

NBER WORKING PAPER SERIES

IMMIGRATION AND INEQUALITY

David Card

Working Paper 14683
<http://www.nber.org/papers/w14683>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2009

Ely Lecture prepared for the January 2009 meetings of the American Economic Association. Some of this findings reported here are taken from an ongoing collaboration with Giovanni Peri and Steven Raphael: I am extremely grateful for their input. I also thank Larry Katz for detailed comments, and Christian Dustmann for helpful discussions. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2009 by David Card. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Immigration and Inequality
David Card
NBER Working Paper No. 14683
January 2009
JEL No. J31

ABSTRACT

Immigration is often viewed as a proximate cause of the rising wage gap between high- and low-skilled workers. Nevertheless, there is controversy over the appropriate framework for measuring the presumed effect, and over the magnitudes involved. This paper offers an overview and synthesis of existing knowledge on the relationship between immigration and inequality, focusing on evidence from cross-city comparisons in the U.S. Although some researchers have argued that a cross-city research design is inherently flawed, I show that evidence from cross-city comparisons is remarkably consistent with recent findings from aggregate time series data. Both designs provide support for three key conclusions: (1) workers with below high school education are perfect substitutes for those with a high school education; (2) “high school equivalent” and “college equivalent” workers are imperfect substitutes; (3) within education groups, immigrants and natives are imperfect substitutes. Together these results imply that the impacts of recent immigrant inflows on the relative wages of U.S. natives are small. The effects on overall wage inequality (including natives and immigrants) are larger, reflecting the concentration of immigrants in the tails of the skill distribution and higher residual inequality among immigrants than natives. Even so, immigration accounts for a small share (5%) of the increase in U.S. wage inequality between 1980 and 2000.

David Card
Department of Economics
549 Evans Hall, #3880
University of California, Berkeley
Berkeley, CA 94720-3880
and NBER
card@econ.berkeley.edu

Immigration has long been a controversial topic among economists.¹ The issue nearly disappeared in the 1960s, but over the past three decades professional interest has picked up as immigrant inflows have surged. The new immigration has attracted attention in part because of its sheer size – approximately 1.25 million people per year over the first half of this decade – and in part because of its composition.² A third or more of the new arrivals are undocumented immigrants from Mexico and Central America with low education and limited English skills (Passel, 2005). These immigrants presumably compete for the same jobs held by the least-skilled native workers, contributing to a *trifecta* of economic factors – technology, trade, and immigration – that are thought to have led to a rise in skill differentials in the U.S. economy since the late 1970s.

This paper presents an overview and synthesis of research on the connection between immigration and wage inequality, focusing on the evidence derived from comparisons across major U.S. cities. The appeal of this research design is illustrated in Table 1, which presents recent data on immigrant densities, education outcomes, and mean salaries for workers in 12 of

¹The founding president of the American Economic Association, Francis A. Walker, wrote a well-known article arguing in favor of restricting immigration (Walker, 1896). Walker believed that a particular problem was the changing nature of the “new” immigrants. He wrote: “Fifty, even thirty years ago, there was a rightful presumption regarding the average immigrant that he was among the most enterprising, thrifty, alert, adventurous, and courageous of the community from which he came. It required no small energy, prudence, forethought, and pains to conduct the inquiries relating to his migration, to accumulate the necessary means, and to find his way across the Atlantic. To-day the presumption is completely reversed”. Walker’s view was disputed by Paul H. Douglas (Douglas, 1919), who argued that cohort-based comparisons showed no decline in immigrant “quality”.

²See U.S. Department of Commerce (2006).

the nation's largest cities.³ The immigrant share of the working-age population in these cities ranges from about 10% in Philadelphia and Detroit to nearly 50% in Los Angeles. These differences in immigrant densities are correlated with large differences in the relative shares of different skill groups. Figure 1, for example, presents a scatter plot of the share of high school dropouts in the working age population of the biggest 124 cities against the corresponding immigrant share. On average, each percentage point rise in the immigrant share is associated with a 0.2% rise in the relative share of dropouts.⁴ Consistent with this general pattern, the dropout share in Los Angeles is *over twice as large* as the share in Philadelphia or Detroit. I believe that there is something to be learned about labor markets in general, and the effects of immigration in particular, by studying how wages and other outcomes respond to this variation.

Nevertheless, as emphasized by Borjas, Freeman and Katz (1997), cross-city comparisons are far from a panacea. Natives and immigrants can easily move between cities, and depending on how sensitive these flows are to differences in local wages, naive cross-city comparisons may reveal a lot or little about the underlying technological parameters that determine the effects of immigration on native opportunities. In fact, Borjas, Freeman and Katz (1997) argue that mobility rates are so sensitive to relative wages that intercity comparisons are essentially uninformative. Instead, they propose the use of aggregate time series models to estimate the underlying parameters of interest.

³Throughout this paper I identify “cities” with Metropolitan Statistical Areas (MSA’s), or in the case of larger urban agglomerations, with the constituent Primary Metropolitan Statistical Areas (MSA’s). See the Data Appendix for details on how the data used in this paper were constructed.

⁴The linear regression coefficient is 0.21 with a standard error of 0.02; the R-squared is 0.38.

This paper makes the case that appropriately identified parameters derived from cross-city and aggregate time series comparisons are in fact mutually consistent. In particular, estimates from both data sources point to three key conclusions. The first is that workers with less than a high school education are perfect substitutes for those with a high school education. This conclusion is important because it means that the impact of low-skilled immigration is diffused across a wide segment of the labor market (the roughly 60% of the population who are counted as “high school equivalents” workers) rather than concentrated among the much smaller dropout population (only 14% of the population). The second conclusion is that “high school equivalent” and “college equivalent” workers are imperfect substitutes, with an elasticity of substitution that appears to be similar whether the relevant labor market is defined as the nation as a whole, or an individual city.⁵ A third conclusion is that within broad education classes, immigrant and native workers appear to be imperfect substitutes, with a large but finite elasticity of substitution. As was first pointed out by Ottaviano and Peri (2006), if immigrants and natives in the same skill category are imperfect substitutes, the competitive effects of additional immigrant inflows are concentrated among immigrants themselves, lessening the impacts on natives.

When the demand side of the national labor market is parameterized with these three

⁵Fisher (1968, 1969) showed that the conditions for existence of a well-behaved aggregate production function – even when all firms face identical prices for all inputs and outputs – are extremely stringent. In particular, the existence of a labor aggregate such as “high school labor” requires the absence of specialization across firms. Nevertheless, simulations reported in Fisher, Solow and Kearl (1977) suggest that economies made up of heterogeneous firms with different CES production functions behave as if they were generated by an aggregate CES. An interesting issue for further research is under what conditions an aggregate economy made up of separated markets – each containing heterogeneous CES firms, and each with potentially different relative prices – behaves as a single aggregate CES, and whether the synthetic parameters at the market and aggregate levels are related in any systematic way.

assumptions, and capital is assumed to be perfectly elastically supplied, the net impact of immigrant inflows over the past two decades on mean wages of different subgroups of native workers are quite small (Ottaviano and Peri, 2008). A similar parameterization of the demand side of the local labor market implies that the effects of immigration on mean wages for different groups of natives in most cities are also small.

Most of the existing research on immigration has focused on between-group inequality. A significant share of the overall rise in U.S. wage inequality is attributable to increases in within-group or residual wage inequality - the variation that remains after controlling for education, age, gender, race and ethnicity (see Lemieux, 2008 for a recent summary). Empirically, residual wage variation among native workers is significantly higher in cities with more immigrants. The relative level of residual wage inequality for natives in different skill groups is uncorrelated with the relative fraction of immigrants, however, suggesting that immigration has a relatively small causal effect.⁶ Taken together with the evidence on between-group wage differentials, I conclude that immigration has had very small impacts on wage inequality among natives.

Nevertheless, when immigrants themselves are counted in the overall population their effect on inequality is clearly positive. Immigrants tend to be concentrated in the upper and lower tails of the skill distribution. Residual wage inequality among immigrants is also higher than among natives. A simple calculation suggests that the presence of immigrants can explain about 5% of the rise in overall wage inequality between 1980 and 2000.

⁶Lemieux's (2008) comparisons of trends in residual wage inequality by sub-period and within narrow occupations also seem to rule out a major role for immigration.

I. Conceptual Framework

a. Overview

The main theoretical apparatus used in studies of wage inequality, and in studies of immigration, is a model of the demand side of the labor market.⁷ In both literatures the supply side is usually simplified by assuming that per capita labor supply is perfectly inelastic, although in models of local labor markets the number of workers can potentially vary (e.g., Card, 2001). Most often capital is assumed be to separable from labor inputs, and perfectly elastically supplied, so the issue becomes one of specifying the relative demand for different skill groups.

As emphasized in international trade theory, some fraction of the overall response of labor demand to relative wages presumably arises through sectoral adjustments (see Kuhn and Wooten, 1991, for an illuminating analysis in the immigration context). In the stark world of the Heckscher-Ohlin model this channel is so important that relative demand curves are flat: changes in the relative supplies of different skill groups lead to expansions and contractions of different industries with no change in relative wage or relative skill utilization within any particular industry. One might guess that sectoral adjustments are particularly important in understanding derived demand in local labor markets (e.g., at the city or state level). Surprisingly, however, this does not appear to be the case. Lewis (2004) showed that differences in relative supplies of different education groups across cities are almost entirely absorbed by within-industry changes

⁷In the inequality context see for example Katz and Murphy (1992), Murphy and Welch (1992), Autor and Katz (1999), and Goldin and Katz (2008), all of which focus exclusively on demand side modeling. In the immigration context see Johnson (1980), Grossman (1982), Borjas (2003), and Ottaviano and Peri (2006), all of which likewise focus on the demand side.

in utilization.⁸ Likewise, Lewis (2008) concludes that sectoral shifts played a relatively small role in the adjustment of the Miami labor market to the Mariel Boatlift. In view of these results it does not appear too unrealistic to adopt a one-sector model of the demand for labor at either the national or local levels.

In a one-sector framework the properties of the relative demands for different skill groups are derived from the properties of the relevant sectoral production function, which I will write in general form as

$$(1) \quad y = f(N_1, N_2, \dots, N_S),$$

where N_s represents labor input from skill group $s=1,2,\dots,S$. It is standard to assume that f exhibits constant returns to scale, implying that the inverse demand functions for each type labor are homogenous of degree 0 in the vector of quantities N_1, N_2, \dots, N_S . This has the important implication that a “skill-balanced” inflows of immigrants – one with the same skill distribution as the existing labor force – has no effect on the relative wage structure.⁹

As a point of departure, it is helpful to think of the case where all labor types are perfect substitutes, so f has the form:

$$(2) \quad f(N_1, N_2, \dots, N_S) = h(\sum_s \theta_s N_s)$$

for some set of (possibly time-varying) productivity weights θ_s . In this case the relative wage

⁸Card and Lewis (2007) present a simplified version of Lewis’s analysis for the case of Mexican immigrants and reach the same conclusion.

⁹Provided that capital is available at a fixed interest rate, such a “balanced” immigration will have no effect on the level of wages for any group either. Researchers sometimes simulate the effect of immigration assuming that the U.S. capital stock is fixed (e.g., Borjas and Katz 2007 present some results with a fixed capital stock). I believe it is plausible to assume that capital is elastically supplied to the U.S. See Ottaviano and Peri (2008) for more discussion.

between any two groups, say 1 and 2, is determined by the ratio θ_1/θ_2 , and is independent of relative supply.¹⁰ Moreover, if capital is freely available at a fixed price immigration has no effect on the absolute level of wages for any group. Although a one-skill model is extreme, there is a long tradition in labor and macro economics of using such a model, and I believe it represents a useful “null hypothesis.” Under this model, relative demand curves are not downward sloping, but are all flat.

