Native mobility, self-selection, and the effect of immigration on wages

Steven Whitlow

September 2018

Abstract

It is well known that studies of the effects of immigration on the wages of U.S. workers that use national time series data have tended to find much greater impacts than those that exploit variation at the state or city level. Less clear is whether national level studies are simply unininformative in determining the effect of immigration on wages or whether measurement error or forces of factor price equalization lead to bias in estimates from area studies. One of the most frequently discussed methods of adjustment is the possibility declines in native net inflows into local labour markets that experience immigration attenuate the supply shock. I note that, to the extent that natives are responsive to immigration, the group of natives who change their work or location choices in response to immigration are likely to be self-selected. In particular, I argue that native workers at the bottom of the wage distribution within skill cells, where immigrants are overrepresented, are the most likely to respond either through internal migration or by exiting the labour force. I implement a standard specification to study the impact of immigration at the state level while correcting for selection into states using a semiparametric two-step method. I find that correcting for selection results in increases in the estimated wage elasticity of at least 30 percent when using two education groups to define the size of the immigrant supply shock facing workers. The impact of the selection correction increases when I add additional selection probabilities to the control function or allow the control function to vary by time. However, the impact of the correction is smaller and less consistent across specifications when using four education groups. The results suggest that state level studies that use coarse skill groups and include a correction for self-selection into states can provide estimates of the impact of immigration on native wages that are similar to those that use a national time series.

1 Introduction

Public perception often links increased levels of immigration to increases in unemployment and decreases in wages, but this is often based on speculation and not evidence. According to a simple supply and demand model with an upward sloping supply curve and a downward sloping demand curve, an increase in the labour force caused by immigration should lead to a decrease in per capita wages and labour supply. However, empirical evidence regarding the impact of immigration on native wages has been mixed. While there is general agreement that the economic impact of immigration on a receiving country is positive on aggregate, there has been considerable debate regarding its distributional effects.

The literature on the labour market impact of immigration can be separated into a few different approaches. The most obvious would be a spatial or "area" approach which exploits disparities in the fraction of immigrants between different regions or cities to look for effects on regional labour markets. One can collapse this regional data into a cross-section and regress average regional wages or regional unemployment on a vector of explanatory variables, such as average educational attainment, as well as the fraction of immigrants in the labour force. The native labour force can be collapsed into finer cells to allow for differing effects of immigration on different sections of the population – allowing one to examine, for instance, the effect of immigration on low-skilled and high-skilled workers within a region.

However, there exist issues with this approach. Immigrants are likely to self-select into the regions with the highest wages, which will result in attenuation of any negative impact of immigration on wages. In addition, native workers are also mobile – not only do they have the same incentive to move to wealthy regions as immigrants, if there is a large positive labour supply shock in one region, they could respond by moving to another region that did not experience that shock. In this way, the impact of migration (whether positive or negative) is diffused across the country.[1] A different approach would be to sidestep these problems by estimating at the national level. Borjas (2003) finds that national level estimates of the wage impact of immigration are two to three times in magnitude as the equivalent state level results, suggesting that forces of factor price equalization cause the wage elasticities estimated in area studies to become biased. Borjas (2006) finds that native outmigration in response to immigration may result in underestimating the wage effect of immigration by around 30 percent in area studies which use U.S. states as boundaries.

My empirical analysis is most related to two recent strands of the literature. The first is that natives and immigrants have previously been found to be imperfect substitutes within skill groups (Card 2009; Ottaviano and Peri 2012). Low-skill native workers in areas exposed to low-skilled immigration tend to work in occupations with relatively more complex task content, which is a potential explanation for imperfect

[1] There are several other channels of adjustment across regions hypothesized in the literature. I discuss several, including the potential impact on native locational decisions, in detail in Section 2.2.2.

substitutability (Peri and Sparber 2009; Amuedo-Dorantes and De La Rica 2011). This is theorized to be a result of the fact that immigrant workers often have limited skills in the primary language in the country to which they are migrating (Peri and Sparber 2009; Imai et al. 2018). If manual tasks require a relatively low degree of language skills, natives will have a comparative advantage in more complex tasks that require a greater degree of communication. Peri and Sparber (2009) argue that if task prices for communication and cognitive tasks are sufficiently higher than those for manual tasks, immigration of low-skilled workers can even increase low-skilled natives' welfare, as natives are freed to shift into occupations with a greater degree of task complexity.

Secondly, attention has been given to investigating whether flows of immigrants into a local labour market crowd out natives. It has long been known that internal migration serves to equilibriate labour markets, at least in part (Blanchard and Katz 1992); however, the prevalence of internal migration in the United States has decreased over time, and low-skilled workers – who should be expected to bear the highest costs as a result of immigration into the U.S. – have found to be less mobile than high-skilled workers (Molloy et al. 2011; Kaplan and Schulhofer-Wohl 2017). Recent results suggest that immigration can play a substantial role in equilibriating labour markets by choosing to live in the regions that offer them the greatest economic opportunities (Røed and Schøne 2012; Cadena 2014; Cadena and Kovak 2016; Amior 2018); however, immigrants may crowd out natives who would have otherwise migrated into these local labour markets (Monras 2015; Dustmann et al. 2017; Amior 2018).

Although most studies of the native migration response to immigration have focused on the change in the size of local skill groups, a natural question that arises in response to the results summarized above is the extent to which the composition of the native workforce *within* skill cells changes in response to immigration. To the extent that (i) tasks can be ordered by complexity and task prices increase with complexity, (ii) workers vary in their ability to complete complex tasks, and (iii) immigrant competition is strongest in occupations with relatively low task complexity, the crowding-out of native inflows is likely to be most strong in those occupations. Further, while it is likely true that a typical low-skilled native worker has a comparative advantage in communication tasks relative to a typical low-skilled immigrant worker, the assumption that abilities are identically distributed within nativity-skill group cells seems overly strong; some natives may have an incentive to a region with less immigrant competition in low-complexity occupations.

This suggests that, to the extent that natives base their locational decisions on utility maximization, native inflows into regions with relatively high levels of immigrant inflows may be *positively* selected and native outflows *negatively* selected, as those working in relatively low complexity occupations and facing the highest levels of immigrant competition choose to live elsewhere. A model of the effect of immigration to a local labour market on native wages that incorporated these concerns would have to account for differences in local labour demand conditions across regions that provide an incentive for both natives and immigrants to relocate and the effect that immigrant locational decisions have on native occupational and locational choice, as well as the decision to participate in the labour market. Simply including controls for occupation or industry would be insufficient, since this approach would conflate the movement of natives between occupations or industries within an area and the movement of natives across regions; the former would not result in differences between area-level and national level studies. Further, to the extent to which this movement across occupations or industries is caused by immigration, this should be considered to be part of the wage impact of immigration.[2]

A simpler approach compatible with the hypothesis stated above is to note that in such a framework, a worker's ability to perform complex tasks is analogous to unobserved ability; to the extent that workers are paid according to their marginal productivity, inferences can be made regarding a worker's unobserved ability directly from their wages. In particular, if native workers who face a greater degree of immigrant competition in the labour market are the most likely to choose to live elsewhere in response to the immigrant supply shock, and immigrant competition is higher at the lower end of the wage distribution within skill cells, then self-selected migration will bias estimates of the wage impact of immigration downward in absolute value, even when ignoring the effect that the displacement has on attenuating the size of the immigrant labour supply shock.

It is natural to consider individuals' decisions regarding where to live and work in the context of a Roy model of locational choice. As in the basic Roy model, to the extent that the returns to unobserved ability differ across locations, geographic sorting of workers will result in the observed mean wage in an area to differ from the underlying population mean, potentially biasing estimates of the wage impact of immigration if immigrant inflows are associated with changes in the distribution of unobserved ability. While many studies have found limited to no native displacement in response to immigration – Card (2005) says "it is very hard to argue that immigration has not had some impact on the fraction of less educated people" in local labour markets – it is important to note that this selectivity bias can exist even if the actual size of the displacement is small.[3]

In this paper, I apply a selection correction method developed by Dahl (2002), who develops an extended Roy model in which workers decide to live and work in one of multiple different areas, choosing the area that provides the highest utility. Utility depends on both the wages and nonwage amenities associated with an area. Dahl

[2]Card (2001) uses discrete observed characteristics at the national level to calculate probabilities for each individual in a city to work in a given occupation, and then sums these probabilities to estimate local supplies of workers; this is designed to account for the fact that native workers may shift out of occupations as a response to increase in supply of workers in that occupation. In a separate section, he examines native inflows and outflows and finds no response to immigration. However, since his occupation classification is quite coarse (six occupation categories, with only two where the mean years of schooling is less than 12 years) and immigrant densities across those occupation categories with lower educational attainment are relatively similar, this approach is unlikely to identify the selection mechanism described here.

