Lured in and crowded out?
Estimating the impact of immigration on natives' education using early XXth century US immigration *

June 2013

Abstract

Immigration can impact educational decisions of natives through two different margins. First, it can increase the potential returns to education by generating a larger supply of unskilled workers, thus raising the relative wages of more educated individuals. Secondly, it can increase the cost of acquiring education through crowding out natives of public schools. We separate these two channels by separately estimating the causal impact of the immigration of adults and children on the educational decisions of natives at the state-cohort level in a context of rapid changes in educational attainment, the United States from 1910 to 1935. We find that more adults leads to an increase in school attendance, high school enrollment and graduation while immigration of children has the opposite effect. We find these two effects to be particularly relevant for urban males of US born parents. We document that parents attempted to counteract the effect of immigration by enrolling their children in private schools but we find little evidence of change in geographic locations. We finally use our estimates to calibrate that only 10 to 15 percent of the "high school movement" can be attributed to the 1920s limitations on immigration and that other factors were obviously more important determinants.

*The usual disclaimer applies. We wish to thank Claudia Goldin for generously sharing her data with us.

The economics literature has focused mostly on the impact of immigrants, in particular, low-skill immigrants on labor market outcomes of natives. While some natives may be unable to readjust their human capital investment in response to immigration shocks, younger generations may have more flexibility if they are "forward-looking" in their educational decision (see Smith (2012) for an example of this for a recent migration wave). However, those same cohorts may also face another, and of opposite size, impact of immigration through the fact that more immigration, in particular of school-age individuals, may also crowd them out of the public education system. In this paper, we separate the two effects using the age at which immigrants arrive to the United States and measure their causal impact at a period where Americans were increasing their high school attendance in large number.

We first present a simple theoretical framework that specifies more formally the two channels through which immigration can impact educational decisions of natives. It can first do so by increasing the return to education since immigration over this period was mostly low-skilled working males[1] and this leads to a larger supply of lower-skilled individuals thus raising the relative wages of high-skilled individuals. But at the same time, more immigrants may increase the cost of acquiring human capital through having larger classrooms, classrooms with more heterogenous students, etc. Thus, the overall impact of immigration on educational attainment is unclear.

This simple model is then used as a basis to develop our empirical strategy. Our key assumption will be that the crowd-out effect should be particularly strong when the immigrants are fellow teenagers but the return effect could be generated by slightly older immigrants. We thus distinguish between immigration flows by age and age at arrival by state to effectively capture the relevant source of variation. However, since immigrants may clearly elect their state of residence based on labor market or schooling market factors, the simple correlation between the two outcomes is unlikely to offer us the causal estimate of immigration. To achieve this, we allocate newly arrived migrants using past location shares of immigrants, as was previously done in Card (2001)and others. Lafortune and Tessada (2010) show that ethnic networks were as important over this period as they are nowadays in determining the location choices of immigrants.

We estimate the causal effect of the immigration of children and adults on educational attainment using two different data sources. The first correspond to state-level aggregates of enrollment and graduation rates, as compiled by Goldin and Katz (1999) from reports "Biennial Survey of Education" and "Reports of the Commissions of Education". These have the advantage of being more precisely measured and of distinguishing between private and public schools. However, they also include both natives and immigrants and thus could generate a spurious relation if immigrants enroll or graduate at a different rate than natives. To alleviate this concern,

[1]Only one percent of immigrants were classified as professionals by immigration authorities. Furthermore, a great majority were males between the ages of 15 and 40 and 56 percent of all immigrants were in the labor force (Carter and Sutch (1997)).

we also measure educational outcomes of natives only using Public-Use Micro Samples of the US Census (Ruggles, Sobek, Alexander, Fitch, Goeken, Hall, King, and Ronnander, 2008). We also measure there geographical mobility as a potential response of native teenagers to immigration flows.

Using our instrumental variables, we find that adult immigration significantly increases school attendance as a teenager and aggregate high school enrollment and graduation rates while child immigration significantly decreases school attendance, high school enrollment and high school graduation rates both at the aggregate level and when focusing on natives only. An increase of one percent in adult immigration at the level of a state increases the aggregate enrollment and graduation rate of about 0.03 percentage points. A similar increase in the number of children immigrating decreases the enrollment and graduation rates by about 5 percentage points, both in the aggregate and micro-data. We also find that these effects are particularly marked for urban and male individuals and for those of US born parents as opposed to second-generation immigrants.

Finally, we also show limited responses of natives in terms of migration to the private system or to suburbs, a phenomena described as "native flight" by Dottori and Shen (2009) in the case of the United States and also found in Danemark by Gerdes (2010) but only using correlations. More children or adult immigrants does not seem to generate a change in location patterns of natives and increase the relative importance of the public sector, suggesting limited response mechanisms of natives in front of the migration shock.

The most similar paper to ours is that of Hunt (2012) who elaborates the same hypothesis as ours but test the results in modern day US (1940-2010). Her instruments are much weaker than ours and because of that, do not allow her to identify separately the role of adult and children migration. However, she does find evidence of the same pattern presented here. She finds that overall, high school completion was increased by immigration and that this effect was larger for Blacks than for Non-hispanic Whites. We think that the historical context in which we study this question may also be more relevant as most natives were likely to attend a schooling level where supply may be more inelastic than in today's tertiary education.

Other studies of modern-day immigration include that of Betts (1998) who studies the period of 1980 to 1990 in the United States. She finds that immigration increases returns but decreases quality. In a separate paper, Betts and Lofstrom (2000) find evidence of "crowding-out" in high schools but not in elementary schools. Hoxby (1998) and Borjas (2004) find evidence that immigration have reduced enrollment of natives in higher education but do not differentiate between ages of immigrants in this finding.

