The Impact of Immigration on the Educational Attainment of Natives *

Jennifer Hunt †

Rutgers University and NBER

jennifer.hunt@rutgers.edu

April 3, 2012

*I thank, without implicating, David Figlio, Tommaso Frattini, James Heckman, John Eric Humphries, Ethan Lewis, Marguerite Lukes, James MacKinnon, Daniel Parent, Steve Pischke and participants in numerous seminars for comments and data advice. I am grateful to the Social Science and Humanities Research Council of Canada for financial support. I am also affiliated with the CEPR (London) and DIW (Berlin).

†Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, N.J. 08901.

Abstract

Using a state panel based on census data from 1940–2010, I examine the impact of immigration on the high school completion of natives in the United States. Immigrant children could compete for schooling resources with native children, lowering the return to native education and discouraging native high school completion. Conversely, native children might be encouraged to complete high school in order to avoid competing with immigrant high–school dropouts in the labor market. I find evidence that both channels are operative and that the net effect is positive, particularly for blacks, though not for Hispanics. An increase of one percentage point in the share of immigrants in the population aged 11–64 increases the probability that natives aged 11–17 eventually complete 12 years of schooling by 0.2–0.3 percentage points, and increases the probability for native–born blacks by 0.4–0.5 percentage points. I account for the endogeneity of immigrant flows by using instruments based on 1940 settlement patterns.

The extent to which the children of low–education or low–income parents are able to achieve their full potential in the United States is a cause for concern when viewed from an international perspective. Contrary to popular mythology, there is less intergenerational mobility in earnings and education in the United States than in continental Europe and Canada, and no more than in the United Kingdom.[1] An important step upward for many children from low socio–economic status families is graduation from high school, yet U.S. high school graduation rates are no longer increasing.[2] In this paper, I contribute to our understanding of the determinants of high school educational attainment by investigating the role of immigration. Increasing immigration in recent decades has led to popular concern that immigration is reducing the quality of K–12 education. If this concern is well founded, rising immigration could reduce native high school graduation rates. Conversely, immigration–induced changes in labor market incentives for educational attainment could have the opposite effect. I seek evidence for these two channels and assess their net effect.

Immigrants and the young children of immigrants generally have a more limited command of English than natives. If immigrants and natives are taught in the same classes, teachers of some subjects may slow the pace of instruction to accommodate non–native speakers. If immigrant students have had worse quality prior education, or have less education than their native classmates, teachers may lower expectations for all students. Immigrant students could also divert financial resources from native students, potentially lowering the quality of their education. For example, Fix and Zimmerman (1993) find that federal Chapter I spending per economically disadvantaged student fell due to the immigration–induced expansion in the number of eligible children. If immigrants are taught separately from natives, resources might also be diverted from natives. For example, federal Title III money for Limited English Proficient education may come at the expense of other federal funding. A lower educational quality for natives will reduce their earnings capacity at a given number of years of education, and this lower return to education in turn may induce natives to complete fewer years of high school. This prediction

[1] Checchi, Ichino and Rustichini (1999); Corak (2006).
[2] Heckman and LaFontaine (2010).

is not unambiguous, however. If high school becomes easier, the fall in marginal cost may outweigh the fall in the marginal benefit and lead to higher native completion rates. Furthermore, if immigrant students are better educated or harder working than their native classmates, they will provide positive peer effects and may relax the resource constraint, and could increase native completion rates.

There exists a second channel through which immigration could increase natives' high school educational attainment (Betts 1998). Incentives to complete high school are influenced by the wage structure, which is in turn affected by the entry of immigrant workers. Immigration will affect wage inequality among natives if the distribution of immigrant skill differs from that of natives. Compared to natives, immigrants to the United States are very disproportionately poorly educated and somewhat disproportionately highly educated. Immigrants are underrepresented among workers with an intermediate level of education, such as a high school diploma. The effect of immigrants entering the labor market should therefore be to increase wage inequality in the lower half of the distribution, particularly the wage gap between high school dropouts and high school graduates. Empirical studies confirm this.[3] The net effect of the changes in the wage structure is likely to be to increase the return to completing high school, and hence native completion rates. Native–born youth are likely to be well informed about the dropout labor market even while still high–school students, since this is the market in which many seek part–time jobs.[4]

Any negative effects on the schooling quality of natives will affect the children of low socio–economic status (SES) parents more than children of high SES parents. Richer parents may more easily move their child to a learning environment with either fewer immigrants or immigrants with better language skills and educational background, by using private schools (Betts and Fairlie 2003; see also Hoxby 1998) or by moving to a different school district (Cascio and Lewis forthcoming). Furthermore, the educational quality of

[3] Borjas and Katz (2007), Ottaviano and Peri (2008).

[4] Smith (2012) presents evidence that adult immigrants with high school or less reduce the employment rate of native high–school students. This reduction could could provide an additional channel for immigration to affect native graduation rates.

a child's school is likely to have a smaller impact on the children of high SES parents, as such parents can compensate in part for a school's deficiencies by providing the child with instruction at home. Any positive effects of immigration on high–school graduation rates are likely to be larger for groups with low graduation rates, rather than near universal graduation rates. Thus, effects through both channels are more likely to affect low SES natives, and consequently to affect minorities more than non–Hispanic whites. Furthermore, native minorities live in closer proximity to immigrants than native non–Hispanic whites, as I show below, increasing their likely responsiveness to immigration. Minority boys, who have particularly low high–school graduation rates (Orfield et al. 2004, Noguera et al. 2011; Noguera 2008), may be particularly sensitive to immigration.

I focus on the impact of immigration on natives' completion of 12 years of schooling, comparing results across ethnicity, race and gender. I use the decennial censuses of 1940–2000 and the pooled 2008–2010 American Community Surveys (ACS) to construct a state panel. I extend two closely related papers, Betts (1998) and Betts and Lofstrom (2000), in several ways. The most important extensions in practice are the distinction between immigrants of different educational attainment and the closer matching of the immigrant inflows to natives' adolescent years. The extension to the use of instrumental variables based on historical immigrant settlement patterns is important in principle but less important in practice. I also confirm the results are robust to considering holders of GEDs to have completed less than 12 years of education. Some of the analysis in Smith (2012) is also closely related to my paper: he examines the effects of adult low–skill immigration on natives' high school enrollment rates. The estimates are imprecise, however, and for this reason that I follow Betts (1998) and Betts and Lofstrom (2000) in examining completed education among several cohorts of older respondents.[5]

[5] Jackson (2011) finds that when a greater share of adult immigrants is unskilled, native college enrollment rises; the effect on contemporaneous completed high school is mixed. Borjas (2006) finds that foreign students do not reduce native enrollment in graduate school. Llull (2010) finds in a structural model that natives increase their years of education in response to immigrants aged 16 and older. Neymotin (2009) finds native SAT scores and probability of applying to top colleges are not negatively affected by the school's share of immigrant test–takers. Gould et al. (2009) uncover negative effects of child immigrants on native examination performance in Israeli schools.

I measure the shares of immigrants in the population when natives are aged 11–17, and I measure native educational attainment at ages 21–27. Unlike Betts (1998) and Betts and Lofstrom (2000), who found a detrimental net effect of immigration on native educational attainment for each native racial and ethnic group, I find the net effect of immigration to be positive for natives generally, and especially for blacks: an increase of one percentage point in the share of immigrants in the population aged 11–64 (0.13 standard deviations) increases the probability natives complete 12 years of schooling by 0.2–0.3 percentage points, and increases the probability for blacks by 0.4–0.5 percentage points. I estimate a detrimental net effect for native–born Hispanics of -0.2 percentage points, although this is imprecisely estimated; the negative sign is attributable to the male natives. All effects are rather small compared to the average native completion rate of 87.8% (81.0% for native blacks; 81.3% for native Hispanics) and given the average immigrant share of 8.9% (8.1% for blacks; 15.5% for Hispanics).

When I distinguish immigrants by age and education, I find that child immigrants reduce native completion rates, while adult dropout immigrants increase them, with both effects larger for native–born blacks than for natives generally. A one percentage point increase in the share of immigrants in the population aged 11–17 (0.22 standard deviations) reduces the probability natives eventually complete 12 years of school by 0.3–0.5 percentage points for all races and ethnicities; a one percentage point increase in the share of immigrants with less than 12 years school in the population aged 18–64 (0.27 standard deviations) increases the eventual native completion rate by 0.7–0.9 percentage points. Effects of more educated adult immigrants are not generally statistically significant. The results for native–born Hispanics are more subtle than for other groups, as the effect of immigrant children depends strongly on the education of the immigrants' parents: child immigrants of more educated parents have beneficial effects, offsetting detrimental effects of child immigrants of poorly educated parents. These subtle results are present for male natives only.

The result that child immigrants reduce native educational attainment suggests the need for reform in immigrant education. Reform could include both increased resources for schools in areas with high immigration (Singer 2008) and the implementation of best practices regarding improving language skills of non–native speakers, remedying educational deficiencies of immigrants, and integrating immigrants with native students (García, Kleifgen and Falchi 2008). Nevertheless, some caution should be exercised in attributing the negative effects of immigration to the school quality channel, as the mechanism could be neighborhood rather than school interaction, and indeed the instrumental variables approach was not successful for the child immigrant variable. My finding that, on balance, immigration causes natives to upgrade their education sheds light on why immigration appears to have little or no negative effect on the wage levels of the native–born.[6]

1 Data and descriptive statistics

The principal data for regression analysis are the IPUMS micro–data samples for the 1940–2000 decennial censuses and the pooled 2008–2010 American Community Surveys (which I refer to as the 2010 ACS data), from which I construct a panel of states.[7] I supplement them with data from the Bureau of Economic Analysis on state personal income per capita. I choose the census and ACS data for the large sample sizes they afford for the measurement of both state immigrant shares and the shares of native cohorts by race and ethnicity attaining a given educational level. Even with the large census samples, I am forced to consider a native cohort as one spanning several birth years. A disadvantage of the data is that they do not contain information about parental education or income, except for children living with their parents.

In order to have a consistent outcome variable over all years, I define the outcome of interest as the completion of 12 years of schooling, with or without the obtention of a high school diploma, as the two may be distinguished only from 1990 onwards.[8] I focus on the native–born who were aged 11–17 in the previous census: this implies current ages 21–27

[6] See also Peri and Sparber (2009, 2011).

[7] Ruggles et al. (2010).

[8] Betts (1998) uses 12 years of schooling for 1980, and obtention of a high school diploma in 1990. This discrepancy drives the negative effect he finds of immigration on native Hispanics.

(20–26 in 2009, 19–25 in 2008). Most covariates are lagged one census, to correspond to the time when natives were aged 11–17. I construct samples of all races and ethnicities pooled, blacks, Hispanics and non–Hispanic whites. Being black and Hispanic are not mutually exclusive, so there is some overlap in the two minority samples. Immigrants are defined as those born abroad, including those born in U.S. territories. I drop the states of Alaska and Hawaii, as their absence from the 1940 and 1950 censuses complicates the use of the instruments and covariates measured in 1940.

Figure 1 depicts the shares of native–born 21–27 year olds who have completed at least 12 years of schooling, by race and ethnicity, for 1940–2010.[9] The shares increase strongly over the early decades then level off around 1990. Minorities begin the period with much lower education, and converge towards non–Hispanic white rates from 1960 until 1980 or 1990. At the start of the period, both blacks and Hispanics (concentrated in different regions) were educated in segregated, inferior schools. As a result of court decisions in the 1940s and 1950s, the Civil Rights Act of 1964 and the Coleman Report (Coleman et al. 1966), educational quality, integration and attainment increased for minorities.[10]

Heckman and LaFontaine (2010) have cautioned that both the increase in high school completion observed in the census and the convergence between whites and minorities mask an increasing share of individuals receiving a General Equivalency Degree (GED). For the purposes of this paper, it would be desirable to know whether any response in native education is coming through time in regular high school, or the propensity to obtain a GED. The measurement difficulty is that, unlike in the ACS, is it impossible to identify GED holders in the census micro–data. In early censuses, when GEDs were uncommon, no specific instructions concerning GEDs were given to the respondents. In 1980, GED recipients were instructed to respond they had completed 12 years of high school, while in 1990 and 2000 they were instructed to respond that they had a high school diploma. I therefore correct the 12–year completion rates using annual published tables on GEDs awarded by state and age, using the Heckman and LaFontaine appendix

[9] For the purposes of the graph, I use 21–27 for all years including 2008–2010.
[10] MacDonald and Monkman (2005), Valencia et al. (2002).

for methodological guidance. I scale down the published figures so as to reflect only native–born GED recipients with the help of the pooled 1999–2001 October supplement to the Current Population Survey and the 2008–2010 pooled ACSs, but the CPS samples sizes do not permit a break–down by race and ethnicity. The published GED data indicate that only a small minority (less than 5%) of GED recipients had completed 12 years of high school, so I reduce the reported number of native completers by the estimated number of native GED recipients to obtain a "true" number of native completers of 12 years of schooling. However, the GED recipients I am subtracting, while not holders of regular high school diplomas, do have the possibility of attending college, so the adjusted measure understates final educational attainment. The adjustment for GEDs becomes increasingly crude as the data get older, as explained in the GED Data Appendix, and I do not attempt to adjust 1970 and earlier years.[11] Appendix Table 1 shows the 12–year completion rates measured in different ways.