[3]In fact, if native inflows are positively selected, this bias could occur without any displacement at all.

proposes a Heckman (1979) type correction method for polychotomous choice models that is semiparametric and simple to implement empirically. The method relies on calculating the *observed* probabilities that an individual of a given set of characteristics chooses to live in each possible location and inserting a polynomial expansion of a subset of these probabilities as additional controls in the regression equation. This technique sidesteps the need to estimate a latent utility function that includes all pecuniary and nonpecuniary factors that influence workers' preferences over locations, as well as the need to specify the joint distribution of the earnings and latent utility error terms.

My results suggest that selectivity bias due to native locational decisions result in substantial underestimation of the wage effects of immigration, though the magnitude of the bias greatly depends on how skill groups are defined. Estimates of the impact of immigration increase by around 30 to 40 percent after including the control function when only two education groups are used in estimation, but the increase is smaller when using finer education categories. The results appear to be driven by relative decreases in the probability that native workers born in immigrant-intensive states choose to stay in their state of birth. I then make small extensions to the selection correction procedure – such as expanding the set of probabilities allowed to enter the control function or allowing the control function to vary over time – which suggest that the impact of correcting for the selectivity bias may be much higher, perhaps as high as 90 percent. The results suggest that native mobility may result in significant selection bias in estimates of the wage impact of immigration obtained by area studies beyond simply attenuating the size of the immigrant labour supply shock.

2 Background

2.1 Canonical model

Most examinations of the effects of immigration treat immigration as a positive labour supply shock. To determine the effect that an increase in the supply of a specific type of worker has on the wages of similar and different types of workers, a theoretical model of the demand side of the labour market is required. Labour supply, however, is typically assumed to be perfectly inelastic, though it has been argued that this is unrealistic.[4]

The standard approach is to use a single-sector framework to model labour demand. Output is produced by a technology with constant returns to scale and with capital being separable from labour. The aggregate production for the economy is given by a

[4]Natives may choose to reduce their hours worked or exit the labour force in response to increased labour supply due to immigration. Further, as will be discussed in detail later, natives may respond to migration into a labour market by moving to another region.

nested CES function where the outermost nest is typically Cobb-Douglas, such as

$$Q_t = A_t L_t^{(\alpha_t)} K_t^{(1-\alpha_t)}, \tag{2.1}$$

where Q_t is aggregate output in the economy, A_t is total factor productivity, K_t is capital, L_t is aggregate labour input, and $\alpha_t \in (0,1)$ is the labour share of output.[5] In this model, changes in the relative supplies of different skill groups can be linked to changes in their relative wages. This approach has a history in labour economics (see Katz and Murphy 1992 or Card and Lemieux 2001), and was first used in the study of the labour market impacts of immigration by Borjas et al. (1992) and Borjas et al. (1996).[6] While in these studies labour demand is modelled at a national level, some studies of immigration have used a similar framework to model local labour demand at the level of local labour markets (Card 2001, 2009).

A key drawback of this model is that it does not allow for capital-skill complementarity. The additional assumption that the outermost nest is Cobb-Douglas is largely innocuous in practice, however, since the separability of capital allows most studies of immigration to sidestep capital in empirical estimation.

In the simple case where labour input L_t is comprised of only one skill group and the capital stock is fixed, then an increase in this single type of worker results in a decrease in wages and an increase in the rental rate. If capital is perfectly elastic, then the capital stock increases in response to immigration, maintaining a constant rental rate r (and a constant capital share of output). Since the former assumption that the capital stock is permanently fixed is clearly unrealistic, the effect estimated under this assumption are best understood as the short-run impact of immigration; the long-run wage impact of immigration is the effect remaining after capital is allowed to adjust to the supply increase.[7]

In the simple one-skill model described here, wages are unchanged in the long run. When the model is extended to allow for more than one skill, immigration can have distributional effects: while *average* wages are unchanged, there can be changes in the *relative* wages of different skill groups. The Cobb-Douglas function in (2.1) has constant returns to scale, which implies that immigration where the inflow of new workers has the same skill distribution as the existing labour force will have no effect on either the average wage or the wage of any skill group. Borjas (2003) referred to this as "balanced" immigration. However, immigration in which the inflow of new workers is distributed differently across skill groups will result in changes in the relative wages

[5]This production function can be used to model output in a national economy, but also at a regional or local level: in the immigration context, see Borjas (2003) or Aydemir and Borjas (2007) at the national level, and Card (2001, 2009) or Cortes (2008) at the local level.

[6]Earlier studies of the labour market allowed for more complicated production technologies (Johnson 1980; Altonji and Card 1991).

[7]The relevance of the fixed capital case has been disputed. Ottaviano and Peri (2012) argue that, because the increase in supply in any given year in the US due to immigration is small relative to the size of the labour force and is typically similar (and thus predictable by firms) from year to year, a considerable degree of capital adjustment in response to immigration occurs even in the short run.

of skill groups, even in the long run, based on which groups are overrepresented in the inflow. However, this unbalanced immigration will result in no change to the average wage, since the marginal productivity of aggregate labour L_t is unchanged after capital adjusts.

The critical issue is thus the construction of the CES labour aggregate L_t. This is not trivial as it means making assumptions about the cross-skill elasticities of various types of workers. While the nested CES function allows for imperfect substitution between different workers of the same broad skill group, workers that share all of the attributes included in the CES nests are perfect substitutes.

For instance, take the simple and common approach where labour is divided into just two intermediate inputs, highly educated ("skilled", L_H) and less educated ("unskilled", L_U):

$$L_t = \left[\theta_t L_{Ht}^{\rho} + (1-\theta_t) L_{Ut}^{\rho}\right]^{1/\rho} \tag{2.2}$$
$$\rho = \frac{(\sigma_s - 1)}{\sigma_s},$$

where $\theta_t \in (0,1)$ is a time-variant weight that determines the relative productivity of skilled and unskilled labour and σ_s is the elasticity of substitution between the two groups. The unskilled worker category consists of high school dropouts and graduates, and the skilled worker category consists of college graduates. Workers with some postsecondary education are split between the two groups.[8] The CES aggregate allows for imperfect substitution between different skill groups; worker types that share a skill group are perfect substitutes. Thus, in this two-skill model, high school dropouts and graduates are assumed to be perfectly substitutable.[9]

While this two-skill model is simple and easy to implement empirically, it has been argued that two skill groups are not enough, especially within the context of determining the labour market impact of immigration. The immigrant share of labour has increased in the US across all levels of educational attainment. However, it is well known that recent immigration has increased the relative supply of high school dropouts in particular: while most native-born workers without any post-secondary education hold a high school degree, most immigrants do not. Therefore, if there is even a small degree of imperfect substitutability between high school graduates and dropouts, a two-skill model will underestimate the negative wage impacts of immigration on the latter group. As a result, many studies prefer a model with four education groups: high school dropouts, high school graduates, workers with some post-secondary education, and college graduates.[10] More generally, and following

[8]The assignment of workers with some postsecondary education to a category is less clear. Ottaviano and Peri (2012) assign all workers with some postsecondary education to the skilled labour group. Card (2009, 2005) divides the supply of these workers equally between each group.

[9]As will be discussed further, since recent immigration to the US has increased the relative supply of dropout labour, this is an important (and disputed) assumption.

[10]For the case of the United States, studies that use a four-skill framework include Borjas (2003), Borjas and Katz (2007), Aydemir and Borjas (2007), Jaeger (2007), and Ottaviano and Peri (2012). Manacorda et al. (2012) uses the same skill groupings to examine the effect of immigration into the United Kingdom.

Borjas (2003),

$$L_t = \left[\sum_{s=1}^{S} \theta_{st} L_{st}^{\rho} \right]^{1/\rho}, \tag{2.3}$$

where L_{st} indexes the labour input of workers with education s at time t, and $\sum_i \theta_{st} = 1$ are time-varying weights which determine the relative productivity of each education group. As before, the elasticity of substitution between different education groups, σ_s, is determined by the equation $\rho = (\sigma_s - 1)/\sigma_s$. An issue that is immediately apparent with this specification is that there is no variation in the elasticity of substitution between different education groups.[11]

Even with four different education categories, there is not enough variation at the national level to meaningfully identify the effect that the immigrant inflow has on wages. However, workers who have the same level of education but differ in experience are unlikely to be perfect substitutes. Indeed, there is evidence that within an education group, workers are closer substitutes for those who have a similar level of experience (Welch 1979; Card and Lemieux 2001). Borjas (2003) notes this and adds that, since the size of the supply shock caused by immigration differs by education group but also by levels of experience within education groups.[12] Classifying skill groups by combining levels of education and levels of experience to form education-experience groups can greatly increase the variation with which to credibly identify the impact of the supply shock. This is particularly useful for analyses that examine a national level labour market, like that in Borjas (2003), which would otherwise have very few data points if skill groups were to be delineated only by education level.[13] Building on (2.4), assume that aggregate labour supply of each education group is given by a CES aggregate of workers with identical education but different levels of experience (Card and Lemieux 2001; Borjas 2003):

$$L_{st} = \left[\sum_{r=1}^{R} \lambda_{sr} L_{sr}^{\eta} \right]^{1/\eta}, \tag{2.4}$$

where L_{srt} indexes the labour input of workers with education s and experience r at time t, and $\sum_r \lambda_{sr} = 1$ are time-invariant weights which determine the relative productivity of each experience group. The elasticity of substitution between different experience

[11] This implies that, for instance, high school graduates are equally substitutable with high school dropouts and college graduates. An alternate specification, proposed by Ottaviano and Peri (2012) proposes an alternate CES form with two nests, the outermost nest being the "skilled" and "unskilled" specification described earlier, and an inner nest consisting of "some college" and college graduate groups for the "skilled" group and high school dropouts and graduates for the "unskilled" group.