For our period of study, few studies have looked at the relationship between education and immigration. While Ramcharan (2002) had already argued that massive immigration of low-skill workers was the most important cause of the explosive growth in secondary education over this

period, it does so by presenting correlations instead of causal relationship. Furthermore, his paper does not separate the potential negative impact that youth immigration may have had on the same process. Our estimates suggest, however, that 10 to 15 percent of the rapid increase in high school enrollment and graduation can be explained by the restrictions on immigration after 1920, highlighting instead the negative impact that immigration may have had on natives' educational attainment.

The rest of the paper is organized as followed. Section 1 summarizes the historical context in question. Section 2 then presents our model which inspires the empirical strategy, which is presented, alongside with the data, in the following section. Section 4 presents our results and the last section concludes.

1 Historical Context

1.1 Educational transformation

The period between 1910 and 1940 is known as the age of the "High School Movement"[2] due to the considerable growth that secondary education experienced in the beginning of the 20th century. A clear sign of this explosive increase is the percentage of American youths who graduated from high school during this period, which shot up from 9% in 1910 to 51% in 1940. As a result of this substantial shift in education levels the country became the world leader in education, pioneering a secondary education system that was both public and universal. The innovative nature behind this system was reflected in the fact that many of the most developed countries in Europe, such as England and Germany, had just begun to increase the coverage of secondary education mid-century, or rather three decades after the United States.

The key questions are: Why did the growth spurt in secondary education occur during this period? Why did this explosion occur in the United States so much earlier than in other developed countries? In order to understand this phenomenon it is necessary to mention that during this period the concept of "High School" in the US experienced a profound change. Having been seen as privilege for the elite during the 19th century, in the 20th century the conceptualization shifted towards being more generalized and accessible to everyone. With respect to this shift Goldin (1998) indicates that secondary education went from being considered specific preparation for those who wanted to continue on to the university level, to being thought of as valuable life training for all. This transformation is what would later allow secondary education to be useful for many professions that required reduced levels of academic preparation. This is evidenced by the fact that the rate of high school graduates that continued on to higher education fell from

[2]Goldin (1998) performs a thorough analysis of the considerable growth that secondary education experienced between 1900 and 1940 , this phenomenon is referred to as the "High School Movement".

49% in 1910 to 25% in 1933.

In order to understand this shift in the focus of secondary education and the increase in the rate of both graduation and registration, it is necessary to analyze what occurred in the labor market during this period, which could have been the trigger that led more people to study in the secondary education system, while decreasing the percentage of students that continued on to higher education. According to Goldin (1998); Goldin and Katz (2007), there was an important increase in the demand of labor sectors that required people with educational preparation, but not necessarily on the level of "college" education. This author adds that this increase in the demand for technical jobs, which required preparation levels below that which is provided in a university setting, was what produced profound changes in the concept of secondary education. In this same vein, the old concept of "high school", which was based upon a “classic” curriculum similar to what was required at the university level, to a much more technical course of study. There is evidence that indicates that in 1900 subjects such as Latin, Greek, and Literature made up the secondary education curriculum, whereas with the transformation described more emphasis was given to technical education, mathematics, English, and foreign languages. An illustrative example of the change in curriculum is that in 1900 51% of secondary education institutions offered Latin courses, whereas this figure had dropped to 30% in 1934 (Goldin, 1998).

1.2 Immigration

On the other hand, with respect to the relationship between levels of immigration and education it is inevitable to question whether there was a causal link between these variables, given that the explosion in secondary education coincides with a significant increase in immigration to the US. It was precisely in this period that the proportion of immigrants versus the total population reached the highest levels in the history of the United States. It is for this same reason that it could be argued that this factor influenced the previously described changes to the educational system. In order to be able to investigate whether there is a causal relationship between these two phenomena, it is necessary to understand the motives behind immigration during this period and the composition of the masses of immigrants that entered the country during this period. Firstly, it must be mentioned that immigrants to the US in the first part of the 20th century were not exposed to restrictions and quotas. Nonetheless, in 1924 the "Immigration Act" was enacted, which severely limited the entrance of foreigners to the country. There would be two types of restrictions: one based on a maximum allowance for the country and the other based on limitations regarding the qualifications of immigrants.

Within this context, it was in 1907 that the maximum flow of immigrants versus the total population was reached, with a total of 1,285,549 people. The historical data shows that in 1910, 13.5 million immigrants lived in the US and the population with a record of having immigrated

represented 20% of the population, a percentage that is considerably higher than current statistics. Additionally, it is important to mention that the composition of immigration to the US changed significantly during the first years of the 20th century. Towards the end of the previous century, immigrants mostly had origins in Western European countries, such as England, Ireland, Scotland, and Germany, however, this would change drastically in the beginning of the 19th century as the influx shifted towards Eastern European countries, such as Romania, Serbia, Bulgaria, Russia, Poland, and Slovakia, in addition to a considerable portion from Italy and the Balkan States.

Until 1930, the majority of immigrants were men and had a level of academic preparation that was below that of the average American citizen. According to the statistics mentioned by an author of this period, Shriver (1913), the majority of immigrants were young (only 4% of immigrants in 1910 were older than 44 years old) and among the working immigrants, 84% were between 14 and 45 years old. Taking into consideration both the characteristics of immigrants during this period and these figures, it is possible to argue that there was an important effect on the local labor market, which could have affected the return to education. According to Shriver, immigration from the South and East of Europe had very little effect on the "skilled" labor market, but in the "unskilled" market the effects were significant. In the same text, Shriver cites a 3-year-long investigation by the Immigration Federal Commission that reports an excess in the supply of "unskilled" jobs and proposes restrictive legislation based on a literacy and reading test to filter entrance into the United States. According to this same author, these immigrants came from countries where labor conditions were undesirable and they were willing to accept pay and conditions that Americans would never agree to.

The previously described statistics give cause to believe that immigration influenced the labor market. Given this relationship, it is possible to argue that there is a relationship between the education levels achieved in that period and the pronounced entrance of immigrants that the country experienced at the beginning of the 20th century. Additionally, Shriver notes that the majority of immigrant children had two choices: go to public schools or go to work in a trade that didn't required little to no educational preparation. An example is cited from the Polish community in which of a total of 6500 school-age children, 4500 attended public schools. This data is useful to support the second channel of impact, which refers to the effect in the classroom of having foreign classmates that, in some ways, worsen the levels of education and increase the cost of the system. Additionally, it could be argued that the increased presence of immigrants in public schools caused displacement of the natives to private schools.