Figure 2 shows the evolution of the share of immigrants over the period, by age group. The share of immigrants in the population of working age, 18–64, traces out a U–shape, falling from 12.1% in 1940 to 6.1% in 1970, before rising to 18.4% in 2010 (top line). The share of immigrants in the school–age population, 11–17, traces a different path, rising almost monotonically from 1.6% in 1940 to a still modest 7.2% in 2010 (bottom line).

Figure 3 shows the time paths of three additional key covariates: the shares of the population aged 18–64 composed of immigrants with less than 12 years of education, exactly 12 years of education, and more than 12 years of education. The share of the lowest education immigrants falls from a high of 10.1% in 1940 to a low of 2.8% in 1980, before rising again to 4.8% in 2010. The shares of the immigrants from the two more educated groups rise monotonically from 0.7-1.1% in 1940 to 5.4% in 2010 for those with exactly 12 years of education and 8.1% for those with more than 12 years education. Appendix Table 2 gives further means of variables measured at the individual level, while Appendix Table 3 give means of variables measured at the state level.

[11] Heckman and LaFontaine report in their appendix that the adjustment for 1970 is small; furthermore, in 1970 GED holders may not have claimed to have completed 12 years of high school.

The Census Bureau produces tabulations of their census data at the school district level. The 1990 tabulations, known as the School District Database (SDDB), may be used to assess which native children are most likely to interact with immigrants in school.[12] The first four columns of Table 1 are based on samples of children of kindergarten, primary or secondary school age, from which I have discarded the small number of school districts with no high school. Panel A column 1 shows that the share of immigrants among such children is 4.2%, but that a native–born child is on average in a school district with only 3.8% immigrants, with corresponding numbers of 2.5% for white non–Hispanic natives, 5.2% for black natives, 10.7% for Hispanic natives, and 14.3% for immigrants.

However, this does not necessarily show that Hispanic natives interact more with immigrants than non–Hispanic white natives within a given state (the relevant question given that my subsequent analysis will rely on within–state variation). This pattern could emerge if Hispanic immigrants and Hispanic natives were concentrated in one region of the country and white natives in another. I therefore compute the panel A numbers for each state, calculate differences between groups for each state, and report the population–weighted average differences across states in panel B. Column 1 shows that while native blacks and native Hispanics are both more likely than native whites to be in school districts with many immigrants, black and Hispanic shares are more similar to each other than panel A indicated. On average (within state), a native black child is in a school district with 2.7 percentage points more immigrants than a native white child, while a native Hispanic child is in a school district with 3.6 percentage points more immigrants. Column 2 panel B shows that the black–Hispanic difference is larger when proximity to Hispanic immigrants is measured, but columns 3 and 4 show there is no sizeable difference in the black–Hispanic exposure to white non–Hispanic immigrants or Asian immigrants. Columns 5 and 6 are based on data on the parents of children of school age. The share of immigrants among parents is higher than among children, but generally similar patterns are found.

These statistics suggest that any effect of child immigration and probably also adult

[12] The data are available at www.nber.org/sddb/, accessed 3 April 2012.

immigration to a given state will be larger for native blacks and especially Hispanics than for non–Hispanic whites, as native minorities interact more with immigrants in their schools and neighborhoods, and probably labor markets. The implication of the high degree of contact between Hispanic natives and immigrants is unclear: immigrants may have less impact on natives similar to themselves, or they could have more impact, for example by encouraging native–born Hispanics to speak more Spanish, possibly at the expense of English, or by straining resources directed at those native–born Hispanics who have limited English proficiency.

2 Estimation

While the main analysis is conducted on a panel of states, I first adjust at the individual level for variation in age and gender structure across states and years. Specifically, for individuals aged 21–27 (20–26 in 2009 and 19–25 in 2008) at time t and born in state s, I run the following linear probability regression for each of the samples:

$$P(E_{ist} \geq 12) = \alpha_0 + \alpha_1 F_{ist} + \sum_{a=20}^{a=27} \gamma_a A_{ist}^a + \sum_s \sum_t \lambda_{st}(\delta_s \times \nu_t) + \eta_{ist}, \tag{1}$$

where i indexes individuals and s individuals' birth state, E represents years of education, F is a gender dummy, A^a are dummy variables for age, δ_s are state dummies and ν_t are year dummies. I experimented with including three dummies for whether the individual could have left school at age 14/15, 16, or 17 given his or her birth state and birth year, but their coefficients were never jointly significant, so I report results without these controls.[13][14]

I weight this regression using weights based on the census weights. The census weights sum to the U.S. population of the census year, while I wish the standard errors to reflect

[13] The coefficients on the school leaving dummies are identified despite the state–year effects, as each year has several birth cohorts. The data sources are Açemoglu and Angrist (2000) and the Digest of Education Statistics, various issues.

[14] The results are unaffected by including three dummies for type of Hispanic and a dummy for black in the regressions based on the Hispanic sample.

the variation in sample sizes from year to year. I adjust the census weights so that the ratios of their sums for each year reflect the ratio of the census sample sizes, resulting in considerably more weight being put on recent years. The average year in the weighted data is 1989 for non–Hispanic whites, 1991 for blacks and 1995 for Hispanics.[15]

In a second step, I use the coefficients $\hat{\lambda}_{st}$ as the dependent variable in a state panel analysis:

$$\hat{\lambda}_{st} = \beta_0 + \beta_1 I^{11-17}_{s,t-10} + \beta_2 I^{E<12}_{s,t-10} + \beta_3 I^{E=12}_{s,t-10} + \beta_4 I^{E>12}_{s,t-10} + \beta_5 X_{s,t-10} + \gamma_s + \nu_t + \epsilon_{st}. \qquad (2)$$

I weight the regressions with the inverse of the squared standard errors on the $\hat{\lambda}_{st}$ in the first step, and cluster the standard errors by state. $I^{11-17}_{s,t-10}$ represents the share of the population aged 11–17 that is foreign–born in the previous census, when the native–born cohort was itself aged 11–17, and is designed to capture natives' exposure to immigrant classmates. Ideally, an additional covariate would capture the presence of immigrants when natives were of elementary school age, but the ten–year spacing of the census precludes this. The immigrant covariates are affected by measurement error due to small samples, and due to being measured for the native's birth state: using birth state avoids endogenous moves of young adults that would plague the use of state of current residence, yet the native may have moved in childhood. The null hypothesis to be tested is that β_1 is negative because immigrant children reduce current school quality (more than they may be expected to increase wage inequality later).

$I^{E<12}_{t-10}$ represents the share of the population aged 18–64 when natives were aged 11–17 that was immigrants with less than 12 years of schooling, and $I^{E=12}_{t-10}$ and $I^{E>12}_{t-10}$ are defined similarly. The null hypothesis to be tested is that β_2 is positive, because the presence of immigrants with less than 12 years education increases the return to completing 12 or more years of education. The necessity of using multi–year birth cohorts is likely to bias β_2 towards zero, since the younger members of the age range 11–17 are likely to base their years of schooling decision on the wage structure, and hence immigration rates, of later

[15] The 1940, 1960 and 1970 censuses are 1% samples, the 1980–2000 censuses are 5% samples, the ACS has a more complicated sampling scheme which results in the pooled 2008–2010 sample being smaller than the 2000 sample. The 1950 census only asked education questions of a subset of the main 1% sample.

years. The signs of β_3 and β_4 are ambiguous, as the inflows of more educated immigrants have opposite effects on the return to exactly 12 years of education versus more than 12 years of education (relative to less than 12 years).

This regression suffers from endogeneity problems, however. Native high school educational attainment and high shares of low–education immigrants in a state may be spuriously negatively correlated. What makes the state economically attractive for immigrants, such as the availability of low–skill jobs, may by the same token mean that natives have a low incentive to complete 12 years of schooling. For example, a downturn in a state's low–skill industries could deter unskilled immigrants from moving to the state and encourage its natives to graduate from high school, leading β_2 to be biased down (the same direction as the measurement error bias). Similar reasoning suggests that β_3 and β_4 could be biased up by endogeneity. β_1 could be biased up if immigrants with children choose states with high educational attainment (the same direction as the measurement error bias), but there may be other biases due to endogeneity in their parents' choice of state if these have not been controlled for properly.

Due to these endogeneities, I implement an instrumental variables strategy using ten–year differences of equation (2):

$$\Delta\hat{\lambda}_{st} = \tau_0 + \beta_1 \Delta I^{11-17}_{s,t-10} + \beta_2 \Delta I^{E<12}_{s,t-10} + \beta_3 \Delta I^{E=12}_{s,t-10} + \beta_4 \Delta I^{E>12}_{s,t-10} + \beta_5 \Delta X_{s,t-10} + \nu_t + \Delta\epsilon_{st}. \quad (3)$$

I estimate this using weights $1/(1/w_{s,t} + 1/w_{s,t-10})$, where w is the weight used in equation (2). I devise instruments for the differenced immigration covariates, based on the flows of immigrants to a state that would have been expected given the 1940 geographic distribution of immigrants from different regions and the subsequent national inflows from those regions.[16] To illustrate, if immigrants from Europe prefer the northeastern United States because it is closer to home and because other Europeans are already there because of geography, and Mexican immigrants prefer the southern border states for analogous reasons, the large national increase since 1940 in the share of immigrants that are Mexican will be associated with an increase in immigration to the southern border states

[16] These instruments are similar to the instrument developed by Card (2001), and also used by Jackson (2009) and Hunt and Gauthier–Loiselle (2010).

relative to the northeast. The predicted flows captured in the instrumental variable will therefore be strongly, though not perfectly, correlated with actual immigrant flows to states. Furthermore, since the national increase in Mexican immigration appears to be the result of increasingly large birth cohorts entering the Mexican labor market[17], and the national decrease in European immigration is due to Europe's having become richer, the decrease in immigration to the Northeast relative to the border states is unrelated to non–immigration factors affecting native education choices.

I define an instrument for each of the education–specific immigration variables as follows. For a state s, the predicted change in the number of immigrants of education level E (aged 18–64), caused by changing origin regions k, can be written as

$$\Delta \hat{M}_s^E = \sum_k \frac{M_{sk}}{M_k} \Delta M_k^E = \sum_k \mu_{sk} \Delta M_k^E, \tag{4}$$

where μ_{sk} is state s's share in 1940 of the national total of immigrants who originate from region k, and ΔM_k^E is the national change in the number of immigrants with education E (aged 18–65) from that region. I use 18 source regions or countries, listed in Appendix Table 3. Because the variables to be instrumented are percentage point changes, I convert $\Delta \hat{M}_s^E$ to percentage points by dividing by the population level (aged 18–65) at the start of the period to which Δ refers, to define the final instrument as:

$$Z_s^E = \frac{\Delta \hat{M}_s^E}{POP_s}. \tag{5}$$

I deliberately base the μ_{sk} on immigrants of all educations (and ages) to emphasize the role of geography and taste and minimize the role of economic factors that might disproportionately attract workers of a specific education level. The instrument will be invalid if non–immigration shocks to high school completion are correlated with 1940 immigrant densities; for example, if improvements to the California and Texas school systems caused a national–level increase in Mexican immigration.[18] By defining an instrument for each education level, I assume that improvements to the California and Texas school systems did not cause a national–level increase in Mexican immigration of any education group.

[17] Hanson and McIntosh (2007).

[18] See Beaudry et al. 2010 for a formal treatment.

It is easy to construct instruments for different immigrant age groups, in particular for ΔI^{11-17}, by replacing the education–specific variables in equations (4) and (5) with age–specific variables, and I do so. However, the intuition of the instruments extends less easily to subdivisions by age group, as immigrant numbers in an age group are strongly influenced by aging as well as immigration, and furthermore, for the 11–17 age group in particular, it is obvious that changes in their inflows will be closely linked to those of adult immigrants.

I choose to use a common first stage for all four race/ethnicity samples, weighting each first stage with the denominator of its dependent variable (the population 11–17 or the population 18–64) in order to improve efficiency. The instruments have weak predictive power when the first stages are weighted by the second stage black or Hispanic weights (essentially the black or Hispanic population aged 21–27). This approach also means the first stage always includes all states and years: some early state–year cells have no native–born blacks or Hispanics aged 21–27.

Immigrant students could have either positive or negative spillovers on their native classmates, depending upon the quantity and quality of their prior education, their English skills, their industriousness, and the extent to which their parents contribute to their education. Parental education is likely to be a proxy for some of these characteristics, and because most children aged 11–17 live with their parents, we can observe their parents' education in the census data. It is therefore possible to split the share of the population 11–17 that is immigrant into immigrants whose parent or parents in the household have less than 12 years schooling, immigrants with at least one parent with 12 years or more, and immigrants with neither parent living in the household. The expectation is that children of more educated parents will make better peers and require fewer resources than children of less educated parents. I discuss possible biases on the coefficients in the results section below. The difficulty with regressions distinguishing child immigrants according to parental characteristics is that the number of endogenous variables becomes too large for the use of 2SLS.[19]

[19] I have experimented with using the (OLS) state–year return to completing 12 years schooling as

Although the natural level at which to examine school quality is the school, there are some reasons to use more aggregated data beyond the limited availability of school–level data. If some natives move out of their school and neighborhood when immigrants move in, analysis at the school or school–district level will not attribute any change in schooling of the native movers to the arrival of the immigrants. If public school data are used, even natives who move to private schools in the same school district will cause the same problem. Also, it is difficult to find an instrument at the school level which accounts for immigrants' potentially endogenous choice of location (and school).

