*Journal of Economic Geography* (2021) **21**, 683–714
Advance Access Publication 23 November 2020
doi: 10.1093/jeg/lbaa033

# The urban–rural education gap: do cities indeed make us smarter?*

*Raoul van Maarseveen*†

Department of Economics, Uppsala University, 752 36 Uppsala, Sweden
†Correspondence to: *raoul.vanmaarseveen@nek.uu.se*

## Abstract

Despite the large urban–rural education gap observed in most countries, little attention has been paid to whether cities actually enjoy a comparative advantage in the production of human capital. Using Dutch administrative data, this paper finds that children growing in urban regions consistently attain higher levels of human capital compared with children in rural regions, conditional on observed cognitive ability and various family characteristics. The elasticity of university attendance with respect to population density is 0.07, which is robust across a variety of specifications. Hence, the paper offers a different explanation to explain the recent success of cities.

**Keywords:** urban–rural education gap, agglomeration economies, human capital
**JEL classifications:** I20, J24, R10
**Date submitted:** 17 June 2020 **Editorial decision:** 26 October 2020 **Date Accepted:** 4 November 2020

## 1. Introduction

The resurgence of the city has received widespread attention from both academics and policymakers in recent decades. Explanations typically highlight the success of cities at attracting well educated and skilled individuals from elsewhere, both by offering higher wages (see Combes et al., 2008) and superior consumption amenities (see Glaeser et al., 2001). Furthermore, cities are thought to have a comparative advantage in employing human capital in the production of goods and services, due to the sharing, matching and learning mechanisms described in Duranton and Puga (2004) and Rosenthal and Strange (2004).

A small literature has analyzed the role of cities in the production of human capital itself. Both Glaeser and Mare (2001) and De La Roca and Puga (2017) find evidence that workers in larger cities accumulate human capital at a faster rate. However, little attention has been paid to the role of cities in human capital accumulation in the period prior to labor market entry. This is surprising, since agglomeration economies are likely to

* The author would like to express his gratitude to the editor and two anonymous referees, as well as Niklas Bengtsson, Cristina Bratu, Matz Dahlberg, Peter Fredriksson, Edward Glaeser, Matthew Lindquist, Björn Öckert, Albert Saiz, Anna Tompsett, Santiago Truffa and Dinand Webbink, and the participants to the 2019 Urban Economics Association conference and VU Amsterdam/Uppsala University seminar series for their excellent comments and suggestions. Furthermore, he would like to thank CPB Netherlands Bureau for Economic Policy Analysis for making the data available for this research, as well as their excellent suggestions. All results in this paper are based on calculations by the author using non-public microdata from Statistics Netherlands. The code used to generate the results is available upon request.

influence educational investment decisions through various channels. In this paper, I investigate whether exposure to urban environments during childhood affects the educational attainment of children and young adults.

Cross-sectional evidence reveals a sizeable rural–urban gap in educational outcomes across a wide range of countries.[1] It is unclear whether these differences are fully explained by the spatial sorting of households or whether they partially reflect an advantage of cities in educating their population. To answer this question, I make use of the particular institutional setting in one country, the Netherlands. The Dutch educational system requires students to make conscious decisions about their desired level of human capital investment both at the end of primary school and at the end of secondary school. Just before making the decisions, students participate in national tests of academic ability. These test scores are highly predictive of future academic outcomes and thus provide an excellent measure of the students' cognitive and academic ability precisely at the moment when they make educational investment decisions. Furthermore, I observe a wide range of household characteristics, including detailed measures of parental education and household income.

The analyses reveal substantial differences in the human capital choices between children growing up in urban areas and rural areas. Conditional on family background and observed cognitive ability, a one log-point increase in population density is associated with a 1.68 percentage point increase in the likelihood that a child enrolls in an academic secondary school, from a base of 23%. Similarly, among high school students who have obtained all prerequisites to enroll for a university degree, a one log-point increase in population density is associated with a 0.8 percentage point increase in the likelihood that a child attends university, from a base of 84%. Children who grow up in more urban environments are thus significantly more likely to select into the schooling tracks that provide higher levels of human capital accumulation. Taken together, a one log-point increase in population density is associated with a 1.4 percentage point increase in the probability that a child attends university, which implies an elasticity of university attendance with respect to population density of 0.07.[2] Expressed differently, growing up in the center of Amsterdam rather than a place at the 25th percentile of the density distribution would increase the probability of attending university by 3 percentage points. Similar to Combes et al. (2008) in the case of the urban wage premium, I find an elasticity about twice as large when individual and family characteristics are not controlled for.

To assess whether differences in unobserved characteristics between urban and rural students might be responsible for differences in human capital investment decisions, I employ the methodology of Altonji et al. (2005) and Oster (2019), which uses the selection on observed variables as guide to the selection on unobserved variables. Under the assumption of Altonji et al. (2005) and Oster (2019) that selection on unobserved characteristics is weakly less than selection on observed characteristics, omitted variable bias cannot account for the results obtained in this paper. In addition, the coefficients are similar across

1 See Appendix A for the urban–rural education gap based on a wide variety of countries.

2 A total of 1.23 pp of the increase is due to the higher probability that a child enrolls in an upper secondary school, taking into account high school dropout rates and that not all students with an academic secondary school degree pursue university studies. A further 0.16 pp is due to the 0.8% increase in university attendance among the children who complete upper secondary school, which contains about 20% of the population. Given that a one-log point increase in density raises university enrollment by 1.4 pp and the mean university enrollment rate of 0.2, an elasticity of 0.07 is obtained.

subgroups, not driven by functional form assumptions and robust to the inclusion of regional fixed effects. Using the historical densities of 1840 as instrument for modern densities to account for the potential endogeneity of urban settlements also leaves the results unaffected.

The findings of this paper contribute to three strands of literature. First, the paper contributes to the urban economics literature by showing that density affects human capital decisions of children. Earlier literature in this field has typically focused on the effect of density at a single decision moment (Frenette, 2006; Newbold and Brown, 2015) or on a single outcome variable (Gibbons and Silva, 2008), whereas this paper estimates the effect of population density on educational outcomes throughout the entire schooling career. In addition, the institutional setting and rich administrative data allow me to condition on a wide range of family characteristics and detailed measures of cognitive ability, which has typically remained unobserved in earlier studies. The closest related work is by Gibbons and Silva (2008), who show that density affects high school test scores in the UK. This paper in comparison takes test scores as given, and focuses on students' educational choices conditional on the primary and secondary school test scores. The resulting elasticity is around three times larger than the elasticity reported by Gibbons and Silva (2008), reflecting that a substantial part of the effect of density on educational attainment operates through differences in educational choices. The finding that children in cities accumulate more human capital also well complements the work by Glaeser and Mare (2001) and De La Roca and Puga (2017), who find that workers in cities accumulate human capital at a faster rate, and the recent work by Eckert et al. (2020), who use a natural experiment to show that refugees who are placed in cities are more likely to start working in skill-intensive occupations.

Second, the results have implications for the long-term economic growth of regions and the persistence of regional differences. It is well known that the stocks of human capital are typically lower in rural regions. The findings of this study suggest that rural regions are in addition at a disadvantage in the expansion of their human capital stock, since children in rural areas invest less in human capital accumulation throughout their childhood, even when they have the same cognitive ability and family background as children in urban areas. While the lower human capital investment in rural regions might be efficient from the perspective of the individual, it does have significant implications for the long-term inequality between urban and rural regions. This is particularly important due to the fact that human capital is the most important predictor of economic growth in regions (Gennaioli et al., 2013),[3] and hence provides a mechanism that could explain the slow observed convergence in incomes between regions within countries as noted by Gennaioli et al (2014) and others.

Finally, the paper contributes to a rapidly growing literature on the spatial (in)-equality of opportunity, following Chetty et al. (2014a) and Chetty and Hendren (2018) in the USA, Alesina et al. (2019) in Africa and Deutscher (2020) in Australia. In this paper, I focus on one specific aspect of places, namely population density, and examine the effects on educational attainment. The results indicate that children who grow up in rural communities do not appear to enjoy or take advantage of the same educational opportunities as children in urban communities, even when they appear to have the same cognitive

3 Although Gennaioli et al. (2013) may overestimate the returns to entrepreneurial education relative to worker education due to selection into entrepreneurial activities, as pointed out by Behrens et al. (2014).

ability and family background as their urban peers. Hence, the findings of this paper highlight that the urban–rural differences in educational outcomes may not just be a result of spatial sorting of households, but may also reflect a lack of opportunities or awareness of opportunities for children residing in rural regions.

Finally, a limitation of this study is that it cannot attribute the overall effect of density on educational attainment to individual mechanisms. Based on the existing literature, there are three main channels through which density is likely to affect educational decisions. First, the returns to education are higher in cities as agglomeration forces mainly complement the productivity of high-skilled workers (Baum-Snow et al., 2018; Autor, 2019), which raises the incentive to invest in human capital in urban areas. Second, the costs of obtaining education are lower in cities, both due to the smaller distance to educational institutes (Frenette, 2006), as well as the possibility that the more diverse school choice in cities improves the match between students their needs and interests and the educational institutes. Third, it might be costlier for youth to acquire information about future educational possibilities in rural areas, due to the absence of university outreach programs as well as the absence of strong network linkages to higher educational institutes in rural communities (Hoxby and Avery, 2012). However, the high correlations between the measures of the individual mechanisms make it difficult to attribute the overall effect of density to the individual mechanisms within the scope of this paper, and a detailed investigation along the lines of De La Roca and Puga (2017) and Dauth et al. (2018) in the case of the urban wage premium is thus left for future research.

The remainder of the paper proceeds as follows. Section 2 reviews the existing theoretical and empirical literature linking density to educational outcomes. Section 3 provides an overview of the context and data. Section 4 discusses the methodology and identification strategy. Section 5 presents the results and the various robustness analyses. Section 6 discusses the implications and concludes.

## 2. Related literature

There are good reasons why children with similar capabilities might make different educational choices depending on whether they live in urban or rural communities. This section provides an overview of the theoretical and empirical support for three potential mechanisms, as well as review the empirical literature which links educational outcomes to urbanization.

### 2.1. Higher returns to education

A key determinant of educational choice in any model of educational investment is the expected returns to schooling. Most models of educational investment, such as the seminal Ben-Porath model, assume that students invest in education until the point where the discounted increase in future wages is equal to the opportunity costs and direct costs of obtaining education, thus providing a tight link between perceived returns to education and human capital investments. Theoretically one would expect highly skilled and educated workers to benefit most from agglomeration economies (see Berry and Glaeser, 2005; Behrens et al., 2014; Davis and Dingel, 2019), which matches the empirical findings that the returns to education are higher in urban environments (Gould, 2007; Combes and Gobillon, 2015; Baum-Snow et al., 2018). This difference has become more pronounced over the last twenty years, during which the urban wage premium for non-college educated workers has essentially disappeared in the US (Autor, 2019). Hence, in the absence of perfect mobility, one would expect the higher returns to education in cities to lead to higher educational investment by children growing up in urban places. However, a key question is whether children and young adults correctly infer the returns to education from their environment and if they respond accordingly. The way in which people form beliefs with regards to the returns to education is still largely unknown, but the experiment by Jensen (2010) suggests that these beliefs are not necessarily deeply held, and that beliefs are updated when new information is presented. Recent work by Adukia et al. (2020) provides evidence for this in the context of India. Rural villages which were connected to cities by roads experienced significant increases in educational attainment, with the effect increasing with the returns to education in newly connected cities.

### 2.2. Costs of schooling

A second channel through which population density can affect educational outcomes is through a reduction in the perceived costs of attending education. At the level of primary and secondary school, the higher density of schools in urban areas means that children on average will spend less time and financial resources on commuting to a school of a given level and quality. In the case of tertiary education, commuting to college or university may become simply unfeasible for children in rural areas, hence necessitating a costly move to attend further education. The empirical literature finds some supporting evidence for the importance of commuting costs on educational decisions, particularly in the case of tertiary education. Card (1993) uses distance to college as an instrument for educational attainment and finds a significant effect in the first stage. Frenette (2006) finds in Canada that children whose families live more than 80 km from the nearest university have a 40% lower probability to enroll at university compared with individuals who grow up within 40 km of a university. However, both studies include only a very limited set of control variables, and as such, it is not clear to what degree the results are driven by the spatial selection of households.

