a college for higher education, while the rest obtained their degree from a technological school or T&T institute.

Appendix table A.1 displays Colombia's spatial distribution of HEI types and private and public colleges by quality. The majority of public colleges are located in small cities rather than in big cities (using the definition of 'cities' as metropolitan areas outlined above). Recall that each department in Colombia has one large public college with multiple branches. Private colleges prevail in big cities, with only 15 percent of private colleges located in small cities. Technological schools and T&T institutes are also much more concentrated in big cities relative to small cities. These numbers show an uneven distribution of HEI types across the country, with a broader variety concentrated in big cities.

To characterize college quality, we classify the selectivity of public and private colleges into quintiles. We generate an average Saber 11 by college using the universe of Saber 11 test scores in 2002, an earlier period, to avoid endogeneity concerns in our analysis. Tier 1 colleges have an average Saber 11 in the upper 20 percent among all colleges, tier 2 in the 60–80 percent range, and tier 5 in the lower 20 percent. Appendix table A.1 shows that the highest quality colleges are more prevalent in big cities. Of the 14 tier 1 public colleges, six are in big cities, five in medium cities and only three are in small cities. Top-tier private colleges are nearly all in big cities, with not a single tier 1 private college located in a small city. Meanwhile, the lowest quality colleges are primarily in small cities, with nine out of ten public colleges in the bottom tier. We anticipate that given this spatial distribution in the quantity and quality of college institutions in Colombia, individuals from small cities would seek to attend college in big cities. However, to a large extent, only the most talented or with sufficient means from small cities would be able to move for college.

Finally, in appendix A, we compare the individuals in our sample with those in a nationally representative Colombian survey. We are confident that our sample captures the population of interest: young college graduates aged 20–35 in Colombia, starting their careers in the formal sector, by far, the most common labor market transition for them.

3. Agglomeration effects

Econometric framework

We assume that $y_{i,t}$, the log wages for individual i at time t, are given by:

$$y_{i,t} = \alpha\, s_{w(i,t),t} + \boldsymbol{x}_i'\boldsymbol{\beta} + \boldsymbol{z}_{i,t}'\gamma + \varepsilon_{i,t}\ , \tag{1}$$

where $s_{w(i,t),t}$ is the log size of the city where i works at time t, measured at time t; $\boldsymbol{x}_i$ is a vector of individual, household, and HEI characteristics determined prior to labor market entry; and $\boldsymbol{z}_{i,t}$ is a vector of individual and job characteristics at time t. The scalar α and the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are parameters, and $\varepsilon_{i,t}$ is an error term. Our coefficient of interest is α, which captures the elasticity of wages with respect to work city size.

Equation (1) is the most straightforward wage specification one could consider to estimate the magnitude of agglomeration effects. We can estimate it by OLS subject to two main caveats. First, $\boldsymbol{x}_i$ does not include every possible individual characteristic realized before labor market entry. As a result, we cannot address the potential sorting of more talented individuals into bigger cities. While we might address this concern by using individual fixed effects, our short panel prevents us from doing so.[12] Second, it ignores dynamic agglomeration effects and assumes that workers experience an immediate increase in wages when relocating to a bigger city (if α turns out to be positive and statistically significant) and lose such premium upon departure.[13]

To mitigate sorting concerns, we avail ourselves of a rich set of individual and HEI characteristics. Individual characteristics include Saber 11 test scores, which proxy for individual ability (defined as college academic readiness), and parental income and education. HEI characteristics include governance type (public or private), type (university or non-university), and quality (mean Saber 11 scores by institution). These variables, determined prior to labor market entry, are generally not available to researchers that use social security administrative data since these data sets track individuals after they enter the labor market. Moreover, dynamic agglomeration effects should be a relatively minor concern in our setting. Since we observe workers at the early stages of their careers, their wages mainly reflect current instead of past agglomeration effects.

Our primary goal is to jointly estimate the elasticity of wages with respect to work city size and college city size. We can extend equation (1) and incorporate the (log) size of the city where i attended college at time $t-j$, $s_{c(i,t-j),t}$, measured at time t:

$$y_{i,t} = \alpha_w s_{w(i,t),t} + \alpha_c s_{c(i,t-j),t} + \boldsymbol{x}_i'\boldsymbol{\beta} + \boldsymbol{z}_{i,t}'\boldsymbol{\gamma} + \epsilon_{i,t} \ . \quad (2)$$

Scalars α_w and α_c are the elasticities of wages with respect to work and college city size, respectively. While equation (1) attributes agglomeration effects solely to work city size, equation (2) attributes them both to work and college city size. Equation (2) entertains the possibility that individuals obtain an agglomeration return not only from the city where they work but also from the city where they attend college.[14]

In order to separately identify α_w and α_c in equation (2), it is necessary to have at least some individuals who work at time t in cities different from those where they attended college at $t-j$. In other words, these individuals need to have moved after college. Individuals who do not move after college, for whom $s_{w(i,t),t} = s_{c(i,t-j),t}$, do not allow us to separately identify α_w and α_c because their work and college cities are the same.[15]

Estimates of agglomeration effects

We begin by estimating equation (1) where we regress log wages on individual and job characteristics. We include the work city population, as a proxy of city size, and indicator variables

[12]As footnote 10 indicates, we observe a subset of workers for a few years, and only 37.1 percent of workers remain in the sample for three or more years. Moreover, wage specifications that include worker fixed effects rely on migrants (a much smaller subset) to identify agglomeration effects. Both data limitations prevent us from adding worker fixed effects in our estimations.

[13]See Combes and Gobillon (2015) for a review of the empirics on the estimation of agglomeration economies.

[14]We could also include the size of the city of high school in equation (2). It turns out that the estimated elasticity, though statistically significant, is of small quantitative importance. We leave it out in this exposition to simplify our analysis, but later consider it in estimation.

[15]We measure both work and college city sizes at time t. This prevents our identification from being driven by city size changes between $t-j$ and t for individuals who do not move after college. Given the short time elapsed between college and work (on average, 2.9 years in the sample), such city size changes would not be consequential.

Table 2: Agglomeration effects for college graduates

	Log wages				
	(1)	(2)	(3)	(4)	(5)
Log work city size	0.0523 (0.0110)***			0.0213 (0.0123)*	0.0200 (0.0120)*
Log college city size		0.0792 (0.0055)***		0.0730 (0.0089)***	0.0655 (0.0117)***
Log high school city size			0.0396 (0.0071)***	-0.0093 (0.0047)**	
Female	-0.1005 (0.0053)***	-0.0969 (0.0049)***	-0.1021 (0.0050)***	-0.0966 (0.0048)***	-0.0972 (0.0049)***
Ethnic origin	-0.1087 (0.0284)***	-0.0776 (0.0261)***	-0.0982 (0.0380)***	-0.0773 (0.0263)***	-0.0713 (0.0246)***
Technological school	-0.0845 (0.0229)***	-0.1007 (0.0228)***	-0.0839 (0.0222)***	-0.1011 (0.0219)***	-0.1013 (0.0220)***
T&T institute	-0.2228 (0.0313)***	-0.2102 (0.0432)***	-0.2218 (0.0335)***	-0.2115 (0.0431)***	-0.2125 (0.0436)***
Experience	0.1128 (0.0047)***	0.1109 (0.0062)***	0.1145 (0.0056)***	0.1101 (0.0060)***	0.1100 (0.0060)***
Age indicators	Yes	Yes	Yes	Yes	Yes
Year indicators	Yes	Yes	Yes	Yes	Yes
Two-digit sector indicators	Yes	Yes	Yes	Yes	Yes
Observations	190,749	190,749	190,749	190,749	190,749
R^2	0.2179	0.2302	0.2121	0.2324	0.2320

Notes: All specifications include a constant. Coefficients reported with robust standard errors in parenthesis, which are clustered by work city. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. T&T stands for technical and technological.

for females, ethnicities, ages, two-digit economic sectors, and calendar years. This specification resembles the standard wage specification used to estimate agglomeration effects. Our first finding is that there is a work city-size wage premium in Colombia. In column (1) of table 2, we estimate an elasticity of wages with respect to work city population of 0.052 for workers with a college degree, in line with previous estimates for the more general population of 0.054 (Duranton, 2016).[16,17]

In columns (2) and (3), we instead examine how wages vary based on the cities where individuals went to college or high school, respectively. In column (2), we obtain a higher elasticity, 0.079, for college city size, while in column (3), we obtain a lower elasticity, 0.040, for high school city size.