3 Results

I examine the impact of immigration on the probability of natives' completing 12 years of schooling, for all native–born, non–Hispanic whites, blacks and Hispanics, first assessing the net impact of immigration, then decomposing the impact into school quality versus labor market channels, distinguishing child immigrants by parental education, checking the robustness of the results to the treatment of GED holders and finally distinguishing natives by gender.

3.1 The net effect of immigrants ages 11–64

In Table 2, the first six columns show the impact of the immigrant share of the population aged 11–64 using the fixed effects specification of equation (2) and increasing numbers of covariates, column 7 shows the impact using the ten–year differenced specification of equation (3), and column 8 shows the results of applying two stage least squares (2SLS) to the column 7 specification. The first row shows that the principal coefficient of interest is positive, and becomes larger and statistically significant as more covariates are added, with a magnitude of 0.20 in column 5. The 2SLS coefficient is a statistically significant 0.30; the F–statistic associated with the instrument in the first stage is 25. A coefficient of 0.3

the independent variable of interest, instrumenting it with actual or predicted immigrant flows. The coefficient on the return is always very imprecisely estimated, and the first–stage immigrant coefficients are often wrongly signed.

implies that a one percentage point increase in the share of immigrants in the population 11–64 increases the native probability of eventually completing 12 years of education by 0.3 percentage points. This is not a large effect considering that the (weighted) mean completion rate is 87.8%, and the share of immigrants in the population 8.9%.

Following existing literature on the determinants of high school completion, the first controls I add in Table 1 column 2 are for the unemployment rate and the cohort size, measured in the previous census when native respondents were aged 11–17. Neither the youth unemployment rate nor the prime–age unemployment rate is statistically significant, possibly because many of the respondents were some years from graduation when the unemployment rate was measured.[20] Consistent with Card and Lemieux (2001), I find that members of larger cohorts have statistically significantly lower educational attainment – I define the cohort size based on natives only, to avoid endogeneity, as the share of the native population which is aged 11–17. However, when I control for region–specific trends in column 3 (dummies for eight BEA regions interacted with a trend), the coefficient on cohort size falls greatly and becomes insignificant.

In column 4, I control for the share of workers 18–64 who were in agriculture in 1940, interacted with a trend. This covariate is included for its effect on minorities, rather than non–Hispanic whites, and its coefficient is insignificant for all natives. In column 5, I control for the 1940 share of non–Hispanic whites aged 21–27 who had less than 12 years education, interacted with a trend, which captures convergence among states and is statistically significantly positive. In column 6, I take into account that one would expect richer states to be able to afford better educational systems. However, states with better educational systems should become richer, so the coefficient on a control for state income would be biased up. The results show that the correlation between log state personal income per capita and completion of 12 years of schooling is indeed positive and statistically significant, and that the coefficient on the immigrant variable, now a lower bound on the true coefficient, is reduced by one third to 0.13. In the differenced

[20] State–level unemployment rates are not available from other sources for earlier decades, so the unemployment rate cannot be matched to the year the respondent was aged 17, for example.

specifications of columns 7 and 8, I control for all covariates except state personal income.

In Table 3, I analyze natives by race and ethnicity, using the same specifications as Table 1, reporting only the coefficient on the immigrant covariate. For reference, I reproduce the coefficients from the first row of Table 1, for all natives, in the first row of Table 3. The coefficients for non–Hispanic whites, in the second row, are always slightly smaller than those for all natives; the coefficient from 2SLS, in column 8, is a statistically significant 0.25. For blacks in the third row, once state–specific trends and convergence are controlled for in columns 4 and 5, there is a robust positive coefficient of 0.38–0.46, compared to a mean completion rate of 81.0% and a black–white completion gap of 8.5 percentage points. For Hispanics, in the fourth row, an effect of 0.38–0.53 is robust in the preferred least squares columns, but it disappears (coefficient of -0.18) with 2SLS in column 8. I conclude that the net effect of immigration on native completion of 12 years of schooling is positive for natives generally, blacks and non–Hispanic whites, and is imprecisely estimated, but possibly negative, for Hispanics.[21]

3.2 Decomposing the impact of immigration into school quality and labor market channels

I now turn to decomposing the impact of immigration into school quality and labor market channels. I return in Table 4 to the sample of all natives, presenting the same specifications as in Tables 1 and 3, except with immigrants split into four categories. The first row shows that the effect on natives of immigrants aged 11–17, likely to have been natives' classmates, is negative: once trends are controlled for, the coefficient is in the range -0.28–-0.46, statistically significant except for the 2SLS coefficient of -0.28 in column 8. The first stage information for the instrumenting of this variable is presented in Table 5, column 2 (in column 1, I present the first stage used for immigrants 11–64 in the tables above).

[21] For blacks and Hispanics, there is a large positive coefficient on the share of agricultural workers in 1940, which captures convergence among states: agricultural states in 1940 had large shares of either blacks or Hispanics in the population, who were poorly educated. I do not control for the educational attainment of blacks and Hispanics in 1940, as they are based on very small samples for many states. White educational attainment in 1940 is statistically insignificant in the regressions for minorities.

Although the predicted share of the population 11–17 has a statistically significant coefficient, it is not more significant than those of the other excluded instruments, suggesting that I have not managed successfully to instrument the share of immigrants in the population 11–17. The Angrist and Pischke (2008) F–statistic for the excluded instruments (an F–statistic adapted for multiple endogenous variables) suggests the same. The bias on the least squares coefficient cannot be confidently signed, but is probably towards zero (see above), so I conclude that the true coefficient is -0.3 or more negative, consistent with the hypothesis that having immigrant classmates reduces high school attainment.

The second row of Table 4 shows that the effect on natives' acquiring 12 years of schooling of immigrants aged 18–64 with less than 12 years of schooling is positive and statistically significant in every specification. The coefficient is in the range 0.73–0.98. This is consistent with the hypothesis that the presence of unskilled immigrants in the labor market alters the wage structure in such a way as to give natives an incentive to complete 12 years of schooling. A comparison of columns 7 and 8 shows that using 2SLS does not increase the coefficient greatly, despite the expectation it would be biased down in least squares. Yet column 3 of Table 5 shows that the predicted share of immigrants with less than 12 years' schooling in the population 18–64 is a strong predictor in the first stage, much stronger than the other excluded instruments, and the Angrist–Pischke F–statistic is very high.

The results of rows three and four of Table 4 indicate that the effects of adult immigrants with exactly 12 years of schooling and of adult immigrants with more than 12 years of schooling are sensitive to the specification and estimation method. The instruments associated with these covariates are fairly strong in their respective first stages (Table 5 columns 4 and 5).

I repeat the exercise of Table 4 for native non–Hispanic whites in Table 6. The least squares results are qualitatively similar to those for all natives, but the absolute values of the coefficients are smaller. The coefficient on immigrants aged 11–17 is in the range -0.16– -0.30 in the preferred specifications of columns 4–8. For this sample, however, 2SLS reduces the point estimate to essentially zero (c.f. columns 7 and 8), suggesting there is no negative effect on native non–Hispanic whites through the schooling channel (or that it is cancelled out by natives' anticipation of their immigrant classmates' future labor market effect). Nevertheless, the WLS and 2SLS coefficients are not statistically significantly different. The coefficient on the adult immigrant dropout covariate is in the range 0.45–0.60. The coefficients on more educated adult immigrants are qualitatively similar to those in Table 4.

In Table 7, I turn to native–born blacks. For this sample, the negative effect of immigrants 11–17 is larger in absolute value than for non–Hispanic whites, and robustly statistically significant except for 2SLS (column 8), ranging from -0.34– -0.57 in the preferred columns 4–8. The positive effect of adult immigrants with less than 12 years of education is also robust and larger than for non–Hispanic whites, ranging from 0.86–1.11 in columns 4–8. The effects of adult immigrants with exactly 12 years education and with more than 12 years are imprecisely estimated. I conclude that, as expected, the school quality and labor market channels through which immigration might operate are present for blacks and are larger than for non–Hispanic whites.

Finally, I examine native–born Hispanics in Table 8. While the sign of the coefficient on immigrants aged 11–17 is always negative, it is imprecisely measured in many specifications (first row). The evidence of a beneficial effect of adult immigrants with less than 12 years of education is not robust, since in differenced specifications the sign is negative. Another difference from results for other native groups is that immigrants adults with more than 12 years of education have relatively large positive effects. The results for Hispanics therefore appear to provide only weak support for the hypotheses being tested. In the next section, however, I show that this is because the Table 8 regressions are misspecified.

3.3 Distinguishing child immigrants according to parental characteristics

Distinguishing among immigrants aged 11–17 according to their parents' education proves helpful for interpreting the effects of immigration on native–born Hispanics, and these results are presented in columns 2 and 3 of Table 9. Column 1 repeats the preferred fixed effects specification of Table 8 (column 5). As anticipated, the coefficient on child immigrants with parents with less than 12 years of education is negative and statistically significant, with coefficients of -2.0 and -1.8 in columns 2 and 3. The coefficient on child immigrants with more educated parents is positive and statistically significant, with a relatively large coefficient of 3.4 in column 2 (2.2 in column 3).

The magnitude of the two coefficients implies a one percentage point increase in the relevant immigrant share changes the 12 year completion rate by 2–3 percentage points. This compares to shares of immigrants in the 11–17 population of 3.3% with parents with less than 12 years schooling and 4.7% for those with a parent with at least 12 years schooling, and a Hispanic completion rate of 81.3%.[22] These numbers imply moderate elasticities of -0.08 for children of unskilled parents and 0.13–0.20 for children of more educated parents.

It appears that in the specification with the undifferentiated child immigrant variable, the unskilled adult immigrant coefficient was capturing the opposing effects of the unskilled adult immigrants and their children: the coefficient on adult immigrants with less than 12 years of education rises from 0.35 in column 1 to a statistically significant 1.65 in column 2, in accordance with both theory and the results for native–born blacks. The coefficient is smaller in the differenced specification of column 3 (0.73), but statistically significant, and very different from the corresponding coefficient of -0.39 in Table 8 column 7. At the same time, the effect of child immigrants with more educated parents was being picked up by the educated adult immigrant variable in the simpler specification: the coefficient on adult immigrants with more than 12 years of education falls from the

[22] The shares of immigrants in the population 11–17 are computed using Hispanic weights, and are hence higher than implied in Figure 2. See Appendix Table 2.

anomalously high 2.27 in column 1 to 0–0.46 (and statistically insignificant) in columns 2 and 3, similar to the values for non–Hispanic native groups.

The coefficient on the share of the 11–17 year–old population that is immigrants with no parent in the household is strongly negative in both columns. There are several possible explanations for this. The first is causal: such immigrants, who tend to have very low enrollment rates, drop out early from school and have very negative peer effects in the short time they are in school in the United States. If this is so, one might expect a similarly negative impact of children of the lowest educated immigrants. However, unreported regressions splitting parental dropouts into those with more or less than nine years of education do not point to clear differences between the two groups. Two other explanations are endogenous, and imply a negative bias on the coefficient. To the extent that the young immigrants move to the United States without their parents in order to work (Oropesa and Landale 2009), they are the immigrants most likely to move to states with (imperfectly controlled) favorable labor market conditions for youth, conditions which would incite the closely substitutable native students to drop out. To the extent that some young immigrants move to the United States with their parents and initially attend school, (imperfectly controlled) favorable youth labor market conditions are likely to entice both immigrants and natives to drop out of school to work, and hence also to leave home.

Since it is not possible to instrument all the immigrant variables in the expanded specifications of columns 2 and 3, it is not possible to make statements about causal effects. Nevertheless, Table 9 suggests that for Hispanic natives, just as for other natives, adult immigrants with less than 12 years schooling have a beneficial effect on completing 12 years of schooling, and that for Hispanic natives, classmates with educated parents have beneficial effects, while classmates with poorly educated parents have detrimental effects. It is likely that the large effect of child immigrants with no parent in the household reflects endogeneity.

In Table 10, I present the corresponding regressions for native non–Hispanic whites (columns 1–3) and native–born blacks (columns 4–6). For these natives, there is no indication that grouping all child immigrants introduced misspecification, as the coefficients on the adult immigrant variables do not change much when the child immigrant variable is split into components. Neither is there a clear distinction between the effects of child immigrants with lower and higher education parents, and indeed, these coefficients are all statistically insignificant.

The coefficient on child immigrants with no parent in the household is not nearly as negative for non–Hispanic whites in Table 10 as for native Hispanics, and is positive for blacks. Nevertheless, for non–Hispanic whites, it is responsible for the statistically significant negative sign when all child immigrants are grouped in column 1 or in the differenced equivalent (as an unreported regressions show). Also, if the Hispanic weights are used for blacks and non–Hispanic whites, the coefficient on child immigrants with no parent in the household becomes much more negative (other coefficients are less sensitive), though less negative than for Hispanics. In recent decades, which are most heavily weighted in the regressions, such child immigrants are very disproportionately Mexican–born, which may explain why their "effect", whether causal or endogenous, is largest when Hispanic states are weighted more or for native–born Hispanics: within a state, Hispanic immigrants are more concentrated in school districts and possibly labor markets with many native–born Hispanics, and they may be closer labor market substitutes for young native–born Hispanics than for other young natives. For blacks and non–Hispanic whites, 2SLS in Tables 6 and 7 should address the endogenous behavior of these immigrants.