The most widely used form for the production function is a “two-group CES” in which workers are partitioned into “high school equivalents” (H) and “college equivalents” (C). Within each subgroup workers are assumed to be perfect substitutes. Thus

$$(3) \quad f(N_1, N_2, \dots, N_s) = [\alpha_H L_H^\rho + \alpha_C L_C^\rho]^{1/\rho}$$

$$L_H = \sum_{s \in H} \theta_s N_s$$

$$L_C = \sum_{s \in C} \theta_s N_s$$

$$\rho = (\sigma_e - 1)/\sigma_e,$$

where σ_e is the elasticity of substitution between high school and college labor. Usually, high school dropouts, high school graduates, and some fraction of people with 1-3 years of post-secondary education are classified as high school equivalents. College graduates and the remaining fraction of those with some college education are classified as college-equivalents (see e.g., Katz and Murphy, 1992).¹¹ This simple specification has become a cornerstone of the wage inequality literature (see e.g. Katz and Murphy, 1992; Autor and Katz, 1999; Acemoglu, 2002).

¹⁰Even with a perfect substitutes assumption it is still possible to have skill biased technical change if the relative productivity of more highly skilled workers rises over time.

¹¹I am not sure of the origins of this specification but it dates at least to Freeman (1976).

Estimates based on U.S. time series data up to the early 1990s suggested that with the addition of a linear trend term (representing skill biased technical change), a model based on equation (3) could provide a reasonable description of education-related wage differences in the economy, with estimates for $1/\sigma_e$ centered around 0.7. Adding more recent data tends to lead to smaller estimates of $1/\sigma_e$: for example Acemoglu(2002) reports a value of 0.56 using data for 1939-96, while Ottaviano and Peri (2008) present a range of estimates between 0.3 and 0.7. Katz and Goldin (2008) argue recent estimates are confounded by a slowdown in the pace of skill-biased technical change (i.e., the trend in α_C/α_H) in the early 1990s. Allowing for such a shift they obtain estimates for $1/\sigma_e$ close to the benchmark 0.7 number.

The assumption embodied in (3) that there are only two distinct skill groups seems relatively strong, and a number of extensions have been proposed. Card and Lemieux (2001) proposed a nested CES structure, allowing for imperfect substitution between different age or experience groups within each of the H and C groups. Their implied estimates of the elasticity of substitution (σ_x) are relatively large (10 or so) and the age structure of the immigrant labor force is not so different from that of natives, so ignoring this form of imperfect substitution makes relatively little difference in the immigration context (see Ottaviano and Peri, 2008 for comparisons of various simulations).

A much more important distinction is whether two education classes is enough. Borjas (2003) and Borjas and Katz (2007) assume there are four: dropouts (D), high school graduates (H), people with some college (S), and college graduates (C), implying a specification like:

$$(4) \quad f(N_1, N_2, \dots, N_S) = [\alpha_D L_D^\rho + \alpha_H L_H^\rho + \alpha_S L_S^\rho + \alpha_C L_C^\rho]^{1/\rho}$$

where L_D , L_H , L_S , and L_C are CES aggregates of labor inputs of different experience groups

within each education class, as in Card and Lemieux (2001). Note that the elasticity of substitution between any two of the groups ($\sigma_e = 1/(1-\rho)$) is assumed to be constant. Thus, the inverse relative demand function for any two education groups j and k has the simple form:

$$(5) \quad \log w_j/w_k = \log \alpha_j/\alpha_k - 1/\sigma_e \log N_j/N_k.$$

In the two-skill case an equation like (5) holds for college-high school wage premium, but not for the relative wage differential between high school graduates and high school dropouts, which depends only on the relative number of efficiency units of high school labor held by the two groups (i.e., by the θ_s 's in equation 3).

Though seemingly innocuous, the switch from a two-skills model like (3) to a four-skills model like (4) has extremely important implications for the potential effect of immigration on native wage inequality. This is because the U.S. immigrant population has a high fraction of dropouts relative to natives (31% versus 11%) but a fairly similar fraction of high school equivalents (63% versus 59%).¹² In a four-skill model the relatively high fraction of dropouts in the immigrant population distorts the relative share of dropouts and lowers the wage of dropouts relative to other education groups. In a two skill model, however, the share of dropouts is irrelevant. What matters instead is the share of high school equivalent labor, which is unaffected by immigration when immigrants have the same relative fractions of high school-equivalent and college-equivalent labor as natives.

¹²In the 2005/2006 American Community Survey the education distribution for working age immigrants is 30.5% dropouts, 24.2% high school graduates, 18.6% some college, and 26.7% college graduates. The corresponding fractions for natives are: 10.8%, 30.4%, 30.9% and 27.8%. I assume that each dropout supplies 0.7 units of high school labor, that one half of those with some college supply 1.2 units of high school labor, and that the other half supply 0.8 units of college equivalent labor.

A second important question is whether immigrants and natives in the same education (and experience) class are perfect substitutes, as is assumed in Borjas (2003), Borjas and Katz (2007), or imperfect substitutes, as was implicitly assumed in the seminal studies by Grossman (1982) and Borjas (1987). Ottaviano and Peri (2006, 2008) and Manacorda, Manning and Wadsworth (2006) propose an additional nest to the CES structure to allow for imperfect substitutability. In these studies there are three elasticities of substitution: one between immigrants and natives with the same age and education (which I will denote by σ_l); one within education classes between different experience groups (σ_x), and the third (σ_e) between education groups.¹³

Interestingly, even a modest degree of imperfect substitutability between natives and immigrants can make a significant difference in the implied impacts of immigration on native wages. Loosely speaking, the higher is $1/\sigma_l$ the greater is the concentration of the wage impacts caused by immigrant inflows on immigrants themselves, and the smaller is the spillover effect on natives. For example, in a two-education group model that ignores differences by age (i.e., $1/\sigma_x=0$) the effect of an inflow of high-school equivalent immigrants on the college-high school wage gap for natives depends on the difference [$1/\sigma_e - 1/\sigma_l$], rather than on $1/\sigma_e$ as in equation (5). (See Manacorda, Manning, and Wadsworth (2006), equation 12).

¹³Dustmann, Frattini, and Preston (2008) point out that the degree of substitution between immigrants and natives will likely depend on how long immigrants have worked in the U.S., and other factors, such as mother tongue and the quality of education in their home country. Arguably these should be taken into account by classifying immigrants into different subgroups. I do not pursue this here.

b. Aggregate Time Series Evidence on the Relevant Elasticities

Before turning to the main evidence that I review here, drawn from cross-city comparisons, it is useful to briefly summarize the state of the literature based on aggregate times series studies. Focusing first on the issue of the number of education groups, my reading is that recent studies support a relatively simple two-group structure. Goldin and Katz (2008, chapter 8) present an analysis of the determinants of the high school graduate wage premium from 1914 to 2005, allowing the inverse elasticity of substitution between high school graduates and dropouts to vary over time. In a model with a simple linear trend interaction they find that the inverse elasticity was 0.35 at the beginning of their sample and has steadily trended downward, reaching 0 in 1985. Confirmatory evidence is presented by Ottaviano and Peri (2008), who use annual data from 1963 to 2006, and obtain estimates for the inverse elasticity of substitution between dropouts and high school graduates in the range of 0 to 0.04. Ottaviano and Peri (2008) also attempt to estimate the elasticity of substitution between workers with 1-3 years of college, and those with a bachelors degree or more.¹⁴ Again, they obtain estimates of the inverse substitution elasticity that are small and statistically insignificant (taking account of serial correlation).

These findings shed some light on the difficulty that Borjas (2003) and Borjas and Katz (2007) had in attempting to estimate a single inverse elasticity of substitution between four education groups. Borjas (2003) reports two estimates of $1/\sigma_e$: 0.74 (with a standard error=0.65) and 0.76 (with standard error=0.58). In a replication of Borjas (2003) that uses 2000 Census data (instead of the 1999-2001 Current Population Surveys used in the original study), Raphael and

¹⁴Strictly speaking the traditional two-education group model assigns some of the workers with some college to the high school group and the remainder to the college graduate group.

Ronconi (2008) obtain an estimate close to 0. Borjas and Katz (2007) modify Borjas' original specification by adding a trend break in the pace of skill-biased technical change in the 1990s and obtain an estimate of 0.41 (standard error=0.31). I conjecture that an important reason for the instability and imprecision of these estimates (apart from the relatively small number of time series observations) is that the data do not support the four skills model. In particular, the wage gap between dropouts and high school graduates appears to be uncorrelated with the relative supply of dropouts. Moreover, judging from Katz and Murphy (1992), Acemoglu (2002) and evidence in Goldin and Katz (2008, table 8.2), the college-high school wage gap appears to depend on the relative number of people with 12 years or less of schooling, not just the number with exactly 12 years of schooling, suggesting again that the four-group model is mis-specified.

There is somewhat less time series evidence on the elasticity of substitution between immigrants and natives (σ_I). Ottaviano and Peri (2006) originally reported estimates of $1/\sigma_I$, on the order of 0.10 to 0.20. Some details of their analysis have been criticized by Borjas, Grogger and Hanson (2008), and in the latest version of their paper Ottaviano and Peri (2008) present estimates for $1/\sigma_I$ that range from 0.04 to 0.08. Though small in magnitude the estimates are quite precise. Manacorda, Manning and Wadsworth (2006) present estimates based on aggregate U.K. data that are larger in magnitude (around 0.15) but also relatively precise.

II. Immigrant Settlement Patterns as a Source of Identifying Information

While the aggregate time series evidence is useful, I believe that additional information on the critical parameters of the demand side of the labor market can be gleaned from cross-city comparisons. As noted above, the main issue in interpreting cross-city comparisons is that the

supplies of labor in a city may respond to relative wages, leading to a classic identification problem. In this paper I present evidence based on instrumental variables (IV) specifications that use earlier immigrant settlement patterns as a source of identifying information. This section briefly reviews the basis for this strategy.

As a starting point, Table 2 summarizes some of the main characteristics of immigrants from different sending countries observed in the 2000 Census. For reference the first two rows compare natives and all immigrants. As noted in the last section, a striking difference is in the distribution of very low versus “middle” levels of education. Immigrants are relatively clustered at the lowest levels of education, while natives are clustered in the center. The shares in the upper quartile of the education distribution are more similar. Comparisons across the different rows of the table show that the education distributions vary widely by source country. Immigrants from Mexico – by far the largest source country – are very poorly educated, as are those from El Salvador and Guatemala. Immigrants from the Philippines and India (the second and third largest source countries, respectively) are better-educated than natives.