[16]Many studies estimate agglomeration effects in two stages (see Combes and Gobillon, 2015, for the advantages of this estimation strategy). In the first stage, they regress wages on individual and job characteristics, worker fixed effects, and city fixed effects. The second stage regresses the estimated city fixed effects on city-size measures, such as density or population size. When we implement this two-stage estimation strategy, we obtain odd values for many fixed effects of small cities since most of them have very few observations in our data. This result is not surprising as the algorithm by Duranton (2015) identifies small, isolated municipalities as independent cities. Therefore, we include in the second stage only cities that have at least 30 workers in the sample and with a population greater than 75,000 inhabitants (54 cities in total). When we incorporate this restriction, we obtain a similar work elasticity of 0.049. The elasticity increases to 0.058 when considering only cities with more than 100,000 people in the second stage.

[17]When we control for individual ability, parental education and income, and college quality, as shown in table 3, the work elasticity declines by 53 percent to 0.025 and remains highly significant. This drop suggests more productive workers sort into bigger cities, in line with the standard results obtained when including worker fixed effects in many studies that estimate agglomeration effects (Combes, Duranton, and Gobillon, 2008).

In column (4), we simultaneously include work city, college city, and high school city sizes. We find that the work elasticity shrinks from 0.052 to 0.021 and becomes only marginally significant, while the college elasticity remains relatively high at 0.073. The high school elasticity becomes negative and quantitatively small. In column (5), we estimate equation (2) not including high school city size as a regressor and again find that college elasticity is higher than work elasticity. Thus, agglomeration effects seem to mainly arise from college city size rather than work city size.[18]

Table 3 shows that the college agglomeration effect persists even as we consider mediating factors that might explain it. First, we consider that the best HEIs are generally established in big cities, as seen in appendix table A.1. Thus, agglomeration returns to college city size might reflect HEI selectivity. In column (2), we use tier indicators for HEI quality and find that, as expected, workers who obtain their college degrees from more selective institutions earn significantly more than others. For instance, workers who graduate from tier 1 institutions with the highest Saber 11 scores earn 29.1 percent ($e^{0.2553}$-1) more than those who graduate from the lowest quality institutions (i.e., the tier 5 omitted category). However, even conditional on measures of HEI selectivity, the college agglomeration effect remains substantial, with an elasticity of 0.053.[19]

Second, we estimate wage specifications controlling for individual ability (Saber 11) in column (3) and individual ability and HEI quality in column (4). As expected, workers who are of higher ability or who attend a higher-quality institution earn more. Nonetheless, the college agglomeration effect persists (at 0.046 in column 4). Third, we include study field indicators (such as engineering, education, and social sciences) in column (5) since some high-paying fields of study might be primarily offered in big cities. We also control for the public status of the HEI and high school attended. The college city agglomeration effect further declines to 0.029 and remains statistically significant. Lastly, in column (6), we consider parental background characteristics and include maternal education and household income indicators, with little effect on the estimated college elasticity (0.031). Note that the work elasticity across all columns remains relatively lower than in table 2 and statistically insignificant.

Our findings raise the question of what explains the substantial college agglomeration effects. As discussed in the Introduction, the micro-foundations of agglomeration—matching, learning, and networking—provide possible explanations (Duranton and Puga, 2004, Puga, 2010). In big cities, which offer a greater number and variety of higher education options, students are likely to find a good match to their abilities and preferences. Further, big cities provide a better setting for knowledge spillovers and more extensive networks for job search. Unobserved ability might also account for the observed college city size premium, as big cities might attract individuals with the talent to thrive in them (Glaeser and Maré, 2001, Combes, Duranton, and Gobillon, 2008).

[18] As already discussed, the elasticities are estimated exclusively based on movers when simultaneously including multiple location variables. It is not our purpose to provide causal estimates for each of these city size' elasticities. At this point, we want to underscore the relatively larger effect that college city size has on wages. To unfold potential drivers for this larger magnitude, we later analyze work movers' location trajectories.

[19] In an analogous specification, we also control for HEI selectivity linearly (average Saber 11 by HEI) and obtain a college elasticity of 0.044. As noted earlier, we do not use our sample to calculate average Saber 11 by HEI. Instead, we rely on the universe of test scores back in 2002, an earlier period, to avoid endogeneity concerns. Note that the number of observations drops in table 2 relative to table 1 because two-digit sectors are missing for ten percent of workers. Further, in table 3, observations decrease by 3.2 percent because Saber 11 data are not available for recently established HEIs. Results are similar if we include observations with missing values and add indicator variables for them.

Table 3: Mechanisms on agglomeration effects

	Log wages					
	(1)	(2)	(3)	(4)	(5)	(6)
Log work city size	0.0201 (0.0123)	0.0176 (0.0112)	0.0156 (0.0114)	0.0149 (0.0108)	0.0131 (0.0104)	0.0132 (0.0104)
Log college city size	0.0663 (0.0119)***	0.0530 (0.0133)***	0.0515 (0.0119)***	0.0460 (0.0130)***	0.0292 (0.0141)**	0.0311 (0.0133)**
Tier 1 (highest quality)		0.2553 (0.0360)***		0.1676 (0.0301)***	0.2004 (0.0343)***	0.1840 (0.0310)***
Tier 2		0.0773 (0.0353)**		0.0395 (0.0346)	0.0612 (0.0343)*	0.0541 (0.0330)
Tier 3		0.0497 (0.0256)*		0.0391 (0.0252)	0.0089 (0.0215)	0.0129 (0.0217)
Tier 4		0.0505 (0.0201)**		0.0477 (0.0174)***	0.0263 (0.0291)	0.0336 (0.0265)
Standardized Saber 11			0.1107 (0.0063)***	0.0890 (0.0036)***	0.0726 (0.0039)***	0.0671 (0.0035)***
Public high school					-0.0085 (0.0054)	0.0156 (0.0056)***
Public college					-0.1150 (0.0272)***	-0.0905 (0.0232)***
Female	-0.0968 (0.0048)***	-0.0853 (0.0051)***	-0.0611 (0.0052)***	-0.0608 (0.0056)***	-0.0427 (0.0054)***	-0.0417 (0.0052)***
Ethnic origin	-0.0688 (0.0250)***	-0.0272 (0.0212)	-0.0046 (0.0204)	0.0103 (0.0191)	0.0186 (0.0197)	0.0088 (0.0198)
Technological school	-0.0980 (0.0177)***	-0.0204 (0.0151)	-0.0424 (0.0145)***	-0.0070 (0.0128)	-0.0125 (0.0218)	-0.0063 (0.0197)
Experience	0.1115 (0.0060)***	0.1070 (0.0042)***	0.0997 (0.0046)***	0.0992 (0.0038)***	0.0937 (0.0036)***	0.0932 (0.0034)***
Study field indicators	No	No	No	No	Yes	Yes
Mother education indicators	No	No	No	No	No	Yes
Household income indicators	No	No	No	No	No	Yes
Observations	184,651	184,651	184,651	184,651	184,651	184,651
R^2	0.2324	0.2553	0.2645	0.2731	0.2982	0.3012

Notes: All specifications include a constant and indicator variables for ages, calendar years, and two-digit economic sectors. Coefficients reported with robust standard errors in parenthesis, which are clustered by work city. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. We use standardized Saber 11 scores to calculate averages for each college and group them in five tiers. Tier 5 (lowest quality) is the omitted category. Study fields include agronomy & veterinary, arts, education, social sciences, economics & business, engineering, and math & sciences.