2 Simple theoretical framework

Having already described the important changes that occurred during that period, it is necessary to create a model that is able to explain those transformations in general terms, as well as to support the analysis that will be carried out in the empirical data section of this paper.

This section is made up of two models, both of which are necessary to understand the way in which immigration impacted educational levels. The way in which the entrance of immigrants might affect the return to education, by means of relative changes in pay rates and costs to the educational system as a consequence of having immigrant classmates, will be analyzed.

The decision to attend school is related to the way in which the investment in human capital affects the wellbeing of the individual. The most direct way in which education impacts people is by increasing their productivity and, therein, their income. For this reason the model will consider that the decision to attend school is completely linked with the topic of cost versus benefit, wherein the benefit is the salary that one is able to obtain given a certain level of education, and the cost is derived from the sacrifices implied in educating oneself. (Cost includes monetary elements, as well as time, transportation, quality, etc).

Formally, an individual chooses his or her level of education (y), maximizing the following equation:

$$\underset{y}{Max}\ W(y_i) - y_i C_i$$

Where y represents a dichotomous variable, with a value of $y = 1$ if the individual decides to finish their secondary education and be "high skilled"(receive a diploma), or a value of $y = 0$ if they leave their schooling before graduating. Within this equation it is assumed that the cost of "certifying" one's education is only incurred in those cases in which the student completes his or her secondary schooling. That is to say that if the individual leaves before graduating, the cost would be null because $y = 0$ in this case.

Additionally, with the objective of incorporating heterogeneity among the participants, the cost of secondary education C will be influenced by the abilities of each individual. Obviously, the decision to complete one's secondary education will be easier for individuals with increased intellectual capacity. The abilities are defined according to the parameter a, which is distributed in a continuous manner according to a distribution function $f(a)$. Without lost of generality, we restricted a to an interval of $[0, 1]$ and the total population is standardized at 1. Given this ability, all of the individuals choose $y = 1$ when $a_i > a*$, where $w(1) - w(0) = C(a*)$ and $y = 0$ when $a_i < a*$

In this way, immigration can affect one's decision to attend school by means of two different channels: It can influence the return to education, which we have defined as the breach between

the salary that a graduate can obtain, versus someone that did not finish their schooling ($w(1) - w(0)$), or it can influence the cost of educating oneself.

In order to understand how immigration can impact the return to education, consider an economy in which there are competitive businesses that produce a fairly homogenous Y, which is produced only with two factors: "skilled" labor and "unskilled" labor. Technology is represented by a constant elasticity of substitution function, or CES, where $\sigma = \frac{1}{1-\rho}$ is the elasticity.

$$Y = A_t[\theta_t S_t^{\rho} + (1-\theta_t)U_t^{\rho}]^{\frac{1}{\rho}}$$

A_t is the total productivity of the factors. The θ_t parameter represents the proportion of skilled labor that is utilized in production. S_t is the quantity of skilled laborers (with a high school education); U_t represents the unskilled laborers (without high school education).

Given this scenario, the businesses maximize their profits and the first order conditions are the following:

$$W_s = A_t[\theta_t S_t^{\rho} + (1-\theta_t)U_t^{\rho}]^{\frac{1}{\rho}}\rho\theta_t S_t^{\rho-1} \tag{1}$$

$$W_u = A_t[\theta_t S_t^{\rho} + (1-\theta_t)U_t^{\rho}]^{\frac{1}{\rho}}\rho(1-\theta_t)U_t^{\rho-1} \tag{2}$$

Dividing (1) with (2) and applying the logarithm we arrive at the following expression:

$$log(\frac{W_S}{W_U}) = log(\frac{\theta_t}{1-\theta_t}) - \frac{1}{\sigma}log(\frac{S}{U}) \tag{3}$$

From the equation (3) we see that a relative increase in unskilled laborers (U), as caused by the large scale immigration wave in the beginning of the 20th century, would generate a decrease in the term ($\frac{1}{\sigma}log(\frac{S}{U})$) and, therein, an increase in the relative salaries of skilled laborers (or in terms of the previous model, and increase in $w(1) - w(0)$). Additionally, the bigger the substitution elasticity between both types of work, the bigger the effect in the relative supply of labor in the relative salary ($\frac{1}{\sigma}log(\frac{S}{U})$).

On the other hand, the cost C increases with the quantity of child immigrants that arrive in each cohort. Considering this concept, we seek to incorporate the idea that there is a negative externality that is produced within the educational institutions, given the fact that those institutions with higher proportions of immigrant students achieved lower education levels. This can be understood as a deterioration in the quality of the level of education that is achieved. Accordingly, it is necessary to make more economic sacrifices in order to acquire quality education. There are studies that support the negative effect of immigrants in the classroom: Friesen

and Krauth (2011) find in the recent US data that having more individuals of ethnic groups that speak different language than English lowered the mathematics score of native students in elementary schools. Along this same line, there are additional studies that were carried out in other countries: Gould, Lavy, and Paserman (2009) uses data from Israel and find that having more immigrant peers (principally from the Soviet Union) in elementary school had a negative effect in the success rate of students taking entrance exams for secondary schools. Additionally, Jensen and Rasmussen (2011) showed that schools with higher concentrations of immigrants have significantly worse results than those schools with lower proportions of immigrants, in both mathematics and language using data from the 2000-2005 period in Denmark.

Based on this information, the two possible channels of impact that we seek to explore in this study have been modeled. As mentioned, these effects may be triggered by immigrants of distinct age groups and training levels, which is why an important part of this research will be the data selection and classification that is carried out for the immigrant groups being studied, allowing us to differentiate between immigration as it relates to children and adults.