A well known fact about immigrants is that they tend to settle in country-specific enclaves. Interesting examples include the clustering of Arab immigrants in Detroit (see Abraham and Shryock, 2000), Polish immigrants in Chicago (Pacyga, 1991), and Mexican immigrants in Los Angeles and Chicago. Figure 2 illustrates the pattern using data for Filipino immigrants in the 124 largest U.S. cities. The x-axis of this figure represents the ratio between the fraction of Filipino immigrants who lived in a specific city in 1980, and the average fraction of all immigrants who were living in this city in 1980. The y-axis represents the same relative share, taken over the set of immigrants observed in the 2000 Census who arrived after 1980. The

graph suggests that cities where Philippine immigrants were clustered in 1980 (relative to the locations of all immigrants) tended to attract an excess share of Filipino immigrants in the following decades. Similar patterns hold across nearly all source countries.¹⁵

As a result of the tendency for new immigrants to move to the same cities as earlier immigrants from the same country, the number and relative skill distribution of immigrants arriving in a city over a given interval of time is fairly predictable. If M_m immigrants arrive from country m to the U.S. as a whole (say between 1990 and 2000), and the fraction of earlier immigrants from country m who lived in city j at some previous date (say 1980) is λ_{mj} , then a naive clustering model predicts that $\lambda_{mj}M_m$ immigrants from country m will move to city j. Letting N_m denote the earlier population of immigrants from country m in the U.S. as a whole, and N_{mj} denote the number living in city j, the predicted inflow rate, as a fraction of the city's population (P_j) at some reference date is

$$(6) \quad (\sum_m \lambda_{mj} M_m) / P_j = \sum_m [N_{mj}/P_j] M_m / N_m$$

which is just a weighted average of the national inflow rates from each source country, with weights that depend on the shares of the country's earlier immigrants in city j.¹⁶ If the national inflow rates from each source country are exogenous to conditions in a specific city, then the

¹⁵The R-squared for the scatter of points in Figure 2 is 0.77. The mean R-squared across 37 source countries/country groups is 0.36. Other source countries with high enclave tendencies include Mexico, Cuba, and Poland. Immigrants from India, China, and Taiwan show relatively low enclave tendencies.

¹⁶If N_{mc} and P_j are measured at the same time point then N_{mj}/P_j is just the fraction of immigrants from source country m in city j at that point. If the two variables are measured at different time points then N_{mj}/P_j is the fraction of immigrants from country m in city j at the date of measurement for N_{mj} , multiplied by the overall growth rate of the population of city j from that date to the date of measurement of P_j .

predicted inflow based on (6) will be exogenous. Assuming that a fraction δ_{ms} of immigrants from country m are in skill group s , the predicted inflow rate of new immigrants in skill group s is

$$(7) \quad (\Sigma_m \lambda_{mj} M_m \delta_{ms}) / P_j = \Sigma_m [N_{mj}/P_j] \delta_{ms} M_m/N_m,$$

which again can be interpreted as a weighted average of the skill-group-specific inflow rates from each source country. Finally, note that a predicted relative inflow rate (for example of college-equivalent versus high school equivalent immigrants) can be calculated by taking the ratio of two expressions like the one in equation (7).

In the analysis in the next section I use predicted relative inflows calculated in this way as instrumental variables for the relative number of dropout workers in a city and the relative number of college-equivalent workers in a city. Given the very large unskilled inflows from Mexico over the 1980s and 1990s, the calculated values of the instruments are significantly correlated with the fraction of Mexican immigrants in a city in 1980, leading to a concern that the instruments may be picking up other features about a city that are also correlated with the initial fraction of Mexicans living there (even if the surge in national inflows of Mexicans was exogenous to conditions in the city). To address this concern I have refit many of the models excluding Mexican immigrants from the calculation of the instrumental variables. This does not change the general pattern of the results in any of the tables reported below, though the point estimates and precision are sometimes affected.

III. Cross-City Evidence on the Impacts of Immigration on Native Wage Differentials

a. Data Overview

I use 1980-2000 Census data along with data from the combined 2005 and 2006 American Community Surveys (ACS) to construct a panel data set of city-level labor market outcomes in 1980, 1990, 2000, and 2005/2006. As noted in the discussion of Table 1, I define “cities” as MSA’s or PMSA’s as of the 2000 Census. I have used information on the changing definitions of MSA’s and PMSA’s to attempt to match the 2000 boundaries as closely as possible in 1990 and 1980. (The ACS public use files use the same geographic coding as the 2000 Census). Some of the boundary changes cannot be incorporated because the smallest geographic units in the public use Census files (the so-called Public Use Micro Areas) are rather large (100,000 people or more) and do not always correspond to the geographic units that define MSA’s and PMSA’s (counties in most of the country and towns in New England). Such problems are likely to be more serious for relatively small cities. For this reason, and because immigrants mainly live in larger cities, I limit attention to the 124 largest MSA’s or PMSA’s in the country as of 2000.

Table 3 presents some summary statistics on the characteristics of native and immigrant men and women in the four sample years. The data on education, experience, and employment rates are based on samples that include people over the age of 18 with between 1 and 45 years of potential experience. The wage outcomes are constructed by dividing annual wage and salary earnings by the product of weeks worked and usual hours per week. To eliminate the influence of outliers I have “Winsorized” hourly wages in each sample year at a lower value of 0.75 times the Federal minimum wage, and an upper value of 50 times the minimum wage. In calculating wages I also exclude people who have positive self-employment income, since there is no information on how they divide their hours of work between self-employment and non-self

employment jobs.

The entries in the first column of Table 3 suggest that both natives and immigrants have become better educated over the past 25 years, with a slightly bigger gain for natives. Working-age natives have also become older (a gain in over 2 years of potential experience since 1980), while working-age immigrants' experience levels have remained constant, reflecting the rapid inflow of younger immigrants. Employment rates of native men have drifted down (see Autor and Duggan, 2003), while those of immigrant men have remained constant, and rates for women have risen. Mean real wages of native men have been roughly constant (using the CPI as a price index) while those of immigrant men have fallen, and real wages of both native and immigrant women have risen.

The last two columns of Table 3 show two simple measures of wage inequality: the variance of log hourly wages, and variance of residual wage inequality.¹⁷ By these measures wage inequality has risen for all groups, with a particularly large jump between 1980 and 1990. Hourly wages are measured with considerable error in the Census, so the levels of wage inequality reported in Table 3 are too high (see Lemieux 2006 for a careful discussion of measurement problems in the context of wage inequality). The trends after 2000 may also be affected by the switch between the Census questionnaire, which asks about earnings in the previous calendar year, and the ACS questionnaire, which asks about earnings in the previous 12

¹⁷Wage residuals are obtained from a series of linear regression models fit separately by gender, immigrant status, and year. The models for natives include a flexible combination of age, education, and ethnicity variables (a total of 55 covariates). The models for immigrants include these controls as well as dummies for each of 38 countries of origin (or country groups), interactions of the origin dummies with a measure of years in the U.S., and interactions of education variables with broad country group dummies (a total of 168 covariates). See the Data Appendix.

months. Presumably these measurement problems affect all groups, so comparisons of the relative trends between immigrants and natives are still informative.

b. Analysis of the Dropout/High School Graduate Wage Gap

Perhaps the single most important issue for understanding how immigrant inflows have affected native wage structures is the degree of substitutability between high school graduates and dropouts. Cross-city comparisons are potentially very useful here because the relative share of dropouts varies widely across cities (see Figure 1). If there is an important degree of imperfect substitution, the relative wage of dropouts should be lower in high-dropout cities. To address the question, I estimate cross-city models of the form

$$(8) \quad r_{Djt} - r_{Hjt} = a + b X_{jt} + c \log [S_{Djt} / S_{Hjt}] + e_{jt}$$

where r_{Djt} represents the mean residual log wage among native male dropout workers in city j and year t, r_{Hjt} represents the mean residual log wage among native male high school graduates, X_{jt} is a vector of city-level control variables, S_{Djt} and S_{Hjt} represent total annual hours of work by all dropouts and high school graduates in city j and year t, and e_{jt} is an error term.¹⁸ As in equation (5), the coefficient c is interpretable as the negative of the inverse elasticity of substitution between dropouts and high school graduates, and is expected to be negative or 0 in the limiting case of perfect substitutes. Note that wages are measured for native men, but supplies are measured over all workers present in city j. Note too that I use residual wages from a model fit

¹⁸The focus on male wages is standard in the literature (e.g., Katz and Goldin, chapter 8), and is meant to abstract from relative trends in female wages that are driven by changes in the relative selectivity of female workers. I have fit all the models in this paper using residual wages for men and women: the results are very similar to those reported here.

to the entire U.S. workforce. Thus, r_{Djt} is interpretable as the mean wage differential for native male dropouts who work in city j , relative to the national labor market, after adjusting for observed characteristics.¹⁹

A concern with (6) is that there may be some unobserved city-specific factor that shifts the relative demand for dropout workers, leading to higher relative wages and higher relative employment and confounding the estimation of the inverse substitution elasticity. I use as an instrumental variable for the relative supply ratio the log of the ratio of the predicted inflows of dropout and high school immigrants to city j over the previous 10 years. Specifically, I calculate expressions like (7) for each city in 2000 using national inflows of immigrants from each of 38 source countries/country groups over the period from 1990 to 2000, and the shares of each group observed in each city in 1980. I use estimates for the skill shares (the δ_{ms} terms in equation 7) derived from the skill shares of the national pool of immigrants from each country who arrived between 1990 and 2000. I then take the log of the predicted relative inflows for dropouts and high school graduates. This instrument has a relatively strong correlation with the actual ratio of dropout to high school labor supply in 2000, with t-statistics in the first-stage equation on the order of 14 (or 10 if Mexicans are excluded from the construction of the instrument).

Estimation results for various versions of equation (8) are presented in Table 4. An issue for the specification is whether one should allow city fixed effects that capture any permanent city-specific factors that account for differences in the mean (adjusted) wage gap between

¹⁹The average log wage differential between native male dropouts and high school graduates in the Census samples was -17.5% in 1980, -22.9% in 1990, -25.4% in 2000, and -26.6% in 2005/2006. Much of the change over time can be attributed to changing composition of the two groups: as discussed in Card (2005), the gap adjusted for differences in experience and ethnicity is slightly declining over the 1980-2000 period.

dropouts and high school graduates. To probe this issue I include specifications with a lagged dependent variable (estimated for the city in 1990).²⁰ If there are permanent differences across cities the coefficient of the lagged dependent variable will be close to 1. As control variables I include in some specifications the log of the city size in 1980 and 1990, the share of college workers in the city in 1980 and 1990, the share of workers employed in manufacturing in 1980 and 1990, and the mean wage residual for all native workers in the city in 1980 and 1990. The latter are meant to pick up any potential correlation between the high school/dropout relative wage, and the average level of wages in a city.

All the specifications point to the same conclusion, which is that elasticity of substitution between dropouts and high school graduates is effectively infinite. The conclusion is similar whether controls are added to the model or not, and whether the model is estimated by OLS or IV.²¹ Although not shown in the table, IV results based on predicted inflows that exclude Mexicans are also very similar (and similarly precise). Figure 3 illustrates the configuration of the data underlying the IV results. This figure plots the residual wage gap between dropouts and high school graduates against the instrumental variable. Clearly there is wide variation across cities in the predicted relative inflow of dropouts to high school graduates. But this variation is uncorrelated with the residual wage gap for native dropouts.

²⁰I fit some of these models using the lagged wage for native women as an instrument for the lagged dependent variable and obtained coefficient estimates very similar to the ones reported here.

²¹The covariates are all individually insignificant except the mean residual wage in 1990, which has a negative coefficient in all model (typical estimate = -0.15, typical standard error =0.05). The estimates imply that the high school/dropout wage gap is higher in high-wage cities. This may reflect the upward pressure of the minimum wage in low-wage cities.