Work vs. college agglomeration effects

As noted before, the work and college elasticities are separately identified by individuals who move for work after college—namely, the 'work movers.' There are approximately 44,000 work movers, accounting for about a quarter of the final sample. Among work movers, about half move to smaller cities than those where they attend college, and the remaining half move to bigger cities. We use the terms 'work movers to smaller cities' and 'work movers to bigger cities' for these two groups.

We now simultaneously estimate the college and work elasticities for work movers in table 4,

Table 4: Agglomeration effects for non-movers and work migrants

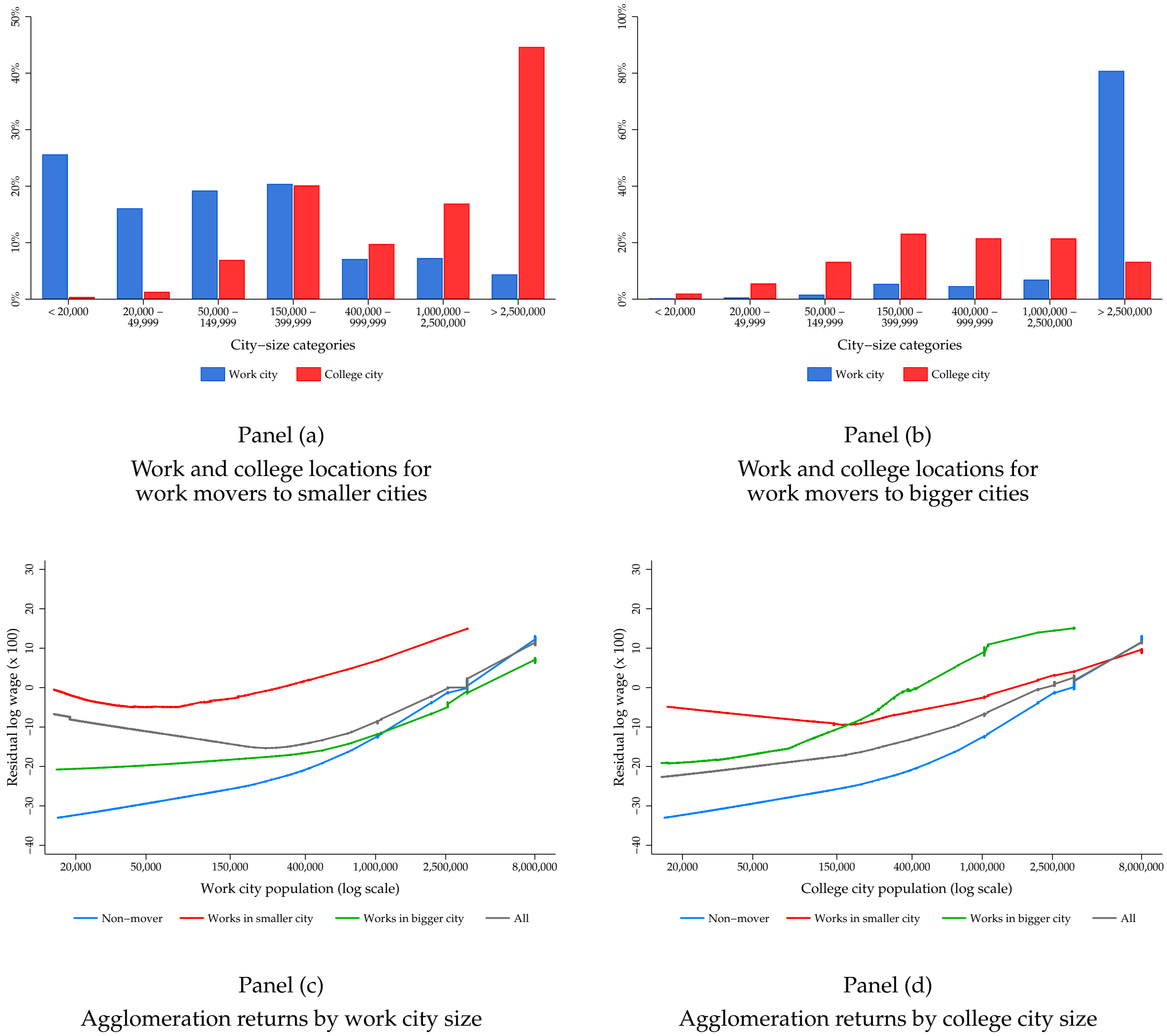
	Log wages				
	(1)	(2)	(3)	(4)	(5)
Log work city size	0.0200	0.0239	-0.0003	0.0413	0.0126
	(0.0120)*	(0.0085)***	(0.0100)	(0.0080)***	(0.0065)*
Log college city size	0.0655	0.0578	0.0522	0.0785	0.0202
	(0.0117)***	(0.0063)***	(0.0064)***	(0.0065)***	(0.0027)***
Non-movers	Yes	No	No	No	No
Work mover to smaller city	Yes	Yes	Yes	No	Yes
Work mover to bigger city	Yes	Yes	No	Yes	Yes
Mechanisms	No	No	No	No	Yes
Observations	190,749	43,666	21,879	21,787	42,840
R^2	0.2320	0.2672	0.2165	0.3375	0.3248

Notes: All specifications include a constant. Coefficients reported with robust standard errors in parenthesis, which are clustered by work city. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. Controls include experience and indicator variables for females, ethnic origins, technological schools, T&T institutes, ages, calendar years, and two-digit economic sectors. Additional mechanisms in column (5) are the same as in column (6) of table 3.

using the same controls as in table 2. Column (1) replicates column (5) of table 2 utilizing the entire sample. In column (2), we restrict the sample to work movers. Column (3) estimates the elasticities for the sample of work movers to smaller cities, and column (4) does it for the sample of work movers to bigger cities. Two main findings emerge from columns (2), (3), and (4). First, estimates for work movers in column (2) are very similar to those for the full sample in column (1) since work movers identify both elasticities. Second, the college elasticity is greater than the work elasticity for each group of work movers.

Figure 1 will help shed some light on the second finding. Panels (a) and (b) classify cities into seven size bins by population. In panel (a), the red bars show the percent of work movers to smaller cities who attend college in cities of each size bin, and the blue bars show the percent of work movers to smaller cities who work in cities of each size bin. Panel (b) does the same classification for work movers to bigger cities. Panels (c) and (d) depict residual wages for four groups of individuals: work movers to smaller cities in red, work movers to bigger cities in green, non-work movers in blue, and the full sample in gray. We recover residuals for individuals in each group from four separate regressions of log wages on the same controls used in table 2, except for work city size. Those controls are basic demographics (gender, ethnicity, and age), work experience, type of HEI attended, and indicators for two-digit economic sectors and years. We apply a locally weighted scatterplot smoothing (*lowess*) to these residual wages, and in panel (c), we show the smoothed residuals as a function of work city size. We apply the same *lowess* smoother to residual wages, and in panel (d), we show the smoothed residuals as a function of college city size. By definition, residual wages are identical for non-work movers (blue lines) in both panels (c) and (d).

Importantly, panels (c) and (d) separately plot residual wages with respect to work and college city sizes, respectively. Since neither work nor college city sizes are included in the regressions used to estimate residual wages, the residuals in panel (c) do not control for college city size, and



Panel (a)
Work and college locations for work movers to smaller cities

Panel (b)
Work and college locations for work movers to bigger cities

Panel (c)
Agglomeration returns by work city size

Panel (d)
Agglomeration returns by college city size

Figure 1: Agglomeration returns for non-work movers and work movers by city size

those in panel (d) do not control for work city size. Thus, the visual approximations discussed below are only suggestive, given that work and college elasticities are estimated simultaneously in table 4.