3 Empirical strategy and Data

Having described the theoretical link that may exist between immigration and levels of secondary education, this section will outline the empirical strategy that will be used to identify and measure the effects that are produced through the distinct channels of impace that have been mentioned. First, the regressions that will be used to measure the distinct effects will be defined, followed by a description of the data, and then the criteria for classification of each group of immigrants. Lastly, there will be an explanation of how the problem of endogeneity that occurs as a consequence of the immigrants being able to choose their location according to the educational conditions of the place.

3.1 Empirical specification

The first regression to estimate is relates adult and child immigration with the educational outcomes of interest. We parametrize the relationship using the following equation:

$$y_{ist} = \beta_1 \ln(N_{st}^C) + \beta_2 \ln(N_{st}^A) + \alpha X_{ist} + \mu_s + \nu_t + \varepsilon_{ist} \tag{4}$$

where y_{ist} represents the educational attainment of individual i, born in state s and in cohort t. This is related to the immigration of children and adults through the parameters β_1 and β_2 respectively. In some cases, we will add individual controls X_{ist} including the age, gender and parents' nativity. With the intention of controlling additional factors that may affect the outcome,

all of the regressions defined here include fixed effects by year and by state to control for changes specific to each state, or changes that took effect year to year and were not necessarily related to immigration. Additionally, the estimation will cluster standard errors by state, in order to allow for errors being correlated within each state over tim.

The idea behind this regression is to separate the two channels previously mentioned. With respect to what has been described in section 2, the child immigration coefficient should reflect the increase in the cost of educating oneself, or the decreased quality of the education as a consequence of having immigrant classmates, and therein should be negative. Accordingly, the adult immigration coefficient should be positive since it seeks to capture the impact of the shift in returning to education.

3.2 Instrumentalization

A central theme in regards to the estimation of the equations previously defined is that immigrants are able to choose their destination according to the education levels said place. For this reason adult immigration, both 'skilled' and 'non-skilled', and child immigration may be endogenous variables. For example, these foreigners may have chosen states that had a higher probability of having better schools, or possibly those places with a larger quantity of public schools. In order to solve this problem of endogeneity, an instrument was created for each of the three variables, following a strategy similar to the one employed by Card (2001). The main idea behind this strategy is based on the idea that people from the same country tend to form networks and immigrate to the same places (Munshi, 2003). There are studies that show that immigrants tend to go to the places where they have more connections and links in order to form networks that, in certain ways, will protect them (Lafortune and Tessada, 2010). This may refer to increased access to information, or other reasons such as contacts, family, protection, etc. In accordance with this theory, past decisions regarding immigration of individuals of the same ethnicity are a good predictor of the decision to immigrate to a specific place.

Following this line of thought, the proportion of each ethnic group by state in 1900 (one decade before the first cohort) is used as a predictor to anticipate the quantity of immigrants from each ethnic group that would arrive to each state. Using this strategy, we can predict the total number of adult and child immigrants by state, which may be exogenous to the local conditions. Therein, the instruments were constructed in the following manner: All of the adults and children were separately assigned to a specific state for each year and ethnic group, according to the concentration that existed in 1900, for said group and place. For example, if 15% of all Russians were located in Wisconsin in 1900, then 15% of all of the children from Russia that entered the country after 1900 are assigned to Wisconsin (the same for 15% of the adults). With this system, a prediction of the migratory flow of the two types of immigration is created by

ethnic group. Adding up for all ethnic groups we can construct a prediction of the flow of each type of immigrant to each state.

The following equations formally define the way in which the instruments were constructed for each type of the endogenous variables:

$$\hat{N}_{st}^{k} = \sum_{J} \frac{N_{js1900}}{N_{j1900}} \times N_{jt}^{k} \tag{5}$$

Where k correspond to the type of immigrant (children or adult), j corresponds to the ethnic group and s corresponds to state. N_{js1900} is the stock of immigrants of the ethnic group j in the state s in 1900 and N_{j1900} is the total stock of immigrants of the ethnic group j in the US in 1900. We grouped countries of origin into 42 different ethnic groups as specified in Lafortune (2013) to reduce the noise in the shares of some small groups.

Note that the allocation of both children and adults follow the same *total* shares and not shares specific to each age group. This is done to alleviate potential concerns that immigrants of a distinct type have selected a given location in the past based on the attractiveness of that location for a given age group. Thus, the reason a given state is predicted to receive a higher number of children in a given year is not because that state was more attractive to immigrant children in 1900 but because it was more attractive in 1900 to a given ethnic group which, over this period, received a larger flow of children immigrants.

The identification of the effects of immigration on the dependent variable will only be achieved if the chosen instrument affects the dependent variable (education levels) exclusively through immigration and not by other factors. If past incoming groups of immigrants chose their states of residence because they were anticipating the future states of education in these places, then there would be an identification problem and the instrument would not be valid. It's important to keep in mind that it is not likely that this is the case, given that these people would not only have to have anticipated future states of education in their chosen home, but they must also be particular for that specific ethnic group. If some immigrants before 1900 preferred a state because they anticipated better conditions for all ethnic groups, but not for any group in particular, then the instrument would still be valid. That is to say that in order for the instruments to lose their validity there must be a situation in which an ethnic group anticipates some type of "shock" that affects them in particular, and in a specific state. This condition would have to persist over time until past 1910 in order for it to affect the decisions of immigrants from that time period.

If the immigrants from 1900 anticipate changes that do not persist over time (do not last until 1910), then the identification strategy continues to be valid. Finally, the condition of relevance for the instruments is backed up by the studies previously mentioned, in which it is proven that the decisions of immigrants is closely related to the past decisions of individuals of the same ethnic

group.