Furthermore, the higher density of educational institutions in cities may also allow for better matching between schools and students, as long as students and schools are heterogeneous in some dimensions. Burgess et al. (2019) show that the majority of students do not list the nearest secondary school as their first preference in the UK, indicating that students do not perceive schools as a homogenous good. Such school-child match specific component can be based on academic preferences of the child, such as the level of instruction (see for instance Bau, 2019) and the focus on certain specialization tracks, or can be based on more personal preferences of the child, such as the religious orientation of a school (Cohen-Zada and Sander, 2008). Models of school choice typically assume that an outside option is available with zero utility, which depending on the context might consist of staying home, working or choosing a different level of schooling. The smaller the number of schools available, the more likely that unfavorable draws of the school-student match quality will result in the student choosing the outside option, which may reduce human capital investment in rural areas. Gibbons and Silva (2008) argue that greater school choice and competition in cites may explain the better performance of urban schools in the UK.

### 2.3. Information and network effects

Third, density may affect human capital decisions through the availability of information on future schooling prospects or due to network effects. Hoxby and Avery (2012) find in the USA that high-performing students who apply to non-selective colleges are disproportionally located in rural areas, as the lower density of high-performing students makes it less profitable for colleges to engage in outreach campaigns. In addition, Hoxby and Avery (2012) suggest that the lower Social-Economic Status (SES) composition of households in rural areas limits the exposure of children to alumni of various educational institutes, as well as reduce the expertise of study counselors on how to advise high-performing students. Hence, the limited information about future educational possibilities in rural areas provides a third potential channel through which density may affect educational investment.

### 2.4. Empirical evidence

Despite theoretical reasons to expect children and young adults in urban areas to invest more in their human capital, the empirical evidence on this subject is limited. Knight and Shi (1996) show large differences in years of educational attainment in China between urban and rural regions, although it remains unclear to what degree this is driven by household heterogeneity. Katz and Goldin (2008, 222) instead report a negative relationship between town size and high school graduation rates in the early 20th century in the USA, which they suggest might be the result of the higher opportunity costs of education in cities. More recently, Newbold and Brown (2015) show that the likelihood that Canadian youth attend university is increasing in city size.

A key challenge in this literature has been dealing with the spatial sorting by households and unobserved heterogeneity between children. Previous work has established that households sort into cities based on parental education ( De la Roca, 2017 ;Ahlin et al., 2018; Bosquet and Overman, 2019) and cognitive skills (Bacolod et al., 2010; Bütikofer and Peri, 2020). Separating spatial selection from area effects is particularly difficult in the case of education, since panel data on educational outcomes is typically unavailable on the individual level. As such, the methods developed to estimate the urban wage premium, which rely heavily on the inclusion of individual fixed effects, cannot easily be applied in the case of educational outcomes.

As a way to overcome the spatial sorting of households, some studies have relied on quasi-experimental variation to quantify the effect of exposure to neighborhood on educational outcomes. Chetty et al. (2016) use the Moving-to-Opportunity program to study neighborhood effects, which provided a randomly selected group of families the opportunity to move to a better neighborhood with less poverty and higher incomes. They find that children who moved to better neighborhoods before the age of 13 years had significantly higher schooling outcomes compared with those who did not receive the opportunity to move. However, such experiments are typically difficult to scale beyond the individual city. In a more general approach, Chetty and Hendren (2018) use the difference in the age of children when they move between regions to identify neighborhood effects. They find a negative relationship between population density and upward mobility as measured by the income at age 26 years, although it remains unclear to what extent differences in human capital accumulation play a role.

## 3. Context and data

In this paper, I utilize the Dutch context to investigate the relationship between population density and educational outcomes, which is particularly well suited for this question.

Tracking in the Dutch educational system starts at the age of 11–12 years, which means that meaningful decisions on human capital investment are taken from an early age onwards. Furthermore, students participate in high-stakes tests of academic ability shortly before making their key educational decisions. Hence, I observe detailed measures of cognitive ability precisely when students make educational decisions, which remains typically unobserved in most other settings. One concern with the Netherlands is that it is relatively urbanized by international standards. However, earlier studies in the Netherlands on the urban wage premium found results in line with other countries (Groot et al., 2014; Verstraten et al., 2019).[4]

### 3.1. Context

Figure 1 provides an overview of the Dutch school system, including the flows between school types.[5] Compulsory education starts at the age 6 years, when all students enroll in primary school. The primary school education lasts for 6 years, at the end of which students participate in a national test measuring their ability in reading comprehension, mathematics and vocabulary, as well as their reasoning and studying ability.[6] Tracking begins at the end of primary school when students can enroll in three different levels of secondary school: upper, middle and lower secondary school. The three levels differ both in length of study, the difficulty of the material and the access to tertiary education which a degree grants. Students are free to apply to any of the three levels of secondary school after finishing primary school, but secondary schools can choose whether to accept students. During the admission decision, secondary schools primarily rely on the scores on the standardized test at the end of the primary school, as well as the recommendation of the primary school teacher. In the years 2003–2015, the score on the national standardized test was considered the leading admission criteria for secondary schools.[7]

Figure 2 shows the distribution of the end of primary school test score for the cohort born in 1996 (panel a) as well as the probability that students with a given test-score enroll in an upper secondary school (panel b). The test score is highly predictive of whether a student enrolls in upper secondary school. Students who score less than 535 (which is about the median test score) are highly unlikely to be enrolled in an upper secondary school 4 years later, whereas among the group of students who obtain the maximum score of 550,[8] 95% enrolls at an upper secondary school. Panel c in Figure 2 shows how these

4 Verstraten et al. (2019) find an elasticity of wages with respect to employment density of 0.02, whereas Groot et al. (2014) find an elasticity of 0.048 without including individual fixed effects. Both are very similar to the estimates cited by Combes and Gobillon (2015) for models with and without person fixed effects in other countries.

5 All data on school enrollment en educational attainment is collected by the Ministry of Education and is made available for research via Statistics Netherlands. The school enrollment data for all levels is taken from the "Onderwijsdeelnemerstab" register, the end of primary school test from the "Citotab" registry and the secondary school final examination grades from the "Examvotab" registry.

6 This test consists of 200 multiple choice questions (100 on reading comprehension and vocabulary, 60 on Mathematics and 40 on studying and reasoning ability). The tests are centrally graded by the Citigroup agency, which also develops the annual test.

7 At the end of 2015, the system was reformed to make the primary school teacher's recommendation binding and secondary schools were no longer allowed to use the end-of-primary school test score. Politicians and teachers felt that the test had become the only selection criteria on which secondary schools evaluated children, which they argued put undue pressure on the children to perform at one specific moment in time.

8 The maximum score of 550 is obtained by 5.7% of the students in the 1996 cohort.

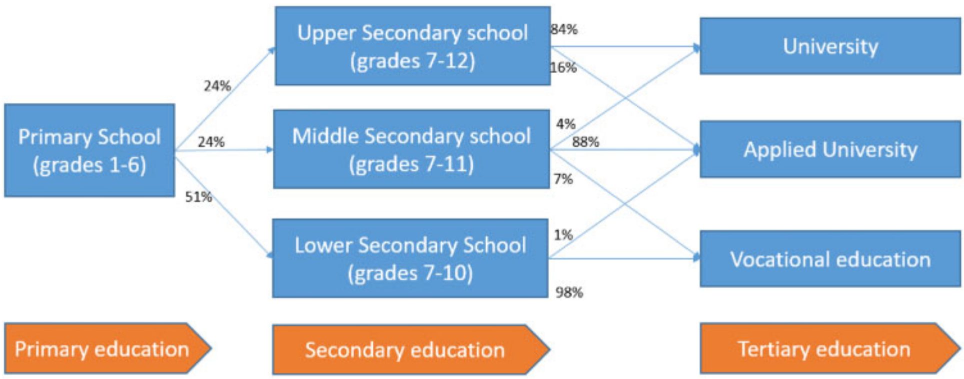


**Figure 1.** Tracking through the Dutch education system for the 1996 cohort. The figure is based on the cohort born in 1996. The secondary school enrollment percentages reflects enrollment 4 years after completing primary school. The tertiary education percentages reflect enrollment within 3 years after finishing the highest secondary school degree. There are some streams within the secondary schooling system (for instance, students dropping from upper secondary school to middle secondary school), which are not displayed here. In addition, some students continue with applied university after obtaining a vocational education degree or with university after obtaining an applied university degree. These streams are not displayed here either, as this usually happens outside of the 3-year window.

probabilities differ between urban and rural regions, providing a first indication of the differences in educational decisions between urban and rural regions.

Once students enroll in a secondary school, they continue for another 4–6 years depending on the selected type of secondary school. At the end of secondary school, all students take a national examination which determines if they are granted a degree. This end of secondary school test is specific for their level of secondary school.[9] Once students obtain a secondary school degree, they continue to tertiary education.[10] The tracking system is designed such that certain levels of secondary school feed into a specific level of tertiary education. As can be seen from Figure 1, 84% of the students who obtain a degree from an upper secondary school are registered at a university 3 years after graduation. Similarly, 88% of the students who obtain a degree from a middle secondary school and 98% of the students completing the lower secondary school are, respectively, registered at an applied university[11] and a vocational education 3 years

9 All students have a basic set of subjects (such as Dutch, English and Math A), as well as a set of subjects related to their chosen subject specialization (health, humanities, natural sciences or social sciences) in which they complete the national examination. The material covered in the examination as well as the difficulty varies depending on the type of secondary school.

10 Upon obtaining a middle or lower secondary school degree, students with excellent grades have the opportunity to take the final two years of the next secondary school level. When discussing degree completion or conditioning on degree, these students are always classified according to the highest degree they obtained.

11 The applied university is similar to the Fachhochschule in the German system. Applied university is more focused on practical skills compared with universities and part of the degree requirement typically involves various long-term internships at companies. For illustrational purposes, individuals with a university/applied university/vocational education on average had an annual income of respectively 50.00/36.000/25.000 euro between 2007 and 2009 (Statistics Netherlands, 2011) .

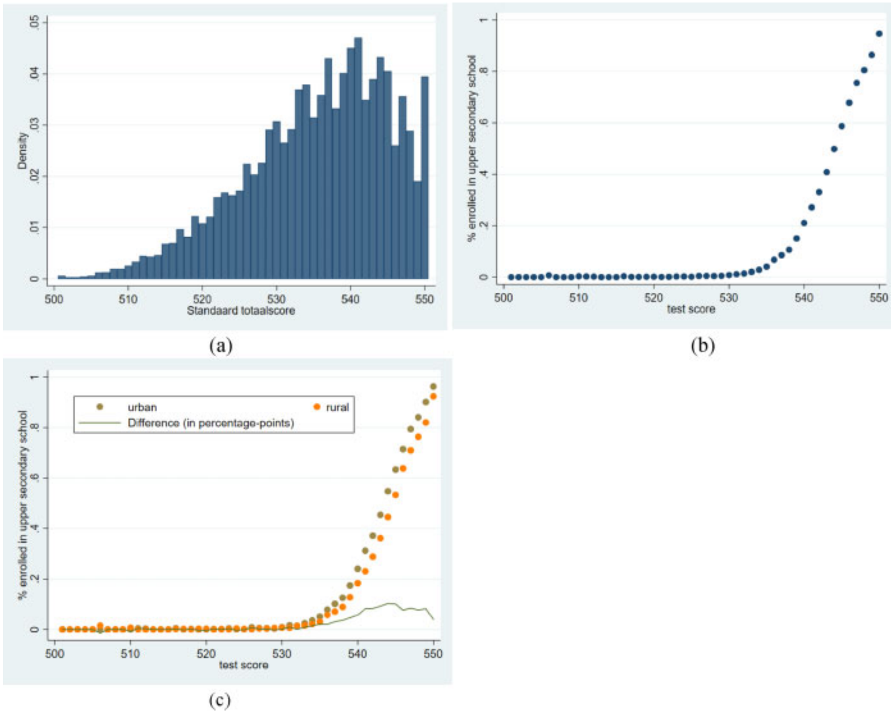


**Figure 2.** Frequency and probability of enrolling in upper secondary school by end-of-primary school test score. (**a**) Displays the frequency for each of the 50 potential scores (501–550) that students could receive at the end-of-primary school test for the cohort born in 1996. (**b**) Displays the percentage of students that are enrolled at an upper secondary school 4 years after taking the test by the end-of-primary school test score. (**c**) Splits the sample between those who grew up in a place with less than the median density (rural) and those who grew up in a place with above-median density (urban). The green line shows the difference in upper secondary school enrollment between urban and rural places in percentage points.

after graduation.[12] Hence, the choice that students make at the end of primary school matters greatly for their future educational prospects.