The estimates in Table 4 strongly support the conclusion from recent studies based on aggregate time series data. Both sources of evidence suggest that “dropout labor” is not an independent factor of production. Rather, wages of dropouts and wages of high school graduates vary in proportion across cities, as predicted by the assumption that the two groups are perfect substitutes. I have fit many different versions of the models in Table 4, including specifications using native female wages, and pooled male and female wages, and other models that exclude the largest 3 or 10 cities in the country. All of these models yield estimates of the inverse substitution elasticity that are close to 0.²²

c. Analysis of the College/High School Wage Gap

Based on the preceding analysis I believe it makes sense to use a simple two-skill model, with college-equivalent and high school-equivalent labor types, to study wage differentials at the city level. Following the literature, I fit models of the form

$$(9) \quad r_{Cjt} - r_{Hjt} = a + b X_{jt} + c \log [S_{Cjt} / S_{Hjt}] + e_{jt}$$

where r_{Cjt} represents the mean residual wage of native male workers with exactly 16 years of education in city j and year t , r_{Hjt} represents the mean residual wage for native male high school graduates, and S_{Cjt} and S_{Hjt} represent the supplies of college equivalent and high school equivalent labor employed in city j in year t . I define high school equivalent hours as the sum of hours

²²Very similar findings are also reported by Card and Lewis (2007) who look at the high school graduate-dropout wage premium for a slightly larger set of cities in 2000 and between 1990 and 2000. They report models that control for overall employment growth between 1990 and 2000, including IV specification that use employment levels in the 1982-90 period as an instrument for employment growth between 1990 and 2000. They find that the high school premium is uncorrelated with employment growth rates.

worked by high school graduates, plus 0.7 times the hours worked by dropouts (assuming that dropouts have 0.7 efficiency units of high school graduates) plus 1.2 times the hours worked by one half of people with 1-3 years of post-secondary schooling (assuming that half of people with some college are high school equivalents, and each has 1.2 efficiency units of high school labor). I define college equivalent hours as the sum of hours worked by college graduates plus 0.8 times the hours worked by one half of people with 1-3 years of post-secondary schooling (assuming that half of people with some college are college equivalents, and each has 0.8 efficiency units of college labor).

Table 5 presents estimation results for a variety of alternative specifications. One immediate and important difference between the results in this table and those in Table 4 is that city-specific values of the college-high school wage are highly correlated over time, with rather complex dynamics. As shown in column (1) of Table 5, when two lagged values (i.e., the wage gaps in 1990 and 1980) are included, the sum of their coefficients is very close to 1, suggesting that an appropriate model may be a first order autoregression in differences:

$$\Delta (r_{Cjt} - r_{Hjt}) \approx \beta \Delta (r_{Cjt-1} - r_{Hjt-1}) + \text{other terms},$$

where β is a number like -0.3. Specifications that impose a first differences on both the lagged dependent variables and the relative supply terms are reported in columns (4) and (8).

When dynamics are ignored, and no other covariates are included, the estimate in column (2) of Table 5 suggests that the college-high school wage gap in a city is positively correlated with the relative supply of college workers. Adding two lags of the dependent variable, lagged relative supply, and controls for city size and the employment share in manufacturing in 1980 and 1990 (column 3) pushes the coefficient on current relative supply into the negative range, but

the point estimate is small in absolute value. Imposing a first differences structure on the lagged dependent variables and the relative supply variable (column 4) leads to estimates that are very close to those from the corresponding unrestricted specification in column 3.

One interpretation of these OLS estimates is that the relative supply of college workers is responsive to relative wages, leading to a positive bias in the estimated inverse elasticity of substitution. This conclusion is consistent with existing work on relative migration flows of different education groups (e.g., Dahl, 1998), which tends to find that mobility rates of college workers are sensitive to group-specific wage levels.

Columns 5-8 of Table 5 present instrumental variables estimation results, using predicted relative inflows of college-equivalent and high school-equivalent immigrants from 1990 to 2000 as an instrument for the relative supply of college-educated workers. This instrument is a lot weaker than the instrument for the relative supply of dropouts used in Table 5, but when lagged supply and control are added (columns 6-8) the t-statistic on the instrumental variable in the first stage equation is over 4.5 (see the entries in the bottom row of Table 5). The IV estimates of the inverse elasticity of substitution between college and high school workers range from 0.26 to 0.41 – not far below the estimates obtained in many recent aggregate time series studies, and consistent with a value for σ_e in the range of 1.5-2.5.

My interpretation of the estimates in Table 5 is that relative demand for college versus high school workers at the city level exhibits about the same elasticity with respect to relative wages as relative demand at the national level. However, relative supply is endogenous at the local level, confounding simple observational comparisons across cities (like the simple univariate model in column 2). This does not mean that cross-city comparisons are

uninformative. Rather, it means that researchers have to address the endogeneity problem to obtain interpretable estimates. For this purpose the enclaving tendency of immigrants, coupled with differences in the education distributions of immigrants from different origin countries, is particularly helpful, although other identification strategies are certainly worth pursuing.²³

d. Analysis of the Immigrant/Native Wage Gap

A third key issue for understanding the impact of immigration is whether immigrants and natives in the same broad skill class are perfect or imperfect substitutes. Here, cross-city comparisons are potentially very useful because there is enormous variation in the relative fraction of immigrants across cities. Assuming there are only two skill groups – high school equivalents and college equivalents – I estimate models separately by skill group of the form:

$$(10) \quad r_{Mjt} - r_{Njt} = a + b X_{jt} + c \log [S_{Mjt} / S_{Njt}] + e_{jt}$$

where r_{Mjt} represents the mean wage residual for immigrant men in a particular skill group in city j and year t , r_{Njt} represents the corresponding mean wage residual for native men, and S_{Mjt} and S_{Njt} represent the total hours of work by all immigrants and all natives in the skill group (i.e., including men and women). The coefficient c in this equation is interpretable as an estimate of the inverse elasticity of substitution between immigrants and natives in the particular skill group (i.e. $c = -1/\sigma_l$). Again there is a obvious concern that unobserved factors in a city may lead to

²³Moretti (2004) presents an analysis of the effect of the college share on wages of different education groups in a city, using the age structure of a city in 1970 as a predictor of the change in the college share in a city between 1980-1990. This is a reasonably powerful instrument. His estimates (Moretti, 2004, Table 5) imply that a 10 percent point increase in the college share reduces the college-high school wage premium by about 0.12 log points. This is equivalent to an elasticity with respect to the log of the relative supply of college-equivalent workers of about -0.25.

higher wages and employment levels of immigrants relative to natives, confounding the estimation of the inverse substitution elasticity. Following the strategy in Table 4 and 5, I use predicted inflows of immigrants in the particular skill group (high school or college) to city j over the 1990-2000 period (based on immigrant shares in 1980) to instrument for the share of immigrants in 2000.

It is worth noting a potential difference between the IV strategy for the estimation of equation (10) and the corresponding strategies for equations (8) and (9). In the latter cases, the instrumental variable is the ratio of predicted immigrant inflows in two skill groups (e.g., high school equivalents versus college equivalents). These ratios are not as strongly correlated with the initial immigrant share in 1980 as the predicted immigrant inflows of high school- or college-equivalent workers used as instruments for the estimation of (10).²⁴ To the extent that initial immigrant shares are correlated with other unobserved features that affect relative wage differentials in a city, enclave-based instruments may be less attractive for estimating $1/\sigma_1$ than for estimating the inverse elasticity across education groups. Assuming that OLS estimates of $1/\sigma_1$ from a specification like (10) are biased toward 0 by endogeneity, and that the IV estimates are not fully purged of endogeneity, both the OLS and IV estimates will tend to underestimate the (absolute) magnitude of $1/\sigma_1$.

Estimation results for equation (10) are shown in Table 6. All the specifications include measures of city size, college share, and manufacturing share in 1980 and 1990, as well as the

²⁴The correlation of the instrument for the ratio of dropouts to high school graduates with the share of immigrants in a city in 1980 is 0.43. The correlation of the instrument for the ratio of college- to high school-equivalent labor with the 1980 immigrant share is -0.55. By comparison the correlation of the instruments for the fractions of immigrants in the high school or college workforces with the immigrant share in 1980 is 0.90.

mean wage residuals for all immigrants and all natives living in the city in 1980. The latter are meant to control for any attributes in a city that lead to persistently higher wages for immigrants and/or natives, although the coefficients of these variables are generally small and statistically insignificant. Wage gaps between immigrants and natives in a city do not exhibit the complex dynamics of the college-high school wage premium, and models with a lagged dependent variable have a relatively small autoregressive coefficient (around 0.15). The estimated relative supply effects are also very similar whether or not the lagged dependent variable is included.

The estimated relative supply coefficients are generally bigger in absolute value when estimated by IV than by OLS, suggesting some endogeneity in relative supply that is mitigated by the IV strategy, although both OLS and IV estimates are uniformly negative. The relative supply coefficients are also bigger (in absolute value) for college workers than high school workers. Taken literally, this means that less-educated immigrants and natives are closer to perfect substitutes ($\sigma_l \approx 40$) than are more-educated immigrants and natives ($\sigma_l \approx 17$).

The estimation results in Table 6 are very consistent with the range of estimates for $1/\sigma_l$ obtained by Ottaviano and Peri (2008), which center around -0.04. Specifications similar to the ones in Table 6 fit for both men and women, and for women only, also confirm this conclusion, with estimates for the inverse substitution elasticity in the same range. IV estimates of the inverse elasticity using predicted inflows that exclude Mexican immigrants are slightly larger in absolute value (e.g., the estimate corresponding to the entry in column 4 is -0.031 with a standard error of 0.008), as are estimates from specifications that exclude the control variables.²⁵

²⁵The IV coefficient estimates from models without controls are -0.027 (standard error=0.006) and -0.080 (standard error=0.010) respectively.

Overall, I conclude that there is strong time series and cross city evidence of a small but detectable degree of imperfect substitution between immigrants and natives.

e. Within-Group Residual Inequality

Existing studies of the impact of immigration on the wages of natives have focused on the effect on mean wage differentials between skill groups. While much of the rise in wage inequality over the past two decades has been driven by increases in between-group inequality – especially the gap between high and low education groups – within-group inequality has also risen substantially (Autor and Katz, 1999; Lemieux, 2006, 2008).

Across major cities the level of within-group or residual wage inequality is strongly correlated with immigrant densities. This is illustrated in Figure 5, which plots the residual variance of wages for college-equivalent and high school equivalent male workers in each of the largest 124 cities in 2000 against the immigrant share.²⁶ A 10 percentage point increase in the immigrant share is associated with a 0.025 point rise in the residual variance of high school equivalent men's wages (standard error = 0.002 points), and a 0.027 point rise in the residual variance of college equivalent men's wages (standard error = 0.003 points).

To the best of my knowledge there is no well-developed theory of how the presence of immigrants (or a more diverse workforce generally) affects residual wage inequality among natives. If one assumes that the causal effect of higher immigration depends on the fraction of

²⁶The same patterns holds using women's wages, or men's and women's pooled. For example, the correlation between immigrant share and the residual variance of college-equivalent women's wages is 0.62, versus 0.64 for men. The correlation between immigrant share and the residual variance of high school-equivalent women's wages is 0.75, versus 0.77 for men.

immigrants in a worker's own skill group , however, then an appropriate empirical specification will have the form

$$(11) \quad V_{Cjt} - V_{Hjt} = a + b X_{jt} + c \log [Imm_{Cjt} / Imm_{Hjt}] + e_{jt},$$

where V_{Cjt} represents the residual variance of wages among college-educated workers in city j and year t, V_{Hjt} represents the corresponding variable for high-school-educated workers, and Imm_{Cjt} and Imm_{Hjt} denote the fractions of immigrants in the college-equivalent and high-school equivalent workforces.

Table 7 presents estimation results for specifications based on equation (11). For ease of interpretation I use the difference in logs of the residual variances for native male college graduates and high school graduates as the dependent variable in these models, although specifications that use the difference in residual variances show the same patterns of sign and significance. The table includes both OLS and IV models that use the ratio of the predicted number of college and high school immigrants moving to the city between 1990 and 2000 as an instrumental variable for the relative immigrant fraction.

Two main conclusions emerge from the results in Table 7. First, the ratio of the residual variances among college and high school workers is fairly persistent over time, with a coefficient on the lagged dependent variable close to 0.5. Second, there does not appear to be a large effect of relative immigrant densities on the relative residual variances of college and high school workers.