3.4 Robustness checks and further results

For the sample of all native–born, I can test the sensitivity of the results to classifying GED holders as having less than 12 years of education. There is no apparent way to do this for the 12–year completion rate adjusted for sex and age, so I use the unadjusted 12–year completion rate as the basis for the state panel analysis. In Table 11, I present results from regressions with a single immigration covariate (upper panel), and from regressions with immigration split by age and education (lower panel). The fixed effects results in columns 1 and 2 show that the failure to adjust for sex and age has no effect in the upper panel, but in the lower panel reduces the absolute value of the coefficients, rendering the coefficient on child immigrants statistically insignificant. Reclassifying the GED holders in column 3 increases the immigration impact from 0.20 to 0.31 in the upper panel; in the lower panel it generally increases slightly both the absolute value of the coefficients and the standard errors, leaving the results qualitatively the same. The use of the differenced specification in column 4 returns the upper panel coefficient to 0.2, and in the lower panel makes the coefficient on child immigrants statistically significantly negative once more. 2SLS in column 5 increases the coefficient in the upper panel to 0.52, larger than the coefficient of 0.30 in Tables 2 and 3 column 8, though not statistically significantly so. In the lower panel, 2SLS flips the sign of the insignificant coefficient on child immigrants and raises the point estimate on adult dropout immigrants to 0.97, a similar magnitude to the 0.79 estimated in Table 4. The coefficients on the two other immigrant variables become large in absolute value. The comparison of Table 11 with earlier tables indicates that natives adjust do schooling through changes in years of regular high school, and schooling adjustment is not driven by changes in the propensity to obtain a GED.

I have also experimented with adding further covariates to the regressions with the original dependent variable. The share of native children 11–17 who are second generation immigrants, distinguished by parental education, and the share of third generation children 11–17 whose parents are high school dropouts yield some interesting coefficients, but do not affect the main results of interest. The student–teacher ratio for all grades in public schools always has a statistically insignificant coefficient (the sources are the Digest of Education Statistics and the Biennial Survey of Education, various years). Finally, matching state characteristics to natives' state of current residence, rather than state of birth, does not change the qualitative picture. For all races and ethnicities together, coefficients are smaller in absolute value, while for minorities there is no clear pattern.

As a check on the effects of immigrant children distinguished by the education of the parent, I can instead distinguish immigrant children by their birth region (these results are not reported; distinguishing by birth region and parental education together causes all relevant coefficients to become insignificant). Child immigrants from Latin America have a stastistically significant negative effect on natives generally (coefficient -0.49), and this effect is statistically significant at the 10% level for each of the three native groups considered. This is consistent with a negative effect of child immigrants from low SES or low education families. The effect of Asian children is negative and significiant at the 10% level for all natives (coefficient -0.56) and insignificant for native groups individually. The effect of European children is small and insignificant for all natives, but has a relatively large significant positive effect for native–born Hispanics (coefficient 1.75), whereas the coefficient for blacks is negative (and is zero for non–Hispanic whites). This is consistent with native–born Hispanics responding positively to high–SES or high–education schoolmates or neighborhood companions.

Native males and native females could be differently affected by immigration, either due to differing degrees of labor market substitutability with immigrants, or due to differing performance in school, or different types or degrees of interaction with immigrants in school (or the neighborhood). I have therefore repeated all the analysis distinguishing natives by gender, returning to the original definition of the dependent variable. I present two panels of results in Table 12, using a single immigration covariate in the upper panel, and distinguishing among immigrants in the lower panel. The analysis for native Hispanics shows that the subtle effects of immigrants age 11–17 come entirely from the effects on males (lower panel, columns 1 and 2). Coefficients for native Hispanic females are very imprecisely estimated. The table also shows that the negative sign of the net effect of immigration on native Hispanics in the 2SLS comes from native males, though standard errors remain large (upper panel, columns 3 and 4).

The results for native–born blacks (columns 5–8) show coefficients for native–born black females are also very imprecisely estimated, indicating that the statistical significance of earlier results was driven by males, even if the general pattern of results is similar for the two genders. The analysis for all natives does not reveal any noteworthy differences, and while the point estimates for non–Hispanic whites hint the negative effect of immigrants age 11–17 might be due to males, the results on the whole are similar for males and females (see Appendix Table 5).

4 Conclusion

In this paper, I have shown that natives' probability of completing 12 years of education is increased by immigration, albeit by a small magnitude, an effect that is larger for blacks than non–Hispanic whites. This positive net effect is composed in part of the additional incentive to complete 12 years provided by the presence of unskilled adult immigrants in the labor market, and of the negative effect of the presence of immigrant children, possibly classmates or neighborhood companions. Both of these effects are larger for native–born blacks than non–Hispanic whites, as would be expected given the lower high school completion rates of blacks and their residence in school districts with more immigrants. The results are robust to considering native–born holders of GEDs to have completed less than 12 years of education.

For native–born Hispanics, by contrast, 2SLS results suggest a small negative effect, though it is imprecisely estimated. Estimates by gender, though also imprecisely estimated, indicate that the negative sign is due to the males in the native sample. The components of the net effect are more subtle than for other native groups. While the presence of unskilled adult immigrants provides an incentive for high school completion similar in magnitude to that for native–born blacks, the effect of child immigrants is much larger, and strongly dependent on the education of the child immigrants' parents: children of parents with less than 12 years of education have a deleterious impact on native completion rates, while children with a parent with 12 or more years of education have a beneficial effect. These subtle effects show themselves among native males only, and not among females. These results are subject to the caveat that 2SLS could not be employed in the necessary more complex specification.

Hispanics are the native group living in school districts with the most immigrants, and black and Hispanic males have the lowest native graduation rates, so it is not surprising that native Hispanic males are most sensitive to child immigration. The relatively large negative effect on native Hispanic males of child immigrants of poorly educated parents may be an indicator that native students are most affected in school when exposed to culturally similar immigrants. However, it is not entirely clear why black males, for example, do not also increase their 12–year completion rates in response to child immigrants of educated parents. The fact that native Hispanics increase their 12–year completion rates in response to European and Canadian (and Australian and New Zealand) child immigrants, while blacks do not, suggests that the explanation can be at most in part related to similarity between native and immigrant Hispanics, and at most in part due to differential exposure to immigrant children of higher SES, since native blacks and Hispanics have very similar exposure to white non–Hispanic immigrant children.

The results suggest the need for reform in accommodating immigrant students, particularly those with less educated parents and in schools with many native–born minorities, and help explain how natives respond to immigration in such as way as to obviate the potential negative impact on native wages.

References

Açemoglu Daron and Joshua Angrist. 2000. "How Large are Human–Capital Externalities? Evidence from Compulsory Schooling Laws". *NBER Macroeconomics Annual*, 15 pp. 9–74.

Angrist, Joshua D. and Jörn–Steffen Pischke. 2008. *Mostly harmless econometrics: an empiricist's companion*, Princeton University Press.

Beaudry, Paul, David A. Green and Benjamin Sand. 2010. "Does Industrial Composition Matter for Wages? An Empirical Evaluation Based on Search and Bargaining Theory". UBC working paper.

Betts, Julian R. 1998. "Educational Crowding Out: Do Immigrants Affect the Educational Attainment of American Minorities?". In Daniel S. Hamermesh and Frank D. Bean (Eds.), *Help or Hindrance? The Economic Implications of Immigration for African-Americans*, New York: Russell Sage Foundation.

Betts, Julian R. and Robert Fairlie. 2003. "Does Immigration Induce 'Native Flight' From Public Schools into Private Schools?", *Journal of Public Economics*, 87 (5–6) pp.987–1012.

Betts, Julian R. and Magnus Lofstrom. 2000. "The Educational Attainment of Immigrants: Trends and Implications". In George J. Borjas ed. *Issues in the Economics of Immigration*, Chicago: University of Chicago Press.

Borjas, George J. 2007. "Do Foreign Students Crowd Out Native Students from Graduate Programs?". In Ronald G. Ehrenberg and Paula E. Stephan eds. *Science and the University*, Madison: University of Wisconsin Press.

Borjas, George J. and Lawrence F. Katz. 2007. "The Evolution of the Mexican–Born Workforce in the United States". In George J. Borjas ed. *Mexican Immigration to the United States*, Chicago: University of Chicago Press.

Card, David. 2001. "Immigrant Inflows, Native Outflows and the Local Labor Market Impacts of Higher Immigration". *Journal of Labor Economics*, 19 (1) pp.22–64.

Card, David and Thomas Lemieux. 2001. "Dropout and Enrollment Trends in the Post War Period: What Went Wrong in the 1970s?". In Jonathan Gruber ed. *An Economic Analysis of Risky Behavior Among Youth*, Chicago: University of Chicago Press.

Cascio, Elizabeth and Ethan Lewis. Forthcoming. "Cracks in the Melting Pot: Immigration, School Choice and Segregation". *American Economic Journal: Economic Policy.*

Checchi, Daniele, Andrea Ichino and Aldo Rustichini. 1999. "More equal but less mobile? Education financing and intergenerational mobility in Italy and the US". *Journal of Public Economics*, 74 pp.351–393.

Coleman, J.S., E.Q. Campbell, C.J. Hobson et al. 1966. *Equality of Educational Opportunity*, Washington D.C.: Office of Education, U.S. Department of Health, Education and Welfare.

Corak, Miles. 2006. “Do Poor Children Become Poor Adults? Lessons for Public Policy from a Cross Country Comparison of Generational Earnings Mobility”. *Research on Economic Inequality*, Vol. 13.

Cragg, J. G., and Donald, S. G. 1993. “Testing Identifiability and Specification in Instrumental Variable Models, *Econometric Theory*, 9, 222240.

Fix, Michael and Wendy Zimmerman. 1993. *Educating Immigrant Children: Chapter I in the Changing City*, Washington: Urban Institute Press.

García, Ofelia, Jo Anne Kleifgen and Lorraine Falachi. 2008. “From English Language Learners to Emergent Bilinguals”. *Equity Matters: Research Review No. I*, New York, NY: Teachers College, Columbia University. www.tc.edu/i/a/document/6468_Ofelia_ELL_Final.pdf, accessed Sept 28, 2009.

Gould, Eric D., Victor Lavy and M. Daniele Paserman. 2009. “Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence”. *Economic Journal*, 119 (540) pp. 1243–1269.

Hanson, Gordon H. and Craig McIntosh. 2007. “The Great Mexican Emigration”. National Bureau of Economic Research Working Paper 13675.

Heckman, James J. and Paul A. LaFontaine. 2010. “The American High School Graduation Rate: Trends and Levels”. *Review of Economics and Statistics* 92 (2) pp. 244–262.

Hoxby, Caroline M. 1998. “Do Immigrants Crowd Disadvantaged American Natives out of Higher Education?”. In Daniel S. Hamermesh and Frank D. Bean (Eds.), *Help or Hindrance? The Economic Implications of Immigration for African-Americans*, New York: Russell Sage Foundation.

Hunt, Jennifer and Marjolaine Gauthier–Loiselle. 2010. “How Much Does Immigration Boost Innovation?”, *American Economic Journal: Macroeconomics*, 2 (2) pp. 31–56.

Jackson, Osborne. 2011. “Does Immigration Crowd Natives Into or Out of Higher Education?”. Northeastern University working paper.

Llull, Joan. 2010. “Immigration, Wages and Education: A Labor Market Equilibrium Structural Model”. CEMFI working paper.

MacDonald, Victoria–María and Karen Monkman. 2005. “Setting the Context: Historical Perspectives on Latino/a Education”. In Pedro Pedraza and Melissa Rivera eds. *Latino Education: An Agenda for Community Action Research*, Mahwah, N.J.: Lawrence Erlbaum Associates.

Neymotin, Florence. 2009. "Immigration and its Effects on the College–going Outcomes of Natives". *Economics of Education Review*, 28 pp. 538–550.

Noguera, Pedro. 2008. *The Trouble With Black Boys: And Other Reflections on Race, Equity and the Future of Public Education.* Hoboken, N.J.: John Wiley.

Noguera, Pedro, Aída Hurtado and Edward Fergus, eds. *Invisible No More: Understanding the Disenfranchisement of Latino Men and Boys.* New York, N.Y.: Routledge.

Orfield, Gary, Daniel Losn, Johanna Wald and Christopher B. Swanson. 2004. *Losing Our Future: How Minority Youth are Being Left Behind by the Graduation Rate Crisis.* Cambridge, M.A.: The Civil Rights Project at Harvard University. Contributors: Advocates for Children of New York, The Civil Society Institute.

Oropesa, R.S. and Nancy S. Landale. 2009. "Why Do Immigrant Youths Who Never Enroll in U.S. Schools Matter? School Enrollment Among Mexicans and Non–Hispanic Whites". *Sociology of Education*, 82 pp.240–266.

Peri, Giovanni and Gianmarco Ottaviano. Forthcoming. "Rethinking the Effects of Immigration on Wages". *Journal of the European Economic Association.*

Peri, Giovanni and Chad Sparber. 2011. "Highly–Educated Immigrants and Native Occupational Choice". *Industrial Relations*, 50 (3) pp. 385–411.