Tertiary education in the Netherlands differs from many other European countries in the sense that few university degrees have a binding constraint on the number of students they accept.[13] Students who have obtained the relevant qualifications can typically register for their study of choice at the university of their choice, without additional entry conditions. Tuition fees are around 1800 euros per year for universities and applied universities and around 1000 for vocational education, with governmental student loans available if needed. Primary schools and secondary schools are free of tuition. Educational institutes

12 These figures only count students who remained in the Netherlands after completing their secondary school. Around 2% of Dutch students register for a bachelor- or master degree abroad at some point during their education (Department of Education, 2016). As robustness check in the main analysis, I exclude all students from the analysis who are not registered at any educational institution at least 2 out of the 3 years between the ages 19 and 21 years.

13 The most important exceptions are medicine, veterinary medicine and dentistry, as these studies typically are comparatively expensive to offer.

at all levels are financed directly by the national government, with the financing based on the number and type of students enrolled.

### 3.2. Data

The data used in the analysis are primarily for the cohorts born between 1994 and 1998.[14] I restrict the sample to individuals born in the Netherlands for whom both parents can be identified. For these individuals, I observe the place of residence between the ages 1 and 20 years, the enrollment status in all types of secondary and tertiary education and the results on the national tests at the end of primary school[15] and secondary school. In addition, a large number of family characteristics are available, including parental income, education, country of birth and year of birth. The summary statistics for all variables are provided in Table 1.

Parental income is defined similarly as in Chetty and Hendren (2018) and consists of the sum of income for both parents when the child is between ages 14 and 18 years, divided by the number of parent-years with non-missing income. Some parents have a negative income, top-coded income or have missing income for more than 5 parent-years. For the parents with negative income and parents with fewer than 5 parent-year income observations, income is set to zero and a dummy is included for both groups in each regression. A dummy is also included for top-coded incomes. The robustness checks reveal that the results are not sensitive to exclusion of these (small) groups, which together contain about 1.5% of the observations. As the coverage of the income tax data may have improved slightly over the years, I allow the coefficient of parental income as well as the coefficients of the three dummy variables to vary across cohorts.

For each parent, I observe 13 possible levels of education.[16] I do not observe education for all parents, particularly for older parents who completed schooling before some of the national education registers started. Discussions with Statistics Netherlands reveal that parents with missing education are more likely to be low educated, since highly educated parents are more likely to be captured by the various educational registers. In the baseline result, missing education is included as a 14th education type. However, the results are robust to the exclusion of this group. To opt for an as flexible approach as possible when controlling for parental education, I include dummies for each of the 196 possible parental education combinations. Similarly, parental country of birth is grouped into ‘The Netherlands’, ‘Europe’ and ‘Non-European’ for each parent, and nine dummies are included to account for each possible parental combination. Finally, the age of the oldest parent at the time of the birth is added as an additional control variable.

14 Some analyses can be carried out using a larger sample. For consistency, results in the main text are based on the cohorts 1994–1998 unless specified otherwise. The main tables also report results for a larger sample as robustness test whenever possible.

15 A small number of students is exempted from making the test due to disabilities and not all schools have agreed to make the test scores available to Statistics Netherlands. Nonetheless, I observe the test score for the large majority of the population (70%). The coefficients on the urbanization measure of columns (1) and (2) of Tables 2 and 3 are virtually unchanged (within 10% in all cases) when estimating it either on the full sample of individuals born between 1994 and 1998 or the 70% sample for whom I observe the test score. Hence, different selection between rural and urban regions into which schools agreed to make the test scores available to Statistics Netherlands is not driving the results.

16 These groups are kindergarten, primary school, some secondary education (low, middle or high), secondary education degree (low, middle or high), some university or applied university, applied university bachelor, university bachelor, university master or applied university master, and doctoral degree. Vocational education is not listed separately, but included in the three levels of secondary education.

**Table 1.** Summary statistics for individuals born in the Netherlands between the years 1994 and 1998

| | *N* | Mean | Sd. Dev. | p1 | p99 |
|---|---|---|---|---|---|
| Child characteristics | | | | | |
| Urbanization measured at age 11 years | 631,815 | 12.07 | 0.87 | 9.99 | 13.64 |
| Year of birth | 631,815 | 1995.99 | 1.42 | 1994 | 1998 |
| Parental characteristics | | | | | |
| Log parental income | 631,890 | 10.36 | 0.83 | 8.82 | 11.68 |
| Insufficient income data (dummy) | 631,890 | 0.001 | 0.03 | 0.00 | 0.00 |
| Negative household income (dummy) | 631,890 | 0.003 | 0.05 | 0.00 | 0.00 |
| Top coded incomes (dummy) | 631,890 | 0.011 | 0.10 | 0.00 | 1.00 |
| Age of oldest parent at the time of birth | 631,890 | 33.46 | 4.97 | 23.00 | 48.00 |
| Country of birth mother (categorical) | 631,890 | 1.22 | 0.58 | 1.00 | 3.00 |
| Country of birth father (categorical) | 631,890 | 1.23 | 0.59 | 1.00 | 3.00 |
| Education level mother (categorical) | 361,625 | 9.49 | 3.45 | 2.00 | 14.00 |
| Education level father (categorical) | 340,446 | 9.10 | 3.55 | 2.00 | 14.00 |
| Secondary school enrollment and graduation variables | | | | | |
| End of primary school test score | 631,890 | 535.43 | 9.66 | 510.00 | 550.00 |
| Upper secondary school enrollment | 631,890 | 0.23 | 0.42 | 0.00 | 1.00 |
| Middle secondary school enrollment | 631,890 | 0.24 | 0.43 | 0.00 | 1.00 |
| Lower secondary school enrollment | 631,890 | 0.50 | 0.50 | 0.00 | 1.00 |
| Upper secondary school graduation | 631,890 | 0.19 | 0.39 | 0.00 | 1.00 |
| Middle secondary school graduation | 631,890 | 0.25 | 0.43 | 0.00 | 1.00 |
| Lower secondary school graduation | 631,890 | 0.45 | 0.50 | 0.00 | 1.00 |
| GPA upper secondary school degree | 121,106 | 6.63 | 0.69 | 5.50 | 8.50 |
| Specialization track upper secondary school degree (categorical variable) | 121,106 | 4.47 | 2.65 | 1.00 | 10.00 |
| Tertiary education enrollment (within 3 years of high school graduation, conditional on graduating) | | | | | |
| University enrollment | 400,305 | 0.20 | 0.40 | 0.00 | 1.00 |
| Applied University enrollment | 400,305 | 0.29 | 0.46 | 0.00 | 1.00 |
| Vocational education enrollment | 400,305 | 0.50 | 0.50 | 0.00 | 1.00 |

*Note*: The table shows the summary statistics for the individuals born between 1994 and 1998 in the Netherlands. Due to the privacy-sensitive nature of the microdata, it is not possible to report minima or maxima, hence the 1st and 99th percentiles values are displayed instead. As explained in the main text, education is not available for all parents. Secondary school enrollment is measured 4 years after completing primary school. The GPA and specialization track of the upper secondary school degree are only observed for students who graduated from upper secondary school. Tertiary education enrollment is based on cohorts born in 1994–1996, as for the cohorts born in 1997/1998 not enough time has passed for all students to measure. Tertiary education enrollment is defined as the highest enrollment within 3 years of high-school graduation. The graduation rate of middle secondary school is higher than enrollment due to the fact that about 10% of the children enrolled in an upper secondary school drop down to a middle secondary school during their final three years in secondary school.

The next step is to construct an index of urbanization. For each individual, the zip code in which he or she resided between the years 1995 and 2018 is observed. The size of the zip codes is relatively small at 8 $km^2$. I define urbanization as the log of the number of people living within 10 km of the centroid of the zip code in which an individual resides (see Appendix B for the details of the procedure).[17] The threshold of 10 km is selected

17 To avoid potential reverse causality problems, I calculate the density based on the spatial distribution of the population in 1995. As the robustness analysis reveals, the results remain virtually unchanged when I use the spatial distribution of 1840 as instrumental variable.

based on the fact that the majority of Dutch children travel to school by bicycle, where 10 km seems a reasonable upper bound for their reach. Furthermore, the choice of a 10-km radius is in line with recent studies of agglomeration economies on wages (De La Roca and Puga, 2017; Verstraten et al., 2019 ). Nonetheless, the number of individuals that live within, respectively, 5, 10 and 20 km is highly correlated, and as such, the exact distance cut-off has little influence on the results.[18] Densities are the highest in the urbanized Western part of the Netherlands and are relatively low in the North and East, as shown in Figure 3.

## 4. Empirical approach

The Dutch educational system provides two key decision moments that can be exploited to analyze the impact of urbanization on educational outcomes.[19] The first decision moment is at the end of primary school, when children select one of the three levels of secondary school. At this point, I observe the cognitive ability of the child as measured by the score on the national end of primary school test, as well as a wide range of family characteristics. This allows me to study whether, conditional on observed academic ability and parental background, children who grow up in urban areas make different educational choices compared with children who grow up in rural environments. The estimating equation is displayed in Equation (1), where the family characteristics contain the variables discussed in Section 3.2 and dummies are included for all possible primary school test scores. The educational outcome measure is whether a child attends an upper secondary school four years after graduation from primary school. The log-linear relationship between density and educational outcomes is well supported by the data, as will be shown in Section 5.2,

$$
\begin{aligned}
P(\text{attends upper secondary school})_i = {} & \beta_1 * \text{Child characteristics}_i \\
& + \beta_2 * \text{Family characterstics}_i + \beta_3 * f(\text{test score}_i) \\
& + \beta_4 * \text{Urbanization}_i + \gamma_t + \varepsilon_{1i}.
\end{aligned} \tag{1}
$$

The second decision moment is at the end of secondary school. Here, I focus on the group of students who obtained a degree from an upper secondary school which provides access to the large majority of university degrees without further conditions. Equation (2) estimates the effect of density on the probability of attending an academic university within 3 years after graduating from high school,[20] conditional having obtained a degree from

18 For instance, the correlation between density within 10 km and the density within 5 km and the density within 20 km is 0.89 in both cases. Alternatively one could also define density based on the number of schools within a given distance as in Gibbons and Silva (2008) or only use the population density of the zip code itself. The results are similar and the conclusions of this paper do not depend on the selected measure of density.

19 In the empirical framework, I treat the decision for secondary school and tertiary school as two separate decisions. It would also be possible to estimate the two decisions in a joint framework, in which the decision for secondary school and tertiary education are made jointly. While a joint framework might be a better representation for the decision process of certain students, it would significantly complicate the exposition and the interpretation of the coefficients. Furthermore, the most important effect of a joint framework would be a change in the timing when a student does or not does decide to attend university, which should leave the joint effect of density on university attendance across the two decision moments unaffected.

20 The reason for allowing a 3-year lag is that some students opt for a gap year between completing secondary school and starting tertiary education.