Tentatively, I conclude that the correlations exhibited in Figure 5 are not a causal effect of immigration, but rather a consequence of unobserved city-wide factors that are associated with higher immigration and a higher level of residual variance for both skill groups. It should be

noted, however, that a “relative” specification such as (11) is not the only possible causal model. If a greater fraction of immigrants in one skill group has equal effects on the residual variance of wages for workers in both groups, then a specification like (11) will fail to estimate the true effect of immigration.

IV. Summary and Discussion

Cross-city and time series comparisons of the effects of relative supplies of different skill groups on relative wages are consistent with three key hypotheses:

- (1) workers with below high school education are perfect substitutes for those with a high school education
- (2) “high school equivalent” and “college equivalent” workers are imperfect substitutes, with an elasticity of substitution on the order of 1.5-2.5
- (3) within education groups, immigrants and natives are imperfect substitutes, with an elasticity of substitution on the order of 20.

These hypotheses imply that the structure of relative labor demand at the city or national level is consistent with a simple nested CES sub-production function for labor inputs with two skill groups at the upper level, each of which is itself an aggregate of immigrant and native labor inputs in the appropriate skill group.

The combination of these assumptions – plus the assumption of perfectly elastic capital supplies – means that at the national level, immigration over the past two or three decades has had very minor effects on relative wage differentials between natives in different skill groups (see Ottaviano and Peri, 2008), and a negligible effect on between-skill group wage variability.

The main explanation for this somewhat surprising conclusion is that under a two-education-group model, what matters for the structure of wages is the relative fractions of immigrants and natives who are high school-equivalent and college-equivalent workers. U.S. immigrants are only slightly under-represented in the college-equivalent group relative to natives (36% versus 41%). Compared to the distribution among natives alone immigrant arrivals have hardly distorted the relative fraction of college-equivalent workers in the economy, and have therefore had little impact on the college-high school wage gap.

The overall impact of immigration on native inequality depends on the effects on between-group differentials and on the effects on within-group inequality. There is little existing research on the latter channel, although cross-city comparisons show a rather strong correlation between immigrant density and the residual variation in native wages. On the other hand, the relative level of within-group inequality between college and high school educated natives is uncorrelated with the relative density of immigrants in the local college and high school workforces. Assuming that any within-group impacts caused by immigrants are concentrated among natives in the same skill group, this suggests that the causal effects of immigration on within-group inequality are small.

Overall, I conclude that immigration has not had much effect on native wage inequality in the U.S. as a whole. Nevertheless, because immigrants are clustered at the high and low ends of the education distribution, and because they also tend to have higher residual inequality than natives (see Table 3), wage inequality over all workers in the economy is higher than it would be in the absence of immigration. Table 8 illustrates this point, showing data on wage inequality in 1980 and 2005/2006 for all workers, and for immigrants and natives. In 1980, for example, the

variance of log hourly wages across all male workers was 0.390, versus 0.385 among native men. Likewise, the variance among all female workers was 0.318, versus 0.317 among native women. Over the past 25 years the gap between the variance of wages in the entire workforce and among natives has widened: thus immigration can be said to have contributed to the rise of inequality in the workforce even if it has had no effect on the inequality of native wages. The effect is relatively small, however. For men, native inequality rose by 0.137 while overall inequality rose by 0.142. If overall inequality had risen as fast as native inequality it would rise by 4% less. For women, native inequality rose by 0.139 while overall inequality rose by 0.148. If overall inequality had only risen as fast as native inequality it would rise by 6% less. These comparisons suggest that the presence of immigration can account for a relatively small share (4-6%) of the rise in inequality in the past 25 or so years.²⁷

While I believe that recent research using time series and cross city comparisons has made significant progress in clarifying the effects of immigration on labor market outcomes for natives, several important issues deserve further attention. First, given the importance of the degree of substitutability between dropouts and high school graduates, more research on the types of jobs held by the two groups and further evidence on how excesses of dropout labor are absorbed in high-immigrant cities would be useful. Additional insights into the degree of substitution between similar immigrants and natives are also needed. Finally, more research is needed to better understand why a simple 2-skill-group CES model seems to work so well at both

²⁷I have not attempted to decompose the rise in inequality within the immigrant workforce into a components attributable to immigrants who were already in the country in 1980 and a component for new immigrants. Assuming imperfect substitutability between immigrants and natives, however, the arrival of new immigrants has presumably lowered the wages of earlier immigrants (Ottaviano and Peri, 2008; Manacorda, Manning, and Wadsworth, 2006).

the national and city levels.

References

- Abraham, Nabeel and Andrew Shryock. 2000. *Arab Detroit: From Margin to Mainstream*. Detroit: Wayne State University Press.
- Acemoglu, Daron, 2002. "Technical Change, Inequality, and the Labor Market." *Journal of Economic Literature*, 40(1): 7-72.
- Autor, David H. and Mark G. Duggan. 2003. "The Rise In The Disability Rolls And The Decline In Unemployment." *Quarterly Journal of Economics*, 118(1): 157-205.
- Borjas, George J. 1987. "Immigrants, Minorities, and Labor Market Competition." *Industrial and Labor Relations Review*, 40(3): 382-392.
- Borjas, George J. 2003. "The Labor Demand Curve is Downward Sloping: Re-examining the Impact of Immigration on the Labor Market." *Quarterly Journal of Economics* , 118(4): 1335-1374.
- Borjas, George J., Richard B. Freeman and Lawrence F. Katz. 1997. "How Much Do Immigration and Trade Affect Labor Market Outcomes?" *Brookings Papers on Economic Activity* 1997(1): 1-90.
- Borjas, George J., and Lawrence F. Katz. 2007. "The Evolution of the Mexican-Born Workforce in the United States." In *Mexican Immigration*, ed. George J. Borjas, 13-56. Chicago: University of Chicago Press (for NBER).
- Borjas, George J., Jeffrey Grogger and Gordon H. Hanson. 2008. "Imperfect Substitution Between Immigrants and Natives: A Reappraisal." Available at:
<http://irpshome.ucsd.edu/faculty/gohanson/ImperfectsubstitutionFinal.pdf>

- Card, David. 2001. "Immigrant Inflows, Native Outflows and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics* 19 (1): 22-64.
- Card, David. 2005. "Is the New Immigration Really So Bad?" *Economic Journal* 115 (4): F300-F323.
- Card, David and Thomas Lemieux. 2001. "Can Falling Supply Explain the Rising Return to College for Younger Men?" *Quarterly Journal of Economics* 116 (2): 705-746.
- Card, David and Ethan G. Lewis. 2007. "The Diffusion of Mexican Immigrants During the 1990s: Explanations and Impacts." In *Mexican Immigration*, ed. George J. Borjas, 193-207. Chicago: University of Chicago Press (for NBER).
- Dahl, Gordon. 2002 "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." *Econometrica*, 70(6):2367-2420.
- Douglas, Paul H. 1919. "Is the New Immigration More Unskilled Than the Old?" *Quarterly Publications of the American Statistical Association*, 16(116): 393-403.
- Dustmann, Christian, Tommaso Frattini, and Ian Preston. 2008. "The Effect of Immigration Along the Distribution of Wages." CREAM Discussion Paper No. 03/08.
- Fisher, Franklin M. 1969. "The Existence of Aggregate Production Functions." *Econometrica* 37(4): 553-577.
- Fisher, Franklin M., Robert Solow and James M. Kearn. 1977. "Aggregate Production Functions: Some CES Experiments." *Review of Economic Studies* 44(2): 305-320.
- Freeman, Richard B. 1976. *The Overeducated American*. San Diego: Academic Press.
- Grossman, Jean, B. 1982. "The Substitutability of Natives and Immigrants in Production." *Review of Economic and Statistics*, 64(3): 596–603.

Goldin, Claudia and Lawrence F. Katz. 2008. *The Race Between Education and Technology*.

Cambridge MA: Harvard University Press.

Johnson, George P. 1997. "Changes in Earnings Inequality: The Role of Demand Shifts."

Journal of Economic Perspectives 11 (2): 41-54.

Katz, Lawrence F. and David H. Autor. 1999. "Inequality in the Labor Market." In *Handbook of Labour Economics*, ed. Orley Ashenfelter and David Card, 1463-1555. Amsterdam and New York: North Holland.

Katz, Lawrence F. and Kevin M. Murphy. 1992 . "Changes in Relative Wages, 1963-1987:

Supply and Demand Factors." *Quarterly Journal of Economics* 107 (1): 35-78.

Kuhn, Peter and Ian Wooten. 1991. "Immigration, Trade, and the Wages of Natives." In *Immigration, Trade, and the Labor Market*, ed. John M. Abowd and Richard B. Freeman, 285–304, Chicago: University of Chicago Press.

Lemieux, Thomas. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *American Economic Review*, 96(3): 461-498.

Lemieux, Thomas. 2008. "The Changing Nature of Wage Inequality." *Journal of Population Economics*, 21(1): 21-48.

Lewis, Ethan G. 2004. "Local Open Economies within the U.S. How do Industries Respond to Immigration?" Federal Reserve Bank of Philadelphia Working Paper 04-1.

Lewis, Ethan G. 2008. "How Did the Miami Labor Market Absorb the Mariel Immigrants?" Federal Reserve Bank of Philadelphia Working Paper 04-1. Revised April 2008.

- Manacorda, Marco, Alan Manning and Jonathan Wadsworth. 2006. "The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain." CREAM Discussion Paper No. 08/06.
- Moretti, Enrico. 2004. "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data." *Journal of Econometrics* , 112 (2004): 175- 212.
- Murphy, Kevin M. and Finis Welch. 1992. "The Structure of Wages." *Quarterly Journal of Economics* 107 (1): 285-326.
- Ottaviano, Gianmarco and Giovanni Peri. 2006. "Rethinking the Effects of Immigration on Wages." NBER Working Paper Number 12497.
- Ottaviano, Gianmarco and Giovanni Peri. 2008. "Immigration and National Wages: Clarifying the Theory and the Empirics." Available at:
www.econ.ucdavis.edu/faculty/gperi/Papers/OP_redux_july_2008.pdf
- Pacyga, Dominic A. 1991. *Polish Immigrants and Industrial Chicago: Workers on the South Side, 1880-1922*. Chicago: University of Chicago Press.
- Passel, Jeffrey S. 2005. "Estimates of the Size and Characteristics of the Undocumented Population." Research Report of the Pew Hispanic Center. Washington DC: Pew Hispanic Center.
- Raphael, Steven and Lucas Ronconi. 2008. "Reconciling National and Regional Estimates of the Effects of Immigration on U.S. Labor Markets: The Confounding Effects of Native Male Incarceration Trends." Unpublished Working Paper, University of California Goldman School of Public Policy.

U.S. Department of Commerce Bureau of the Census. 2006. "Table 4. Cumulative Estimates of the Components of Population Change for the United States, Regions and States, April 1, 2000 to July 1, 2006." NST-EST2006-04.

Walker, Francis A. 1896. "Restriction of Immigration." *The Atlantic Monthly* 77(464): 822-829.