3.3 Data and classification of immigrants

Immigration data come from the 1910-1930 IPUMS ((Ruggles, Sobek, Alexander, Fitch, Goeken, Hall, King, and Ronnander, 2008)). Immigrants are classified by their country of origin (which is mapped to an ethnic group), their age, their year of arrival and their current state of residence. For native children, born between 1891 and 1915, we form 5-year "age cohorts" between 1910 to 1930, such that each is allocated to the 5-year multiple in which he turns 15-19 year old. These native teens are matched to two different flows of immigrants. First, immigrant children consist of the flow of immigrants of the same birth years as the natives and arriving in years before the "cohort" is based. Immigrant adults are selected such that they were born in the 20 year interval earlier than the children and arrived within 5 years of the year the cohort is based. The location shares of the distinct ethnic groups in 1900 that were utilized to construct the instruments were obtained from the tables of the 1900 census ((Office, 1901)).

On the other hand, the aggregate education data comes from Goldin and Katz (1999). These figures are compiled by state from 1910 to 1938 (every 4 years, and in some cases every 2) in the following manner: in the years preceding 1918, the data is compiled in reports known as "Reports of the Commissioner of Education"; in the following years a source called "Biennial Surveys of Education" were used. This was the way in which graduation and registration rates were obtained for secondary education. In both cases, total rates and rates for public education were obtained, which permitted the analysis of the impact of immigration on the choice between public and private education. The graduation rate is measured as the number of graduates per year over the total population over 17 years old, which is calculated by state. On the other hand, the rate of school enrollment is measured as the number of registered people by state over the total population of children between the ages of 14 and 17 (this is measured for public schools, private schools, and jointly). We use the observations in 1910, 1914, 1920, 1924 and 1930 to match our cohorts from the IPUMS data such that the 1914 data is matched to the 1915 cohort and the 1924 data is matched to the 1925 cohort.

We supplement this using micro-data on educational attainment and school attendance of natives from the 1910 to 1970 IPUMS ((Ruggles, Sobek, Alexander, Fitch, Goeken, Hall, King, and Ronnander, 2008)). We take two different samples in this case. In the first one, we focus on teenagers and restrict the sample to the census years where we can observe the individuals aged 10 to 19. In that sample, we measure whether these individuals were attending school and where they were residing (urban, metropolitan, central city, different state than their birth state). For these individuals, however, we cannot obtain the number of years of schooling as this information only becomes available in the censuses of 1940 onwards. We thus use in this

case adults aged 40 to 59 and study their probability of having enrolled in high school, having graduated from high school and their highest grade ever attained.

Table 1 presents the summary statistics of our outcomes of interest. A majority, 58 percent, of teenagers were attending school at the moment of the census. About half were living in a urban setting, 40 percent in a metropolitan area and 26 percent in the central city of that area. Almost 16 percent of teenagers had already migrated to another state than their state of birth. The average years of schooling of our sample is 9.4 years and while 52 percent completed more than 8 years of schooling, only 33 percent ever graduated from high school. The increase in high school attendance and graduation is even larger among this native population than the overall statistics from the aggregate data suggest: for cohorts born between 1891 and 1895, 39 percent attended and 23 graduated from high school while for cohorts born between 1911 and 1915, this rises to 64 and 42 percent respectively.

4 Results

4.1 First stage

Table 2 presents the results of the first stage. It clearly demonstrates that, in this period, we have more than sufficient power to separately identify the causal effect of adults and children immigrating. A predicted change of 1 percent in the number of adult immigrants to a given state translates into a real change of about 0.7 while a similar change in the predicted number of children lead to a 0.85 percent change in the actual number of teens. Whether the weights are given by our sample of adult outcomes, children outcomes or aggregate outcomes does not significantly alter the strength of the instruments. The F-tests are all much higher than the critical values of Stock and Yogo (2005). Furthermore, each instrument separately predicts our one of our endogenous variable, implying that we are not really using one instrument for two endogenous variables (See (Shea, 1997)). The Shea partial R-square of our regression is relatively high and suggests that our empirical strategy is not masking a lack of exogenous instruments.

Table 3 explores how robust our first stage is. In the first Panel, we exclude the first two cohorts. Our instrument continues being able to strongly predict the location choices of adults but the precision of our estimator for children becomes very small. Panel B, on the other hand, excludes the last two cohorts and show that our results are robust to that exclusion. Finally, Panel C omits the state of New York, the largest immigrant-receiving state, with little effect on the predictive power of our instruments, although, in this case, the predicted number of children also helps explaining the location pattern of adults.

4.2 Main outcomes

We first present the results of our regressions using aggregate educational data from historical sources to measure our outcome variables. Table 4 presents the OLS and instrumental variable coefficients of our regression of interest where the outcome variable is either the state-level graduation rate (in the first two columns) and enrollment rate (in the last two). The correlation between the number of immigrants and the graduation rates are very small and insignificant. For enrollment rates, a higher number of immigrant children is associated with a higher enrollment rate. However, these correlations are likely to be biased upwards for immigrant children as these immigrants are likely to elect states that have the best educational systems and thus we should expect the causal effect to be smaller than the observed correlation. This is exactly what we observe in columns (2) and (4) where the number of immigrant children, once instrumented, now is estimated to lower the aggregate enrollment and graduation rates. Formally, an increase of one percent in the number of immigrants causes a decrease in the enrollment rate of about 6 percentage points and a decrease in the graduation rate of about 5. An increase in one percent in the number of adult immigrants, on the other hand, raises both rates by 3 percentage points but these results are not significant.[3]

This is consistent with our hypothesis but could be due to a mechanical relationship in the case, at least of immigrant children. The reason is that the enrollment and graduation rates were computed using all teenagers. If immigrant children arrive in the country and are less likely to attend and graduate from high school than their native counterpart, their arrival would naturally lower the aggregate rates without any crowding-out effect on natives. To study whether this can explain the results we obtain in the aggregate data, we now turn to our micro-data estimates where we are able to restrict the sample to natives. These results are presented in Table 5 where Panel A presents the OLS coefficients and Panel B, the IV ones. The first column corresponds to a linear probability model of being in school at the moment of the Census and is restricted to individuals aged 10 to 19. The middle two columns correspond to a linear probability model of having attended high school (that is, the highest grade ever attained being 9 or more) and the probability of graduating from it (that is, the highest grade ever attained beign 12 or more). Finally, the last column is simply the highest grade ever attained by the individual. We note that the estimates of Panel A are, as before, biased towards zero compared to the causal estimates obtained in Panel B. The results on being currently enrolled in school in column (1) mirror the results of the previous table not only in significance but also in magnitude. An increase of one percent in the number of immigrant children arriving to a state decreases the probability that a native is currently enrolled in school by 5 percent. A similar increase in the number of adult