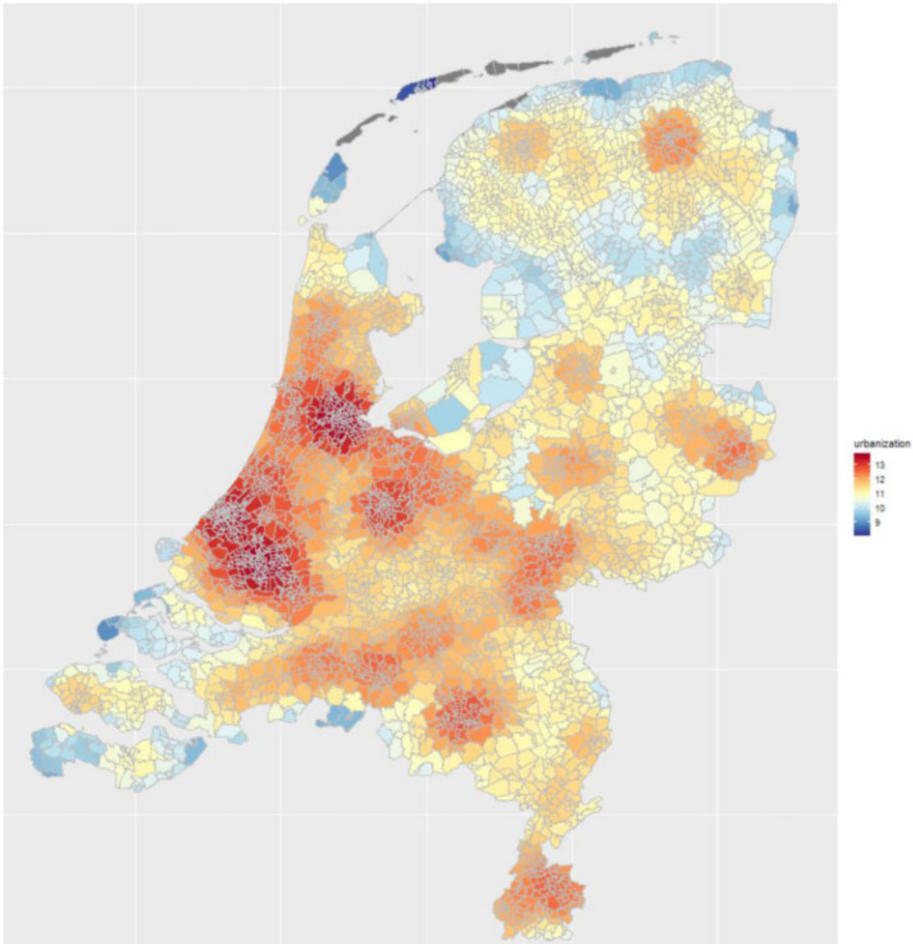


**Figure 3.** Urbanization measure per zip code. The map displays the log of number of individuals within 10 km, based on the 1995 population distribution. For graphical purposes, areas with an urbanization grade less than 9 are gray (containing 230 out of the 631.890 observations in the baseline sample). Appendix B provides details on the construction of the urbanization measure.

an upper secondary school, high school Grade Point Average (GPA) and family characteristics.

$$P(\text{attends university}|\text{upper secondary school degree})_i \\ = \alpha_1 * \text{Child characteristics}_i + \alpha_2 * \text{Family characteristics}_i + \alpha_3 * f(\text{GPA high school}_i) \\ + \alpha_4 * \text{Urbanization}_i + \gamma_t + \varepsilon_{2i}. \tag{2}$$

Under the assumption that the covariance between $\varepsilon_{it}$ and the urbanization measure is zero, Equations (1) and (2) will correctly identify the effect of growing up in an urban environment on educational outcomes. One concern is that parents move to places best fitted to realize the potential outcomes of their children, in which case the estimates of $\alpha_4$ and

$\beta_4$ would be biased. However, I find no evidence of this in the Dutch setting. For the cohorts born between 1994 and 1998, only 5.4% of the families with children between the ages 6 and 17 years make a substantial move (more than 20 km), thus limiting the degree to which families respond to the realized potential of their children. In addition, within the group of children who move, there is no correlation between the change in density and the observed academic ability of the children.[21] Nonetheless, I show that the results of Equations (1) and (2) are robust to estimating it on either the full sample or on the group of children who do not move across municipalities between the ages 6 and 17 years.[22]

A second concern is that parental characteristics may vary between urban and rural places in ways that are not fully captured by the control variables. For instance, it might be that parents in urban areas are more ambitious than parents in rural areas and that this is not fully captured by the differences in wages. Separating spatial sorting from area effects has been a key challenge in the urban literature and solutions to this have typically relied heavily on individual-fixed effects (Combes et al., 2008; De La Roca and Puga, 2017). However, such approaches are typically not possible in the case of educational outcomes due to the lack of consistent national panel data on the educational outcomes of children.[23] Section 5.2 explores this potential threat to identification, using the methodology of Altonji et al. (2005) and Oster (2019) to assess the potential importance of sorting on unobserved characteristics.

Finally, conditioning on the test scores in Equation (1) and (2) might risk overcontrolling for the effect of density on educational attainment, as test scores themselves may also be affected by density. Gibbons and Silva (2008) find that students in urban schools in the UK show more rapid improvements in test scores compared with rural students. Hence, to the degree that urban–rural differences in test scores reflect better learning outcomes in urban areas rather than heterogeneity in cognitive ability between students, the reported coefficients may somewhat underestimate the full effect of population density on educational outcomes.

## 5. Results

### 5.1. Baseline results

Table 2 presents the results of Equation (1), analyzing how the decision of children to enroll in an upper secondary school depends on population density. Column (1) shows the regression when only the urbanization measure is included, whereas column (2) adds child and family characteristics and column (3) adds dummies for the end of primary school test scores. The coefficient of urbanization is statistically significant in all three specifications

21 Result not separately included in the paper but available from the author upon request.

22 A different approach would be to use the Chetty and Hendren (2018) identification strategy and use children who move between regions at different ages to generate urban exposure effects. However, within the group of movers there is a substantial decline in the average end-of-primary school test score by the age of moving, even after the age of 12 years when the test score is essentially predetermined. This differential selection into migration by the age of children hence makes it difficult to apply the Chetty and Hendren (2018) identification strategy within this context.

23 Such approach is difficult to implement for educational outcomes, as most countries lack repeated outcomes for children on a national level, which would allow fixed-effect estimations similar to Combes et al. (2008) with grades instead of wages. Individual-fixed effect approaches have been successfully applied within schools in the economics of education literature to identify effects of school level variables, such as teacher quality (see Chetty et al., 2014b).

**Table 2.** Effect of urbanization on probability of enrolling in upper secondary school

| | Baseline estimates | | | Sensitivity analysis | | | Area fixed effects | | IV |
|---|---|---|---|---|---|---|---|---|---|
| | Ind. level controls (1) | Ind. + Fam. controls (2) | Ind. + Fam controls + test-score (3) | Movers excluded (4) | Different cohorts (1999–2002) (5) | Full parental education (6) | Municipality-fixed effects (7) | Provincial-fixed effects (8) | IV estimates (9) |
| Urbanization at age 11 years | 0.0342*** | 0.0260*** | 0.0167*** | 0.0180*** | 0.0167*** | 0.0141*** | 0.0123** | 0.0196*** | 0.0178*** |
| | (0.0067) | (0.0037) | (0.0017) | (0.0019) | (0.0020) | (0.0015) | (0.0045) | (0.0020) | (0.0027) |
| Individual controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Family controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Test score dummies | No | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.005 | 0.13 | 0.50 | 0.50 | 0.52 | 0.53 | 0.51 | 0.50 | 0.50 |
| Number of observation | 631.731 | 631.731 | 631.731 | 554.826 | 503.200 | 242.156 | 631.731 | 631.731 | 628.396 |

*Note*: All results apart from column (5) are based on individuals born between 1994 and 1998. Dependent variable is whether a child is enrolled in an upper secondary school 4 years after completing primary school. Individual controls include a gender dummy and cohort dummies. Family controls include dummies for parental education combinations (168 dummies), parental nationality combinations (9 dummies), year-of-birth of oldest parent (40 dummies), family income (interacted with cohort-fixed effects), low-income dummy (interacted with cohort-fixed effects), top-coded income dummy (interacted with cohort-fixed effects) and dummy for insufficient income data (interacted with cohort-fixed effects). For a detailed explanation of the construction of the urbanization measure or the family controls, see Section 3. Column 4 excludes all children who moved municipalities during school-going age. Column 5 instead estimates the model on the cohorts born between the years 1999 and 2002. Column 6 removes all parents for whom uncertainty exists about the education level of one or both parents. Column 7 adds municipality-fixed effects for the municipality in which children live at age 11 years (430 dummies), whereas column (8) adds province-fixed effect for the province in which children live at age 11 years (12 dummies). Column (9) instruments for modern densities (based on 1995 population distribution, see Appendix B) with historical densities. All standard errors are clustered on the municipality level. Columns (1–8) are estimated by OLS, column (9) by 2SLS. First stage results of column (9) are reported in Appendix C3 (first stage $F$-statistic: 2349.94).

at the 1% level. In economic terms, the baseline result of column (3) indicates that an increase in density by one log-point (which coincidentally is close to one standard deviation, see Table 1) is associated with a 1.68 percentage point increase in the probability that a child attends upper secondary school. Given that the mean percentage of children attending the upper secondary school is about 23%, this increase is substantial. When comparing the individual columns, the coefficient drops somewhat when family characteristics are added, which is largely driven by the differences in parental education between urban and rural regions. The coefficient declines further when the test scores are added, which might reflect a lower initial academic ability of children in rural regions, or alternatively might be the effect of overcontrolling to the degree that growing up in an urban environment may also affect end-of-primary school test scores.[24]

Columns 4–6 test the robustness of the results by excluding various subgroups. Column (4) excludes children who moved between municipalities during their school-going age, as their families may have sorted themselves into places based on the potential outcomes of their children. Column (5) changes the sample and instead estimates the effect on the cohorts born between 1999 and 2002, for whom information is also available. Finally, column (6) reduces the sample to the individuals for whom the exact education level of both parents is known. The coefficient is very stable across the various subgroups, suggesting that these groups are not driving the results.

One concern might be that children in urban environments enroll in classes that are too difficult for their level of ability, resulting in higher drop-out rates in urban areas.[25] Around 21% of the students who are enrolled in an upper secondary school four years after finishing primary school eventually do not obtain an upper secondary school degree, which is far from negligible.[26] To control for the possibility that differences in dropout rates between urban and rural areas are driving the results, Table 3 instead analyzes the probability that a student obtains a degree from an upper secondary school. The results are very similar to Table 2, indicating that a misallocation of students at the end of primary school in urban communities is unlikely to drive the results. Furthermore, Table C1 shows the results when directly analyzing drop-out rates, by estimating the probability that a student obtains a degree from an upper secondary school, conditional on being enrolled in an upper secondary school 4 years after finishing primary school. The results indicate that children in urban areas are actually slightly less likely to drop out from an upper

24 Children in urban environments indeed have higher average test scores on the end-of-primary school tests, even when conditioning on household and child characteristics. A one-log point increase in population density raises the average end-of-primary school test score by 0.2 points and increases the probability that a child has at least a score of 545 by about 1 percentage point (mean 15 pp).

25 Alternatively, one may also worry that children in rural areas obtain a degree from a middle secondary school, then continue with some years of applied university before switching to university. In this case, children from rural areas would simply follow a different track to end up at university. I do find some evidence for this, as children who obtain a middle secondary school degree are slightly more likely to enroll in university in the seven years after graduating from middle secondary school if they grow up in a rural area rather than an urban area. However, the increased usage of this alternative route in rural areas is far too small to compensate for the negative effects described in Tables 2 and 4. Overall, a one-log point increase in density reduces the probability that a student enrolls for university by 0.6% among the group of middle secondary school graduates (25% of the sample, see Table 1). Hence, taking this path to university into account reduces the overall effect of density on university enrollment by $0.25 * 0.6 = 0.15\%$. Hence, this alternative route can compensate for about 10% of the total effect of density on university enrollment of 1.4%.

26 The majority (62%) of these students instead obtains a degree from a middle secondary school. The drop-out figure of 21% is in line with the statistics reported for this period by Statistics Netherlands (2019).