[12] For example, Figure I of Borjas (2003) shows that while the overall immigrant share of high-school dropouts increased from 1960 to 2000, this increase is disproportionately among young dropouts: about half of dropouts with less than 20 years of potential experience were immigrants in 2000, compared to 30% of those with 35 to 40 years. In 1960, this pattern was reversed, with the immigrant share of dropouts increasing with potential experience.

[13] The division of each education group into 8 education-experience groups provides Borjas with 160 observations instead of just 20.

groups, σ_r, is determined by the equation $\eta = (\sigma_r - 1)/\sigma_r$.[14] In the Appendix, I show that the wage elasticity – the change in wages in an education-experience cell that occur after an increase in labour supply in that cell due to immigration – can be given by

$$\begin{aligned}\frac{dW_{srt}}{dm_{srt}} &\approx \left[(\alpha_t - \rho)\frac{\alpha_{srt}}{\alpha} + (\rho - \eta)\frac{\alpha_{srt}}{\alpha_{st}} + (\eta - 1)\right] \\ \frac{dW_{srt}}{dm_{srt}} &\approx \left[(1 - \rho)\frac{\alpha_{srt}}{\alpha} + (\rho - \eta)\frac{\alpha_{srt}}{\alpha_{st}} + (\eta - 1)\right],\end{aligned} \tag{2.5}$$

in the short run and long run respectively, where α_t is the labour share of income, define α_{st} as the share of total income accruing to workers of education s, and α_{srt} as the share of total income accruing to workers of education s and experience r. Note that these are *partial* effects: they only represent the effect of immigration of a workers' own type, ignoring the complementarities that arise from immigration of workers of different skill. The total effect of immigration is thus likely to be more positive, but impossible to estimate in reduced-form except in the simple case where there are only two skill groups. I discuss this in more detail in Section 2.2 and the Appendix.

Of course, one could make arguments to divide workers into finer skill groups based on characteristics other than education and experience. If immigrant and native workers are imperfect substitutes, then a disproportionate amount of the negative wage impacts of an immigrant inflow would fall on existing immigrants (Cortes 2008; Card 2009; Ottaviano and Peri 2012); if the degree of imperfect substitution was large, this would explain a lack of a large wage impact being detected among native workers as a result of immigration.[15] There is also a large body of empirical evidence for this idea of positive spillovers from human capital at the level of local labour markets, and many studies have found that skilled immigration is associated with increased employment opportunities for natives of a similar level of skill (Peri et al. 2015; Moser et al. 2014; Kerr et al. 2015).[16]

Peri and Sparber (2009) construct indices for occupations that evaluate their content of manual and communication intensive tasks, and argue that, due to language barriers, low-skilled immigrants have an absolute disadvantage in communication intensive tasks and thus a comparative advantage in manual tasks relative to low-skilled natives. They find that in cities with higher levels of immigration, natives tend to specialize more in communication intensive occupations. However, natives and immigrants with high levels of education have generally been found to be more imperfectly substitutable in production than natives and immigrants of low skill.[17]

[14] Note that the productivity parameters at the lowest level of the CES nest are not time-varying; in the empirical estimation, the impact of immigration will be identified off changes over time in education-experience skill cells.

[15] Borjas et al. (2008, 2012) contest these findings, arguing that the findings of imperfect substitutability in Ottaviano and Peri (2012) are sensitive to the removal of employed high-school students. Others, such as Jaeger (2007) have also found that natives and immigrants are perfect substitutes.

[16] Kerr et al. (2017) provides a review of the literature on this topic.

[17] An exception is Cortes (2008) who finds an elasticity of subsitution between low-skill natives and

2.2 Previous literature

2.2.1 Wage impact of immigration

The model in the previous section suggests that immigration changes wages of native workers by altering the relative scarcity of labour skill groups. In the 1990s, there was a growing awareness of growth in income inequality in the previous decade and, in particular, growth in the earnings gap between high-school and college educated workers (Bound and Johnson 1992; Katz and Murphy 1992). In this context, attention began being directed towards increases in unskilled immigration as a partial cause of these changes in the wage structure (Altonji and Card 1991; LaLonde and Topel 1991).

The immigration literature can be divided into categories based upon the level of aggregation that they define as a labour market. The "area approach" to empirical estimation of the impact of immigration defines labour markets at the subnational level; there is variation in the level of exposure to immigration by region, and this should be able to be used to determine its labour market impact.Immigrants are likely to choose to live in the cities with the best economic opportunities, however, which if left unaddressed will result in biased estimates. Altonji and Card (1991) and Card (2001, 2009) instrument the immigrant share variable using a shift-share variable that relies on the tendency of immigrants to cluster into enclaves - immigrants from a certain country often choose to settle into the same areas as previous immigrants from that country. The shift-share instrument uses the settlement patterns of immigrants from each source country to city k at some reference date at time $t - r$, and predicts the inflow rate of immigrants into city k at time t if immigrants from each given source country had identical settlement patterns as they did at time $t - r$.[18] This immigrant shift-share instrument has since become standard in studies that examine spatial variation in immigrant density that is not plausibly exogenous.

Except in the simple two-skill case, reduced-form estimates of the total impact of immigration will only be valid if the skill distribution of immigrants is equal across time and space, which in general is not the case. As a result of this issue, area studies of immigration typically seek to identify the partial effect of immigration, (2.5), which determines the direct effect of an increase in the supply of a worker's own type. The results from these studies that measure the total effect of immigration are not comparable to those that measure the partial effect, since what they attempt to measure is different. There have been many area studies of immigration that attempt to measure this effect.[19] The wage impact of immigration in these studies is limited at most; Card

immigrants of just 1.24. This results in a 10 percent increase in the number of low-skill immigrants reducing the wages of existing low-skill immigrants by 8 percent, but a reduction in wages of low-skill natives of only 0.6 percent. Most other studies show a much smaller degree of imperfect competition between these workers, however. For instance, Card (2009) finds an value of 40 for this elasticity, and an elasticity of 17 between high-skill immigrants and natives.

[18]This instrument is very similar to the Bartik instrument (Bartik 1991), which uses changes in industry employment growth at the national level to identify exogenous variation in local employment.

[19]Friedberg and Hunt (1995) provide a review of early attempts in the literature. Comprehensive

(2001) describes the literature by saying "a 10-percentage-point increase in the fraction of immigrants is estimated to reduce native wages by no more than 1 percentage point"; Kerr and Kerr (2011) describe the estimated elasticities as "small and clustered near zero".

A different approach would be estimate the impact of immigration using a national time series. This avoids the potential endogeneity concerns associated with area studies, but giving up regional variation poses problems for identification. Borjas et al. (1992) and Borjas et al. (1996) and Jaeger (2007) impose a nested CES function similar to that in the previous section, with imperfect substitutability between skilled and unskilled labour; after imposing this technology, one can then estimate elasticities and use these to simulate the impact of an increase in the labour supply caused by immigration. With perfect substitutability between experience groups, the nested CES function allows the large number of cross-skill elasticities to be expressed as a function of only two variables (see footnote 54).[20] The total impact of immigration for a given skill group is simply the sum of the cross-skill elasticities – which depend on the elasticities of substitution between education groups, the elasticity of substitution between labour and capital, and the shares of income accruing to the various skill groups – multiplied by the percentage change in the labour supply for each respective skill group during the period of interest. In other words, this approach takes the assumptions implied by the CES function and estimates of elasticity of substitution between education groups and estimates what wages would have been in the case of zero immigration. Initial results from this approach suggested that the wage impact of immigration was underestimated by area studies.[21]

Borjas (2003), following Card and Lemieux (2001), allowed for imperfect substitutability within skill groups between workers of differing experience. This change in assumptions allows for a sufficient number of observations to perform regressions similar to those in area studies at the national level. The results are striking: the national level estimates of the wage impact of immigration are two to three times the size in absolute value as the equivalent state level results. This suggests that the more negative results provided by earlier national level studies were not the consequence of their implicit theoretical assumptions, but instead the result of forces of factor price equalization causing the wage elasticities estimated in area studies to become biased. Further studies by Aydemir and Borjas (2007, 2011) reached similar conclusions, roughly a 10 percentage point increase in skill cell labour supply due to immigration is found to decrease native wages by 3 to 4 percent.

reviews are also provided by Longhi et al. (2005, 2008), Okkerse (2008), and Kerr and Kerr (2011).