[3]In the case where we exploited all the data points offered by the aggregate data instead of focusing only on years that matched our IPUMS samples, we obtained the same coefficients in terms of magnitude but lower standard errors, in particular in the case of adult immigration, where the impact was then significant.

immigrants, on the other hand, increases the probability that the native is in school by 3 percent. The subsequent columns show that this translates into a lower probability that natives have ever enrolled in high school of about 6 percent in response to a one percent increase in the number of children but the impact of adult immigrants is now muted. Furthermore, immigrant children do not seem to prevent graduation from high school, suggesting that the results of Table 4 were partially reflecting the mechanical relationship between the arrival of less qualified migrants and the aggregate graduation rates. Finally, the last column suggests that the effect of immigration on educational attainment was very non-linear as the number of children immigrants arriving did not influence the eventual number of years of education obtained by natives.

4.3 Heterogeneous effects

Having shown that the arrival of children and adult immigrants significantly altered the high school enrollment and graduation rates, we next study whether these effects are concentrated within a sub-population or more generalized. First, we can explore whether the effects of immigration were more concentrated for one gender than another. Goldin (1998) shows that, in this period, females were consistently more likely to enroll and graduate from high school than their male counterparts. However, their labor market participation, in particular once married, were significantly lower than that of men. Thus, we could expect both the crowd-out effect and the labor market incentives generated by increased immigration to be larger for males than for females. Panel A of Table 6 demonstrates that this is exactly the case. In this table, we focus on our three measures related to high school, mainly whether the individual attended school in the year of the Census, whether the individual ever enrolled in high school and whether the individual graduated from high school. The results presented in Table A suggest that immigration did not have a significant effect, in general, on school choices for females, except in the case of high school enrollment where the crowd-out effect is large. The response of males, however, is significantly larger than that of females, in particular through the labor market channel. An increase in one percent in the number of adult immigrants raises the probability that a male was enrolled and graduated from high school and also increases the likelihood that he was in school at the time of the census. Males appear to be less likely than females to not be in school at the time of the Census when more children immigrants are locating in the state.

Panel B of Table 6 further explores the heterogeneity in responses to immigration by distinguishing between second-generation and natives. Second-generation (defined here as having at least one parent foreign-born) were 9 percent less likely to enroll or graduate from high school. They were also 6 percent less likely to report being in school in the year of the Census. Thus, we may expect that these individuals would be less influenced by immigration since they were not as likely to be pursuing their education as far as children of native parents. However, one could also think that second-generation individuals would be living closer to immigrants than others

and potentially be competing more directly with immigrants than natives. The results suggest that the first effect dominates and that native individuals are more likely to increase their enrollment and graduation from high school in response to a larger number of adult immigrants and to decrease their participation in the high school market in response to a higher number of immigrant children.

Immigrants were much more likely to settle in urban zones. Thus, the effects we have documented so far should be observed more strongly in urban instead of rural zones. Unfortunately, the census micro-data does not allow us to know whether the individual was born in a rural or urban zone. Furthermore, it is likely that changes in educational attainment would alter the likelihood that an individual migrates to the city the causing difficulty in the interpretation of the interaction. Finally, immigration itself could induce immigration, which complicates furthermore the analysis. However, we still present that interaction in Panel C of Table 6 for the first outcome which is measured when individuals are in their teens and thus potentially less subject to a problem of ex-post migration. The results suggest that only individuals living in urban areas respond to immigration. The response of rural individuals is close to 0 for both channels.

Finally, we study whether the change documented here occurred most importantly in the private or public educational sector. The high school movement was a "public school" phenomena and we also expect that most immigrants would have attended public school and not private school given their income levels. To measure this difference, we must use the state-level aggregates and present results for the graduation and enrollment rates in Table 7. The results suggest that the impact of both types of immigration had the most impact on the enrollment and graduation rates in the public sector. The coefficients for the private sector are much smaller although often similarly significant as those in the public sector.

4.4 Interpreting the results

The results presented above, while significant in magnitude, could simply reflect a lower bound if parents and/or teenagers take some additional measures to avoid the impact of immigration. If that was the case, experiencing such an immigration shock but preventing the readjustments would generate even larger impacts on the schooling decisions of natives. We find only limited evidence of that pattern in the data. In Table 7, columns (3) and (6) look at the fraction of graduation and enrollment rates in the public sector. We find that the arrival of more children immigrants increased the fraction of high schoolers enrolled and graduating from the public system, thus showing limited "crowding-out" of natives towards the private sector. There is also very limited evidence of relocation of natives in response to these flows, as demonstrated by Table 8. The outcomes presented in this table is a dummy identifying whether the native teen was living in a urban setting, a metropolitan area and a central city within the metropolitan area.

Finally, the last column uses an indicator variable which is equal to 1 if the state of residence of the teen is not her state of birth. The results suggest very limited evidence of an impact of immigration on any of these variables. None of the coefficients in the first three columns are significant and most of them are extremely small. In the last column, there is some evidence that teens who were born in states where more adult immigrants arrived in the last 5 years were less likely to be currently living out of their state of birth. The effect is, however, relatively small and significant only at the 10 percent significance level.