**Table 3.** Effect of urbanization on probability of obtaining a degree in upper secondary school

| | Baseline estimates | | | Sensitivity analysis | | Area-fixed effects | | IV |
|---|---|---|---|---|---|---|---|---|
| | Ind. level controls | Ind. + Fam. controls | Ind. + Fam controls + test score | Movers excluded | Full parental education | Municipa-lity-fixed effects | Provincial-fixed effects | IV estimates |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Urbanization at age 11 years | 0.0296*** | 0.0226*** | 0.0146*** | 0.0156*** | 0.0130*** | 0.0113** | 0.0153*** | 0.0158*** |
| | (0.0060) | (0.0030) | (0.0013) | (0.0014) | (0.0013) | (0.0039) | (0.0016) | (0.0016) |
| Individual controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Family controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Test score dummies | No | No | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.004 | 0.12 | 0.44 | 0.45 | 0.47 | 0.45 | 0.44 | 0.44 |
| Number of observations | 631.731 | 631.731 | 631.731 | 554.826 | 242.156 | 631.731 | 631.731 | 628.396 |

*Note*: All results are based on individuals born between 1994 and 1998. Dependent variable is whether a child graduates from an upper secondary school. Individual controls include a gender-dummy and cohort-dummies. Family controls include dummies for parental education combinations (168 dummies), parental nationality combinations (nine dummies), year-of-birth of oldest parent (40 dummies), family income (interacted with cohort fixed effects), low-income dummy (interacted with cohort-fixed effects), top-coded income dummy (interacted with cohort-fixed effects) and dummy for insufficient income data (interacted with cohort-fixed effects). For a detailed explanation of the construction of the urbanization measure or the family controls, see Section 3. Column (4) removes all parents for whom uncertainty exists about the education level of one or both parents. Column (5) adds municipality-fixed effects for the municipality in which children live at age 11 years (430 dummies), whereas column (6) adds province-fixed effect for the province in which children live at age 11 years (12 dummies). Column (7) instruments for modern densities (based on 1995 population distribution, see Appendix B) with historical densities. All standard errors are clustered on the municipality level. Columns (1–7) are estimated by OLS, column (8) by 2SLS. First stage results of column (8) are reported in Appendix C3 (first stage *F*-statistic: 2349.94).

**Table 4.** Effect of urbanization on probability of enrolling at university/applied university, conditional on having an upper secondary school degree

| | Baseline estimates | | | Sensitivity analysis | | | Area-fixed effects | | IV |
|---|---|---|---|---|---|---|---|---|---|
| | Ind. level controls | Ind. + Fam. controls | Ind. + Fam controls + GPA | Movers excluded | Cohorts of Table 2 (1994–1998) | Full parental education | Municipality-fixed effects | Provincial-fixed effects | IV estimates |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Urbanization at age 11 years | 0.0232*** | 0.0083*** | 0.0080*** | 0.0084*** | 0.0087*** | 0.0072*** | 0.0098* | 0.0135*** | 0.0071** |
| | (0.0022) | (0.0016) | (0.0017) | (0.0018) | (0.0019) | (0.0018) | (0.0047) | (0.0022) | (0.0025) |
| Individual controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Family controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| GPA and specialization track | No | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.003 | 0.03 | 0.06 | 0.07 | 0.06 | 0.06 | 0.07 | 0.07 | 0.06 |
| Number of observation | 289.109 | 289.109 | 289.109 | 253.430 | 139.559 | 110.921 | 289.109 | 289.109 | 287.763 |

*Note*: All results apart from column (5) are based on individuals born between 1989 and 1998. Dependent variable is whether a child attends university within 3 years of graduating from upper secondary school. Individual controls include a gender dummy and cohort dummies. Family controls include dummies for parental education combinations (168 dummies), parental nationality combinations (nine dummies), year-of-birth of oldest parent (40 dummies), family income (interacted with cohort-fixed effects), low-income dummy (interacted with cohort-fixed effects), top-coded income dummy (interacted with cohort-fixed effects) and dummy for insufficient income data (interacted with cohort-fixed effects). For a detailed explanation of the construction of the urbanization measure or the family controls, see Section 3. Column (4) excludes all children who moved municipalities during school-going age. Column (5) limits the sample to the cohorts born between 1994 and 1998, in line with the baseline sample of Table II. Column (6) removes all parents for whom uncertainty exists about the education level of one or both parents. Column (7) adds municipality-fixed effects for the municipality in which children live at age 11 years (430 dummies), whereas column (8) adds province-fixed effect for the province in which children live at age 11 years (12 dummies). Column (9) instruments for modern densities (based on 1995 population distribution, see Appendix B) with historical densities. All standard errors are clustered on the municipality level. Columns (1–8) are estimated by OLS and column (9) by 2SLS. First stage results of column (9) are reported in Appendix C3 (first stage $F$-statistic: 1028.11).

secondary school, again showing that differential dropout rates between urban and rural areas are not driving the results of Table 2.

The second key educational decision moment is after children obtain an upper secondary school degree, when they have to decide on the level of tertiary education. Table 4 shows the estimates for Equation (2), analyzing the effect of urbanization on the probability that a student enrolls at a university within 3 years of obtaining an upper secondary school degree, which provides access to the large majority of university–subject combinations. Column (1) only includes the urbanization measure, whereas column (2) adds student and family characteristics and column (3) adds controls for the specialization track and the Grade Point Average (GPA) on the national examination at the end of upper secondary school.[27] Since I condition on the national end of high school examination scores (instead of end of primary school test scores), it is possible to use a larger sample and to also include the cohorts born between 1989 and 1993.[28]

The results in Table 4 indicate that children who grow up in urban areas are more likely to enroll at a university, conditional on their academic ability and family characteristics. The preferred specification in column (3) indicates that children who grow up in an area where log density is one-point higher (about one standard deviation) are 0.8 percentage points more likely to enroll in university, from a base of 84%. The coefficient remains very similar when excluding children who moved between municipalities during their childhood (column 4), limiting the sample to the cohorts born between 1994 and 1998 (column 5) or limiting the sample to the children for whom no uncertainty exists over parental education (column 6).

The next question that arises is which alternatives these children select when they decide not to attend university. One alternative is to enroll at the lower rated applied universities. As Figure 1 shows, about 16% of the students who obtained an upper secondary school degree decide to enroll at an applied university, despite having the necessary qualifications to start at a university. Table C2 shows the effect of density on the probability that a student enrolls at an applied university, conditional having obtained a degree from an upper secondary school. The coefficient is significant and negative, indicating that students growing up in rural communities are more likely to instead enroll at an applied university. The coefficients are nearly identical to the coefficients in Table 4, suggesting that diversion of rural students into the lower rated applied universities fully explains the lower enrollment rates at universities found in Table 4.

### 5.2. Robustness

The results in Section 5.1 indicate that population density may play a role in the educational decisions of children. Naturally, some concerns arise over the interpretation and the robustness of this result. This section will discuss four potential threats to identification: the log-linear functional form assumption, the influence of general regional differences, the endogeneity of population density and spatial sorting on unobserved characteristics.

27 The GPA is based on the national examinations at the end of high school. In the final 2 years of secondary school, students can enroll in one of the possible four tracks (humanities, social sciences, biology or natural sciences). The track determines the subjects in which the students take their final examinations.

28 For the students born in 1994–1998, it is possible to include both the secondary school test score as well as end of primary school test score. However, adding the end of primary test score results offers little explanatory power over the high school GPA and leaves the coefficient virtually unchanged. Hence, the loss of this variable due is more than outweigh by the higher precision obtained due to the larger sample.

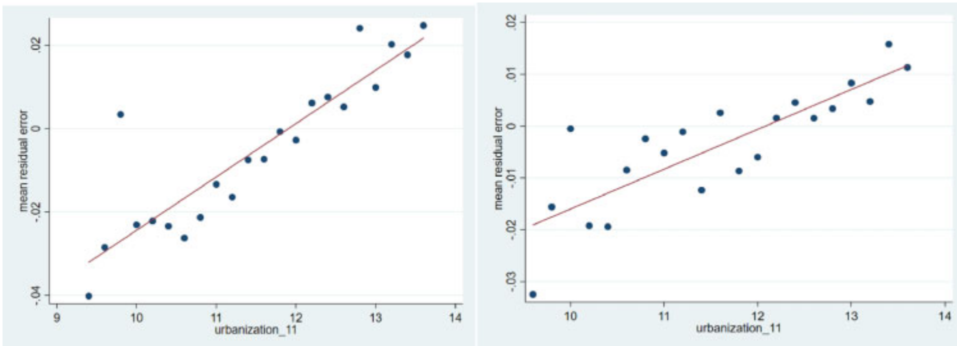


**Figure 4.** Residuals of baseline estimates plotted against urbanization measure. Average residuals of columns (3) of Tables 2 (left) and 4 (right) without the urbanization measure plotted against the urbanization measure. Each dot corresponds to the average residual in a 0.2 log-point bin. Bins containing fewer than 500 observations are not depicted.

### *5.2.1. Functional form assumption*

All estimations have been based on the assumption of a log-linear relationship between population density and educational outcomes, which may or may not accurately represent the true functional form. If the functional form is misspecified, it might bias the coefficients. To test this possibility, Figure 4 plots the residuals of the baseline estimations (columns 3 in Tables 2 and 4) without the urbanization measure against the urbanization measure itself. The residuals are averaged over 0.2 log point intervals. As Figure 4 shows, the log-linear functional form assumption seems to fit the data quite well and hence is unlikely to bias the results. In addition, the residuals reveal that the results do not depend on any specific part of the density distribution: the log-linear relationship seems to describe the actual relationship well throughout the observed density distribution.

### *5.2.2. Influence of broad regional differences*

Second, the map of the urbanization measure displayed in Figure 3 shows that urbanization is highest in the West and relatively low in the North and East of the country. As a result, the urbanization measure may at least partially reflect broader economic or cultural regional differences within the country, rather than urbanization *per se*. To test this possibility, columns (7) and (8) in Tables 2 and 4 add provincial and municipality fixed effects to the model. The coefficient of the urbanization measure in this case is identified on the variation in population density within provinces and municipalities. However, the coefficients change relatively little and remain statistically significant in all cases, even though the standard errors increase substantially as most of the variation in the urbanization measure is discarded. Hence, the results are not driven by broader differences between the various regions.[29]

29 Kelly (2019, 2020) shows that spatial autocorrelation in the dependent and independent variables means that the standard errors might be too small in many spatial regressions. Using the standard error correction proposed by Kelly (2020) increases the standard errors of column 3 in Table 2 by a factor of 1.8, which thus remains significant at the 0.1% level. The correction for Table 4 standard errors is somewhat more complicated due to the low number of observations in many zip codes, which results in very noisy estimates of the zip code fixed effect. If we assume the same spatial autocorrelation in Table 2 as in Table 4 then the estimates in Table 4 remain significant at the 5% level with a *t*-statistic of 2.6.

#### *5.2.3. Endogeneity of Urbanization*

Third, one might be concerned that population density itself is endogenous. Factors that attract population to certain areas and hence contribute to city formation may also directly affect educational outcomes. Furthermore, reverse causality can play a role if individuals migrate to areas with good schools or favorable schooling policies. Combes et al. (2008) highlight the potential of such contemporary factors to bias estimations in the case of the urban wage premium and use historical densities as an instrumental variable (IV). To alleviate concerns that the results in this paper are driven by the endogeneity of the urbanization measure, I follow Combes et al. (2008) and instrument current population densities with the population density based on the 1840 Dutch census.[30,31] The details of this procedure and a map of the population densities of 1840 are contained in Appendix C3. Even though it is not a perfect historical instrument, as 4 of the 13 Dutch universities predate 1840, it should substantially reduce any bias arising from the endogeneity of the urbanization measure. The first stage is significant as shown in Appendix C3, which indicates that the instrument is relevant. The results of the 2SLS estimations are provided in column (9) in Tables 2 and 4. In all three cases, the coefficients hardly change when using the 2SLS estimator. Hence, the endogeneity of urbanization is not driving the results.

#### *5.2.4. Sorting on unobserved variables*

Finally, a key concern for many urban and regional studies is the spatial sorting of households on unobserved characteristics. Despite observing some of the key factors of importance for educational decisions, such as parental education and cognitive ability of the child, some other factors such as ambition and non-cognitive skills remain unobserved. One method to assess the importance of such sorting on unobserved variables is provided by Altonji et al. (2005) and Oster (2019). Their procedure relies on the assumption that the sorting on observed variables is informative of sorting on unobserved variables, and in particular that the sorting on unobserved characteristics is weakly less than the sorting on observed characteristics. As such, the degree to which the coefficient of interest and the $R^2$ change when control variables are added can provide an indication of the potential importance of omitted variable bias.