Peri, Giovanni and Chad Sparber. 2009. "Task Specialization, Immigration and Wages". *American Economic Journal: Applied Economics*, 1 (3) pp.135–169.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder and Matthew Sobek. 2010. *Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database].* Minneapolis: University of Minnesota.

Singer, Audrey. 2008. "Reforming U.S. Immigration Policy: Opening New Pathways to Integration". *Opportunity 08: A Project of the Brookings Institution.* www.brookings.edu/papers/2007/0228demographics_singer_Opp08.aspx, accessed 29 March 2012.

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

Stock, James H. and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression". In James H. Stock and Donald W.K. Andrews eds. *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, Cambridge University Press.

U.S. Department of Education, National Center for Education Statistics. *Digest of Education Statistics*, various issues 1965–2010.

U.S. Department of the Interior, Bureau of Education Bulletin. *Biennial Survey of Education*, various issues 1939–1958.

Valencia, Richard R., Martha Menchaca and Rubén Donato. 2002. "Segregation, desegregation, and integration of Chicano students: old and new realities". In Richard R. Valencia ed. *Chicano School Failure and Success: Past, Present and Future*, London: Routledge.

GED Data Appendix

Official GED Data

The basic data on GEDs are taken from the annual GED Statistical Reports ("Who Took the GED?", available at www.acenet.edu/Content/NavigationMenu/ged/pubs/GED_Archived_Annual_.htm, accessed 23 March 2012). From 1989 onwards, statistics are available for the age distribution by state of GEDs awarded. A few missing and implausible values are filled in with linear interpolation, and for a number of states that began reporting the statistics only later, linear extrapolation back to 1989 is used. From 1974 to 1988, the age distribution is only available for those tested, rather than those who were awarded a GED. For these years, I compute for each age group the pass rate in the most recent year for which both the distribution of test–takers and the distribution of credentials awarded are available (1989 for most states), and use the pass rates to adjust the distribution of test–takers in 1974–1988 to obtain the age distribution of GEDs awarded. New York begins reporting information by age only from 1984, and I impute 1974–1983 values by extrapolating the 1984–1988 trends. Before 1974, no information is provided by age and state, and for 1969–1973, I assume the age distribution is the same as in 1974.

The age categories are not the same in all years, and the age categories for test–takers are not always the same as those for GEDs awarded. When the age categories change (in general they become finer over time), I split the coarser categories based the distribution in the nearest year with finer categories. Even the finest divisions do not give the age distribution for each year of age (above age 19), so after I harmonize the data at the finest categorization, I assume that GEDs are uniformly distributed within an age group. I assume that recipients 16 or younger are age 16. The final age distributions are combined with statistics for total GEDs awarded by state and year to obtain GEDs awarded by state and year for each year of age.

I use these numbers to calculate how many 21–27 years olds held GEDs in 1980, 1990, and 2000, assuming that by the census date, one third of the year's GEDs had been received. For "2010", I average the age 19–25 stock in 2008, the age 20–26 stock in 2009, and the age 21–27 stock in 2010.

Birthplace of GED recipients

The official GED statistics give no breakdown by birthplace, a breakdown necessary since I seek statistics for natives only. The October supplement to the CPS allows those whose highest degree is a GED to be identified, and from 1994 onwards, the basic monthly CPSs contain birthplace. The ACS contains both the GED and birthplace information. I therefore use the pooled 1999–2001 October CPSs and the pooled 2008–2010 ACSs to estimate the immigrant share among those who completed a GED but had no college. The sample sizes are too small, however, to do this for each state (and too small to allow the calculation of the shares of minorities, which I do not attempt). I instead distinguish three groups of states in the CPS and six groups in the ACS, based on the share of immigrants in the population aged 21–27.

Combining the data sources

For 2000 and 2010, it is straightforward to adjust the total GED stocks calculated from official GED data with the share of immigrants from the census and ACS data to reflect natives only. For 1980 and 1990, I assume that the share of immigrants among GED holders aged 21–27 changes proportionately to immigrants' share in the 21–27 population. I calculate the latter shares for the three state groups using the 1980 and 1990 censuses.

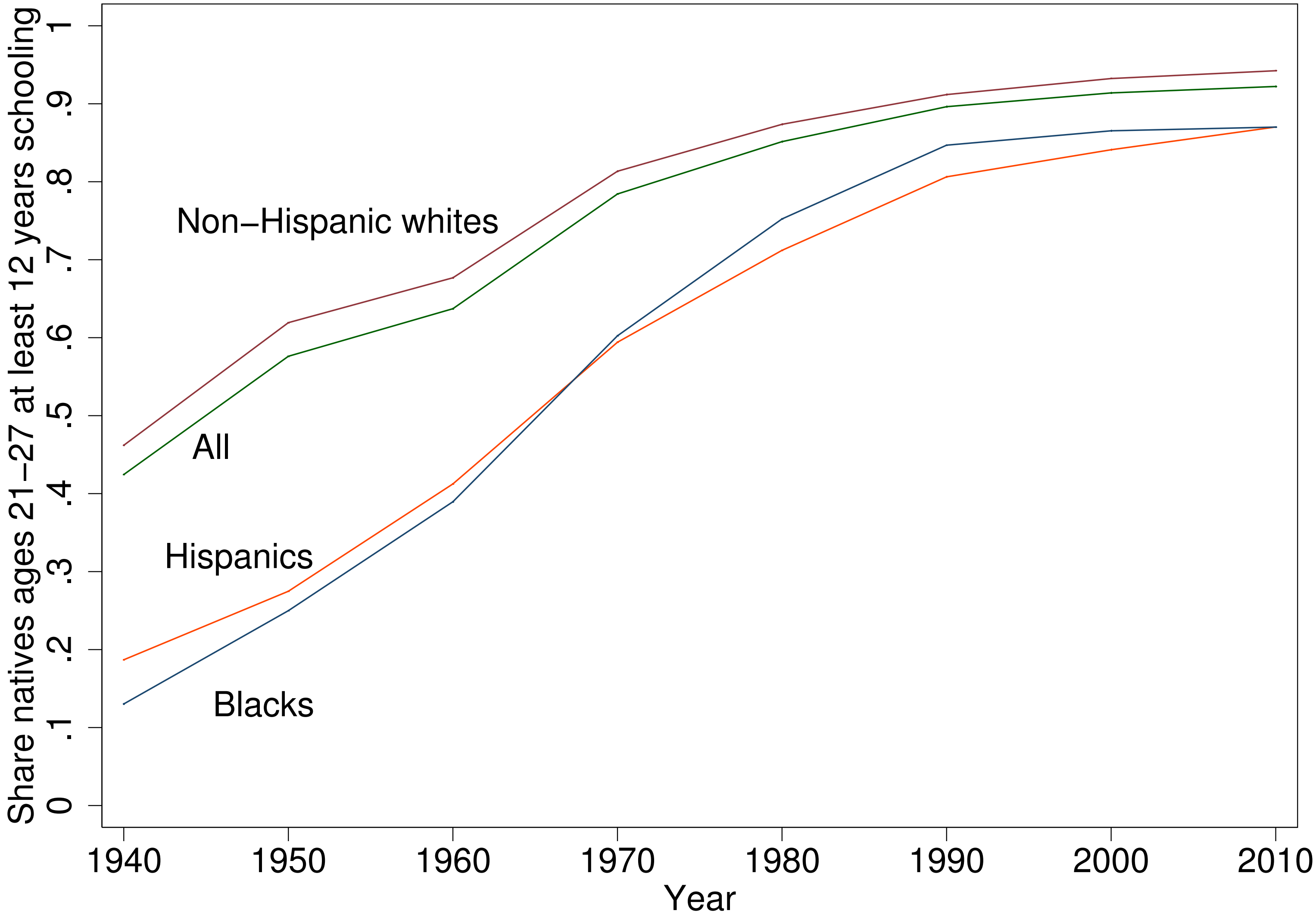
I then calculate the share of the native population aged 21–27 that holds a GED by dividing the GED stocks of natives 21–27 by the native population aged 21–27 (or the equivalent for 2008–2010), the latter computed by summing the census weights for natives, including those whose native status is imputed. I subtract this fraction from the fraction of natives with 12 years of schooling or more, which had been calculated based on a sample with non–missing (and non–imputed) education and birthplace.

GED–adjusted statistics

The first row of Appendix Table 1 shows the national raw native 12–year completion rates, while the second row shows the rates adjusted to count GED holders as non–completers. The adjustment reduces the completion rate by five percentage points in 1980, and by more in later years. Nevertheless, the adjusted completion rate is estimated to have increased between 1980 and 2010, with stagnation between 1990 and 2000. The third row provides the 12–year completion rate as directly measured in the 2010 ACS. This rate of 87.8% lies between the rates of the first two rows (84.9% adjusted, 92.3% unadjusted), indicating that too many GEDs have been subtracted in the adjustment. One reason for this would be that some of the GED holders subtracted in fact went on to obtain more education, and do not appear as GED holders in the 2010 ACS. In panel B, I consider high school graduation rates, which must be lower than 12-year completion rates. I measure this rate directly in the pooled 2008–2010 ACSs as 86.3%.

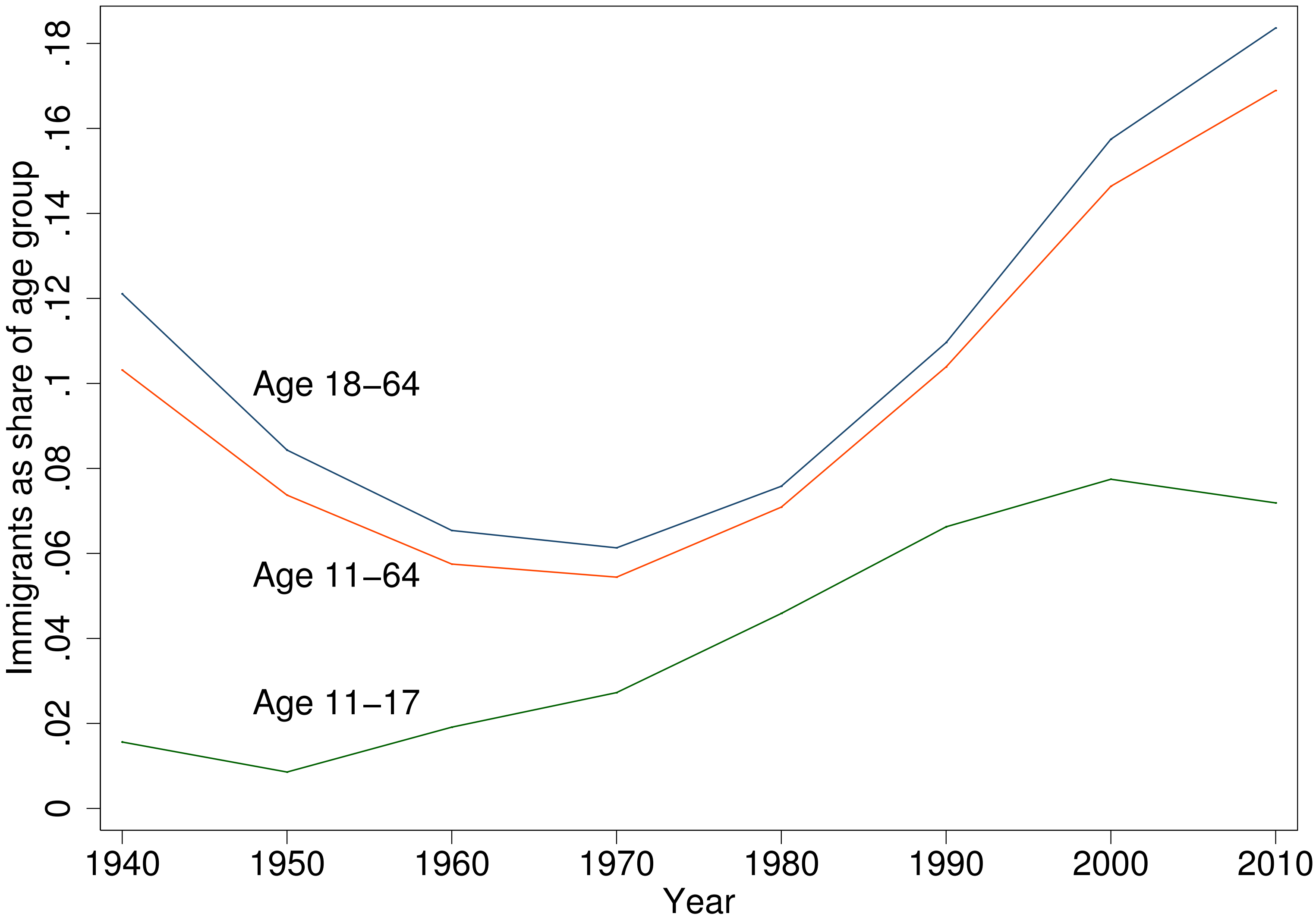
In panel C, I present for comparison the Heckman and LaFontaine (2010) high school graduation estimates for 1970–2000. They use age groups that do not quite match mine, and remove only recent immigrants from their GED counts. I assume that their raw numbers for 1970 and 1980 include 12–year completers, since they cannot be distinguished from high school graduates, and that their raw numbers from 1990 onwards exclude them. This would account for the lack of progress between 1980 and 1990 in their graduation rates, in constrast to my adjusted 12–year completion rates in panel A. Our 1980 rates should be similar, which they are. The graduation rate for 2010 measured directly (panel B) appears high compared to the estimated Heckman and LaFontaine rates for 2000.