[20] "Skilled labour" is either defined as in the previous section, or alternatively defined as those with at least a high school degree with the "unskilled" category consists solely of high school dropouts.

[21] Ottaviano and Peri (2012) and Manacorda et al. (2012) show that under slightly different assumptions - namely, allowing natives and immigrants to be imperfect substitutes, this approach can give similar results to area studies. However, this does not address the fact that, as mentioned below, estimates of the *partial* effect of immigration are larger in absolute value when estimated at the national level.

2.2.2 Channels of factor price equalization

As a result of the disparities between national and area studies, attention has been given to determining the extent to which trade and migration of native workers result in arbitrage of the labour market impacts of immigration. There are various channels in which labour markets can be integrated, either through trade in goods or through factor mobility, that would allow markets to adjust to immigration without changes in relative wages. If we allow local markets to produce multiple goods with differing levels of skill intensity, which can be traded between markets, then a Heckscher-Olin model would suggest that adjustment to changes in the skill mix in a region caused by immigration can be accomplished simply by an increase in the production of goods that more intensively use workers of a given skill level, without any changes in wages or within-sector skill utilization (Rybczynski 1955). However, available evidence within the context of U.S. immigration suggests that most adjustment to skill-mix changes occurs within, rather than between, industries (Lewis et al. 2004, 2005; Card and Lewis 2007).

However, it is fair to say the mobility of native workers is the method of adjustment that has received the most attention, and is the most related to this paper. Natives may respond to an inflow of immigrant labour into a region by choosing to live in a different area that did not experience that inflow. This would result in the the labour supply shock resulting from immigration being diffused to labour markets that did not experience an inflow of immigrants, but instead an inflow of native migrants, and thus bias estimates of native wage impacts from area studies.

The importance of native internal migration as an adjustment channel to immigration is disputed (Filer 1992; Card and DiNardo 2000; Card 2001; Borjas 2006; Wozniak and Murray 2012; Glitz 2012). There are various issues that make empirical estimation of the native migration response difficult, but the most obvious is shared with the empirical estimation of the wage impact on natives – immigrants have an incentive choose to live in areas which offer them the best economic opportunities. This means that controlling for local demand conditions is crucial to identifying the native migration response to immigration. Other issues may arise based on the empirical specification used. If the count or share of the native workforce of a given skill type is used as the dependent variable, then there is the possibility of conflating cohort effects with migration effects.[22] Peri and Sparber (2011) present simulation evidence that some specifications may exhibit bias which is increasing in the standard deviation in the sample of the log population size. Similarly, specifications that examine net or gross native inflows or outflows are likely to experience similar bias issues.[23]

[22] Hunt (2017) finds that immigration to a state is associated with increased high school graduation rates, suggesting that the cohort effect is not negligible.

[23] Internal migration flows have to balance; a movement between a highly populated region to a less populated region, for instance, will require a higher inflow rate for the latter than the outflow rate for the former, even though the number of workers moving to and from each region is identical. Variation in population sizes between regions suggest that a similar bias may exist for inflow and outflow rates, especially if immigrants move to high population states (or cities).

Most studies find that the native outflow response is limited; dynamic models of internal migration have found high fixed costs of moving to a new state (Kennan and Walker 2011); this suggests that outflows of natives in response to immigration are likely to be limited unless there are many natives who are already close to being indifferent about moving. Recently, however, more attention has been given to the relationship between immigration to regions and native inflows. This is related to the hypothesis presented by Borjas (2001): that new and past immigrants are less attached to specific regions than natives and thus "grease the wheels" of a national labour market by arbitraging differences in wages across space. Since low-skill native workers in general have a low propensity to migrate – either because of higher psychic costs of moving or because the wage gains from doing so are less than a fixed moving cost – immigrants could improve labour market efficiency if the internal migration of natives is insufficient to equalize wages. For instance, Cadena and Kovak (2016) find that the exposure of low-skill native workers to negative local demand shocks during the Great Recession was greatly reduced by the reallocation of Mexican immigrants, who have a much greater propensity to migrate. Cadena (2014) finds that reallocation of low-skill immigrants to states with lower minimum wages attenuates the negative employment impacts of minimum wage increases.

Reallocation of immigrant flows can compensate for low responsiveness of natives to geographic differences in economic opportunity, but a natural question that arises is whether immigration crowds out native internal migration. Recent evidence suggests that some degree of crowding out may occur. Monras (2015) finds that, in the 1990s, native internal migrants in the United States tended to avoid states with high exposure to Mexican immigration. Dustmann et al. (2017) find that reductions in native employment after Czech workers were allowed to work in Germany were almost entirely driven by reduced inflows.

If native workers responded to immigration solely by outflows from the effected region, not only would immigration have some degree of a negative wage impact on effected skill groups – which could be recovered from estimation at the national level – moving costs would fall on a certain segment of the native workforce. Most studies, however, find positive or only weakly negative native outflows in response to immigration. The effect that decreased native inflows have on native welfare is more ambiguous. While immigration into a region would reduce the wages of effected skill groups, in the absence of immigration, some degree of factor price equalization could have taken place through native inflows. The effect on the wages of those who did not move to a region but would have in the absence of immigration is unknown, since workers' wages are only observed for the region in which they currently live and work.

3 Selection correction

The findings summarized above describe the existing findings about the native wage response to immigration and the channels through which this wage impact can be arbitraged. While the the extent to which native workers respond to immigration through internal migration has been researched in detail, the extent to which this locational response may result in changes in the distribution of unobserved ability across regions has not. To address this, I adopt the approach of correction for selection on unobservables across regions proposed by Dahl (2002). Dahl presents an extended Roy (1951) model of mobility and earnings and, building on earlier work on polychotomous choice models by Lee (1982, 1983) and Ahn and Powell (1993), develops a semiparametric two-step correction method that is tractable even in the case of a large number of choices and simple to implement empirically. Dahl then applies this correction to determine the effect that Roy-style sorting across space on state-specific returns to education; the application that I propose of Dahl's method is quite similar.

Note that while I allow for natives to vary in their unobserved ability in my empirical estimation, I continue to assume the distribution of immigrants' unobserved ability is identical across states.

Section 3.1 briefly describes the Roy model that underlies Dahl's paper. Section 3.2 describes the empirical implications of the model and the assumptions that allow for a simple selection estimator. Section 4 describes how this selection estimator can be easily added to the basic specification of area studies of the wage impact of immigration.

3.1 Roy model of locational choice

In this section, I very briefly describe the model of selection presented by Dahl (2002), which extends the original Roy model as it allows for multiple choices of sector and for a worker's decision of location to be motivated both by wage and nonwage factors. For ease of exposition, I drop all time subscripts; in my empirical implementation, I first assume that the control function is time invariant, but I allow it to vary by year in Section 6.2.2.

Earnings and preferences Individuals live for two periods; in the first period, individuals are born in a random state. In the second period, they choose a location in which to live and work based upon which state will provide them with the highest utility; utility is comprised of both earnings and nonwage amenities. Assume that the log earnings, w_{ik} of an individual i who was born in state j and lives in state k is given by

$$w_{ik} = \beta_{0k} + x_i\beta_{1k} + u_{ik}, \qquad (k = 1, \ldots, K) \tag{3.1}$$

where β_{0k} is a scalar constant for state k, x_i is a vector of individual characteristics which includes level of education and potential experience, and u_{ik} is an error term. Since wages are only observed in the state k that individual i chooses to work, the sample is self-selected and the error term u_{ik} will generally not be mean-zero. Individuals have preferences both over earnings and non-earnings utility:

$$V_{ijk} = w_{ik} + t_{ijk}, \qquad (j = 1, \ldots, J), (k = 1, \ldots, K) \tag{3.2}$$

where V_{ijk} indexes utility for individual i born in state j who moves to state k, and t_{ijk} is a taste vector which comprises all factors other than log earnings that influence the utility of living in state k, given that the individual's was born in j. This includes pecuniary components of utility such as moving costs, but also non-pecuniary components such as the psychic cost of moving from j to k, differences in the amenities associated with each location, climate, and so forth. Note that the nonwage component of utility t_{ijk} is allowed to depend on state of birth j while wages are not; this assumption will later be used as an exclusion restriction to identify the control functions . The components of (3.2) can be expressed in terms of deviations from the underlying population mean:

$$w_{ik} - \mathbb{E}(w_{ik}|x_i) = u_{ik} \tag{3.3}$$

$$t_{ijk} - \mathbb{E}(t_{ijk}|z_i) = \epsilon_{ijk}, \tag{3.4}$$

where z_i is a vector which contains x_i but also all other variables that could influence location decisions, including the individual's state of birth, and ϵ_{ijk} is an error term which captures deviations from mean the mean tastes of individuals with identical characteristics.[24] We can rewrite (3.2) using (3.3) and (3.4) as

$$V_{ijk} = \underbrace{\mathbb{E}(w_{ik}|x_i) + \mathbb{E}(t_{ijk}|z_i)}_{V_{jk}} + \underbrace{u_{ik} + \epsilon_{ijk}}_{e_{ijk}}, \tag{3.5}$$

where V_{jk} is referred to by Dahl as the subutility function and is analogous to the conditional value function in the dynamic discrete choice literature, and e_{ijk} is an individual error term.