Using figure 1, we first examine why the college elasticity is greater than the work elasticity for work movers to smaller cities. These movers work mainly in small cities—the blue bars in panel (a) show 81 percent work in localities below 400,000 people—and earn very high residual wages across all city sizes, as seen in the upper red curve of panel (c). Thus, a linear approximation to the red curve over the relevant range of work city sizes (i.e., cities between 20,000 and 400,000 where most of these individuals work) would result in a relatively low work elasticity—almost zero, in fact, as the best linear approximation over that range would be close to a horizontal line.

At the same time, 71 percent of work movers to smaller cities attend college in medium and big cities, and 91 percent attend college in cities with over 150,000 people (as shown in the red bars in panel a). They earn residual wages that increase with college city size over the relevant college city range (as depicted in the red curve in panel d). In this case, a linear approximation to the red curve over the relevant college city range (i.e., cities over 150,000 where most of these individuals attend college) would result in a positive college elasticity. These visual approximations help us understand the results in column (3) of table 4, where the estimated work elasticity for work movers to smaller cities is zero and not statistically significant, and the college elasticity is positive and significant at 0.052.

We then examine why the college elasticity is greater than the work elasticity for work movers to bigger cities. These movers work predominantly in big cities—the blue bars in panel (b) show that 81 percent work in cities above 2.5 million people. Their residual wages increase with work city size, though at a slower pace than for non-work movers, as shown by the green and blue lines in panel (c). A linear approximation to the green curve in panel (c) over the relevant work city range (i.e., cities above 2.5 million) would generate a positive elasticity that is nonetheless moderate relative to that of non-work movers. In turn, 66 percent of work movers to bigger cities attend college in small and medium cities, mainly cities between 150,000 and 2.5 million people, and earn residual wages that increase rapidly with college city size, as depicted by the green curve in panel (d). A linear approximation to this green curve over the relevant college city range would lead to a large college elasticity relative to that of non-work movers. These visual approximations help us understand the results in column (4) of table 4, which show that the work elasticity for work movers to bigger cities is positive and significant (0.041), and the college elasticity is much larger (0.079) and also significant.

Residual wages for work movers to smaller and bigger cities help explain the higher college than work elasticity for the entire sample in column (1) of table 4. The work elasticity in column (1), 0.020, is a (weighted) average of the work elasticities in columns (3) and (4), −0.000 and 0.041, respectively. The estimate is very close to that in column (2), 0.0239, as we would expect given that the estimates in column (1) are identified by the work movers included in column (2). A similar pattern holds for the college elasticity in column (1). We can visually approximate both elasticities for the entire sample in the gray curves of panels (c) and (d) of figure 1. A linear approximation to the gray curve in panel (c) generates a relatively low work elasticity given the curve's U-shaped pattern. This U-shape emerges from the high residual wages that work movers to smaller cities earn in small localities. In contrast, a linear approximation to the gray curve in panel (d) generates a much larger college elasticity. Since 90 percent of individuals in the full sample attend college in medium and big cities, the college elasticity—more so than the work elasticity—is the one that captures agglomeration returns to city size.

The mobility patterns that we present for work movers to smaller and bigger cities are consistent with a comparative advantage narrative where individuals choose to work in the city size that pays them the most. Table 5 compares individual, household, and HEI characteristics for work movers to those of workers in their origin (college) and destination (work) cities. Column (1) reports the characteristics of work movers to smaller cities. Column (2) describes the workers with whom work movers to smaller cities would compete were they to stay in their college city, and column (3) describes the workers with whom work movers effectively compete as they move to small cities.

Table 5: Summary statistics by work move type

	Movers to smaller cities	Workers in medium or big cities	Other workers in small cities	Movers to bigger cities	Workers in small or medium cities	Other workers in big cities
	(1)	(2)	(3)	(4)	(5)	(6)
Standardized Saber 11	0.75	1.02	0.40	0.85	0.69	1.11
High household income	0.34	0.44	0.16	0.32	0.30	0.49
Mother, college degree	0.26	0.33	0.14	0.26	0.24	0.35
Tier 1 (highest quality)	0.39	0.43	0.21	0.36	0.34	0.46
Observations	24,209	170,097	17,642	24,160	69,663	118,125

Notes: Column (2) includes non-work movers from medium or big cities and work movers to bigger cities who move to medium or big cities. Column (3) includes non-work movers from small cities and work movers to bigger cities who move to a small city. Column (5) includes non-work movers from small or medium cities and work movers to smaller cities who move to small or medium cities. Column (6) includes non-work movers from big cities and work movers to smaller cities who move to a big city. High income households earn more than three times the annual minimum wage.

Similarly, column (4) reports the characteristics of work movers to bigger cities. In comparison, column (5) describes the workers with whom work movers to bigger cities would compete were they to stay in their college city, and column (6) describes the workers with whom work movers effectively compete as they move to big cities.

Columns (1) to (3) reveal that work movers to smaller cities are negatively selected compared to individuals from their origin cities but positively selected in their destination cities. Work movers to smaller cities have lower standardized test scores, come from less advantaged households, and attend less selective institutions than other workers in medium and big cities. However, work movers to smaller cities have much higher standardized test scores, come from more advantaged households, and attend selective institutions at a much higher rate than other workers in small cities. In other words, work movers to smaller cities are the big fish in the small pond.

Columns (4) to (6) suggest that work movers to bigger cities are slightly positively selected compared to individuals from their origin cities but negatively selected in their destination cities. Compared to their peers who attend college and remain for work in small and medium cities, work movers to bigger cities have higher standardized scores and are marginally more likely to come from advantaged households and attend selective institutions. Nevertheless, compared to other workers in big cities, work movers to bigger cities have substantially lower standardized scores, come from less advantaged households, and attend selective institutions at a much lower rate. In other words, work movers to bigger cities are relatively talented in their small localities, but once they relocate to big cities, they become the small fish in the big pond.

Therefore, our findings suggest substantial agglomeration effects of college city, which are larger in magnitude than work city agglomeration effects. At least in part, both effects are related to workers' sorting across cities based on their comparative advantages. In section 4, we characterize an earlier round of sorting, when individuals choose a college location. However, we first address some potential sources of concern in our estimates of agglomeration effects.

Robustness

One source of concern is that individuals may move to small cities to work in mining, oil, and other extractive industries. These sectors pay high wages and are heavily concentrated in small or rural localities. When we exclude workers from these sectors—who account for less than 5 percent of the sample—the results are virtually unaltered. We obtain work and college elasticities of 0.021 and 0.066.

Another related concern is that the high residual wages that work movers to smaller cities earn may be driven by individuals who return home to their small cities, where they may have a family business or a network advantage. In alternative estimations, available upon request, we let the work and college elasticities vary for returnees and other movers to smaller cities. The estimated elasticities are not statistically significantly different between the two groups. We conclude that return migrants do not drive the high observed wages for work movers to smaller cities.

Meanwhile, recall that table 3 shows that the college agglomeration effect persists even as we consider mediating factors that might explain it, such as HEI quality, individual ability, study field, and parental background. We have also shown that work movers to smaller and bigger cities drive the identification of the higher college than work elasticity for the entire sample. We thus need to validate if mediating factors can account for the larger college elasticity in the sample of work movers. We estimate a specification analogous to column (6) of table 3, but restrict the sample to work movers in column (5) of table 4. We find that the work and college elasticities are similar across both specifications. More importantly, the college elasticity remains larger than the work elasticity among the sample of work movers and persists when we control for a comprehensive set of mediating factors.

In supplementary estimations, available upon request, we examine how much of the college agglomeration effect can be explained by the mediating factors in each sample of work movers. We find that mediating factors altogether (those in column 6 of table 3) account for 64 percent of the estimated college elasticity among movers to bigger cities but only 46 percent among movers to smaller cities. This finding suggests that more than half of the agglomeration returns to college city size in smaller cities cannot be explained by conventional mediating factors, hinting that college city size does indeed generate returns.