Figure 1: Share of natives with at least 12 years of schooling, by race and ethnicity



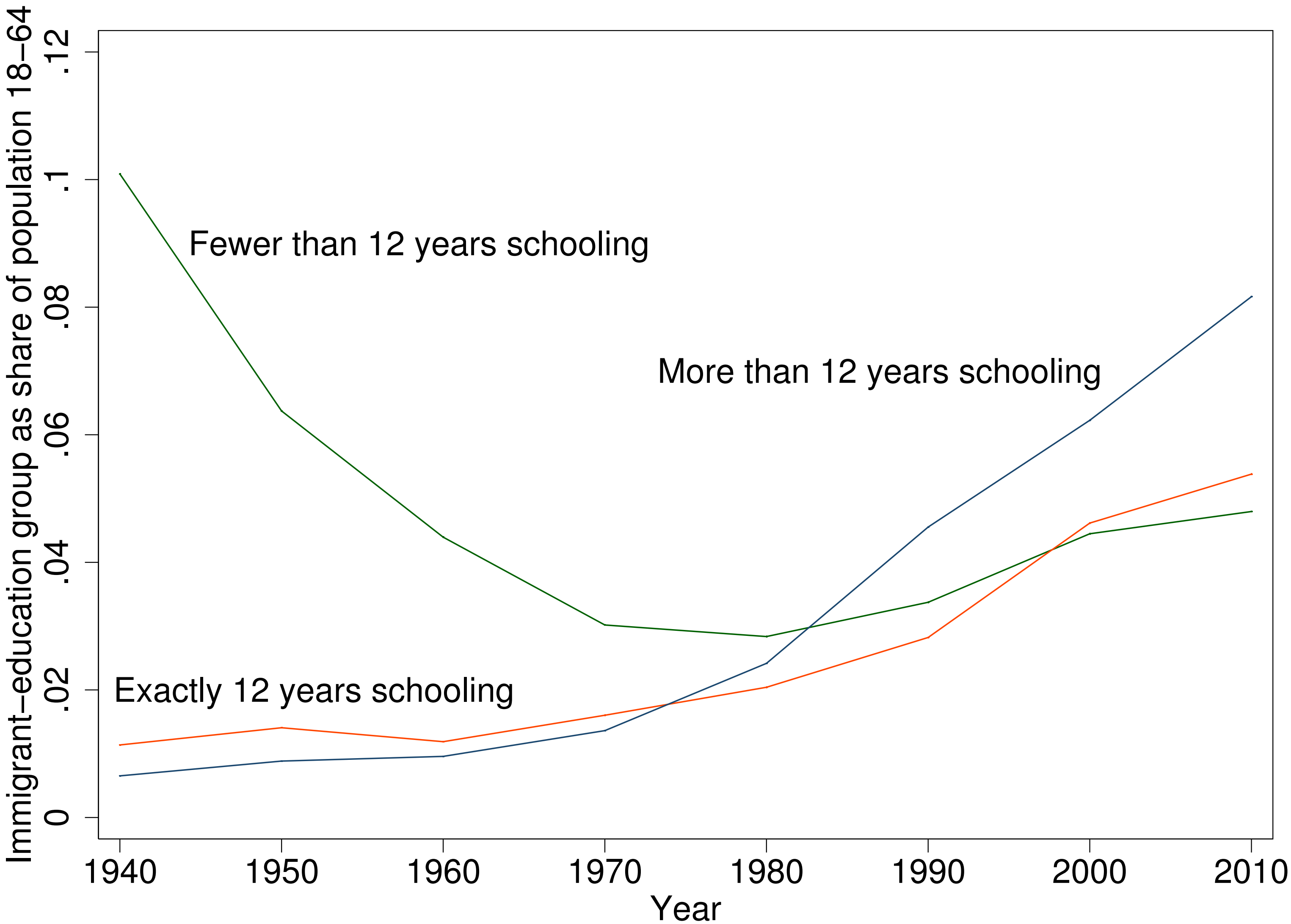
Note: The share of natives aged 21–27 who have completed at least 12 years of schooling.
Source: U.S. Census 1940–2000, American Community Survey 2008–2010.

Figure 2: Immigrant share in various age groups



Note: Immigrants as a share of each age group.
Source: U.S. Census 1940–2000, American Community Survey 2008–2010.

Figure 3: Immigrant education groups as share of the population, ages 18–64



Note: Immigrants 18–64 with various education levels as share of the total population aged 18–64.
Source: U.S. Census 1940–2000, American Community Survey 2008–2010.

Table 1: Interaction of natives by ethnicity with immigrants, school district level 1990

	Immigrants	Hispanic immigrants	White non-Hispanic immigrants	Asian immigrants	Immigrant parents	Hispanic immigrant parents
	(1)	(2)	(3)	(4)	(5)	(6)
A. National level shares						
Overall	0.042	0.023	0.0055	0.010	0.102	0.044
In native's district	0.038	0.020	0.0052	0.009	0.078	0.031
In native white's district	0.025	0.011	0.0045	0.008	0.069	0.025
In native black's district	0.052	0.028	0.0060	0.011	0.120	0.055
In native Hispanic's district	0.107	0.074	0.0084	0.017	0.201	0.139
In immigrant's district	0.143	0.088	0.0138	0.028	0.312	0.159
B. Difference in state shares						
Native black-native white	0.027	0.016	0.0014	0.005	0.061	0.033
Native Hispanic-native white	0.036	0.027	0.0017	0.004	0.072	0.049
Native Hispanic-native black	0.009	0.011	0.0003	-0.002	0.012	0.015

Source: School District Database (tabulations of the 1990 census at school district level), provided by the NBER.

Notes: Based on school districts including a high school: 10,935 school districts for children and 12,503 for parents of children (the reason for the difference is unknown). In panel B, the shares by race/ethnicity are calculated at the state level, then averaged across states weighting with total enrollment in the state. An immigrant is a child of school age born abroad; an immigrant parent is someone born abroad who is the parent of a school-aged child (born anywhere). White denotes non-Hispanic white.

Table 2: Effects of immigrants in population 11-64 on native probability of completing 12 years education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Fixed effects				10-year differences	
							WLS	2SLS
Share population 11-64 which is immigrant; t-10	0.03	0.07	0.11	0.15*	0.20**	0.13*	0.18**	0.30**
	(0.13)	(0.08)	(0.09)	(0.08)	(0.06)	(0.07)	(0.05)	(0.11)
Unemployment rate age 18-24; t-10	--	0.69*	0.01	-0.06	-0.11	-0.05	-0.08	-0.11
		(0.38)	(0.18)	(0.18)	(0.17)	(0.16)	(0.10)	(0.10)
Unemployment rate age 25-54; t-10	--	-0.68	0.07	0.18	0.31	0.26	0.31	0.34
		(0.79)	(0.38)	(0.40)	(0.40)	(0.36)	(0.24)	(0.23)
Share of native population which is age 11-17; t-10	--	-1.36**	-0.37	-0.30	-0.24	0.05	0.04	0.07
		(0.36)	(0.25)	(0.24)	(0.25)	(0.22)	(0.17)	(0.18)
Share workers in agriculture 1940*year	--	--	--	0.006	0.006	0.003	0.048	0.060*
				(0.004)	(0.003)	(0.003)	(0.032)	(0.032)
Share white natives 21-27 less than 12 years school 1940*year	--	--	--	--	0.007**	0.006**	0.063**	0.071**
					(0.003)	(0.003)	(0.027)	(0.027)
Log personal income per capita; t-10	--	--	--	--	--	0.104**	--	--
						(0.026)		
BEA regions*year (p-value)	--	--	0.00	0.00	0.00	0.00	0.00	0.00
R^2	0.91	0.92	0.97	0.97	0.97	0.98	0.85	--
Observations			343				294	

Notes: The dependent variable is the share of natives ages 21-27 who have completed 12 years of education, adjusted at the individual level for age and sex. Estimation is by weighted least squares columns 1-7, and two-stage least squares column 8, with weights w the inverse of the squared standard errors on the state-year interaction coefficient in the individual regression for columns 1-6, and $1/(1/w_t+1/w_{t+10})$ for columns 7 and 8. All specifications include year dummies; fixed effects specifications also include state dummies. The dependent variable is based on 1950-2010 data, the independent variables on data from 1940-2000. The instrument in column 8 is based on the 1940 distribution of immigrants from different regions (see text). Standard errors are clustered by state and reported in parentheses.

Table 3: Effects of immigrants in population 11-64 on native probability of completing 12 years education, by race and ethnicity

Native-born sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Fixed effects				10-year differences	
							WLS	2SLS
All	0.03	0.07	0.11	0.15*	0.20**	0.13*	0.18**	0.30**
	(0.13)	(0.08)	(0.09)	(0.08)	(0.06)	(0.07)	(0.05)	(0.11)
Non-Hispanic whites	-0.00	0.01	0.07	0.08	0.15**	0.11	0.14**	0.25**
	(0.08)	(0.06)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)	(0.12)
Blacks	0.03	0.12	0.20	0.43**	0.47**	0.38**	0.44**	0.46**
	(0.15)	(0.11)	(0.14)	(0.10)	(0.10)	(0.11)	(0.09)	(0.15)
Hispanics	0.12	0.21	0.42**	0.51**	0.53**	0.38**	0.32**	-0.18
	(0.26)	(0.18)	(0.20)	(0.20)	(0.20)	(0.13)	(0.15)	(0.23)
Unemployment rates; t-10	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Share of native pop which is age 11-17; t-10	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BEA regions*year	--	--	Yes	Yes	Yes	Yes	Yes	Yes
Share workers in agriculture 1940*year	--	--	--	Yes	Yes	Yes	Yes	Yes
Share white natives 21-27 less than 12 years school 1940*year	--	--	--	--	Yes	Yes	Yes	Yes
Log personal income p.c.; t-10	--	--	--	--	--	Yes	--	--
Observations (all/whites/blacks/Hispanics)			343/343/324/332				294/294/270/283	

Notes: The dependent variable is the share of natives ages 21-27 of a racial/ethnic group who have completed 12 years of education, adjusted at the individual level for age and sex. Estimation is by weighted least squares columns 1-7, and two-stage least squares column 8, with weights w the inverse of the squared standard errors on the state-year interaction coefficient in the individual regression for columns 1-6, and $1/(1/w_t+1/w_{t+10})$ for columns 7 and 8. All specifications include year dummies; fixed effects specifications also include state dummies. The dependent variable is based on 1950-2010 data, the independent variables on data from 1940-2000. The instrument in column 8 is based on the 1940 distribution of immigrants from different regions (see text). Standard errors are clustered by state and reported in parentheses. Each coefficient is from a different regression.

Table 4: Effects of immigrants by age and education on native probability of completing 12 years education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Fixed effects				10-year differences	
							WLS	2SLS
Share pop 11-17 which is immigrant; t-10	-0.55**	-0.62**	-0.46**	-0.46**	-0.45**	-0.38**	-0.30**	-0.28
	(0.22)	(0.15)	(0.11)	(0.11)	(0.10)	(0.11)	(0.07)	(0.17)
Share pop 18-64 which is immigrant < 12 yrs edu; t-10	0.90**	0.98**	0.94**	0.92**	0.88**	0.80**	0.73**	0.79**
	(0.24)	(0.16)	(0.17)	(0.16)	(0.15)	(0.14)	(0.14)	(0.15)
Share pop 18-64 immigrant 12 years edu; t-10	0.46	0.63	-0.41	-0.40	-0.45	-0.40	-0.18	0.33
	(0.62)	(0.57)	(0.36)	(0.35)	(0.28)	(0.27)	(0.24)	(0.70)
Share pop 18-64 immigrant > 12 years edu; t-10	-0.31	-0.29	0.49	0.53*	0.67**	0.53*	0.43*	-0.01
	(0.63)	(0.48)	(0.31)	(0.31)	(0.28)	(0.30)	(0.22)	(0.64)
Unemployment rates; t-10	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Share of native pop which is age 11-17; t-10	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BEA regions*year	--	--	Yes	Yes	Yes	Yes	Yes	Yes
Share workers in agriculture 1940*year	--	--	--	Yes	Yes	Yes	Yes	Yes
Share non-Hispanic whites <12 years education 1940*year	--	--	--	--	Yes	Yes	Yes	Yes
Log personal income p.c.; t-10	--	--	--	--	--	Yes	--	--
R^2	0.94	0.95	0.98	0.98	0.98	0.98	0.87	--
Observations			343				294	

Notes: The dependent variable is the share of natives ages 21-27 who have completed 12 years of education, adjusted at the individual level for age and sex. Estimation is by weighted least squares columns 1-7, and two-stage least squares column 8, with weights w the inverse of the squared standard errors on the state-year interaction coefficient in the individual regression for columns 1-6, and $1/(1/w_t+1/w_{t+10})$ for columns 7 and 8. All specifications include year dummies; fixed effects specifications also include state dummies. The dependent variable is based on 1950-2010 data, the independent variables on data from 1940-2000. The instruments in column 8 are based on the 1940 distribution of immigrants from different regions (see text). Standard errors are clustered by state and reported in parentheses.