Data Appendix

A. Processing of 1980, 1990, 2000 Census and 2005/2006 ACS

1. I begin with the 5% public use files of the 1980-2000 Censuses and the two ACS's. I define "immigrants" as all individuals whose citizenship status is either non-citizen or naturalized citizen – thus people born abroad to U.S. parents are not considered immigrants. From the Census files I extract 100% of all immigrants and a 50% random sample of natives over the age of 18. From the ACS files I extract all individuals over the age of 18.
2. I adjust the person weight variables for the 1990 and 2000 Censuses to account for the under-sampling of natives. I define a 1980 weight equal to 1 for immigrants and 2 for natives. For the ACS files I use the person weight.
3. For the 1980 and 1990 Census files, I make two adjustments to the MSA/PMSA code. First, I assign MSA or PMSA to county groups (1980) or PUMA's (1990) that are not assigned an MSA/PMSA code if more than one-half of the residents live in a single MSA/PMSA. Second, I attempt to add or drop county groups or PUMA's to adjust for boundary changes between 1980 and 2000, for the 124 largest cities (as of 2000).
4. For the 2000 Census I assign MSA or PMSA to PUMA's that are not assigned an MSA/PMSA code if more than one-half of the residents live in a single MSA/PMSA. For the ACS files, I assign the same MSA/PMSA to each PUMA as is assigned in 2000 Census (after adjusting for unassigned MSA/PMSA codes).
5. In the 1980 Census, I define "education" based on highest grade completed. For the 1990 and 2000 Census and the ACS, I define "education" based on the reported grade variable. An important issue in this coding decision for the 2000 Census and ACS files is how to classify people who report having attending college but not completing at least one year. Mean wages for this group are relatively similar to those who have "some college" (e.g., for native men in 2000, the mean log wage for this group is 2.672, versus 2.726 for those with some college, and 2.576 for those with exactly high school). Thus, I classify these people as having some college. This choice also makes the distribution of education outcomes between 1990 and 2000 more similar than if the people who went to college but failed to complete a year are coded as high school graduates.
6. I define "experience" as age-16 for dropouts, age-19 for high school graduates, age-21 for people with some college, and age-23 for people with a BA or higher. I then restrict the sample to people with 1-45 years of experience.

Copies of the programs used to read and process the public use data sets are available on request.

B. Wages and Residual Wages

1. I define an hourly wage by dividing wage and salary earnings by the product of hours per week and weeks worked. I set the wage to missing for anyone with positive self-employment or farm income. I “Winsorize” hourly wages at a lower bound of 0.75 times the minimum wage and an upper bound of 50 times the minimum wage.
2. I then perform linear regressions, separately by year, gender and immigrant/native status, of log wages on a set of covariates. The residuals from these regressions are used to construct the “adjusted” or “residual” wages used in the paper, and also to construct measures of “residual wage inequality”.
3. The wage models for natives include the following variables:
 - cubic in experience
 - years of completed education
 - complete interactions of dummies for 4 education groups (dropout, high school graduate, some college, college graduate) with dummies for 5-year experience bins
 - a dummy for people with an advanced degree
 - a dummy for those who attended school last year
 - dummies for people who worked full time (>1400 hours) or low hours (under 1000)
 - interactions of dummies for black, Hispanic and Asian ethnicity with years of completed education and a dummy for college completion
 - a dummy for living in an MSA/PMSA that is at least partially in one’s state of birth, interacted with dummies for 4 education groups
 - dummies for living in a smaller MSA/PMSA (not one of the top 124) or in a non-metropolitan area.
4. The wage models for immigrants include the following variables:
 - cubic in experience
 - years of completed education
 - complete interactions of dummies for 4 education groups (dropout, high school graduate, some college, college graduate) with dummies for 5-year experience bins
 - a dummy for people with an advanced degree
 - a dummy for those who attended school last year
 - dummies for people who worked full time (>1400 hours) or low hours (under 1000)
 - interactions of years of education and a college dummy with 4 broad immigrant groups (Mexicans, Europeans, “high education” Asian source countries, “middle education” Asian source countries) and an Hispanic indicator.
 - dummies for 38 source countries and source country groups
 - interactions of source country dummies with a quadratic in years in the U.S.
 - dummies for living in a smaller MSA/PMSA (not one of the top 124) or in a non-metropolitan area.