4. Sorting before labor market entry

Location trajectories for non-work movers and work movers (i.e., those who move for work after college) may conceal an earlier round of sorting. Individuals may choose to pursue college education in localities with greater supply, variety and quality of HEIs, and for that reason, sort into bigger cities upon completing high school. Therefore, some individuals labeled as non-work movers or work movers in our earlier estimations might have moved earlier to attend college in a bigger city.

Table 6 reports individual, household, and HEI characteristics for non-work movers and work movers, conditional on whether they did not move for college (columns 1a to 1c) or moved for college (columns 2a to 2c). We also report average log wages for each group to illustrate the labor

Table 6: Summary statistics by migration path and work city size

	Did not move for college			Moved for college		
	Stayed	Moved for work smaller city	Moved for work bigger city	Stayed	Moved for work smaller city	Moved for work bigger city
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Standardized Saber 11	0.97	0.75	0.85	1.08	0.76	0.85
High household income	0.43	0.31	0.33	0.41	0.37	0.30
Mother, college degree	0.31	0.22	0.26	0.36	0.30	0.26
Tier 1 (highest quality)	0.41	0.38	0.34	0.47	0.39	0.42
Log wages	14.09	14.16	14.16	14.18	14.07	14.18
Observations	142,361	12,796	18,350	21,218	11,413	5,810

Notes: High income households earn more than three times the annual minimum wage.

market outcomes of these individuals. Column (1a) describes non-work movers, those who never left their high school city. Columns (1b) and (1c) describe individuals who attended college in their high school city but moved to a smaller or bigger city for work, respectively. Column (2a) describes individuals who moved for college to a bigger city after high school and remained in that city for work.[20] Columns (2b) and (2c) describe individuals who also moved for college but then subsequently moved for work to a smaller or bigger city, respectively. From here on out, we refer to these six categories as 'migration paths.'

Several aspects are noteworthy across these migration paths. First, mobility rates are low: 67 percent of college graduates in our sample never move (column 1a), 82 percent do not move for college (columns 1a to 1c), and 55 percent of those who move for college do not move for work (the number of observations in column 2a over the sum of observations from columns 2a to 2c).

Second, big cities attract many individuals from smaller cities despite low mobility rates. Among the movers' pool, i.e., those who move either for college or work (columns 1b to 2c), 65 percent start their careers working in a bigger city than their high school city. As a result, while only 52 percent of individuals attend high school in big cities, a higher fraction—62 percent—attend college in big cities, and even more (65 percent) start their careers in big cities. In other words, individuals gravitate towards big cities as they pursue college and start their work careers.

Third, big cities attract the most talented high school graduates, who then stay and work in a big city upon college graduation. Individuals who move for college to a bigger city and stay for work (column 2a) earn the highest wages, have the highest standardized test scores, are the most likely to have a mother with a college degree, and attend a top-tier institution at the highest rate. Most of the individuals (73 percent) who follow this migration path move to one of the three big cities for college.

[20] Recall that our sample does not include individuals who move for college to a smaller city than their high school city, as these account for a very small fraction of the sample. Hence, all the college moves we analyze are to bigger cities than the high school city.

Fourth, work movers who moved for college are not consistently different from work movers who did not move for college. Focusing on workers who move to bigger cities for work, those who moved for college (column 2c), and those who did not (column 1c), have similar standardized test scores, wages, and household backgrounds. That said, those who moved for college do attend top-tier institutions at higher rates (42%) compared to those who did not move for college (34%). In turn, workers who move to smaller cities for work (columns 1b and 2b) exhibit similar standardized test scores and attend top-tier institutions at similar rates. While those who moved for college come from more affluent backgrounds (column 2b), those who did not have higher wages (column 1b). Therefore, the selection patterns that we highlighted in section 3, i.e., work movers to smaller (bigger) cities are positively (negatively) selected in their destination cities, do not appear to vary by whether individuals moved previously for college.

Next, we formally examine individual characteristics as determinants of the college migration decision. To that end, we estimate a probit model where the dependent variable equals one if the individual moves for college and zero otherwise. We consider the determinants discussed in table 6 as well as measures of individuals' higher education access. These measures include the number of public and private colleges available within 50 km of an individual's high school city and their interactions with her ability and household income. We include these interactions because the mere existence of HEIs within the 50 km radius does not guarantee admission to the student. In effect, she must meet an admission standard, which depends on her ability. She must also pay for tuition and other expenses, which depends on her household income.

Table 7 reports the probit estimates for the college migration decision. Columns (1) and (2) report parameter estimates, while columns (1a) and (2a) report the marginal effects corresponding to columns (1) and (2), respectively. In this discussion, we focus on coefficients that are significantly different from zero. According to column (1), more able individuals, with more educated mothers, from households with very high incomes, and who attend high school in small or medium cities are more likely to move for college. Further, individuals with less public or private colleges available within 50 km of their high school city are more prone to move. Public colleges' availability is a more substantial deterrent of mobility than private colleges' availability, given that tuition is generally much lower at public than private colleges.

Marginal effects in column (1a) indicate that individuals from small cities are about 42 percentage points more likely to move for college than individuals from big cities. Individuals from the highest income households are 7 percentage points more likely to move for college than those from the lowest income households. The availability of an additional public college decreases the probability of moving by 3 percentage points, while the effect of a private college is weaker (1.8 percentage points.) All these margins are greater in magnitude than those of other characteristics, including ability, which has a positive but small effect. High school city size is the strongest mobility determinant, even controlling for college availability. For instance, an individual from a small city would need about 14 public colleges to be indifferent between moving and not. The effect for an individual from a medium city is lower at five public colleges but still substantial. Therefore, individuals from small and medium cities perceive a value in moving for college beyond mere college access.

Table 7: Determinants of college mobility

	Move for college			
	Probit		Marginal effects	
	(1)	(2)	(1a)	(2a)
# of public colleges	-0.2924 (0.1132)***	-0.5479 (0.1176)***	-0.0297 (0.0147)**	-0.0363 (0.0161)**
# of private colleges	-0.1802 (0.0559)***	-0.0908 (0.0581)	-0.0183 (0.0075)**	-0.0228 (0.0080)***
Standardized Saber 11	0.1713 (0.0210)***	0.2085 (0.0252)***	0.0174 (0.0054)***	0.0113 (0.0030)***
Very high household income	0.5283 (0.0648)***	0.5055 (0.1077)***	0.0698 (0.0189)***	0.0208 (0.0122)*
High household income	0.1194 (0.0681)*	0.0036 (0.0852)	0.0112 (0.0067)*	-0.0275 (0.0133)**
Moderate household income	-0.0316 (0.0580)	-0.1154 (0.0698)*	-0.0026 (0.0049)	-0.0294 (0.0117)**
Low household income	-0.0756 (0.0342)**	-0.1571 (0.0462)***	-0.0060 (0.0035)*	-0.0175 (0.0086)**
# of public colleges × Saber 11		0.0296 (0.0254)		
# of private colleges × Saber 11		-0.0187 (0.0070)***		
# of public colleges × very high hh income		0.3931 (0.1123)***		
# of public colleges × high hh income		0.3064 (0.0847)***		
# of public colleges × moderate hh income		0.1870 (0.0559)***		
# of public colleges × low hh income		0.1335 (0.0456)***		
# of private colleges × very high hh income		-0.1489 (0.0288)***		
# of private colleges × high hh income		-0.1109 (0.0209)***		
# of private colleges × moderate hh income		-0.0668 (0.0142)***		
# of private colleges × low hh income		-0.0377 (0.0110)***		
Mother, college degree	0.3039 (0.0520)***	0.3284 (0.0497)***	0.0325 (0.0108)***	0.0423 (0.0111)***
Mother, associates degree	0.1068 (0.0391)***	0.1018 (0.0372)***	0.0096 (0.0044)**	0.0109 (0.0048)**
Mother, high school degree	0.0378 (0.0248)	0.0390 (0.0233)*	0.0032 (0.0023)	0.0040 (0.0026)
Medium city during high school	0.9406 (0.3823)**	0.7928 (0.2926)***	0.1494 (0.0773)*	0.1367 (0.0622)**
Small city during high school	2.0242 (0.4154)***	1.8609 (0.3274)***	0.4151 (0.1106)***	0.4065 (0.0921)***
Observations	211,948	211,948	211,948	211,948
Pseudo R^2	0.4470	0.4573		

Notes: All specifications include a constant, the number of technological schools and T&T institutes within 50 km, and indicator variables for females, ethnic origins, public high schools, birth order, ages, and years. Coefficients and marginal effects reported with robust standard errors in parenthesis, which are clustered by high school city. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. Marginal effects evaluated at the mean values of the independent variables. Predicted probabilities for missing interactions in column (2a) are shown in figure 2. 'Mother, less than high school', 'Very low household income,' and 'Big city during high school' are the excluded categories.