Table 5: First stage of two-stage least squares

	(1)	(2)	(3)	(4)	(5)
	Share of immigrants in age group		Share of population aged 18-64 which is immigrant with		
	11-64	11-17	less than 12 years education	exactly 12 years education	more than 12 years education
Predicted share pop 11-64 which is immigrant	0.40** (0.07)	--	--	--	--
Predicted share pop 11-17 which is immigrant	--	0.67** (0.24)	0.03 (0.06)	-0.04 (0.02)	0.04 (0.05)
Predicted share pop 18-64 which is immigrant less than 12 years edu;	--	0.02 (0.07)	0.77** (0.05)	0.01 (0.04)	0.08 (0.07)
Predicted share pop 18-64 which is immigrant 12 years edu	--	-1.75** (0.66)	-0.72** (0.24)	0.35** (0.08)	-0.19 (0.15)
Predicted share pop 18-64 which is immigrant more than 12 years edu	--	0.85** (0.31)	0.40** (0.12)	0.12 (0.06)	0.53** (0.10)
F statistic for joint significance	33.5	7.1	109.7	42.5	38.7
Angrist-Pischke F statistic	--	14.1	279.3	54.7	29.3
Partial R^2	0.11	0.25	0.16	0.11	0.14
Observations	294	294	294	294	294

Notes: Estimation is by weighted least squares on 10-year differenced data, with weights $1/(1/h_t+1/h_{t+10})$, where h is the denominator of the dependent variable in each regression: population age 11-64 in column 1, population age 11-17 in column 2, population age 18-64 with less than 12 years education column 3, population age 18-64 with exactly 12 years education column 3, population age 18-64 with more than 12 years education column 5. All specifications include year dummies and the (differenced) non-immigrant covariates of Table 1 column 8. The excluded instruments (predicted shares) are based on historical patterns of settlement by age and education. The dependent variable and excluded instruments are based on data from 1940-2000 (i.e. from t-10), while most other covariates are based on data from 1950-2010. Standard errors are clustered by state and reported in parentheses.

Table 6: Effects of immigrants by age and education on native non-Hispanic whites' probability of completing 12 years education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Fixed effects						10-year differences	
							WLS	2SLS
Share pop 11-17 which is immigrant; t-10	-0.49**	-0.50**	-0.30**	-0.30**	-0.29**	-0.27**	-0.16**	0.01
	(0.17)	(0.14)	(0.11)	(0.10)	(0.10)	(0.11)	(0.07)	(0.18)
Share pop 18-64 which is immigrant < 12 yrs edu; t-10	0.59**	0.62**	0.59**	0.60**	0.53**	0.50**	0.45**	0.45**
	(0.17)	(0.13)	(0.17)	(0.17)	(0.15)	(0.16)	(0.14)	(0.13)
Share pop 18-64 immigrant 12 years edu; t-10	0.34	0.47	-0.36	-0.36	-0.41**	-0.39*	-0.22	0.72
	(0.49)	(0.45)	(0.35)	(0.36)	(0.22)	(0.21)	(0.17)	(0.51)
Share pop 18-64 immigrant > 12 years edu; t-10	-0.11	-0.15	0.39	0.37	0.58**	0.54**	0.37	-0.31
	(0.47)	(0.41)	(0.29)	(0.31)	(0.26)	(0.29)	(0.23)	(0.55)
Unemployment rates; t-10	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Share of native pop which is age 11-17; t-10	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BEA regions*year	--	--	Yes	Yes	Yes	Yes	Yes	Yes
Share workers in agriculture 1940*year	--	--	--	Yes	Yes	Yes	Yes	Yes
Share white natives 21-27 less than 12 years school 1940*year	--	--	--	--	Yes	Yes	Yes	Yes
Log personal income p.c.; t-10	--	--	--	--	--	Yes	--	--
R^2	0.94	0.95	0.98	0.98	0.98	0.98	0.88	--
Observations	343						294	

Notes: The dependent variable is the share of native-born non-Hispanic whites ages 21-27 who have completed 12 years of education, adjusted at the individual level for age and sex. Estimation is by weighted least squares columns 1-7, and two-stage least squares column 8, with weights w the inverse of the squared standard errors on the state-year interaction coefficient in the individual regression for columns 1-6, and $1/(1/w_t+1/w_{t+10})$ for columns 7 and 8. All specifications include year dummies; fixed effects specifications also include state dummies. The dependent variable is based on 1950-2010 data, the independent variables on data from 1940-2000. The instruments in column 8 are based on the 1940 distribution of immigrants from different regions (see text). Standard errors are clustered by state and reported in parentheses.

Table 7: Effects of immigrants by age and education on native blacks' probability of completing 12 years education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Fixed effects				10-year differences	
							WLS	2SLS
Share pop 11-17 which is immigrant; t-10	-0.73**	-0.74**	-0.64**	-0.57**	-0.56**	-0.54**	-0.34**	-0.46
	(0.29)	(0.19)	(0.14)	(0.16)	(0.16)	(0.16)	(0.14)	(0.28)
Share pop 18-64 which is immigrant <12 yrs edu; t-10	0.64**	0.73**	1.28**	1.11**	1.09**	1.05**	0.86**	1.12**
	(0.34)	(0.23)	(0.25)	(0.24)	(0.21)	(0.24)	(0.20)	(0.23)
Share pop 18-64 immigrant 12 yrs edu; t-10	-0.26	0.18	0.13	0.11	0.04	0.04	0.37	0.87
	(0.62)	(0.68)	(0.64)	(0.68)	(0.64)	(0.64)	(0.57)	(1.29)
Share pop 18-64 immigrant > 12 yrs edu; t-10	0.65	0.57	0.50	0.89	1.00*	0.96*	0.65	-0.01
	(0.70)	(0.62)	(0.51)	(0.58)	(0.57)	(0.57)	(0.47)	(1.07)
Unemployment rates; t-10	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Share native pop age 11-17; t-10	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BEA regions*year	--	--	Yes	Yes	Yes	Yes	Yes	Yes
Share workers in agriculture 1940*year	--	--	--	Yes	Yes	Yes	Yes	Yes
Share white natives 21-27 less than 12 years school 1940*year	--	--	--	--	Yes	Yes	Yes	Yes
Log personal income p.c.; t-10	--	--	--	--	--	Yes	--	--
R^2	0.96	0.97	0.97	0.98	0.98	0.98	0.86	--
Observations			324				270	

Notes: The dependent variable is the share of native-born blacks age 21-27 who have completed 12 years of education, adjusted at the individual level for age and sex. Estimation is by weighted least squares columns 1-7, and two-stage least squares column 8, with weights w the inverse of the squared standard errors on the state-year interaction coefficient in the individual regression for columns 1-6, and $1/(1/w_t+1/w_{t+10})$ for columns 7 and 8. All specifications include year dummies; fixed effects specifications also include state dummies. The dependent variable is based on 1950-2010 data, the independent variables on data from 1940-2000. The instruments in column 8 are based on the 1940 distribution of immigrants from different regions (see text). Standard errors are clustered by state and reported in parentheses.

Table 8: Effects of immigrants by age and education on native Hispanics' probability of completing 12 years education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Fixed effects				10-year differences	
							WLS	2SLS
Share pop 11-17 which is immigrant t-10	-0.18 (0.27)	-0.39 (0.29)	-0.70** (0.28)	-0.57** (0.27)	-0.54** (0.27)	-0.32 (0.31)	-0.32 (0.21)	-0.89* (0.51)
Share pop 18-64 which is immigrant < 12 yrs edu; t-10	0.74 (0.48)	1.18** (0.51)	0.88** (0.47)	0.43 (0.44)	0.35 (0.46)	0.18 (0.47)	-0.39 (0.41)	-0.13 (0.52)
Share pop 18-64 immigrant 12 years edu; t-10	2.40 (1.35)	1.15 (1.17)	-1.77* (0.90)	-0.68 (0.84)	-0.67 (0.83)	-0.57 (0.88)	-0.09 (0.90)	-4.41** (2.14)
Share pop 18-64 immigrant > 12 years edu; t-10	-1.68 (1.28)	-0.74 (1.06)	2.34** (0.85)	2.22** (0.80)	2.27** (0.79)	1.75** (0.84)	2.01** (0.84)	4.10** (1.73)
Unemployment rates; t-10	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Share of native pop which is age 11-17; t-10	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BEA region trends	--	--	Yes	Yes	Yes	Yes	Yes	Yes
Share agricultural workers 1940*year	--	--	--	Yes	Yes	Yes	Yes	Yes
Share white natives 21-27 less than 12 years school 1940*year	--	--	--	--	Yes	Yes	Yes	Yes
Log personal income p.c.; t-10	--	--	--	--	--	Yes	--	--
R^2	0.91	0.91	0.94	0.94	0.94	0.94	0.59	--
Observations			332				283	

Notes: The dependent variable is the share of native—born Hispanics age 21-27 who have completed 12 years of education, adjusted at the individual level for age and sex. Estimation is by weighted least squares columns 1-7, and two-stage least squares column 8, with weights w the inverse of the squared standard errors on the state-year interaction coefficient in the individual regression for columns 1-6, and $1/(1/w_t+1/w_{t+10})$ for columns 7 and 8. All specifications include year dummies; fixed effects specifications also include state dummies. The dependent variable is based on 1950-2010 data, the independent variables on data from 1940-2000. The instruments in column 8 are based on the 1940 distribution of immigrants from different regions (see text). Standard errors are clustered by state and reported in parentheses.

Table 9: Effects of youth immigrants by parental education on native Hispanics' probability of completing 12 years education

	Fixed effects		10-year differences
	(1)	(2)	(3)
Share population 11-17	-0.54**	--	--
which is immigrant; t-10	(0.27)		
Parents less than 12 years edu	--	-2.20**	-1.80**
		(0.67)	(0.59)
One parent 12 or more years ed	--	3.44**	2.22**
		(0.85)	(0.70)
No parent in household	--	-4.67**	-2.37**
		(1.80)	(1.62)
Share population 18-64			
which is immigrant; t-10			
Less than 12 years education	0.35	1.65**	0.73**
	(0.46)	(0.47)	(0.48)
12 years education	-0.67	-1.32**	-0.66
	(0.83)	(0.70)	(0.90)
More than 12 years education	2.27**	-0.00	0.46
	(0.79)	(0.72)	(0.78)
Other covariates	Yes	Yes	Yes
R^2	0.94	0.95	0.61
Observations	332		283

Notes: The dependent variable is the share of native—born Hispanics age 21-27 who have completed 12 years of education, adjusted at the individual level for age and sex. Estimation is by weighted least squares, with weights w the inverse of the squared standard errors on the state-year interaction coefficient in the individual regression for columns 1 and 2, and $1/(1/w_t+1/w_{t+10})$ for column 2. All specifications include year dummies; fixed effects specifications also include state dummies. Other covariates are those in Table 1: unemployment rates, cohort size, BEA region trends, 1940 agriculture share trend, 1940 share white natives with less than 12 years education trend. The dependent variable is based on 1950-2010 data, the independent variables on data from 1940-2000. Standard errors are clustered by state and reported in parentheses.

Table 10: Effects of youth immigrants by parental education on native blacks' and non-Hispanic whites' probability of completing 12 years education

	Native non-Hispanic whites			Native blacks		
	Fixed effects		10-year diffs	Fixed effects		10-year diffs
	(1)	(2)	(3)	(4)	(5)	(6)
Share population 11-17	-0.29**	--	--	-0.56**	--	
which is immigrant; t-10	(0.10)			(0.16)		
Parents less than 12 years edu	--	0.11	-0.03	--	-1.07*	-0.91
		(0.18)	(0.16)		(0.64)	(0.61)
One parent 12 or more years	--	-0.27	-0.09	--	-0.18	-0.52
		(0.29)	(0.20)		(0.68)	(0.49)
No parent in household	--	-1.81**	-0.78	--	0.18	1.84
		(0.58)	(0.52)		(1.62)	(1.82)
Share population 18-64						
which is immigrant; t-10						
Less than 12 years education	0.53**	0.52**	0.46**	1.09**	1.19**	0.84**
	(0.15)	(0.17)	(0.16)	(0.21)	(0.30)	(0.27)
12 years education	-0.41**	-0.42*	-0.23	0.04	-0.08	0.27
	(0.22)	(0.23)	(0.19)	(0.64)	(0.76)	(0.65)
More than 12 years edu	0.58**	0.67**	0.37	1.00*	0.75	0.67
	(0.26)	(0.29)	(0.26)	(0.57)	(0.61)	(0.41)
Other covariates	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.98	0.98	0.88	0.98	0.98	0.86
Observations	343		294	324		270

Notes: The dependent variable is the share of native—born non-white Hispanics (columns 1-3) or blacks (columns 4-6) age 21-27 who have completed 12 years of education, adjusted at the individual level for age and sex. Estimation is by weighted least squares, with weights w the inverse of the squared standard errors on the state-year interaction coefficient in the individual regression for columns 1, 2, 4 and 6 and $1/(1/w_t+1/w_{t+10})$ for columns 3 and 6. All specifications include year dummies; fixed effects specifications also include state dummies. Other covariates are those in Table 1: unemployment rates, cohort size, BEA region trends, 1940 agriculture share trend, 1940 share white natives with less than 12 years education trend. The dependent variable is based on 1950-2010 data, the independent variables on data from 1940-2000. Standard errors are clustered by state and reported in parentheses.