Selection rule Define d_{ijk} as a dummy variable which equals one if a worker i who is born in state j chooses to live in state k and zero otherwise and assume that the set of total state utilities, $V_{ij1}, \ldots, V_{ijK}$, has a unique maximum. Thus it is clear that

$$d_{ijk} = 1 \Leftrightarrow V_{jk} + e_{ijk} > V_{jq} + e_{ijq} \qquad \forall q \neq k, \tag{3.6}$$

since individuals choose to live in the state which maximizes their utility. We can rearrange (3.6) as:

$$d_{ijk} = 1 \Leftrightarrow V_{jq} + e_{ijq} - V_{jk} - e_{ijk} < 0 \qquad \forall q \neq k. \tag{3.7}$$

[24]Clearly, identification will come from the exclusion of variables in z_i from xi. I describe my exclusion restriction in more detail in Section 4.

Selectivity bias Wages are only observed in state j for those individuals where all K selection equations in (3.7) are satisfied. This implies that the error mean term will in general be nonzero, since

$$\mathbb{E}\left[u_{ik}|d_{ijk}=1\right]=\mathbb{E}\left[u_{ik}|V_{jq}+e_{ijq}-V_{jk}-e_{ijk}<0,\forall q\neq k\right], \tag{3.8}$$

so an OLS regression in which the variable of interest is correlated with $\mathbb{E}\left[u_{ik}|d_{ijk}=1\right]$ will provide biased estimates. Dahl refers to this error mean term as the *selectivity bias* for individual i. In the original Roy model with only two sectors and no nonwage determinants of utility, the error mean term is positive - workers choose the sector that gives them the highest wages. In this extended model, individuals may choose to live in an state that offers lower wages but higher nonwage amenities.

3.2 Selection correction procedure

The model presented in Section 3.1 would be very difficult to implement empirically without additional assumptions, since the sign and magnitude of the bias depends on the joint distribution of u_{ik} and the error terms of the selection equations $e_{ij1}-e_{ijk},\ldots,e_{ijK}-e_{ijk}$. There are two key problems with correcting the earnings equation for selection bias. The first is that a parametric correction method would require an assumption on the functional form of this joint distribution and the estimation of a $(K-1)$–dimensional integral, which becomes infeasible as K increases. Secondly, the subutility function V_{jk} depends both on earnings but also on individual tastes, so the researcher would be forced to determine all amenity variables that could determine the utility associated with a location and obtain data on these in order to estimate V_{jk}. Dahl uses a reformulation originally proposed by Lee (1983) to reduce the dimensionality of the problem, then states an assumption under which the error mean term can be described as a function of only the first-best probability in the earnings equation. For purposes of brevity, I only give an intuitive description of this procedure; readers should consult Dahl (2002) for a formal exposition.

The error mean term in (3.8) requires information on the joint distribution of both the earnings error, u_{ik}, but also all the individual utility errors, $e_{ij1},\ldots,e_{ijK}$, which determine which state is chosen. Lee (1983) notes that the selection rule in (3.7) can be simplified by using maximum order statistics. First, notice that

$$W_{ijk}\text{ observed}\Leftrightarrow\max_{q=(1,\ldots,K)}(V_{jq}+e_{ijq}-V_{jk}-e_{ijk})\leq 0, \tag{3.9}$$

where equality holds only in the trivial case where $q=k$. If we define

$$h_{ijk}=\max_{q=(1,\ldots,K)}(V_{jq}+e_{ijq}-V_{jk}-e_{ijk}),\qquad\qquad\forall q\neq k \tag{3.10}$$

then we can express (3.8) in simpler terms:

$$\mathbb{E}\left[u_{ik}|d_{ijk}=1\right]=\mathbb{E}\left[u_{ik}|h_{ijk}<0,V_{j1},\ldots,V_{jK}\right], \tag{3.11}$$

since (3.9) and (3.10), together with the fact that the set $V_{j1}, \ldots, V_{jK}$ has a unique maximum, imply that $h_{ijk} < 0 \Leftrightarrow d_{ijk} = 1$. We are thus able to replace the individual errors, $e_{ij1}, \ldots, e_{ijK}$, with a single random variable, h_{ijk}, in this expectation. The error mean term now depends on the joint distribution of the earnings error u_{ik} and the maximum order statistic h_{ijk}, as well as the K subutility functions. We could proceed and include this control function directly in our estimation if we were prepared to impose assumptions on the joint distribution of u_{ik} and h_{ijk} and how they depend on the K subutility functions.[25] However, if K is sufficiently large, any such parametric procedure would be impacted by the curse of dimensionality. Dahl proposes an alternate method to correct for selectivity bias that avoids parameterizing the selection mechanism, beginning by noting that

$$\mathbb{E}\left[u_{ik}|h_{ijk}<0,V_{j1},\ldots,V_{jK}\right]=\mathbb{E}\left[u_{ik}|h_{ijk}<0,p_{ij1},\ldots,p_{ijK}\right], \tag{3.12}$$

where p_{ijk} is the probability that someone with the same characteristics as individual i moves from state j to state k. Since the individual errors e_{ijk} are orthogonal to the subutility functions V_{jk}, the probabilities p_{ijk} increase monotonically with V_{jk} and the two sets contain the same information. Then let

$$\mathbb{E}\left[u_{ik}|h_{ijk}<0,p_{ij1},\ldots,p_{ijK}\right]=\psi_{jk}(p_{ij1},\ldots,p_{ijK}), \tag{3.13}$$

where $\psi_{jk}(\cdot)$ is an unknown correction function. Since any attempt to estimate this function would be infeasible due to the dimensionality of the problem, Dahl proposes estimating it semiparametrically through a polynomial expansion of the migration probabilities, finding that a quadratic of the migration probabilities is sufficient and the addition of higher order terms are not necessary.[26]

Implementation of the correction, however, requires estimates of the migration probabilities, $p_{ij1}, \ldots, p_{ijK}$. It is desirable to obtain these without imposing a functional form on the selection equations, since doing so would require all of the pecuniary and nonpecuniary variables that impact the utility offered to an individual by a location. However, note that the individual characteristics that determine locational decisions can be described in terms of discrete variables. Individuals can be assigned to cells

[25] For instance, Lee (1983) uses the multinomial logit model to estimate the locational choice decision, so the individual errors are Type I extreme value. A translation of the distribution of the maximum order statistic allows u_{ik} and h_{ijk} to be bivariate normal. Regardless of the joint distribution imposed, it is assumed implicitly that the individual utility errors $e_{ij1}, \ldots, e_{ijK}$ are independent. Dahl makes a similar assumption for the selection correction process (the "index sufficiency assumption") below, but it is more flexible in that it allows some choices other than the observed choice enter into the control function.

[26] Other options exist for the unknown correction function; Dahl suggests Fourier series as a possibility. However, empirical applications of Dahl's technique have generally used a quadratic.

that have the identical (discrete) characteristics. In this case, the estimate of a given migration probability, p_{ijk}, is simply the *observed* probability that an individual with the same type as individual i, born in state j, will move to state k. If the variables that make up an individual's type include all attributes that determine an individual's preferences over the pecuniary and nonpecuniary amenities associated with a location, each individual of a given type will have identical values for the subutility functions associated with each location, $V_{j1}, \ldots, V_{jK}$. Using the observed migration probabilities for each given type sidesteps the need for the imposition of a functional form on the selection equations, as well as the need to estimate them directly. I describe the variables with which I define an individual's type in Section 4.

Unfortunately, with large values of K, estimation of the control functions $\psi_{jk}(\cdot)$ remains infeasible without making further assumptions. Since there is a control function for each potential origin and destination, there are K^2 control functions , each of which depend on K migration probabilities. Dahl proposes two additional assumptions in order to allow feasible estimation of the control functions.

Assumption I (State independence of correction function) Dahl proposes the assumption that the joint distribution of the subutility differences and the earnings error is independent of birth state, except for those who remain in their state of birth:

$$\psi_{jk}\left(p_{ij1}, \ldots, p_{ijK}\right) = \psi_{k}\left(p_{ij1}, \ldots, p_{ijK}\right) \qquad \forall j \neq k, \tag{3.14}$$

and thus the control function is allowed to vary within an state only by two broad categories: those who chose to remain in their state of birth, k (henceforth "stayers") and those who were born outside state k but chose to move there (henceforth "movers"). This assumption has the effect both of reducing the number of distinct control functions while also noting that, in a model where K is large and movement between some states k is limited, some of the control functions would be poorly identified.

However, this assumption may not be enough to ensure feasibility in my case. Dahl estimates state-specific regressions that include only young men in only one Census year. My sample is much larger, so I follow several recent papers that have utilized Dahl's framework that make the following more restrictive assumption (Beaudry et al. 2012; Sand 2013):

$$\psi_{jk}\left(p_{ij1}, \ldots, p_{ijK}\right) = \psi\left(p_{ij1}, \ldots, p_{ijK}\right), \tag{3.15}$$

which states that the control function does not vary by state of residence.