Relative to column (1), estimates in column (2) add college availability interactions with student ability and household income. The interaction of ability and the number of private colleges indicates that higher ability—holding the number of private colleges constant—makes individuals less inclined to move for college, presumably because ability raises admission chances to those colleges. The interactions of household income and the number of public colleges show that public colleges' existence is less of a mobility deterrent for individuals from high-income households than for their lower-income counterparts, likely because the former can afford to move. In contrast, household income interactions with the number of private colleges show that private colleges' availability is more of a mobility deterrent for individuals from high-income households than for their lower-income counterparts, likely because the former can afford local private colleges.

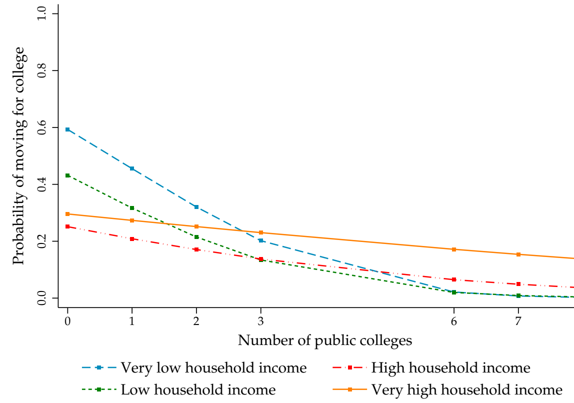
Column (2a) reports the marginal effects associated with column (2). The margins that change the most relative to column (1a) are those related to household income. We plot predicted probabilities of moving for college in figure 2 to highlight the relative roles of college availability, household income, and individual ability. The figure's panels show the predicted probability of moving for college as a function of the number of public or private colleges available for individuals of various household income levels and abilities. In general, the panels portray the decline in the probability of moving for college as the number of available public or private colleges increases. Interestingly, the deterrent effect of college availability on mobility varies more with household income than ability, as is evident from comparing the upper and lower panels. Regardless of the number of private or public colleges available, higher ability individuals are marginally more likely to move because more institutions admit them.

Altogether, our findings on the determinants of college mobility, a round of sorting that precedes labor market entry, reveal that high school city size has a sizeable effect on mobility, even when controlling for individuals' access to colleges in their localities. Parental background, measured by household income and maternal education, is also a critical determinant of college mobility. Young high school graduates from small and medium cities are eager to move to bigger cities for college, especially those from affluent, highly educated households. Further, it is the most capable who stay upon college graduation and start their work lives in big cities.

5. The spatial distribution of skills and implications of initial place

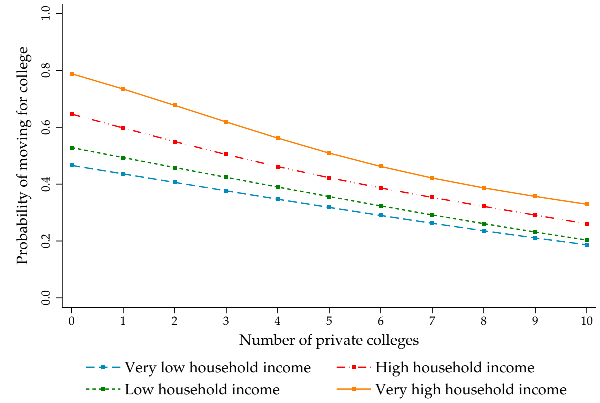
As a result of individual migration decisions, big cities absorb both quantity and quality of young college graduates. Individuals self-select into the work city size that maximizes their labor market prospects and wages, given their comparative advantage (or lack thereof) over individuals in their cities of origin or destination. In other words, individuals move in pursuit of better opportunities. We now explore further implications of these mobility patterns.

Figure 3 provides additional evidence of the concentration of high-ability individuals in big cities. The red lines display high school students' ability distributions in small, medium, and big cities. The ability distributions for individuals in big and medium high-school cities have a mild overlap. However, the distribution of ability in big cities lies clearly to the right of that in small cities. Meanwhile, the blue lines show the distributions of ability by work city size. The ability



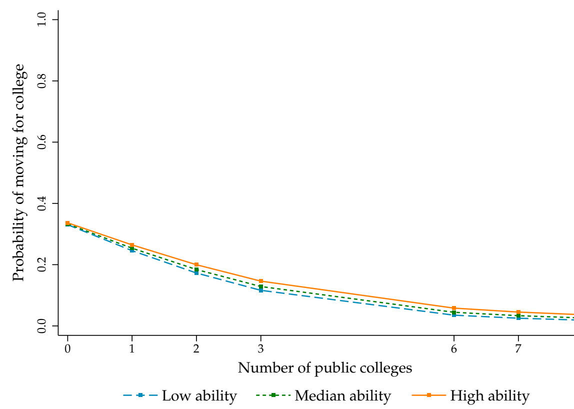
Panel (a)

Public colleges' availability and income



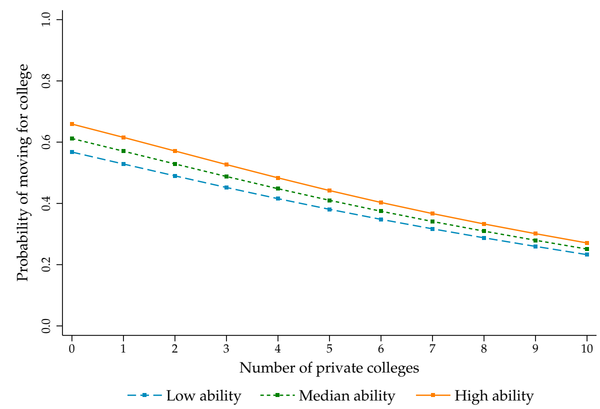
Panel (b)

Private colleges' availability and income



Panel (c)

Public colleges' availability and ability



Panel (d)

Private colleges' availability and ability

Figure 2: Marginal effects for college availability

distribution in big cities lies to the right of medium cities, and both are clearly to the right of small cities. In other words, there is a clear hierarchy where workers sort into city sizes based on ability upon labor market entry.

We formally compare the ability distributions in figure 3 by following the methodology developed by Combes, Duranton, Gobillon, Puga, and Roux (2012*a*). To illustrate the methodology, consider the ability distribution in big and small cities during high school. We approximate the ability distribution in big cities by shifting the ability distribution in small cities by an amount A and dilating it by a factor D. The estimated parameters, $\hat{A}$ and $\hat{D}$, minimize the mean quantile difference between the actual ability distribution in big cities and the shifted-and-dilated distribution in small cities.