Table 11: Sensitivity of results to treatment of GED holders

	GED holders have 12 years schooling		GED holders have less than 12 years schooling		
	Sex-age adjusted	Not sex-age adjusted	Not sex-age adjusted		
	Fixed effects	Fixed effects	Fixed effects	10-year diffs	10-year diffs 2SLS
	(1)	(2)	(3)	(4)	(5)
Share pop 11-64 which is immigrant; t-10	0.20**	0.20**	0.31**	0.23**	0.52**
	(0.06)	(0.07)	(0.10)	(0.10)	(0.19)
Share pop 11-17 which is immigrant t-10	-0.45**	-0.18	-0.15	-0.26**	0.48
	(0.10)	(0.16)	(0.15)	(0.12)	(0.35)
Share pop 18-64 which is immigrant < 12 yrs edu; t-10	0.88**	0.44**	0.56**	0.51**	0.97**
	(0.15)	(0.13)	(0.18)	(0.18)	(0.19)
Share pop 18-64 immigrant 12 years edu; t-10	-0.45	-0.39	-0.57	0.40	4.24**
	(0.28)	(0.30)	(0.43)	(0.55)	(1.62)
Share pop 18-64 immigrant > 12 years edu; t-10	0.67**	0.53*	0.81**	0.05	-3.40**
	(0.28)	(0.29)	(0.40)	(0.52)	(1.53)
Average age of natives	--	Yes	Yes	Yes	Yes
Other covariates	Yes	Yes	Yes	Yes	Yes
Observations	343		294		

Notes: The dependent variable is the share of native—born age 21-27 who have completed 12 years of schooling; in column 1 only this is adjusted at the individual level for age and sex. In columns 1 and 2, GED holders are considered to have completed 12 years of schooling, while in columns 3-5 they are not. Estimation is by weighted least squares: column 1 weights are the inverse of the squared standard errors on the state-year interaction coefficient in the individual regression, columns 2-3 weights k are the native population aged 21-27, column 4-5 weights are $1/(1/k_t+1/k_{t+10})$. All specifications include year dummies; fixed effects specifications also include state dummies. Other covariates are those in Table 1: unemployment rates, cohort size, BEA region trends, 1940 agriculture share trend, 1940 share white natives with less than 12 years education trend. The dependent variables are based on 1950-2010 data, adjusted in columns 3-5 with official GED statistics, the independent variables on data from 1940-2000. Standard errors are clustered by state and reported in parentheses.

Table 12: Results by gender for native-born Hispanics and blacks

	Native Hispanics				Native blacks			
	Fixed effects		10-year differences 2SLS		Fixed effects		10-year differences 2SLS	
	Women	Men	Women	Men	Women	Men	Women	Men
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share pop 11-64 which is	1.68**	0.51**	0.30	-0.27	0.42	0.42**	0.55	0.63**
immigrant; t-10	(0.62)	(0.23)	(0.75)	(0.26)	(0.35)	(0.12)	(0.61)	(0.19)
Share population 11-17	--		--	--	-0.56	-0.79**	-0.89	-0.34
which is immigrant; t-10					(0.78)	(0.22)	(1.64)	(0.40)
Parents less than 12 years edu	0.36	-2.53**	--	--	--	--	--	--
	(2.55)	(0.70)						
One parent 12 or more years	0.75	4.09**	--	--	--	--	--	--
	(2.92)	(1.03)						
No parent in household	5.06	-5.45**	--	--	--	--	--	--
	(13.11)	(2.11)						
Share population 18-64								
which is immigrant; t-10								
Less than 12 years education	2.03	1.80**	--	--	0.53	1.43**	1.05	1.50**
	(1.29)	(0.60)			(0.66)	(0.29)	(1.02)	(0.29)
12 years education	-3.17	-1.29	--	--	1.27	-0.52	-3.93	2.43
	(2.69)	(0.85)			(1.95)	(0.85)	(6.61)	(2.23)
More than 12 years edu	3.18	-0.37	--	--	0.03	1.39*	3.68	-1.11
	(2.13)	(0.87)			(1.91)	(0.74)	(5.72)	(1.89)
Other covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	326	324	276	274	314	317	263	264

Notes: The dependent variable is the share of native-born Hispanics (columns 1-4) or blacks (columns 4-8) age 21-27 of the gender specified who have completed 12 years of education, adjusted at the individual level for age and sex. Estimation is by weighted least squares, with weights *w* the inverse of the squared standard errors on the state-year interaction coefficient in the gender-specific individual regression for columns 1, 2, 5 and 6 and $1/(1/w_t+1/w_{t+10})$ for columns 3, 4, 7 and 8. All specifications include year dummies; fixed effects specifications also include state dummies. Other covariates are those listed in Table 2. The dependent variable is based on 1950-2010 data, the independent variables on data from 1940-2000. Standard errors are clustered by state and reported in parentheses.

Appendix Table 1: Native high school completion and graduation rates 1970-2010

	1970	1980	1990	2000	2010
A. 12-year completion rate, age 21-27					
Unadjusted – GEDs have 12 years education	0.784	0.852	0.897	0.914	0.923
Adjusted with official GED statistics	--	0.803	0.831	0.830	0.849
Adjusted using micro-data to exclude GEDs	--	--	--	--	0.878
B. High-school graduation rate, age 21-27					
Only regular diploma-holders are graduates	--	--	--	--	0.863
C. High-school graduation rate (Heckman and LaFontaine)					
Adjusted with official GED statistics, age 20-23	0.807				
Adjusted with official GED statistics, age 20-24	--	0.786	0.794	0.771	--
Adjusted with official GED statistics, age 25-29	--	0.810	0.779	0.792	--

Notes: 1970-2000 statistics are based on census data, while 2010 statistics are based the pooled ACS years 2008-2010. For 2008 both statistics and the GED adjustment are based on ages 19-25, for 2009 ages 20-26 (panels A and B). Hawaii and Alaska are included. Heckman and LaFontaine's (2010) high school graduation rates in panel C refer to natives and immigrants in the United States for more than 10 years (for ages 20-24) or 15 years (for ages 25-29). In 1990-2000 they do not include completers of 12 years of education who did not receive a diploma; in 1970-1980, they presumably do, since the distinction cannot be made in the micro-data.

Appendix Table 2: Means of individual level variables

	All natives				Male natives	Female natives
	12 or more years education completed	Female	Age	Observations	12 or more years education completed	12 or more years education completed
	(1)	(2)	(3)	(4)	(5)	(6)
A. All						
1950	0.576	0.49	24.0 (2.0)	47,225	0.545	0.608
2010	0.919	0.50	22.9 (2.2)	568,840	0.906	0.933
1950-2010	0.878	0.51	23.8 (2.1)	3,978,143	0.870	0.887
B. Non-Hispanic whites						
1950	0.619	0.49	24.0 (2.0)	40,948	0.585	0.654
2010	0.940	0.49	23.0 (2.2)	397,184	0.932	0.949
1950-2010	0.895	0.50	23.8 (2.1)	3,095,866	0.889	0.902
C. Blacks						
1950	0.250	0.53	24.1 (2.0)	4936	0.216	0.281
2010	0.865	0.51	22.8 (2.2)	68,652	0.830	0.899
1950-2010	0.810	0.54	23.7 (2.1)	463,844	0.786	0.831
D. Hispanics						
1950	0.273	0.52	23.9 (1.9)	956	0.252	0.292
2010	0.869	0.50	22.8 (2.2)	67,613	0.850	0.889
1950-2010	0.813	0.51	23.5 (2.1)	263,446	0.804	0.823

Notes: Weighted with census weights adjusted so that the sum of weights for each year reflects the sample size of the census in that year. 2010 refers to the pooled 2008-2010 ACSs. The sample contains natives aged 21-27, except for 2008, when they are aged 19-25, and 2009, when they are aged 20-26. Standard deviations are in parentheses.

Appendix Table 3: Means of state-level covariates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Non-Hispanic white weights			Black weights			Hispanic weights		
	1940	2000	1940-2000	1940	2000	1940-2000	1940	2000	1940-2000
Unemployment rate	0.161	0.106	0.107	0.100	0.114	0.108	0.153	0.114	0.106
ages 18-24	(0.053)	(0.019)	(0.030)	(0.041)	(0.019)	(0.026)	(0.037)	(0.015)	(0.022)
Unemployment rate	0.071	0.038	0.047	0.053	0.040	0.046	0.075	0.041	0.044
ages 25-54	(0.019)	(0.007)	(0.014)	(0.015)	(0.008)	(0.012)	(0.015)	(0.006)	(0.010)
Share native population	0.140	0.105	0.113	0.150	0.104	0.112	0.138	0.111	0.111
which is aged 11-17	(0.011)	(0.007)	(0.016)	(0.009)	(0.008)	(0.017)	(0.011)	(0.008)	(0.013)
Share employment in	0.092	0.093	0.090	0.167	0.107	0.116	0.110	0.081	0.086
agriculture 1940	(0.068)	(0.062)	(0.063)	(0.064)	(0.0074)	(0.075)	(0.047)	(0.051)	(0.051)
Share native whites with	0.547	0.530	0.532	0.626	0.565	0.574	0.512	0.486	0.488
<12 years education 1940	(0.086)	(0.095)	(0.092)	(0.069)	(0.094)	(0.094)	(0.074)	(0.083)	(0.082)
State personal income per	582	29,947	16,252	378	29,556	17,148	534	31,190	20,622
capita (nominal)	(195)	(4161)	(9801)	(159)	(4562)	(9939)	(180)	(3907)	(9927)
Share population 11-17									
which is immigrant; t-10									
Parents high s dropouts	0.011	0.017	0.015	0.007	0.016	0.013	0.015	0.036	0.034
	(0.007)	(0.017)	(0.018)	(0.004)	(0.016)	(0.017)	(0.013)	(0.019)	(0.022)
One parent high s grad	0.003	0.041	0.028	0.001	0.043	0.030	0.003	0.059	0.047
	(0.002)	(0.026)	(0.023)	(0.001)	(0.027)	(0.025)	(0.002)	(0.023)	(0.026)
No parent in household	0.001	0.006	0.004	0.001	0.006	0.004	0.002	0.011	0.009
	(0.001)	(0.004)	(0.005)	(0.001)	(0.004)	(0.005)	(0.001)	(0.004)	(0.006)
Observations	49	49	343	45	49	320	47	49	330

Notes: The weights are the inverse of the squared standard errors on the state-year interaction coefficient in the individual regression for attainment of 12 years of education, by race/ethnicity. The individual regressions are weighted with census weights adjusted so that the sum of weights for each year reflects the sample size of the census in that year.

Appendix Table 4: 1940 shares of national-level immigrants from various origins, all ages and educations

Origin	(1) Share
United Kingdom	0.050
Ireland	0.059
Italy	0.062
Germany	0.052
Poland	0.060
Russia	0.064
Other Europe	0.050
Mexico	0.043
Puerto Rico	0.096
Canada	0.038
Central America	0.052
South America	0.069
Other Caribbean	0.073
Cuba	0.047
China	0.058
India	0.047
Other Asia	0.045
Rest of world	0.054

Appendix Table 5: Results by gender for native-born non-Hispanic whites and all races and ethnicities

	Native non-Hispanic whites				Natives of all races and ethnicities			
	Fixed effects		10-year differences 2SLS		Fixed effects		10-year differences 2SLS	
	Women	Men	Women	Men	Women	Men	Women	Men
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share pop 11-64 which is	0.23**	0.16**	0.36**	0.25**	0.20**	0.19**	0.31**	0.28**
immigrant; t-10	(0.10)	(0.07)	(0.18)	(0.11)	(0.07)	(0.06)	(0.11)	(0.11)
Share population 11-17	0.06	-0.25**	-0.05	0.13	-0.44**	-0.46**	-0.36*	-0.22
which is immigrant; t-10	(0.27)	(0.11)	(0.30)	(0.18)	(0.10)	(0.12)	(0.18)	(0.21)
Share population 18-64								
which is immigrant; t-10								
Less than 12 years education	0.43**	0.42**	0.31	0.37**	0.94**	0.81**	0.88**	0.71**
	(0.17)	(0.15)	(0.22)	(0.14)	(0.16)	(0.14)	(0.15)	(0.17)
12 years education	-0.37	-0.55**	0.96	1.05	-0.27	-0.63*	-0.03	0.67
	(0.50)	(0.26)	(1.27)	(0.56)	(0.28)	(0.33)	(0.81)	(0.80)
More than 12 years edu	0.20	0.75**	0.07	-0.57	0.50	0.85**	0.30	-0.29
	(0.39)	(0.28)	(1.12)	(0.56)	(0.29)	(0.32)	(0.70)	(0.75)
Other covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	343	343	294	294	343	343	294	294

Notes: The dependent variable is the share of native—born non-white Hispanics (columns 1-4) or all races and ethnicities (columns 4-8) age 21-27 of the gender specified who have completed 12 years of education, adjusted at the individual level for age and sex. Estimation is by weighted least squares, with weights w the inverse of the squared standard errors on the state-year interaction coefficient in the gender-specific individual regression for columns 1, 2, 5 and 6 and $1/(1/w_t+1/w_{t+10})$ for columns 3, 4, 7 and 8. All specifications include year dummies; fixed effects specifications also include state dummies. Other covariates are those listed in Table 2. The dependent variable is based on 1950-2010 data, the independent variables on data from 1940-2000. Standard errors are clustered by state and reported in parentheses.