Finally, we attempt to use our measure to ask how much of the "high school movement" may be due to the tightening of the immigration restrictions in 1924, or more generally, to the decrease in immigration over this period. According to our data, we find that comparing cohorts before 1920 and after 1920, the United States experienced a decrease in the log number of children immigrants of about 0.9 and a decrease in the log number of adult immigrants of about 0.67 percent in an average state. Our estimates would suggest that this change would translate into an increase in the graduation rate of 0.030 and an increase in the enrollment rate of about 0.036. This corresponds to about 10 percent of the change in enrollment and 15 percent of the change in graduation rate between 1910 and 1930. Thus, while the contribution of shifts in immigration pattern was clearly relevant, one could not say that the major barrier to the "high school movement" were the high levels of immigrations before 1920.

5 Conclusion

This paper thus has shown that the age at which immigrants arrive to an economy can differentially impact the education outcomes of natives: while immigrant children may "crowd-out" natives out of public schools, adult immigrants may make education more attractive to natives by altering relative supply of skills in the economy. Our results suggest significant and economically relevant effects of immigration on native outcomes although the change in immigration levels appears to explain a relatively small fraction of the "high school movement" in the United States.

A question that arises from our results is what would thus be the optimal immigration policy of an economy that wants to increase the stock of its qualified labor force. While in the past, a lot of focus has been given the educational level of the selected migrants, our results suggest that age patterns may be equally relevant in that design.

Finally, the limited response in terms of educational systems and geographical mobility does suggest that either natives do not attempt to avoid the impacts generated by immigrants or are unable to. More research is warranted on this topic.

References

Betts, J. (1998): "Educational Crowding Out: Do Immigrants Affect the Educational Attainment of American Minorities?," University of California at San Diego, Economics Working Paper Series qt8vt7f1bh, Department of Economics, UC San Diego.

Betts, J. R., and M. Lofstrom (2000): "The educational attainment of immigrants: Trends and implications," in *Issues in the Economics of Immigration*, pp. 51–116. University of Chicago Press.

Borjas, G. J. (2004): "Do Foreign Students Crowd Out Native Students from Graduate Programs?," Working Paper 10349, National Bureau of Economic Research.

Card, D. (2001): "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 19, 22–64.

Carter, S. B., and R. Sutch (1997): "Historical Perspectives on the Economic Consequences of Immigration into the United States," Working Paper 106, National Bureau of Economic Research.

Dottori, D., and I.-L. Shen (2009): "Low-Skilled Immigration and the Expansion of Private Schools," IZA Discussion Papers 3946, Institute for the Study of Labor (IZA).

Friesen, J., and B. Krauth (2011): "Ethnic enclaves in the classroom," *Labour Economics*, 18(5), 656–663.

Gerdes, C. (2010): "Does Immigration Induce 'Native Flight' from Public Schools? Evidence from a Large Scale Voucher Program," IZA Discussion Papers 4788, Institute for the Study of Labor (IZA).

Goldin, C. (1998): "America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century," *The Journal of Economic History*, 58(2), pp. 345–374.

Goldin, C., and L. F. Katz (1999): "Human capital and social capital: The rise of secondary schooling in America, 1910-1940," *Journal of Interdisciplinary History*, 29(4), 683–723.

Goldin, C., and L. F. Katz (2007): "The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials, 1890 to 2005," NBER Working Papers 12984, National Bureau of Economic Research, Inc.

Gould, E. D., V. Lavy, and M. D. Paserman (2009): "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence," *Economic Journal*, 119(540), 1243–1269.

Hoxby, C. (1998): "Do Immigrants Crowd Out Disadvantaged American NAtives Out of Higher Education?," in *Help or Hindrance? The Economic Implications of Immigration for AfricanAmericans*, ed. by D. Hamermesh, and F. Bean. Russell Sage Foundation, New York.

Hunt, J. (2012): "The impact of immigration on the educational attainment of natives," Discussion paper, National Bureau of Economic Research.

Jensen, P., and A. W. Rasmussen (2011): "The effect of immigrant concentration in schools on native and immigrant children's reading and math skills," *Economics of Education Review*, 30(6), 1503–1515.

Lafortune, J. (2013): "Making Yourself Attractive: Pre-marital Investments and the Returns to Education in the Marriage Market," *American Economic Journal: Applied Economics*, 5(2), 151–78.

Lafortune, J., and J. Tessada (2010): "Smooth(er) Landing? The Dynamic Role of Networks in the Location and Occupational Choice of Immigrants," Mimeo, University of Maryland.

Munshi, K. (2003): "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," *Quarterly Journal of Economics*, 118(2), 549–597.

Office, U. S. C. (1901): *Twelfth census of the United States, taken in the year 1900. Population / William R. Merriam, director ; prepared under the supervision of William C. Hunt*. United States Census Office Washington.

Ramcharan, R. (2002): "Migration and Human Capital Formation: Theory and Evidence from the U.S. High School Movement," IMF Working Papers 02/123, International Monetary Fund.

Ruggles, S., M. Sobek, T. Alexander, C. A. Fitch, R. Goeken, P. K. Hall, M. King, and C. Ronnander (2008): "Integrated Public Use Microdata Series: Version 4.0 [Machine-readable database]," .

Shea, J. (1997): "Instrument relevance in multivariate linear models: A simple measure," *Review of Economics and Statistics*, 79(2), 348–352.

Shriver, W. (1913): *Immigrant Forces: Factors in the New Democracy*, Forward mission study courses. Missionary Education Movement of the United States and Canada.

Smith, C. L. (2012): "The Impact of Low-Skilled Immigration on the Youth Labor Market," *Journal of Labor Economics*, 30(1), 55 – 89.

Stock, J. H., and M. Yogo (2005): "Testing for Weak Instruments in IV Regression," in *Identifcation and Inference for Econometric Models: A Festschrift in Honor of Thomas Rothenberg*, ed. by D. Andrews, and J. Stock. Cambridge University Press, Cambridge.

Table 1: Summary statistics

	Mean	Standard Deviations
	(1)	(2)
	Panel A: Teenage outcome sample (N=912,990)	
Attending school	0.666	0.471
Urban	0.434	0.496
Metropolitan area	0.364	0.481
Central city	0.242	0.428
Lives in a different state than birth	0.151	0.358
	Panel B: Adult outcome sample (N=384,078)	
Years of education	9.436	3.525
Ever attended high school	0.528	0.499
Graduated from high school	0.337	0.473

All statistics are weighted by Census person-weight.