Figure 1: Immigrant Presence and Dropout Share

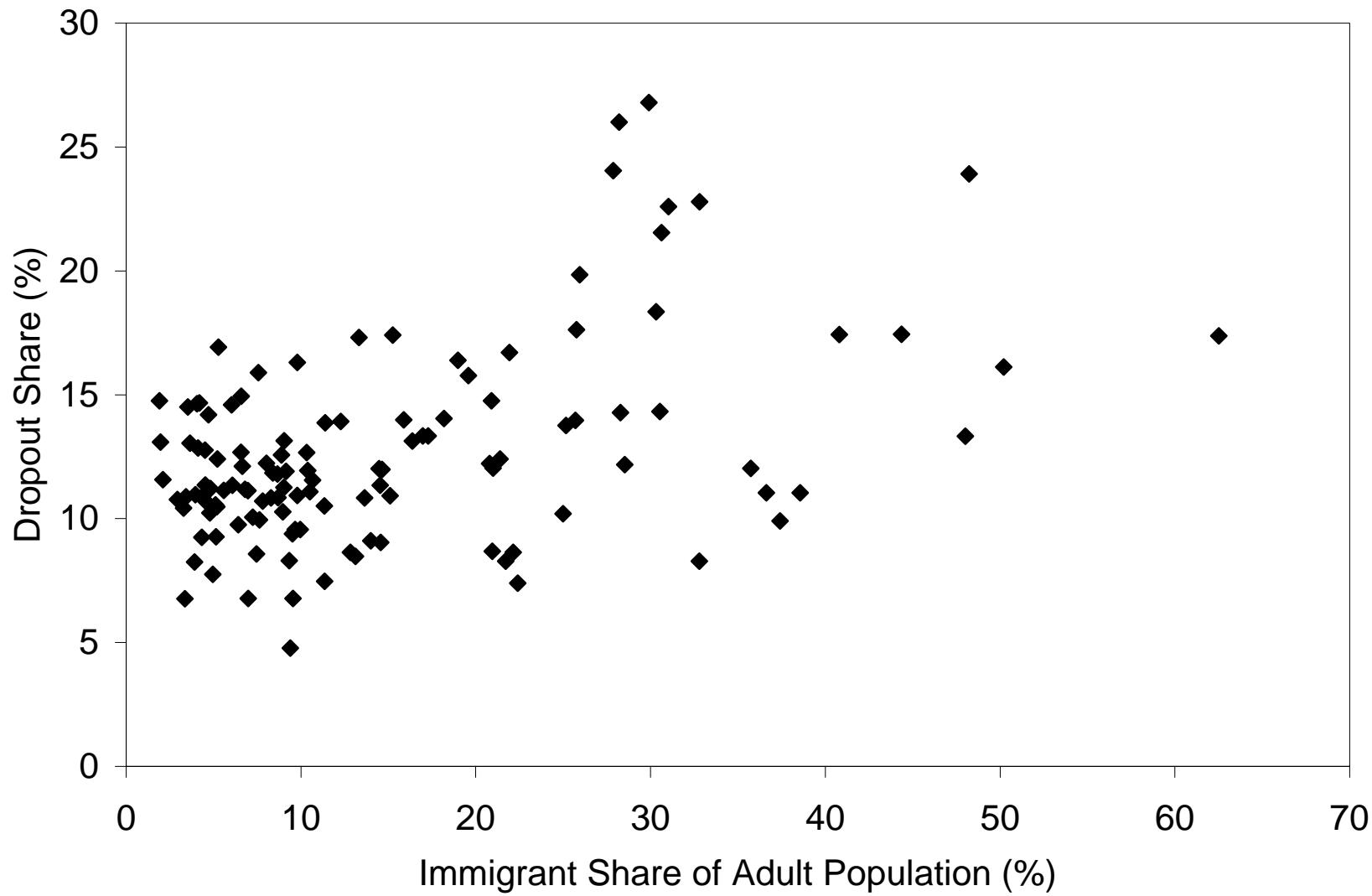


Figure 2: Relative Shares of Filipino Immigrants in Major Cities

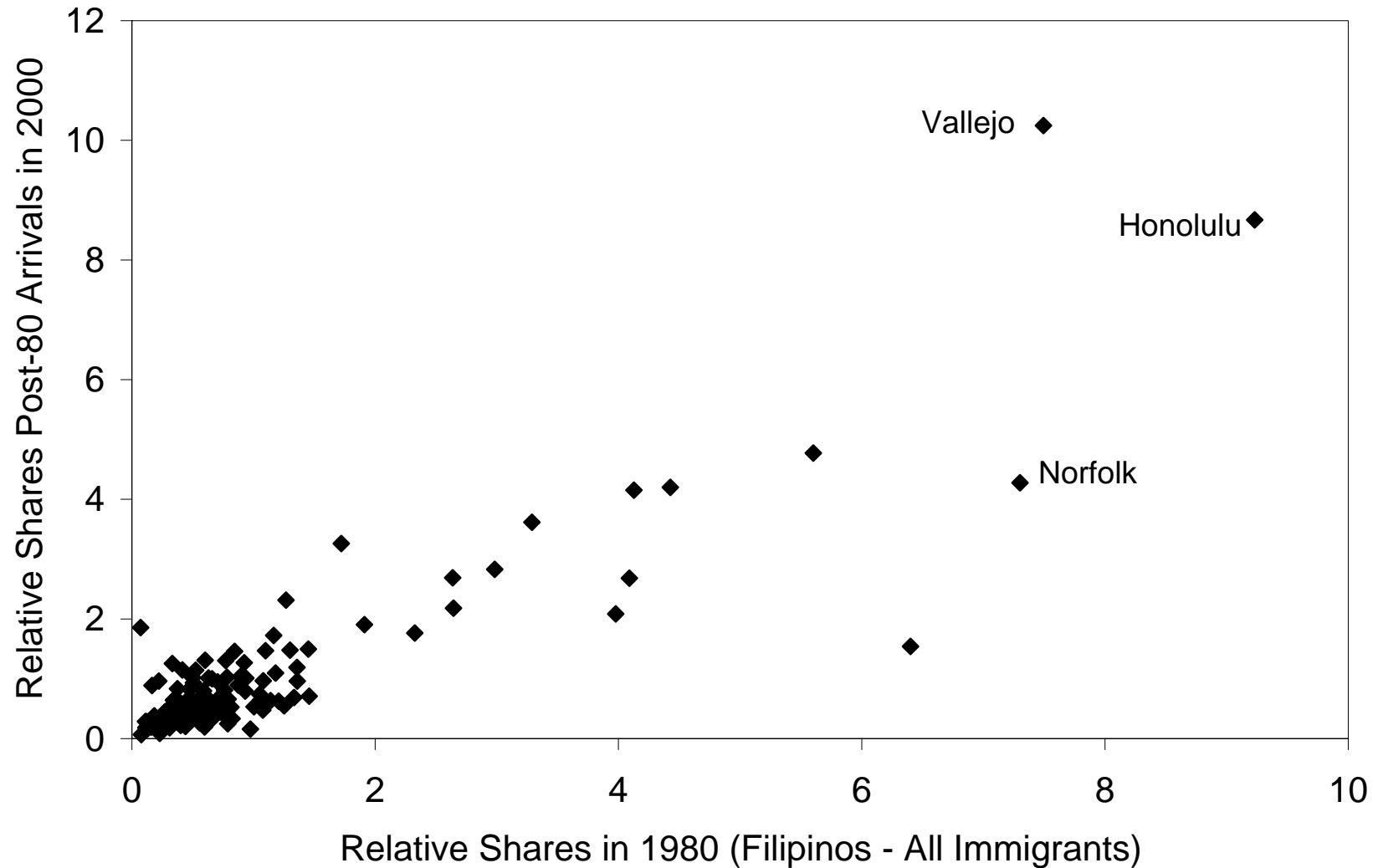


Figure 3: Reduced Form Relationship: Inflow IV and Dropout Wage Gap

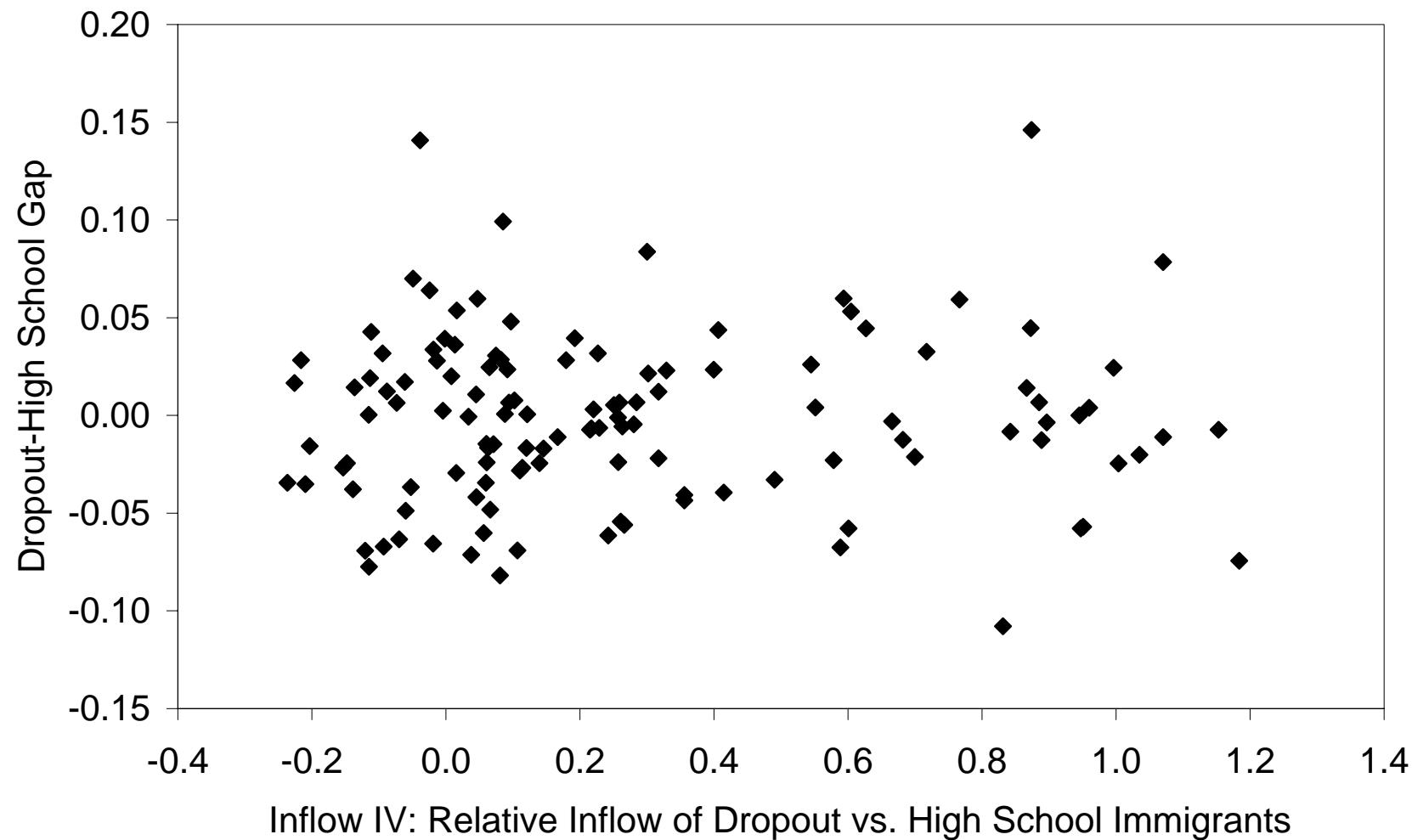


Figure 4: Reduced Form: Inflow IV and Immigrant Wage Gap (College)

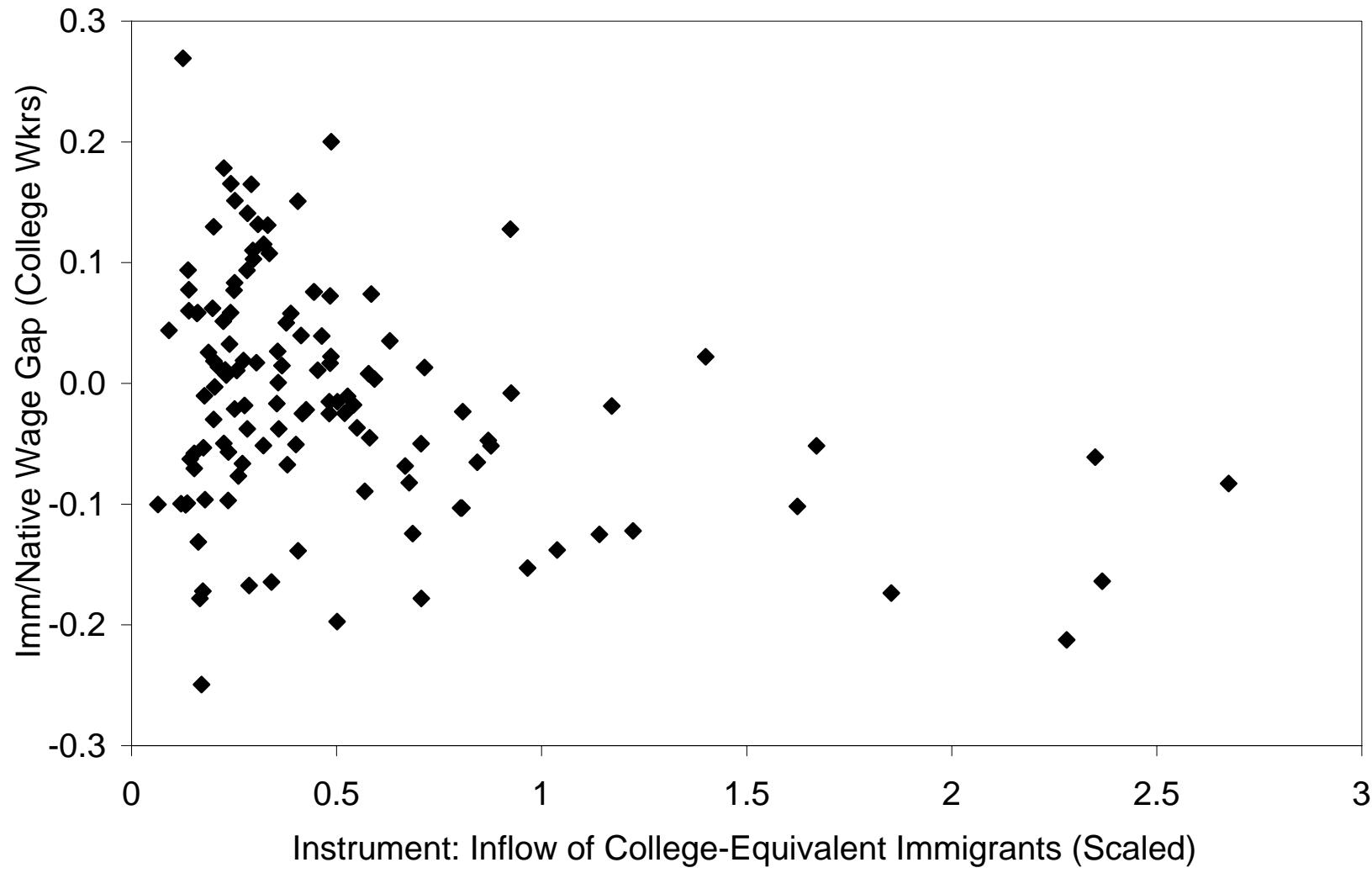


Figure 5: Immigrant Density and Residual Wage Inequality (Native Men)

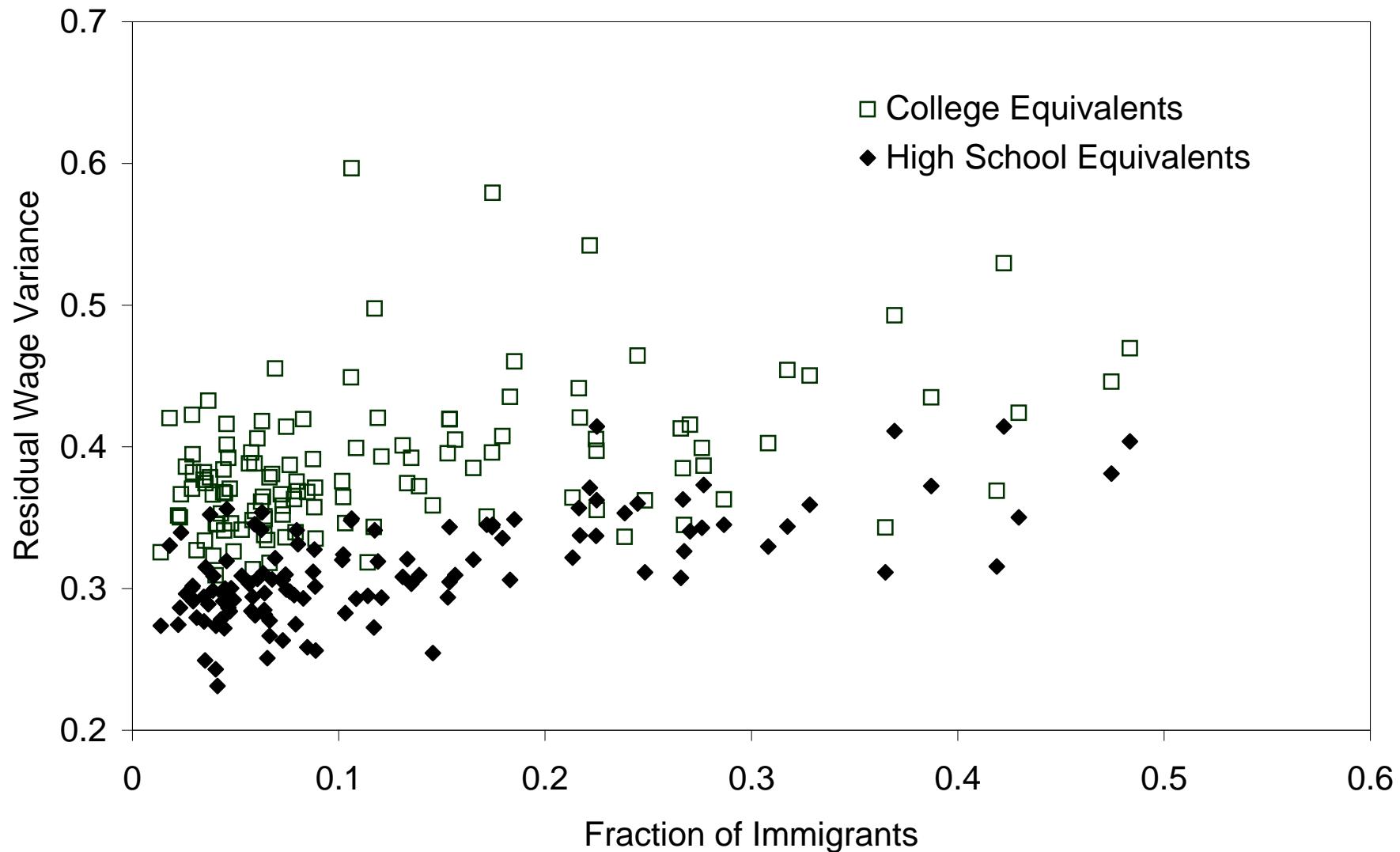


Table 1: Immigrant Presence, Education and Earnings in Large U.S. Cities

	Working Age Population (thousands)	Share of US Pop. (percent)	Percent Immigrants	Education Distribution		
				Less Than High School	College or More	Mean Salary
All US	174,870	100.0	16	14	28	39,806
Larger Cities (124)	116,748	66.8	21	14	31	43,102
Rest of Country	58,122	33.2	7	15	21	32,988
Los Angeles	5,828	3.3	48	24	27	41,260
New York City	5,687	3.3	44	17	35	49,613
Chicago	5,114	2.9	25	14	34	46,522
Washington DC	3,359	1.9	25	10	45	56,076
Atlanta	3,055	1.7	17	13	34	44,110
Philadelphia	3,017	1.7	11	11	33	46,149
Houston	2,904	1.7	31	22	27	41,688
Detroit	2,634	1.5	11	12	27	43,937
Dallas	2,516	1.4	26	20	30	42,545
Phoenix	2,348	1.3	22	17	26	40,949
Riverside County Ca	2,266	1.3	31	23	17	37,409
Boston	2,055	1.2	22	9	46	52,584

Notes: based on tabulation of 2005 and 2006 American Community Surveys. Working age population includes people age 18 or older with 1-45 years of potential experience. Cities are primary metropolitan statistical areas (PMSA's).