Altonji et al. (2005) and Oster (2019) argue that sorting on unobserved variables in most cases is less severe than sorting on observable characteristics for two reasons. First of all, they argue that in the most extreme case, researchers observe a random set of variables since they typically have no direct influence over the data collection in surveys and administrative data. In such cases, the observed variables are a random subset, and as such, sorting on observed and unobserved variables should be equally important in generating bias. However, Altonji et al. (2005) argue that researchers typically direct their effort toward obtaining and including the control variables which are seen as being most likely to generate omitted variable bias. Hence, the included control variables in all likelihood contribute more to the omitted variable bias compared with the unobserved variables. For this reason, both Altonji et al. (2005) and Oster (2019) argue that a reasonable upper

30 I would like to express my gratitude to Paul Verstraten (CPB Netherlands Bureau for Economics Policy Analysis) for providing the shape files containing the boundaries of the 1840 municipalities as well as the digitalized 1840 census.

31 A small number of zip codes cannot be matched to the densities of 1840, as they are located on land that has been reclaimed from the sea after 1840. Nonetheless, the instrument is available for 99.6% of the observations.

bound for the degree of selection on unobserved variables is the selection on observed variables. Furthermore, Altonji et al. (2005) argue that when there is a lag between observing the explanatory variables and the outcome measures, any idiosyncratic shocks that occur between observing the explanatory variables and the outcome measure cannot bias the effect of the predetermined explanatory variable. Hence, to the extent that (unobserved) idiosyncratic shocks are present, they provide an additional reason why sorting on unobserved variables may be less important than sorting on observed variables in generating omitted variable bias.

A key decision when applying the methodology of Altonji et al. (2005) and Oster (2019) is the choice for the upper bound on $R^2_{\max}$, that is, how much of the remaining variance the model variables which are unobserved by the researcher would explain. Oster (2019) argues that an upper bound of 1 on the $R^2$ is too restrictive, as measurement error and the true idiosyncratic error term contained in most models would prevent the researcher from reaching an $R^2$ of 1, even in cases where all relevant variables are observed. Instead, based on a simulation exercise, Oster (2019) suggests using an $R^2$ 1.3 times larger than the $R^2$ of the regression with full controls, or to find a reasonable upper bound based on earlier research. For instance, in the case of child outcomes, Oster (2019) argues that sibling correlations may provide a reasonable upper bound for the explanatory power of environmental and family characteristics. Given the decision with respect to the $R^2_{\max}$, it is then possible to rescale the coefficient to account for the sorting on unobserved variables, using the simplified estimator provided in Oster (2019)[32]:

$$\beta^* = \tilde{\beta} - \left[\dot{\beta} - \tilde{\beta}\right] \frac{R_{\max} - \tilde{R}}{\tilde{R} - \dot{R}}, \tag{3}$$

where $\dot{R}$ and $\dot{\beta}$ are obtained from the short regression of density on educational outcomes reported in columns (1) of Tables 2 and 4, and $\tilde{R}$ and $\tilde{\beta}$ are obtained from the full regression model specified in columns (3) of Table 2 and 4. Table 5 shows the adjusted coefficients of Table 2 and 4 when applying the correction for unobserved variables of Oster (2019), under the assumption that selection on observed variables is equal to selection on unobserved variables. The upper half of Table 5 displays the original coefficients and $R^2$s taken from Tables 2 and 4, which form the inputs for Equation (3). The second half of Table 5 provides the adjusted coefficients obtained from Equation (3) under two different sets of $R^2_{\max}$. In the case of the choice for secondary school level (Table 2), the coefficient remains statistically significant and economically relevant when adjusting for the omitted variable bias, both when using the $R^2$ based on sibling correlations in the data as well as Oster's suggested value of $1.3 * \tilde{R}$. The second column of Table 5 adjusts the coefficients based on university enrollment (Table 4). In this case, the sibling correlations are hard to interpret as they are only available for the small subset of siblings where both children graduate from an upper secondary school. When the coefficient is adjusted for omitted variable bias using Oster's suggestion of $1.3 * \tilde{R}$, the coefficient diminishes substantially, but remains positive. Taken together, the results in Table 5 indicate that under the assumption that selection on unobserved variables is weakly less than selection on observed

32 The reported coefficients in Table 5 are based on the general estimator as reported in Oster (2019, Section 3.2), which allows for a more general covariance structure of the unobserved variable with the observed explanatory and outcomes variables. However, the adjusted coefficients are virtually identical when using the simplified and general estimator in this study.

**Table 5.** Importance of selection on unobserved variables

| | Table 2 | Table 4 |
|---|---|---|
| Original coefficients taken from Tables 2/4 | | |
| $R^2$ of short regression (column (1) in Tables 2/4): ($\dot{R}$) | 0.005 | 0.003 |
| Coefficient of short regression (column (1) in Tables 2/4): ($\dot{\beta}$) | 0.0342*** | 0.0232*** |
| $R^2$ of long regression (column (3) in Tables 2/4): ($\tilde{R}$) | 0.502 | 0.064 |
| Coefficient of long regression (column (3) in Table 2/4): ($\tilde{\beta}$) | 0.0167*** | 0.0080*** |
| Coefficient adjusted for unobserved variables following Oster (2019) | Table 2 | Table 4 |
| $R_{\max}$ obtained from sibling regressions | 0.54 | – |
| Adjusted coefficient ($\beta^*$) using $R_{\max}$ based on sibling regressions | 0.0153*** | – |
| $R_{\max}$ as $1.3 * \tilde{R}$ | 0.65 | 0.083 |
| Adjusted coefficient ($\beta^*$) using $R_{\max}$ of $1.3 * \tilde{R}$ | 0.0106*** | 0.0028 |

*Note*: Adjustment for unobserved variables based on Oster (2019). The $R^2$ based on sibling regression comes from a regression of the model in Table 2, column (3) on the subset of siblings, with the secondary educational choice of the sibling added as control variable. To provide a conservative inference, the standard errors of Table 2/4 have been used to calculate significance.

variables, omitted variable bias cannot explain the majority of the effect of population density on educational outcomes.

## 6. Discussion and conclusion

The results in this paper show evidence that population density affects the educational investment decisions of children in the context of the Netherlands. Conditional on observed ability and parental background, children who grow up in urban areas consistently choose to invest more in their education compared with children who grow up in more rural environments. This result is robust across various specifications, subgroups and spatial scope, and cannot be accounted for by the endogeneity of the urbanization measure or sorting on unobserved variables under the assumptions of Oster (2019). Taken together, the results imply that conditional on family characteristics and academic ability, a one log-point increase in density is associated with a 1.4 percentage point increase in the probability that a child attends university. Given the mean university attendance rate of 20%, this implies an elasticity of university attendance with respect to density of 0.07.[33] Expressed differently, moving from a place at the 25th percentile of the density distribution to Amsterdam increases the probability that a child attends university by 3 percentage points.

It is surprising that the potential for agglomeration economies to affect human capital decisions of children and young adults has not received more attention in the literature, particularly given the implications for long-term regional and national growth. Children in rural communities do not seem to enjoy or take the same educational opportunities as children who grow up in urban communities, even in a country such as the Netherlands where the rural areas are relatively accessible from an international perspective. Hence, the findings suggest that differences in educational attainment between urban and rural

33 1.23% of this increase is due to the increase 1.46% in probability that a child graduates from upper secondary school as seen in Table 3, taking into account that on average 84% continues with university. A further 0.15% of the increase is due to the 0.8% increase in university attendance among the children who complete upper secondary school, which contains about 20% of the population (see Table 1).

communities observed in a wide range of countries may reflect more than just the spatial sorting of households.

Finally, one question that remains is which mechanisms drive the increased human capital accumulation of children in urban environments. Sections 2.1 and 2.2 indicated that based on existing literature, one would expect the agglomeration mechanisms to play a role in the early-life human capital decisions, either through an increase in the returns to education in cities or by increasing the availability of schooling and by reducing the commuting or moving costs to attend further education. Furthermore, there might also be a role for network effects as highlighted in Section 2.3. However, a detailed exposition of specific mechanisms in the spirit of De La Roca and Puga (2017) or Dauth et al. (2018) in the case of the urban wage premium is beyond the scope of this paper and will be an interesting avenue for future research.

## Acknowledgments

The author wishes to acknowledge the statistical offices that provided the underlying data making this research possible: National Institute of Statistics and Censuses, Argentina; National Statistical Service, Armenia; Bangladesh Bureau of Statistics, Bangladesh; Ministry of Statistics and Analysis, Belarus; National Institute for Statistics and Economic Analysis, Benin; National Institute of Statistics, Bolivia; Central Statistics Office, Botswana; Institute of Geography and Statistics, Brazil; National Institute of Statistics and Demography, Burkina Faso; National Institute of Statistics, Cambodia; Central Bureau of Census and Population Studies, Cameroon; Statistics Canada, Canada; National Bureau of Statistics, China; National Administrative Department of Statistics, Colombia; National Institute of Statistics and Censuses, Costa Rica; National Statistics Office, Dominican Republic; National Institute of Statistics and Censuses, Ecuador; Central Agency for Public Mobilization and Statistics, Egypt; Bureau of Statistics, Fiji; National Institute of Statistics and Economic Studies, France; Federal Statistical Office, Germany; National Institute of Statistics, Guatemala; Institute of Statistics and Informatics, Haiti; Central Statistical Office, Hungary; BPS Statistics Indonesia, Indonesia; Central Statistics Office, Ireland; Central Bureau of Statistics, Israel; National Institute of Statistics, Italy; Statistical Institute, Jamaica; National Bureau of Statistics, Kenya; National Statistical Committee, Kyrgyz Republic; Statistics Bureau, Laos; Bureau of Statistics, Lesotho; Institute of Statistics and Geo-Information Systems, Liberia; National Statistical Office, Malawi; Department of Statistics, Malaysia; National Directorate of Statistics and Informatics, Mali; National Institute of Statistics, Geography, and Informatics, Mexico; National Institute of Statistics, Mozambique; Central Bureau of Statistics, Nepal; National Institute of Information Development, Nicaragua; Statistics Division, Pakistan; Central Bureau of Statistics, Palestine; Census and Statistics Directorate, Panama; National Statistical Office, Papua New Guinea; General Directorate of Statistics, Surveys, and Censuses, Paraguay; National Institute of Statistics and Informatics, Peru; Central Statistics Office, Poland; National Institute of Statistics, Portugal; National Institute of Statistics, Romania; Federal State Statistics Service, Russia; National Institute of Statistics, Rwanda; Government Statistics Department, Saint Lucia; Statistics Sierra Leone, Sierra Leone; Statistical Office of the Republic of Slovenia, Slovenia; Statistics South Africa, South Africa; National Bureau of Statistics, South Sudan; Central Bureau of Statistics, Sudan; Bureau of Statistics, Tanzania; National Statistical Office, Thailand; National Institute of Statistics, Togo; Bureau of Statistics, Uganda; State Committee of Statistics, Ukraine; Bureau of the Census, United States; National Institute of Statistics, Venezuela; General Statistics Office, Vietnam; Central Statistics Office, Zambia and National Statistics Agency, Zimbabwe.

## References

Adukia, A., Asher, S., Novosad, P. (2020) Educational investment responses to economic opportunity: evidence from Indian road construction. *American Economic Journal: Applied Economics*, 12: 348–376.

Ahlin, L., Andersson, M., Thulin, P. (2018) Human capital sorting: the "when" and "who" of the sorting of educated workers to urban regions. *Journal of Regional Science*, 58: 581–610.

Alesina, A., Hohmann, S., Michalopoulos, S., Papaioannou, E. (2019) Intergenerational Mobility in Africa (No. w25534). National Bureau of Economic Research.

Altonji, J. G., Elder, T. E., Taber, C. R. (2005) Selection on observed and unobserved variables: assessing the effectiveness of Catholic schools. *Journal of Political Economy*, 113: 151–184.

Autor, D. (2019) Work of the past. Work of the future. American Economic Association: Papers and Proceeding.

Bacolod, M., Blum, B. S., Strange, W. C. (2010) Elements of skill: traits, intelligence, education, and agglomeration. *Journal of Regional Science*, 50: 245–280.

Bau, N. (2019) *Estimating an Equilibrium Model of Horizontal Competition in Education*. London: Centre for Economic Policy Research.

Baum-Snow, N., Freedman, M., Pavan, R. (2018) Why has urban inequality increased? *American Economic Journal: Applied Economics*, 10: 1–42.