Table 2: First stage

	Log imm. adults	Log imm. children
	(1)	(2)
	Panel A: Teenage outcome sample (N=912,990)	
Log pred. adults	0.666***	0.008
	(0.198)	(0.092)
Log pred. children	0.150	0.855***
	(0.188)	(0.127)
	Panel B: Adult outcome sample (N=384,078)	
Log pred. adults	0.763***	-0.019
	(0.195)	(0.089)
Log pred. children	0.091	0.835***
	(0.235)	(0.123)
	Panel C: Aggregate outcome sample (N=245)	
Log pred. adults	0.570***	-0.141
	(0.181)	(0.205)
Log pred. children	0.136	0.982***
	(0.209)	(0.140)

All regressions include fixed effects for year and state. Clustered standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 3: First stage robustness tests

	Log imm. adults	Log imm. children
	(1)	(2)
	Panel A: Excluding the first two cohorts (N=741,620)	
Log pred. adults	0.945***	-0.147
	(0.336)	(0.147)
Log pred. children	-0.006	0.471
	(0.761)	(0.326)
	Panel B: Excluding the last two cohorts (N=259,843)	
Log pred. adults	0.680***	0.052
	(0.148)	(0.080)
Log pred. children	0.806**	0.737***
	(0.394)	(0.166)
	Panel C: Excluding New York (N=834,827)	
Log pred. adults	0.761***	0.027
	(0.171)	(0.093)
Log pred. children	0.300**	0.879***
	(0.146)	(0.124)

All regressions include fixed effects for year and state. Clustered standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Effect of immigration on aggregate graduation and enrollment rates

	Graduation rate (N=245)		Enrollment rate (N=245)	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
log imm. adults	-0.005	0.037	0.001	0.031
	(0.005)	(0.023)	(0.006)	(0.023)
log imm. children	0.005	-0.061***	0.006	-0.055**
	(0.012)	(0.022)	(0.022)	(0.025)

Clustered standard errors in parentheses. All regressions include fixed effects for year and state.
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Effect of immigration on high school attendance and graduation of natives

	Attending school	Attended high school	Graduated from high school	Highest grade achieved
	(1)	(2)	(3)	(4)
		Panel A: OLS		
log imm. adults	0.017**	0.003	0.000	-0.012
	(0.008)	(0.003)	(0.002)	(0.014)
log imm. children	-0.013	0.007	0.011	0.031
	(0.010)	(0.008)	(0.007)	(0.034)
		Panel B: IV		
log imm. adults	0.029*	0.011	0.014	-0.081**
	(0.016)	(0.013)	(0.014)	(0.041)
log imm. children	-0.051**	-0.056***	-0.025	0.035
	(0.021)	(0.021)	(0.018)	(0.065)
N	912,990	384,078	384,078	384,078

Clustered standard errors in parentheses. All regressions include fixed effects for year and state as well as controls for being male, for having both parents being natives and a linear control for age at the time of the census.
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Effect of immigration on high school attendance and graduation of natives-heterogenous effects

	Attending school (1)	Attended high school (2)	Graduated from high school (3)
		Panel A: By gender	
log imm. adults	0.011	0.006	0.007
	(0.015)	(0.012)	(0.014)
log imm. children	-0.032	-0.053**	-0.019
	(0.021)	(0.021)	(0.018)
log imm. adults	0.043***	0.011*	0.013*
*male	(0.014)	(0.006)	(0.007)
log imm. children	-0.042***	-0.006	-0.012
*male	(0.015)	(0.006)	(0.008)
N	912,990	384,078	384,078
		Panel B: By nativity of parents	
log imm. adults	0.017	-0.027*	-0.018
	(0.016)	(0.015)	(0.015)
log imm. children	-0.022	0.064	-0.052**
	(0.021)	(0.021)	(0.018)
log imm. adults	0.013***	0.041***	0.035***
*native parents	(0.004)	(0.010)	(0.009)
log imm. children	-0.013*	-0.049***	-0.040***
*native parents	(0.007)	(0.015)	(0.011)
N	912,990	384,078	384,078
		Panel C: By urbanicity	
log imm. adults	0.076***		
	(0.020)		
log imm. children	-0.102***		
	(0.024)		
log imm. adults	-0.089***		
*native parents	(0.027)		
log imm. children	0.103***		
*native parents	(0.029)		
N	912,990		

Clustered standard errors in parentheses. All regressions include fixed effects for year and state as well as controls for being male, for having both parents being natives and a linear control for age at the time of the census. In Panel C, a dummy for living in a rural area was also included.
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Effect of immigration on graduation and enrollment rates-by sector

	Graduation rate			Enrollment rate		
	Public	Private	Public Share	Public	Private	Public Share
	(1)	(2)	(3)	(4)	(5)	(6)
ln_adults	0.027	0.010*	-0.009	0.029	0.002	0.013
	(0.020)	(0.005)	(0.033)	(0.023)	(0.008)	(0.029)
ln_children	-0.048**	-0.013***	0.081***	-0.039	-0.017***	0.061**
	(0.020)	(0.005)	(0.026)	(0.025)	(0.006)	(0.029)
N	245	245	245	245	245	245

Clustered standard errors in parentheses. All regressions include fixed effects for year and state.
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Effect of immigration on natives' location choice when teenagers

	Urban	Metropolitan Area	Central city	Moved states
	(1)	(2)	(3)	(4)
log imm. adults	0.023	0.004	0.014	-0.029*
	(0.014)	(0.011)	(0.013)	(0.017)
log imm. children	0.027	0.005	0.018	0.015
	(0.020)	(0.015)	(0.013)	(0.026)
N	9.13e+05	9.13e+05	9.13e+05	9.13e+05

Clustered standard errors in parentheses. All regressions include fixed effects for year and state.
*** p<0.01, ** p<0.05, * p<0.1