We first focus on ability distributions in high school cities. We compare the ability distributions in small or medium high-school cities (dashed red lines) to the distribution in big high school cities (solid red line). For small high school cities, the estimated shift and dilation parameters are

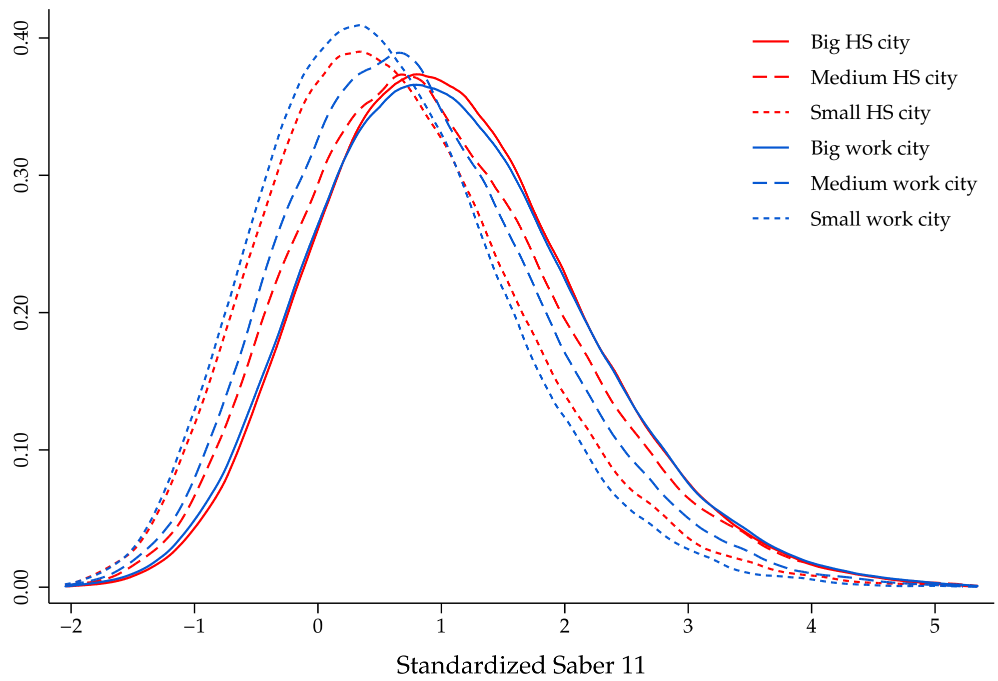
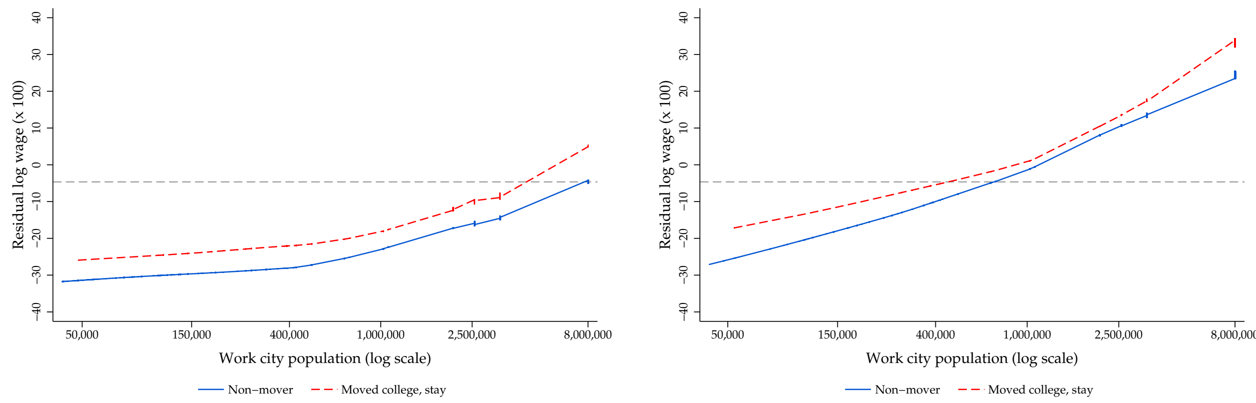


Figure 3: Ability distributions by high school and work location

$\hat{A} = 0.462$ and $\hat{D} = 1.026$, indicating that mean ability in big cities is higher than that in small ones by 0.46 standard deviations, and that ability in big cities is slightly more dispersed than in small ones. For medium high school cities, the estimated shift and dilation parameters are $\hat{A} = 0.139$ and $\hat{D} = 0.965$. Thus, the estimated mean ability difference between medium and big high school cities is much lower, at 0.14 standard deviations.

Next, we focus on ability distributions in work cities. We compare the ability distributions in small or medium work cities (dashed blue lines) to the distribution in big work cities (solid blue line). For small work cities, the estimated shift and dilation parameters are $\hat{A} = 0.544$ and $\hat{D} = 1.110$. These parameters indicate that, due to the individuals' sorting for college and work, the mean ability difference between big and small cities increases from 0.46 to 0.54 standard deviations, and ability in big cities becomes more dispersed than in small ones. The pattern is similar for medium work cities, for which the estimated shift and dilation parameters are $\hat{A} = 0.271$ and $\hat{D} = 1.036$. Hence, the mean ability gap between medium and big work cities almost doubles from 0.14 to 0.27 standard deviations after the sorting takes place. Furthermore, ability in big work cities becomes more dispersed than in medium work cities, reversing the observed pattern for high school cities.

This formal analysis confirms that, as people sort into college and work locations, the ability distribution in big cities detaches from the ability distribution in smaller cities. Since the red lines reflect the distribution that would prevail without mobility while the blue lines reflect the distribution arising from mobility choices, a comparison of both illustrates how college and work mobility generate a concentration of higher-ability individuals in big cities and lower-ability indi-



Panel (a)
Low ability (bottom quartile of Saber 11)

Panel (b)
High ability (top quartile of Saber 11)

Figure 4: Agglomeration returns by migration path and skills

viduals in smaller cities. Big cities start with an advantage in terms of ability and further attract high-ability individuals. In contrast, small cities start at a disadvantage in terms of ability. They lose high-ability individuals to bigger cities, and, while they do gain some migrants, these are not of the same ability as those they lose. The same story goes with medium cities.

We now illustrate the relative roles of ability, high school city size (which contributes to self-selection into college and hence work location), and work city size (which determines work city agglomeration effects) by depicting wage residuals for low- and high-ability individuals as a function of work city size in figure 4. In panels (a) and (b), we show residual wages for individuals in the bottom ('low-ability') and top ('high-ability') quartiles of the ability distribution, respectively. We calculate residual wages as in the lower panels of figure 1 and plot them for two groups of individuals: those who never leave their high school city in blue (i.e., those for whom high school and work city size are the same) and those who move to a bigger city for college and stay in dashed red. As expected, wages are higher for high-ability workers in every city size, especially in the biggest cities.

To illustrate the role of high school city size, consider a low-ability individual who is a non-mover from the biggest city, Bogota (on the right end of the blue line in panel a). Following the horizontal dashed gray line from panel (a) to panel (b), we find that this low-ability individual earns (in the sense of a wage residual) the same as a high-ability individual from a medium city of about 700,000 people. The low-ability individual benefits from living in Bogota, possibly due to greater college access and agglomeration economies. In contrast, the high-ability individual from the medium city has less of this advantage. In other words, high school city size—perhaps more broadly, place of origin—is decisive. The disadvantage of a small place of origin may be more salient in a setting such as Colombia, where young individuals face tight mobility constraints.

Mobility can lessen the importance of the place of origin. For example, a low-ability individual from a small city of 150,000 people can potentially increase her wages by about 30 percent if she moves to Bogota for college and remains there—i.e., a move along the dashed red line in panel (a) from a city of 150,0000 to Bogota at 8,000,000. Even more dramatically, a high-ability individual from the same small city of 150,000 people can potentially increase her wages by about 45 percent when allowed to move to Bogota for college and remain there (i.e., a similar move along the dashed red line in panel b). From an efficiency standpoint, mobility is crucial to maximizing individual earnings' potential.

6. Summary and conclusions

We study the dynamic sorting of young college graduates in Colombia, which leads to differences in skills across cities of different sizes, and its consequences on the estimation of agglomeration effects for work and college city. We find an elasticity of wages with respect to work city population of 0.052, in line with previous estimates. More notably, we find a substantial effect of college city size on wages. This effect is much larger than that of high school or work city size and persists even when we control for a comprehensive set of mediating factors, such as individual ability, college selectivity, parental income and education, and study field while in college.