Assumption II (Index sufficiency) If K is sufficiently large, estimation is still infeasible when the control function still depends on the full set of the selection probabilities,

$p_{ij1}, \ldots, p_{ijK}$. Dahl proposes the following assumption:

$$\psi\left(p_{ij1}, \ldots, p_{ijK}\right) = \psi\left(p_{ijk}, p_{ijj}\right), \tag{3.16}$$

which states that the control function depends only on an individual's probability of moving to the state where they are actually observed, p_{ijk}, and their probability of staying in their state of birth p_{ijj}. If an individual stays in their state of birth, then $p_{ijk} = p_{ijj}$ and their control function depends on only one migration probability. Dahl refers to this as an *index sufficiency assumption* because implicit in (3.16) is that p_{ijk} and p_{ijj} contain all relevant information about the index of migration probabilities, $p_{ij1}, \ldots, p_{ijK}$, and thus the index of subutilities $V_{j1}, \ldots, V_{jK}$.[27] This assumption does *not* imply that the bivariate covariance between the earnings error u_{ik} and the selection errors $e_{ijq} - e_{ijk}$ has to be equal across states; this covariance can vary between states and take any sign. However, if the index sufficiency assumption holds, this covariance must be a function of p_{ijj} and p_{ijk}. To frame this more concretely, Dahl proposes the following specification for the earnings function error term:

$$u_{ik} = \tau_k b_i + d_{ik}, \tag{3.17}$$

where b_i is a fixed, mean zero individual ability term that remains constant regardless of state of residence, and d_{ik} is a mean zero individual-state specific ability term which is uncorrelated across states, and b and d are uncorrelated. In this setup, b_i represents unobserved ability and d_{ik} is an state-specific match term. If τ_k equals some constant c for all states k, then the index sufficiency assumption is appropriate. However, if τ_k varies across states – meaning that unobserved ability is rewarded to differing degrees across states – then the earnings error u_{ik} will be correlated across states.

Dahl's empirical exercise – examining state-specific returns to education – is broadly similar to mine; he experiments with adding more migration probabilities and finds they do not significantly change his results. Further, Monte Carlo evidence suggests that even in cases where the index sufficiency assumption is strongly violated, the selection correction procedure is still able to remove approximately half the selectivity bias (Dahl 2002, Appendix C; Bourguignon et al. 2007).

However, the index sufficiency assumption presents particular difficulties in the empirical application I propose. In particular, consider the case where immigration within a skill cell in state q increases the return to unobserved ability in that skill cell, τ_k. If this is the case, it would result in changes in the bivariate covariance between u_{ik} and $e_{ijq} - e_{ijk}$ for $k \neq q$. In this case, the selection correction procedure will not remove all of the bias.

[27] An intuitive explanation is given as follows: assume two individuals from state j move to state k. If one of the individuals has a higher estimated probability of moving to some other state q, this fact should contain no information as to the relative wages of the two individuals beyond that provided by p_{ijk} and p_{ijj}.

The only way to test the index sufficiency assumption in this framework is to relax it. The advantage of this semiparametric approach is that it is simple to relax this assumption, by allowing additional selection probabilities to enter the control function and determine if they result in changes in the coefficient of interest. However, since adding all the probabilities is infeasible, selection of relevant probabilities is important. I discuss ways in which to relax the index sufficiency assumption which are appropriate within the context of estimating the labour market impact of immigration in Section 6.2.1.

4 Empirical strategy

Specification A base regression equation for area studies which estimate the *partial* wage impact of immigration while allowing for imperfect substitution between workers of the same education but differing experience is (Borjas 2003, 2006; Aydemir and Borjas 2007):

$$\begin{aligned} w_{srkt} = {} & \alpha_0 + \gamma Y_{skrt} + \alpha_1(s) + \alpha_2(r) + \alpha_3(k) + \alpha_4(t) + \alpha_5(k,t) \\ & + \alpha_6(s,t) + \alpha_7(r,t) + \alpha_8(k,s) + \alpha_9(k,s,r) + u_{skrt}, \end{aligned} \tag{4.1}$$

where w_{srkt} represents the average log wage of education group s with *potential* experience r, living in state k at time t, and $\alpha(\cdot)$ represent fixed effects. The estimate of the wage impact of immigrant competition is γ, which is the coefficient associated with Y_{skrt} – the *percentage share* of immigrants in skill group-state-time cell (s, r, k, t).

Recall that, in section 2, the change in wages due to own-skill immigration in the nested CES model was shown to approximately depend on changes in m_{srt} – adding state subscripts, m_{srkt} – the ratio of immigrant labour supply over native labour supply within a cell, which is interpretable as the increase in labour supply within a cell due to immigration. However, this is only an approximation; the effect of m_{skrt} on wages is in fact nonlinear. In the model described in Section 2, the initial stock of immigrants was set to zero; however, there is great variation in the degree of immigrant penetration across skill groups and regions in the United States. Therefore, Borjas (2003, 2006) and Aydemir and Borjas (2007) argue to instead use the percentage share of immigrants in the cell, $Y_{skrt} = (M_{skrt})/(M_{skrt} + N_{skrt}) \approx \ln m_{skrt}$, as the measure of immigrant competition. It can then be shown that the wage elasticity of immigration in (2.5) can be expressed within the reduced-form framework of (4.1) as

$$\frac{dW_{srkt}}{dm_{srkt}} = \frac{\gamma}{(1 + m_{srkt})^2}, \tag{4.2}$$

which simply shows that the effect of further immigration on native wages decreases as the initial stock of immigrants within a cell increases.[28] In my data, the mean value

[28]It should be noted that (2.5) also suggests that the wage impact of immigration will be higher for

of $1/(1+m_{srkt})^2$ is approximately 0.816; thus, the wage elasticity can be obtained by multiplying each coefficient by this value.

The specification (4.1) includes a full set of education, experience group, state, and time fixed effects, as well as several interaction terms. The first set of interaction terms – α_5, α_6, and α_7 – control for changes in state-specific productivity, the return to schooling, and the experience wage profile, over time. Differences in state-specific returns to education are encompassed by α_8, and α_9 allows the experience wage profile within an state to vary by schooling group. Thus, after including the full set of fixed effects and interaction terms $\alpha_1, \ldots, \alpha_9$, the variation in w_{srkt} left over for identification of γ is the *changes over time* in education-experience-state cells (s, r, k).[29]

In my empirical estimation, I experiment with different groupings of education and experience groups. This will help to determine the degree to which the native selection response to immigration can be detected at finer or coarser definitions of skill groups.

Definition of skill cells As discussed in detail in Section 2.1, the definitions of skill groups is disputed. I thus experiment with two different definitions of education, a "high skill" and "low skill" classification and a four-category classification (high school dropout, high school graduate, some college, and college graduate). As in Borjas (2003), Aydemir and Borjas (2007), and Ottaviano and Peri (2012) I examine workers with one to 40 years of experience; I experiment with 10 and 20 year bands for experience groups.[30]

I use states as a level of spatial aggregation for the areas k; most U.S. studies use cities or states as the level of aggregation. There are two principal concerns raised regarding area studies of the labour market impact of immigration: the first are the channels of factor price equalization described in Section 2.2.2, and the second is the possibility of attenuation bias. The first is clearly not a concern in my case, since the goal is to determine how important native self-selection is to the lower effects of immigration estimated by area studies.

However, attenuation bias may occur due to sampling error in the estimation of the immigrant share, Y_{skrt}, and Aydemir and Borjas (2011) show that this sampling error greatly increases as the number of observation in each skill cell s, k, r, t declines, suggesting that any benefit from increased variation obtained from examining a closer level of aggregation such as cities may be lost due to the increase in attenuation bias. The use of Dahl's selection correction method would be made more difficult: the US Census does not provide information on city of birth, and although it is possible to adopt the method for selection correction at the city level, this requires additional

skill cells that contribute a larger amount to the total labour share of income, which is another reason to include a rich set of interaction terms for the fixed effects.

[29] An alternative approach is to omit the last interaction term, α_9, in which case identification is driven by changes over time in the wages of education-state groups.

[30] Although the aforementioned papers use five-year bands for experience groups, they perform their analysis at the national level; the attenuation bias associated with doing so at the state level negates any benefit from examining a finer experience classification.

assumptions on the selection process which may effect the amount of bias that can be removed by the procedure.[31] Further, the average number of observations in each region at a given time t will be lower when k is indexed by city rather than state; Monte Carlo evidence presented in Dahl (2002, Appendix C) suggests that the ability of the selection correction procedure to remove bias begins to decline after the number of observations for each area begins to decline below $10,000$. For this reason, I use states as the level of aggregation instead of cities.