Behrens, K., Duranton, G., Robert-Nicoud, F. (2014) Productive cities: sorting, selection, and agglomeration. *Journal of Political Economy* 122: 507–53.

Berry, C. R., Glaeser, E. L. (2005) The divergence of human capital levels across cities. *Papers in Regional Science*, 84: 407–444.

Bosquet, C., Overman, H. G. (2019) Why does birthplace matter so much? *Journal of Urban Economics*, 110: 26–34.

Burgess, S., Greaves, E., Vignoles, A. (2019) School choice in England: evidence from national administrative data. *Oxford Review of Education*, 1–21.

Bütikofer A Peri G. How cognitive ability and personality traits affect geographic mobility. *Journal of Labor Economics, forthcoming*.

Card, D. (1993) Using geographic variation in college proximity to estimate the return to schooling (No. w4483). National Bureau of Economic Research.

Chetty, R., Hendren, N., Kline, P., Saez, E. (2014a) Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics*, 129: 1553–1623.

Chetty, R., Friedman, J. N., Rockoff, J. E. (2014b) Measuring the impacts of teachers I: evaluating bias in teacher value-added estimates. *American Economic Review*, 104: 2593–2632.

Chetty, R., Hendren, N. (2018) The impacts of neighborhoods on intergenerational mobility I: childhood exposure effects. *The Quarterly Journal of Economics*, 133: 1107–1162.

Chetty, R., Hendren, N., Katz, L. F. (2016) The effects of exposure to better neighborhoods on children: new evidence from the Moving to Opportunity experiment. *American Economic Review*, 106: 855–902.

Cohen-Zada, D., Sander, W. (2008) Religion, religiosity and private school choice: implications for estimating the effectiveness of private schools. *Journal of Urban Economics*, 64: 85–100.

Combes, P. P., Duranton, G., Gobillon, L. (2008) Spatial wage disparities: sorting matters! *Journal of Urban Economics*, 63: 723–742.

Combes, P. P., Gobillon, L. (2015) The empirics of agglomeration economies. In G. Duranton, J. V. Henderson, W. C. Strange (eds) *Handbook of Regional and Urban Economics*, Vol. 5, pp. 247–348. Amsterdam: Elsevier.

Dauth, W., Findeisen, S., Moretti, E., Suedekum, J. (2018). Matching in cities (No. w25227). National Bureau of Economic Research.

Davis, D. R., Dingel, J. I. (2019). A spatial knowledge economy. *American Economic Review*, 109: 153–70.

Deutscher, N. (2020) Place, peers, and the teenage years: long-run neighborhood effects in Australia. *American Economic Journal: Applied Economics*, 12: 220–249.

Duranton, G., Puga, D. (2004) Micro-foundations of urban agglomeration economies. In G. Duranton, J. V. Henderson, W. C. Strange (eds) *Handbook of Regional and Urban Economics*, Vol. 4, pp. 2063–2117. Amsterdam: Elsevier.

Eckert, F., Hejlesen, M., Walsh, C. (2020) *The Return to Big City Experience: Evidence from Refugees in Denmark*. NBER Working Paper.

Frenette, M. (2006) Too far to go on? Distance to school and university participation. *Education Economics*, 14: 31–58.
Gennaioli, N., La Porta, R., Lopez-de-Silanes, F., Shleifer, A. (2013) Human capital and regional development. *The Quarterly Journal of Economics*, 128: 105–164.
Gennaioli, N., La Porta, R., De Silanes, F. L., Shleifer, A. (2014) Growth in regions. *Journal of Economic Growth*, 19: 259–309.
Gibbons, S., Silva, O. (2008) Urban density and pupil attainment. *Journal of Urban Economics*, 63: 631–650.
Glaeser, E. L., Kolko, J., Saiz, A. (2001) Consumer city. *Journal of Economic Geography*, 1, 27–50.
Glaeser, E. L., Mare, D. C. (2001). Cities and skills. *Journal of Labor Economics*, 19: 316–342.
Gould, E. D. (2007) Cities, workers, and wages: a structural analysis of the urban wage premium. *The Review of Economic Studies*, 74: 477–506.
Groot, S. P., de Groot, H. L., Smit, M. J. (2014) Regional wage differences in the Netherlands: micro evidence on agglomeration externalities. *Journal of Regional Science*, 54: 503–523.
Hoxby, C. M., Avery, C. (2012) The missing "one-offs": the hidden supply of high-achieving, low income students (No. w18586). National Bureau of Economic Research.
Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, 125: 515–548.
Katz, L. F., Goldin, C. D. (2008) *The Race between Education and Technology*. Cambridge: Harvard University Press.
Kelly, M. (2019) *The Standard Errors of Persistence*. London: CEPR
Kelly, M. (2020) *Understanding Persistence*. London: CEPR
Knight, J., Shi, L. (1996) Educational attainment and the rural–urban divide in China. *Oxford Bulletin of Economics and Statistics*, 58: 83–117.
Newbold, K. B., Brown, W. M. (2015) The urban–rural gap in university attendance: determinants of university participation among Canadian youth. *Journal of Regional Science*, 55: 585–608.
Minnesota Population Center. (2019). *Integrated Public Use Microdata Series*, International: Version 7.2 [dataset]. Minneapolis (MN): IPUMS (https://doi.org/10.18128/D020.V7.2).
Oster, E. (2019) Unobservable selection and coefficient stability: theory and evidence. *Journal of Business & Economic Statistics*, 37: 187–204.
De La Roca, J., Puga, D. (2017) Learning by working in big cities. *The Review of Economic Studies*, 84: 106–142.
De la Roca, J. (2017). Selection in initial and return migration: evidence from moves across Spanish cities. *Journal of Urban Economics*, 100, 33–53.
Rosenthal, S. S., Strange, W. C. (2004) Evidence on the nature and sources of agglomeration economies. In G. Duranton, J. V. Henderson, W. C. Strange (eds) *Handbook of Regional and Urban Economics*, Vol. 4, pp. 2119–2171. Amsterdam: Elsevier.
Statistics Netherlands. (2011) Studeren loont. Inkomens van afgestudeerden in het mbo, hbo en wetenschappelijk onderwijs [Online publication].
Statistics Netherlands. (2019). Statline. https://opendata.cbs.nl/statline/#/CBS/en/
Verstraten, P., Verweij, G., Zwaneveld, P. J. (2019) Complexities in the spatial scope of agglomeration economies. *Journal of Regional Science*, 59: 29–55.

# Appendix

## A. The Urban–Rural Educational Gap

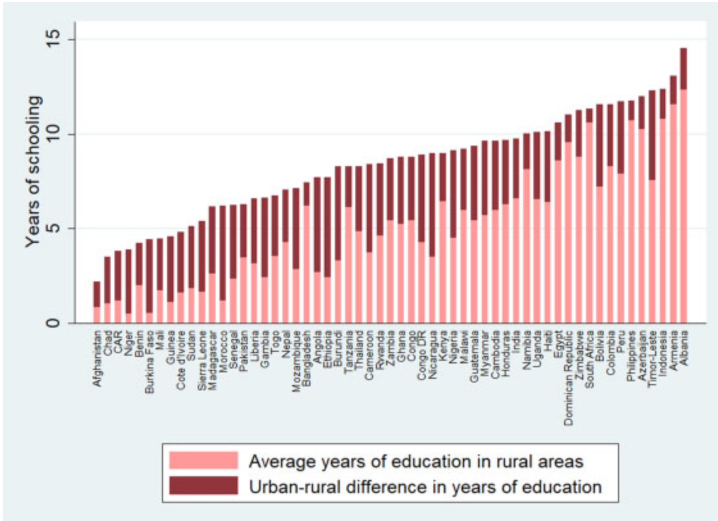


**Figure A1.** Urban–rural education gap based on 25-year-old males in DHS data. The figure is based on 57 countries included in the Demographic and Health Survey (DHS). Reported are the average years of schooling in urban and rural areas for 25-year-old males. The urban–rural definition follows the definition of DHS, which is based on the urban–rural definition of the country in question. Countries with fewer than 100 18-year old males for either the rural or urban region of the country have been excluded from the original sample of 75 countries.

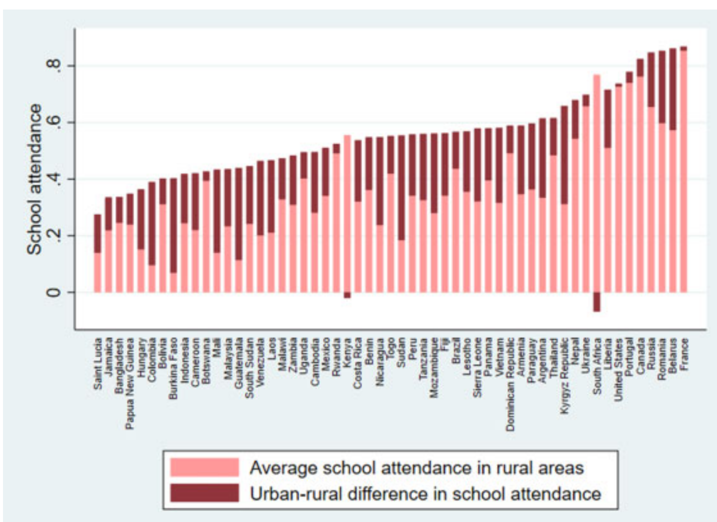


**Figure A2.** Urban–rural school attendance gap based on 18-year-old males in the IPUMS International database (Minnesota Population Center, 2019). The figure is based on the latest census for 55 countries included in the IPUMS International data which include an urban/rural definition in their census (which excludes most developed countries). The *y*-axis shows the percentage of children that indicate that they were attending school at the time of the census in urban and rural areas for 18-year-old males. The urban–rural definition is based on the definition of the statistical office of the country in question.

## B. Construction of Urbanization Measure

The administrative data include information on place of residence from January 1995 onward. Hence, I start by calculating the number of individuals registered per zip code on the 5th January 1995. As the next step, I determine the centroid of each zip code using GIS software. The map of the Netherlands with all the centroids is displayed in Figure A1. For each centroid, I determine which other zip codes lie within a 10-km radius of the centroids and add up the population of these zip codes. For zip codes that lie partially in the 10-km radius, I multiply the share of the zip code area covered by the 10-km radius with the population of that particular zip code.[34,35] The average zip code has 43 other zip-codes within a 10-km radius and even the first percentile of zip codes has 8 other zip-codes within a 10-km radius. The correlation between the number of individuals living within 5, 10 and 20 km is fairly high at 0.89. Finally, I take the log of the number of individuals within the 10-km radius as urbanization measure. Figure 3 in the main text shows the resulting map of the urbanization measure. In addition, Figure B2 shows a histogram of the urbanization measure.


**Figure B1.** Map of Dutch zip codes and centroids.

34 This means that I implicitly assume that the population is spread equally across zip codes. However, as zip codes are fairly small (average of 8 km$^2$) compared to the 10-km radius (314 km$^2$) radius, this assumption has little effect on the relative differences in urbanization between regions.

35 For some very isolated zip codes (for instance, on the islands in the North), there are no other zip codes within 10 km, and hence the population with 10 km is simply the zip codes own population.

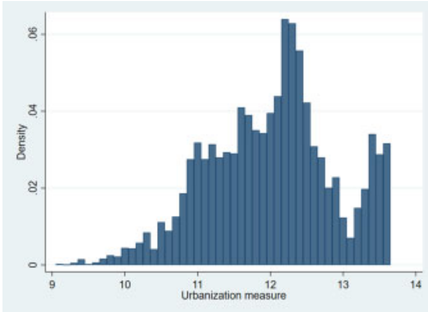


**Figure B2.** Histogram of urbanization measure at the age 11 years. The histogram of urbanization measure for the cohorts born between 1994 and 1998. Areas with an urbanization score below 9 (containing 230 out of the 631.815 observations in the baseline sample) not displayed here.