To understand the much larger effect of college city size, we document mobility patterns of workers who move for work after college to smaller or bigger cities. These patterns are consistent with a comparative advantage narrative where individuals choose to work in the city size that pays them the most—namely, where individuals move in search of better opportunities. Work movers to smaller cities are negatively selected compared to individuals from their origin cities but positively selected in their destination cities. Meanwhile, work movers to bigger cities are slightly positively selected compared to individuals from their origin cities but negatively selected in their destination cities. In other words, work movers to smaller cities are the big fish in the small pond. In turn, work movers to bigger cities are relatively talented in their small localities, but once they relocate to big cities, they become the small fish in the big pond.

Through this pre-labor market sorting, we show big cities gain quantity and quality of young college graduates. Big cities start with an advantage in terms of ability and further attract higher-ability individuals who primarily move there to attend college and stay for work. In contrast, small cities start at a disadvantage in terms of ability. They lose high-ability individuals to bigger cities, and, while they do gain some migrants, these are not of the same ability as those they lose.

Our study also highlights the role of parental location decisions, or to some extent, luck. Individuals born in big cities such as Bogota, or who happen to be in a big city for high school, do not have to move to access college. In contrast, those who spend their youth in small cities with limited access to colleges have to move if they want to pursue a college degree. Talented individuals in small cities can potentially raise their future earnings by moving to bigger cities for college. However, this career path may not be accessible to everyone. Policies expanding college access, particularly for talented students from low-income families, might help mitigate spatial inequality in Colombia.

References

Ahlin, Lina, Martin Andersson, and Per Thulin. 2017. Human capital sorting: The "when" and "who" of the sorting of educated workers to urban regions. *Journal of Regional Science* 58(3): 581–610.

Bosquet, Clément and Henry G. Overman. 2019. Why does birthplace matter so much? *Journal of Urban Economics* 110: 26–34.

Brunello, Giorgio and Lorenzo Cappellari. 2008. The labour market effects of *Alma Mater*: Evidence from Italy. *Economics of Education Review* 27(5): 564–574.

Busch, Oliver and Benjamin Weigert. 2010. Where have all the graduates gone? Internal cross-state migration of graduates in Germany 1984–2004. *The Annals of Regional Science* 44: 559–572.

Carlsen, Fredrik, Jørn Rattsø, and Hildegunn E. Stokke. 2016. Education, experience, and urban wage premium. *Regional Science and Urban Economics* 60: 39–49.

Carranza, Juan Esteban and María Marta Ferreyra. 2019. Increasing higher education access: Supply, sorting, and outcomes in Colombia. *Journal of Human Capital* 13(1): 95–136.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2008. Spatial wage disparities: Sorting matters! *Journal of Urban Economics* 63(2): 723–742.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux. 2012*a*. The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica* 80(6): 2543–2594.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux. 2012*b*. Sorting and local wage and skill distributions in France. *Regional Science and Urban Economics* 42(6): 913–930.

Combes, Pierre-Philippe and Laurent Gobillon. 2015. The empirics of agglomeration economies. In Gilles Duranton, J. Vernon Henderson, and William Strange (eds.) *Handbook of Regional and Urban Economics*, volume 5. Amsterdam: Elsevier, 247–348.

D'Costa, Sabine and Henry G. Overman. 2014. The urban wage growth premium: Sorting or learning? *Regional Science and Urban Economics* 48: 168–179.

De la Roca, Jorge and Diego Puga. 2017. Learning by working in big cities. *Review of Economic Studies* 84(1): 106–142.

Duranton, Gilles. 2015. A proposal to delineate metropolitan areas in Colombia. *Revista Desarrollo y Sociedad* 75: 223–264.

Duranton, Gilles. 2016. Agglomeration effects in Colombia. *Journal of Regional Science* 56(2): 210–238.

Duranton, Gilles and Diego Puga. 2004. Micro-foundations of urban agglomeration economies. In J. Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: Elsevier, 2063–2117.

Eliasson, Kent, Mika Haapanen, and Olle Westerlund. 2020. Regional concentration of university graduates: The role of high school grades and parental background. *European Urban and Regional Studies* : 1–17.

Glaeser, Edward L. and David C. Maré. 2001. Cities and skills. *Journal of Labor Economics* 19(2): 316–342.

Haapanen, Mika and Hannu Tervo. 2012. Migration of the highly educated: evidence from residence spells of university graduates. *Journal of Regional Science* 52(4): 587–605.

International Labour Organization. 2014. Trends in informal employment in Colombia: 2009–2013. Report, International Labour Organization, FORLAC Programme for the Promotion of Formalization in Latin America and the Caribbean.

Matano, Alessia and Paolo Naticchioni. 2016. What drives the urban wage premium? Evidence along the wage distribution. *Journal of Regional Science* 56(2): 191–209.

Puga, Diego. 2010. The magnitude and causes of agglomeration economies. *Journal of Regional Science* 50(1): 203–219.

Roy, Andrew D. 1951. Some thoughts on the distribution of earnings. *Oxford Economic Papers* 3(2): 135–146.

Stenholt Sørensen, Elise and Anders Holm. 2020. Human capital and urbanization—the causal effect of high school on moving out of rural area. Processed, University of Copenhagen.

Suhonen, Tuomo. 2013. Are there returns from university location in a state-funded university system? *Regional Science and Urban Economics* 43(3): 465–478.

Venhorst, Viktor A. and Frank Cörvers. 2018. Entry into working life: Internal migration and the job match quality of higher-educated graduates. *Journal of Regional Science* 58(1): 116–140.

Winters, John V. 2011. Why are smart cities growing? Who moves and who stays. *Journal of Regional Science* 51(2): 253–270.

Appendix A.

Spatial distribution of HEIs

Table A.1: Spatial distribution of HEIs in Colombia

	Big cities	Medium cities	Small cities
HEI categories			
Public colleges	9	11	38
Private colleges	41	27	12
Technological schools	97	20	24
T&T institutes	56	15	19
Public colleges by quality			
Tier 1 (highest)	6	5	3
Tier 2	2	5	7
Tier 3	0	1	6
Tier 4	0	0	3
Tier 5	1	0	9
Private colleges by quality			
Tier 1 (highest)	10	1	0
Tier 2	8	3	0
Tier 3	9	6	3
Tier 4	9	9	4
Tier 5	4	8	4

Notes: HEI stands for higher education institution and T&T for technical and technological. Quality indicators in the middle and lower panels are not available for colleges established after 2002.

Sample representativiness

We compare demographics in our sample to those in the Socio-Economic Database for Latin America and the Caribbean (SEDLAC), a nationally representative household Colombian survey. We restrict the SEDLAC data to workers with a bachelor's degree employed in the formal sector aged 20-35. Our sample is remarkably similar to the SEDLAC sample. For instance, in the SEDLAC sample, 56 percent are female (vs. 57 percent in our sample), aged 29 on average (vs. 26.6 in our sample), and earn average annual wages of 1,692,957 Colombian pesos (vs. 1,647,592 in our sample). Crucially, the spatial distribution of the SEDLAC sample is similar to ours: 63 percent of individuals work in a big city (vs. 65 percent in our sample), 22 percent in a medium city (vs. 17 percent in our sample), and 15 percent in a small city (vs. 18 percent in our sample).

We also use SEDLAC data to examine labor force participation and work in the formal sector among young college graduates. The majority (77.5%) of those aged 20-35 with a bachelor's degree in Colombia work in the formal sector, only 2.5 percent are in the informal sector, while 20 percent are not employed. It is notable that among the 20 percent not employed, the female proportion is relatively high at 76 percent, while the female proportion among active workers is 57 percent (as in our sample).

Thus, we are confident that our sample is highly representative of the population of interest: young college graduates aged 20-35 in Colombia. Although Colombia has a large informal sector ranging between 55 to 59 percent of the workforce in 2009–2013 (International Labour Organization, 2014), informality incidence is substantially lower in our sample. As a result, we observe the great majority of college graduates when they start their careers.