IV strategy As noted in Section 2.2.1, since immigrant locational choice is endogenous, it is common to use a shift-share instrument for immigrant density which exploits the tendency of immigrants to sort into enclaves (Altonji and Card 1991; Card 2001, 2009). Although the use of a large number of fixed effects and interactions may limit the extent to which higher labour demand in states that attract a higher number of immigrants will result in biased estimates of γ, including a specification where immigrant inflows are instrumented ensures that my results are not driven by high-ability natives being attracted to states with higher wages. An appropriate instrument for the estimation in (4.2) would be given by

$$\begin{aligned} \hat{Y}_{skrt} &= \frac{\hat{M}_{srkt}}{\hat{M}_{srkt} + N_{srkt}} \\ \hat{M}_{srkt} &= M_{s,k,r,t-q} + \sum_{l} \left[(\Delta M_{l,t})(\zeta_{k,l,t-q})(c_{s,r,l,t-q}) \right], \end{aligned} \tag{4.3}$$

where $\Delta M_{l,t}$ is the total national increase in labour supplied by immigrants from country group l from $t-q$ to t, $\zeta_{k,l,t-q}$ the percentage of immigrants from country group l who settled in state k in $t-q$, and $c_{s,r,l,t-q}$ the percentage of immigrants from country group l who were part of skill group (s,r) in $t-q$. The shift-share instrument takes the total immigrant inflow from each country group l from $t-q$ to t and gives the increase in immigrant density that would prevail in state-education-experience cell (s,r,k,t) if immigrant locational choices, educational attainment, and potential experience within country groups had remained the same as it was in $t-q$. In constructing the instrument, I use 1960 as the base year and divide immigrant origin countries into 37 country groups that remain consistent over time.

Selection correction There are multiple options available for implementation of the selection correction procedure. In general, estimates of the labour impact of immigration have not been derived from individual-level regressions – the dependent variable in (4.2) is the *average wage* in a skill cell over time. It would be possible to correct (4.2) for selectivity bias in the estimated wages of natives in a two-stage procedure. In the first stage regression, I could estimate selection, obtaining estimates of $\mathbb{E}\left[u_{ikt}|d_{ijkt}=1\right]$. In the second stage, I could then estimate (4.1) replacing mean log wages in a cell with

[31] See, for instance, the adaptation of Dahl's procedure presented in Beaudry et al. (2012).

mean log wages minus the mean selectivity bias, $w_{ikt} - \mathbb{E}\left[u_{ikt}|d_{ijkt} = 1\right]$.

However, it is more straightforward to estimate a regression at the individual level. In this case, the effect of correcting for native self-selection will depend on the correlation between the error mean term, $\mathbb{E}\left[u_{ikt}|d_{ijkt} = 1\right]$, and the immigrant share, Y_{skrt}:

$$\begin{aligned} w_{ik} &= x_{skrt}\beta + \gamma Y_{skrt} + u_{it} \\ w_{ik} &= x_{skrt}\beta + \gamma Y_{skrt} + \mathbb{E}\left[u_{ikt}|d_{ijkt} = 1\right] + \omega_{it} \\ w_{ik} &= x_{skrt}\beta + \gamma Y_{skrt} + \psi(p_{ijk}, p_{ijj}) + \omega_{it}, \end{aligned} \tag{4.4}$$

where x_{skrt} is a vector containing the fixed effects in (4.1), $\psi(p_{ijk}, p_{ijj})$ is a quadratic in p_{ijk} and p_{ijj}, and ω_{it} is an error term which assumed to be mean zero conditional on selection. The measure of immigrant penetration, Y_{skrt}, remains identical for each members of a given state-education-experience cell s, k, r, t. I take this approach in my empirical estimation.

As discussed in Section 3, the selection correction procedure outlined by Dahl (2002) requires individuals to be assigned to cells based on vectors of their discrete characteristics, x_i and z_i. For movers, x_i the is comprised of four education and eight potential experience categories), and four racial categories (white, black, Hispanic, and other). As in Dahl (2002), there are more stayers than movers, so I include additional variables in the vector x_i based upon family characteristics, in particular an individual's marital status (married or nonmarried), and the presence of children under 18.

Identification, of course, depends on an exclusion restriction: including variables in z_i that do not appear in x_i that effect the subjective utility offered to an individual by moving to a state but having no effect on state-specific wages. Here, identification is driven by the inclusion of state of birth in z_i. The intuitive explanation of this assumption can be given as follows: two individuals with identical characteristics x, who were born in different states but currently both live in state k, should earn the same amount except if they differ in unobserved ability. For instance, take two worker living in Arizona with identical observed characteristics, except for state of birth: one worker was born in New Mexico and the other in Vermont. If the worker from Vermont has higher wages, we can conclude that Arizona provided them with a better job "match" that incentivized them to move a greater distance.

Thus, in my empirical implementation I divide individuals into mutually exclusive cells based on their discrete characteristics z_i, which consists of the individual's state of birth and the vector x_i described in the previous paragraph. Thus, within a state, each mover from a given origin state is divided into one of 128 cells, and each stayer is divided into one of 512 cells.[32]

[32] I allow the District of Columbia to be a "state of birth" and a migration probability, but I exclude D.C. from all of my regressions. A small number of native-born residents were not born in one of the 50 states or the District of Columbia; most of this group were either born in Puerto Rico or stated they were born in the United States but did not specify a state. I allow these groups to have a moving probability but exclude the staying probability from their control functions. For native-born citizens born outside of the United States and its territories, I calculate moving probabilities by assigning them to one of the 37 country groups that were defined earlier in the construction of the shift-share instrument, since the

5 Data

5.1 Sample construction

The data used in this paper are sourced from the 1960, 1970, 1980, 1990, and 2000 U.S. Decennial Census Public Use Microdata Samples (PUMS), and the 2008-2010 3-year American Community Survey (ACS), downloaded from the IPUMS website (Ruggles et al. 2017). For brevity, I refer to the ACS data as the 2010 data. For my regressions, I use the data for the 1970-2010 data; the 1960 data is used only for construction of the shift-share IV instrument. My sample consists of men aged 18-64 who participate in the labour force and worked a positive number of weeks in the previous year, do not live in group quarters, and are not in school or the military.[33] To prevent my results from being driven by outliers, I remove workers whose income, usual hours worked, and weeks worked last year imply an hourly wage less than 0.75 times the prevailing federal minimum wage from the sample, following Card (2009).

Workers are grouped into four education categories: high school dropouts, high school graduates, workers with some college education, and college graduates.[34] Potential experience is then calculated based on subtracting the worker's age from the age they are assumed to enter the labour market: 17 for high school dropouts, 18 for high school graduates, 21 for workers with some college education and 23 for college graduates. Workers are then grouped into experience groups based on their potential experience. The finest experience grouping consists of four 10-year bands, starting with one to ten years of potential experience and ending with 31 to 40 years. Coarser experience groupings are constructed by combining these into two 20-year bands.

Workers are defined as "natives" if they are neither naturalized U.S. citizens or non-U.S. citizens and "foreign-born" otherwise. N_{skrt} and M_{skrt}, respectively the total native and foreign-born labour supplied within a given state-education-experience cell, are calculated by aggregating the total labour hours supplied by each worker in the previous year, adjusting for their Census personal sample weight. Total labour hours supplied is calculated by multiplying the number of weeks worked last year by the worker's usual number of hours worked in a week. These are then used to calculate the foreign-born share of each state-education-experience cell, which is given by $Y_{s,k,r,t}$.

Following the literature (Borjas et al. 2012; Ottaviano and Peri 2012), I use a more selective sample in my wage regressions than in determining the measure of immigrant competition: I drop those workers who worked less than 27 weeks of work the previous year and those who typically work less than 35 hours a week, preventing my results

number of native-born citizens born in each individual country is small.

[33]The removal of women from the sample is driven by the fact that, especially in earlier sample years, women are likely to have lower labour force attachment. For married women in particular, this is likely to result in the locational choice decision being less driven by their own comparative advantage in the job market.

[34]I define the "some college" category as workers who do not have a bachelor's degree, but have one full year to four years of college education.

from being driven by workers who have weak labour market attachments entering or leaving the labour market. I also drop all self-employed workers. My dependent variable is log weekly wages; an individual's wage is calculated by dividing their annual income by the number of weeks they work in a year. Wages in each year are deflated using the relevant values of the CPI (all items) and adjusted for Census topcoding; topcoded wages are multiplied by 1.5. All regression results below are weighted by the appropriate Census personal sample weight.[35]

5.2 Descriptive statistics

Before proceeding to the empirical results, it is useful to quickly examine the data on immigrant settlement across states and the impact that this may have on the in- and out-migration decisions of natives. Since immigrant density patterns are well known, and the effect of immigrant inflows on native locational decisions has been an area of active research, the purpose of this section is solely to motivate and provide context to the analysis that follows.