## C. Additional Robustness Analyses

### *C.1. Probability of upper secondary school graduation, conditional on being registered at an upper secondary school*

**Table C1.** Effect of urbanization on probability of obtaining a degree in upper secondary school, conditional on enrollment in upper secondary school

| | Baseline estimates | | | Sensitivity analysis | | Area-fixed effects | | IV |
|---|---|---|---|---|---|---|---|---|
| | Ind. level controls | Ind. + Fam. controls | Ind. + Fam controls + test score | Movers excluded | Full parental education | Municipality-fixed effects | Provincial-fixed effects | IV estimates |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Urbanization at age 11 years | 0.0075 (0.0041) | 0.0060*** (0.0022) | 0.0068** (0.0021) | 0.0069** (0.0022) | 0.0062* (0.0025) | 0.0071 (0.0059) | −0.0005 (0.0028) | 0.0071* (0.0034) |
| Individual controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Family controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Test score dummies | No | No | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.000 | 0.03 | 0.08 | 0.08 | 0.09 | 0.08 | 0.08 | 0.08 |
| Number of observation | 146.879 | 146.879 | 146.879 | 130.634 | 60.945 | 146.879 | 146.879 | 146.248 |

All results are based on individuals born between 1994 and 1998. Dependent variable is whether a child graduates from an upper secondary school, conditional on being enrolled at the third grade of upper secondary school. Individual controls include a gender dummy and cohort dummies. Family controls include dummies for parental education combinations (168 dummies), parental nationality combinations (nine dummies), year-of-birth of oldest parent (40 dummies), family income (interacted with cohort-fixed effects), low-income dummy (interacted with cohort-fixed effects), top-coded income dummy (interacted with cohort-fixed effects) and dummy for insufficient income data (interacted with cohort-fixed effects). For a detailed explanation of the construction of the urbanization measure or the family controls, see Section 3. Column (4) excludes all children who moved municipalities during school-going age. Column (5) removes all parents for whom uncertainty exists about the education level of one or both parents. Column (6) adds municipality-fixed effects for the municipality in which children live at age 11 years (430 dummies), whereas column (7) adds province-fixed effect for the province in which children live at age 11 years (12 dummies). Column (8) instruments for modern densities (based on 1995 population distribution, see Appendix B) with historical densities. All standard errors are clustered on the municipality level. Columns (1–7) are estimated by OLS and column 8 by 2SLS. First stage results of column (8) are reported in Appendix C3 (first stage $F$-statistic: 596.93).

### C.2. Applied university enrollment

**Table C2.** Effect of urbanization on applied university attendance

| | Baseline estimates | | | Sensitivity analysis | | | Area-fixed effects | | IV |
|---|---|---|---|---|---|---|---|---|---|
| | Ind. level controls | Ind. + Fam. controls | Ind. + Fam controls + GPA | Movers excluded | Cohorts of Table 2 (1994–1998) | Full parental education | Municipality-fixed effects | Provincial-fixed effects | IV estimates |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Urbanization at age 11 years | −0.0233*** | −0.0094*** | −0.0089*** | −0.0095*** | −0.0101*** | −0.0081*** | −0.0090* | −0.0129*** | −0.0089*** |
| | (0.0023) | (0.0016) | (0.0016) | (0.0017) | (0.0017) | (0.0017) | (0.0041) | (0.0018) | (0.0025) |
| Individual controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Family controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| GPA and specialization track | No | No | Yes | Yes | Yes | Yes | Yes | No | Yes |
| $R^2$ | 0.00 | 0.03 | 0.07 | 0.07 | 0.06 | 0.07 | 0.07 | 0.07 | 0.07 |
| Number of observation | 289.109 | 289.109 | 289.109 | 253.430 | 139.559 | 110.921 | 289.109 | 289.109 | 287.763 |

*Note*: All results apart from column (5) are based on individuals born between 1989 and 1998. Dependent variable is whether a child attends applied university as highest education within 3 years of graduating from upper secondary school. Individual controls include a gender dummy and cohort dummies. Family controls include dummies for parental education combinations (168 dummies), parental nationality combinations (nine dummies), year-of-birth of oldest parent (40 dummies), family income (interacted with cohort-fixed effects), low-income dummy (interacted with cohort-fixed effects), top-coded income dummy (interacted with cohort-fixed effects) and dummy for insufficient income data (interacted with cohort-fixed effects). For a detailed explanation of the construction of the urbanization measure or the family controls, see Section 3. Column (4) excludes all children who moved municipalities during school-going age. Column (5) limits the sample to the cohorts born between 1994 and 1998, in line with the baseline sample of Table 2. Column (6) removes all parents for whom uncertainty exists about the education level of one or both parents. Column (7) adds municipality-fixed effects for the municipality in which children live at age 11 years (430 dummies), whereas column (8) adds province-fixed effect for the province in which children live at age 11 years (12 dummies). Column (9) instruments for modern densities (based on 1995 population distribution, see Appendix B) with historical densities. All standard errors are clustered on the municipality level. Columns (1–8) are estimated by OLS and column (9) by 2SLS. First stage results of column (9) are reported in Appendix C3 (first stage $F$-statistic: 1028.11).

### *C.3. IV first stage*

Figure C3.1 displays the population densities of Dutch municipalities in 1840. The number of municipalities (1340) is relatively large compared the modern day number of municipalities (380). The map below is used as input to construct the IV measure, namely the number of individuals in 1840 living within 10 km of all current zip codes, using the procedure outlined in Appendix B. Notice that I make the implicit assumption that the population in 1840 was spread homogenously within municipalities, as the 1840 population statistics are only available at the municipality level. Figure C3.2 shows the resulting IV-density measure for the zip codes.

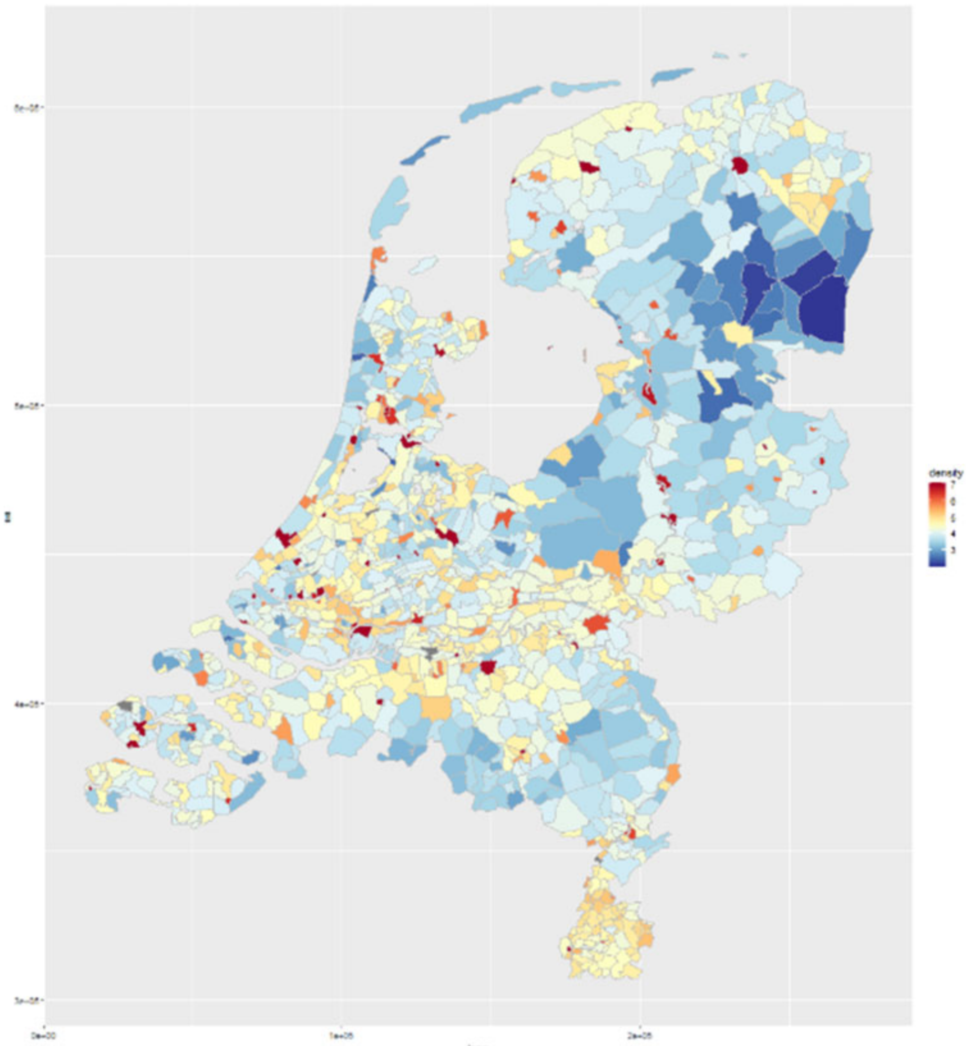

**Figure C3.1.** Population density per municipality in 1840. When calculating the densities based on the 1840 population map, I implicitly assumed that population is spread homogenously within municipalities. Notice that the Netherlands contained a substantially larger amount of inland water compared with the contemporary Netherlands, due to the land reclamation programs in the 19th and 20th centuries.

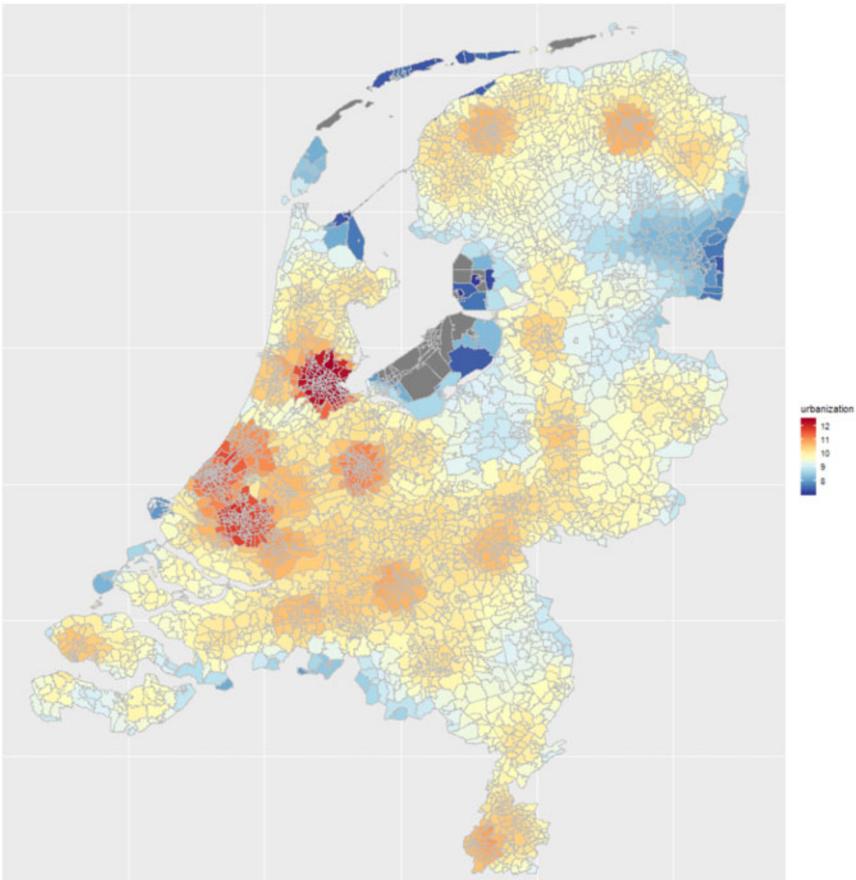


**Figure C3.2.** Urbanization measure based on 1840 population distribution. The map displays the urbanization measure for each zip code (log of number of people living within 10 km), based the 1840 population distribution. See Appendix B for the construction of the measure and Figure C3.1 for the 1840 population distribution. The gray areas have an urbanization measure below 8. The gray mass in the center of the Netherlands consists of land that has been reclaimed since 1840, and hence had very few to no individuals living with 10 km in 1840, as can be also seen in Figure C3.1.

**Table C3.** First stage of regressing urbanization based on the 1840 densities on contemporary urbanization measure (based on 1995 densities)

| | First stage |
|---|---|
| Urbanization based on 1840 densities | 0.635*** |
| | (0.0045) |
| *F*-statistic | 2349.94 |
| $R^2$ | 0.58 |
| Number of observation | 628.396 |

*Note*: First stage of 2SLS regression of baseline urbanization measure (based on the 1995 population distribution) on the historical urbanization measure (based on the 1840 population distribution). The first stage of Tables 4 and C2 are virtually identical and not separately reported here, with a slightly smaller *F*-statistics (1028.11) due to the smaller number of observations.