Table 2: Characteristics of Immigrants in 2000

	Working Age Population (thousands)	Share of All Immigrants (percent)	Fraction Arrived:		Educational Attainment:			
			After 1980	After 1990	Mean Years Completed	Dropouts	12-15 Years	College or More
Natives	141,272	--	--	--	13.3	14.2	60.6	25.2
Immigrants	23,627	100.0	70.5	39.9	11.6	37.4	38.8	23.8
<u>By Country of Origin:</u>								
Mexico	7,267	30.8	75.1	43.8	8.6	69.8	26.5	3.7
Philippines	1,078	4.6	66.1	31.5	14.1	9.2	43.7	47.0
India	838	3.5	78.4	51.4	15.6	9.6	20.2	70.2
Viet Nam	806	3.4	75.3	39.7	11.7	34.6	45.8	19.6
China	715	3.0	82.0	50.1	13.6	24.2	29.2	46.7
El Salvador	698	3.0	85.1	37.0	8.9	65.0	30.6	4.4
Korea	664	2.8	66.4	33.1	14.0	10.6	45.8	43.6
Cuba	586	2.5	52.3	29.1	12.5	30.0	48.3	21.7
Dominican Rep.	536	2.3	74.2	38.1	10.8	48.8	41.9	9.3
Canada	517	2.2	47.6	31.9	14.3	8.9	49.8	41.3
Germany	455	1.9	32.6	21.0	13.9	8.3	59.3	32.4
Jamaica	429	1.8	66.7	27.3	12.6	23.8	57.8	18.4
Columbia	400	1.7	71.9	40.5	12.5	24.7	53.3	21.9
Guatemala	400	1.7	84.0	45.9	8.8	64.5	30.4	5.1
Haiti	333	1.4	75.1	34.5	11.8	35.2	51.3	13.5
Poland	310	1.3	74.5	42.3	13.3	16.3	58.2	25.6

Notes: based on tabulation of 2000 Census. Working age population includes people age 18 or older with 1-45 years of experience.

Table 3: Summary Statistics for Samples from 1980, 1990, 2000 Census and 2005/2006 ACS

		Mean Education	Years of Experience	Employment Rate (%)	Mean Wage	Variance(Log Wage) Overall	Variance(Log Wage) Residual
<u>Native Men</u>	1980	12.5	18.8	90.1	25.07	0.385	0.288
	1990	12.9	18.7	88.5	23.72	0.462	0.322
	2000	13.2	20.4	86.8	25.86	0.487	0.353
	2005/6	13.4	21.4	86.2	25.35	0.522	0.361
<u>Native Women</u>	1980	12.2	19.6	65.4	16.75	0.317	0.269
	1990	12.8	19.4	74.7	17.05	0.382	0.295
	2000	13.3	20.7	77.1	19.51	0.408	0.313
	2005/6	13.5	21.8	76.8	19.74	0.456	0.335
<u>Immigrant Men</u>	1980	11.6	19.1	87.5	24.49	0.444	0.321
	1990	11.4	18.0	86.5	21.73	0.517	0.347
	2000	11.6	18.8	86.5	23.21	0.557	0.390
	2005/6	12.0	19.9	90.6	21.45	0.544	0.352
<u>Immigrant Women</u>	1980	11.0	20.6	60.0	17.15	0.343	0.291
	1990	11.2	19.9	65.0	16.94	0.414	0.318
	2000	11.7	20.0	64.8	19.27	0.484	0.367
	2005/6	12.2	20.9	67.2	18.58	0.515	0.356

Note: samples include people age 18 or older with 1-45 years of potential experience. Wages are reported in 2007 dollars. Residual wage variance is based on linear prediction models, fit separately by year, gender, immigrant status.

Table 4: Estimated Models for the Relative Wage Gap Between Native Male Dropouts and High School Graduates

	Estimated by OLS:				Estimated by IV:		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Relative Supply of Dropout vs. High School	--	0.00 (0.01)	0.01 (0.01)	0.01 (0.01)	0.02 (0.01)	0.02 (0.01)	0.02 (0.01)
Lagged Dependent Variable	0.29 (0.09)	--	--	0.25 (0.08)	--	--	--
Controls for Log City Size, College Share, Manufacturing Share, and Mean Wage Wage Residuals for All Workers in 1980 and 1990	no	no	yes	yes	no	yes	yes
R-squared	0.09	0.01	0.24	0.29	0.01	0.24	0.29
First-stage t-statistic	--	--	--	--	14.82	14.03	14.16

Notes: standard errors in parentheses. All models are estimated on cross section of 124 larger cities in 2000.

Estimates are weighted OLS or IV, using the 1990 population of the city as weight. Dependent variable is the difference between the mean adjusted log wage of high school dropouts, and the mean adjusted wage of high school graduates. Log relative supply measure is based on annual hours of all dropouts and all high school graduates (men and women, natives and immigrants). Instrumental variable for models in columns 5-7 is the log of the ratio of predicted inflows of dropout immigrants and high school graduate immigrants over the 1990-2000 period, based on national inflows of 38 source country groups and shares of each group in a city in 1980.

Table 5: Estimated Models for the Relative Wage Gap Between Native Male College and High School Graduates

	Estimated by OLS:				Estimated by IV:			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Relative Supply of College vs. High School	--	0.08 (0.02)	-0.04 (0.05)	-0.01 (0.05)	-0.42 (0.28)	-0.41 (0.15)	-0.26 (0.12)	-0.28 (0.12)
Log Relative Supply Lagged 10 years	--	--	0.12 (0.08)	0.01 (--)	--	0.60 (0.21)	0.34 (0.17)	0.28 (--)
Log Relative Supply Lagged 20 years	--	--	0.01 (0.04)	--	--	-0.10 (0.07)	0.01 (0.07)	--
Relative Wage Gap Lagged 10 years	0.66 (0.08)	--	0.66 (0.07)	0.68 (0.06)	--	0.46 (0.12)	0.51 (0.10)	0.56 (0.09)
Relative Wage Gap Lagged 20 years	0.29 (0.06)	--	0.26 (0.06)	0.32 (--)	--	0.45 (0.10)	0.35 (0.09)	0.44 (--)
Controls for Log City Size and Mfg Share in 1980 and 1990	no	no	no	yes	no	no	yes	no
R-squared	0.59	0.09	0.68	0.37	0.02	0.60	0.71	0.24
First-stage t-statistic	--	--	--	--	1.98	4.73	4.87	4.66

Notes: standard errors in parentheses. All models are estimated on cross section of 124 larger cities in 2000. Estimates are weighted OLS or IV, using the 1990 population of the city as weight. Dependent variable is the difference between the mean adjusted log wage of college graduates and the mean adjusted wage of high school graduates. Log relative supply measure is based on annual hours of all college equivalent and high school equivalent workers (men and women, natives and immigrants). Models in columns (4) and (8) are fit in first difference form. Instrumental variable for models in columns 5-8 is the log of the ratio of predicted inflows of college equivalent and high-school equivalent immigrants over the 1990-2000 period, based on national inflows of 38 source countries and shares of each group in a city in 1980.

Table 6: Estimated Models for the Relative Wage Gap Between Immigrants and Natives within Skill Group

	High School Equivalent Workers				College Equivalent Workers			
	Estimated by OLS		Estimated by IV		Estimated by OLS		Estimated by IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Relative Supply of Immigrants/Natives	-0.019 (0.006)	-0.019 (0.006)	-0.023 (0.008)	-0.022 (0.008)	-0.036 (0.011)	-0.029 (0.011)	-0.060 (0.014)	-0.055 (0.015)
Lagged Dependent Variable	--	0.159 (0.060)	--	0.159 (0.060)	--	0.200 (0.100)	--	0.128 (0.106)
Controls for Log City Size, College Share and Mfg. Share in 1980 and 1990, and Mean Wage Residuals for All Natives and All Immigrants in 1980	yes	yes	yes	yes	yes	yes	yes	yes
R-squared	0.308	0.349	0.301	0.338	0.444	0.463	0.455	0.471
First-stage t-statistic	--	--	11.67	11.68	--	--	13.76	12.85

Notes: standard errors in parentheses. All models are estimated on cross section of 124 larger cities in 2000. Estimates are weighted OLS or IV, using the 1990 population of the city as weight. Dependent variable is the difference between the mean adjusted log wage of male immigrants and natives who are classified as "high school equivalent" workers (columns 1-4) or "college equivalent" workers (columns 5-8). Log relative supply measure is based on annual hours of all high school equivalent or college equivalent workers (men and women). Instrumental variable for models in columns 3-4 is the predicted inflow of high school equivalent immigrants between 1990 and 2000, divided by city population in 2000. Instrumental variable for models in columns 7-8 is the predicted inflow of college equivalent immigrants between 1990 and 2000, divided by city population in 2000.

Table 7: Estimated Models for the Effect of Immigration on Relative Within-Group Residual Variance for Native Men

	Estimated by OLS:				Estimated by IV:		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Relative Fraction of Immigrants (College/HS)	--	0.00 (0.03)	0.01 (0.03)	-0.03 (0.05)	0.07 (0.05)	0.03 (0.04)	0.06 (0.22)
Lagged Log Relative Fraction of Immigrants (1990)	--	--	--	0.04 (0.04)	--	--	-0.02 (0.15)
Lagged Dependent Variable	0.51 (0.06)	--	0.47 (0.06)	0.45 (0.07)	--	0.48 (0.07)	0.49 (0.11)
Controls for Log City Size, College Share and Manufacturing Share in 1980 and 1990	no	yes	yes	yes	yes	yes	yes
R-squared	0.34	0.21	0.46	0.47	0.21	0.46	0.46
First-stage t-statistic	--	--	--	--	9.46	9.89	2.32

Notes: standard errors in parentheses. All models are estimated on cross section of 124 larger cities in 2000.

Estimates are weighted OLS or IV, using the 1990 population of the city as weight. Dependent variable is the log of the ratio of the residual wage variance for native men with college or more versus exactly 12 years of education.

Instrumental variable for models in columns 5-7 is the log of the ratio of predicted inflows of college-equivalent and high school-equivalent immigrants over the 1990-2000 period, based on national inflows of 38 source countries and shares of each group in a city in 1980.

Table 8: Summary of Changes in Variance of Log Wages for All Workers and Natives Only

	Variance of Log Hourly Wages:			Percent Immigrants
	All Workers	Natives	Immigrants	
Male Workers				
1980	0.390	0.385	0.444	6.9
2005/2006	0.532	0.522	0.544	18.0
Change	0.142	0.137	0.100	11.1
Female Workers				
1980	0.318	0.317	0.343	6.7
2005/2006	0.466	0.456	0.515	13.9
Change	0.148	0.139	0.172	7.2

Notes: samples include people age 18 or older with 1-45 years of potential experience who have positive wage and salary earnings and no self-employment earnings.