The map in Figure 1 shows the overall immigrant density across all workers in the employment sample in 2010; immigrant density is highest in densely populated states and states along the southern border. This pattern of immigrant settlement is highly stable across time; Table 1 gives a summary of the states with the highest mean overall immigrant density across all workers in the employment sample from 1970-2010. Since the states with the highest immigrant densities relative to the rest of the U.S. also tend to be relatively highly populated, these states account for the vast majority of the population of immigrant workers in the U.S. The ten states with the highest immigrant densities accounted for 69.7 percent of the total immigrant population; the top 15 states accounted for 79.1 percent.[36]

Figure 2 graphs the density of all immigrants along the wage distribution of natives *within* each immigrant's state-education-experience cell separately by skill level.[37] The horizontal line represents the native wage distribution; values above it indicate that immigrants are more concentrated at a given part of the native within-cell wage distribution than natives. If the phenomenon of immigrant downgrading in the United

[35]Weighting the regressions by the Census personal sample weight results in lesser weight towards cells with smaller populations, reducing the potential for measurement error in the measure of immigrant density.

[36]This is also consistent over time: the mean values for 1970-2010 are 74.5 percent and 81.9 percent respectively. The fact that these mean values are higher than in 2010 alone suggests that immigrants are less concentrated in these states in 2010 relative to earlier years.

[37]To create this figure, I first calculate the percentile of the native log real weekly wage distribution within state-education-experience-year cells that is associated with each immigrant's wage, where education and experience are divided into two categories each. I then calculated the (weighted) percentage of immigrants falling within each percentile across all years. The figure itself is a lowess-smoothed plot of the densities with a bandwith of 0.1 and the top and bottom 2 percentiles omitted, so each skill level has 96 observations. This figure is very similar to that found in Figure 1 of Dustmann et al. (2012), except that they plot the kernel density of all recent immigrants along the wage distribution of all native workers. Since we are looking within state-skill cells, this will control for the tendency of immigrants to locate in states with both high wages and high costs of living.

States was primarily driven by highly educated immigrants working in low-wage jobs, we would expect to see high-skill immigrants being more concentrated at the lower end of the native within-cell wage distribution, but the opposite is the case here. Low-skill immigrants are far more clustered at the low end of the native within-cell wage distribution than high-skill immigrants. If the degree of immigration wage competition is higher for natives at points in the native within-cell wage distribution with relatively high densities of immigrants, we would expect these natives to have a greater incentive to change locations.[38]

In Figures 3 and 4, the relationship between native locational choice decisions and immigrant density (in levels) is examined. Figure 3 plots the estimated probability of a native worker of a given level of education and experience staying in their state of birth in 2010 against the percentage of workers of the same level of education and experience who are immigrants in their state of birth. Figure 4 plots the percentage of native workers within a state of a given level of education and experience in 2010 who were born outside the state against the percentage of workers of the same level of education and experience who are immigrants in that state; note that education and experience are divided into two categories each, giving a total of 200 observations for each graph.[39]

Next, I examine the relationship between changes over time in immigrant densities across education-experience cells and changes in the propensity of natives to remain in, or move into, these cells. Figure 5 plots the change in the probability of a native workers of a given education-experience skill group remaining in their state of birth from 1970-2010 against the change in the immigrant density in that education-experience cell in their state of birth during the same time period. Figure 6 plots the change in the percentage of native workers in a given state-education-experience cell who were born outside that state – the "mover density" – against the change in the immigrant density in that state-education-experience cell.

The results are strikingly similar. In Figure 5, we see that the increase in the tendency of native workers to remain in their state of birth is much more pronounced in states that experienced relatively low levels of immigration. In Figure 6, we see that the change in the mover density in a cell tends to become more negative as the immigrant density in that cell increases.[40] Note the two high-leverage observations on the bottom right of the graph represent low-skilled workers in California; the fact that the decline in native inflows to that state coincide with increases in immigrant inflows is similar to that made in Figure 5 of Borjas (2006), which uses workers' state of residence five years

[38]If this was the case, the relative density of immigrants at the low end of the native wage distribution would increase further as low-earning natives choose to leave.

[39]The figures would look similar if workers were only divided by state and not further divided into state-skill cells.

[40]It should be noted that decreases in the mover density may occur over time as the population of workers born in the state increase over time, even if the size of the flow is unchanged. In cells with growing populations, however, increases in the stayer probability can only be achieved by increasing the flow of workers out of the state.

ago to calculate net inflows.

Figures 3 to 6 suggest that the total value of the economic opportunities and amenities offered by immigrant-intensive states to natives are higher than in states with lower levels of immigration, as both the stayer probabilities and the mover densities in immigrant-intensive cells is high. However, the degree to which these immigrant-intensive states are more attractive to natives appears to be *diminishing* over time – and as the immigrant density increases. If in fact there are changes in patterns of locational decisions in response to immigration, there is the potential for selectivity bias in estimates of the wage impact of immigration.

6 Results

6.1 Main results

Base model I now proceed to my main empirical results. Table 2 shows the estimates of γ in (4.1), estimated by OLS, for four different skill groupings with and without correction for selectivity bias. Note that in all regressions I weight each worker by their Census person weight, and adjust standard errors for clusters in state-education-experience-year.[41]

Before discussing the effects of the selection correction procedure, I will first discuss the results without the correction terms. Column 1 displays the results where workers within each state are divided into combinations of two education groups and two experience groups. Similarly, in column 2 workers are divided into two education groups and four experience groups; in column 5 workers are divided into four education groups and two experience groups; and in column 6 workers are divided into four education groups and four experience groups.[42]

Immigrant density is significantly negatively associated with native wages within state-skill cells across all the alternate skill groupings. Two conclusions are immediately apparent: the wage impact is estimated to be larger when using only two education groupings instead of four, and for a given division of education groups, the wage impact is estimated to be larger when using two experience groupings instead of four. The estimate of the wage impact, γ, using 16 skill groups (four education and four experience groups) is thus the least negative at -0.094, while the estimate using four skill groups is almost three times larger in magnitude, at -0.270.[43] However, even the

[41]Each regression is performed on 7.4 million observations and includes several hundred fixed effect terms. For this reason, I use the Stata package `reghdfe`, which greatly increases the speed of estimation of regressions with large numbers of fixed effects, to obtain the results in the paper. The exact process is described in Correia (2016); I performed the regressions in Table 2 with the Stata `regress` command and obtained identical estimates of coefficients and standard errors.

[42]Note that the fixed effects vary based on the skill grouping. The education and experience categories in each fixed effect term are the same categories used to calculate the immigrant share variable, so the number of fixed effect terms increases as finer skill groups are used.

[43]As previously mentioned, all estimates can be converted into wage elasticities by multiplying by 0.816. For instance, $d\ W_{srkt}/d\ m_{srkt} = -0.270 \times 0.816 = -0.220$. The interpretation attached to this figure

largest coefficient is only just over half of equivalent elasticities found by national level studies; Aydemir and Borjas (2011) obtain an estimate -0.489 at the national level for the U.S. using 1960-2000 data and five education and eight experience groups; Borjas (2006) finds a value of -0.533 for the same period when using four education and eight experience groups.[44]

There are several possible reasons for the magnitude of the estimates to increase as the skill groupings become more coarse. One possibility, as argued by Aydemir and Borjas (2011), is attenuation bias. To the extent that estimates of immigrant density are more accurate in coarser skill cells, this may result in more negative estimates of the native wage impact. An alternate explanation – as argued by Card (2009) and Ottaviano and Peri (2012) – would be that four education categories is an inappropriate specification of skill groupings and that high school dropouts and graduates are perfect substitutes. To the extent this is the case, the level of immigrant competition faced by native high school dropouts will be greatly overestimated when using four skill categories (and the level of competition by native graduates underestimated), as in my sample, low skill immigrants are far more likely to be dropouts than graduates relative to low skill natives.

The focus of this paper, however, is on the direction of γ after controlling for native self-selection, rather than the absolute magnitude. With this in mind, Columns 3, 4, 7, and 8 provide the results for a modified version of the OLS regressions in Columns 1, 2, 5, and 6 which add the quadratics of the stayer and mover probabilities.

In a standard Heckman correction, the null of no selection can be tested by testing the significance of the coefficient on the inverse Mills ratio term. With this selection correction method, there are multiple correction terms, so a Wald test for joint significance of all correction terms is analogous; I find that the selection terms are jointly highly significant in all regressions in Table 2.

The addition of the correction terms result in an increase in OLS estimates of the magnitude of the wage impact of immigration on natives for all four skill group definitions. The estimates of immigration in Columns 3, 4, 7, and 8 increase in magnitude by 30.7 percent, 37.6 percent, 22.5 percent, and 10.6 percent respectively relative to the OLS regression without selection correction. F tests show that the coefficients are all significantly different after correcting for selection.[45] Importantly, although the magnitude of the estimates increase in all four specifications, the increase is higher both in relative and absolute terms for the regressions with only two education groups. I discuss this in more detail below.

---

is that an immigration-driven increase in labour supply in a skill cell of 10 percent reduces native wages in the cell by 2.2 percent.

[44]Estimating a regression at the national level using two education and two experience groups would not be very useful, since such a regression would only have $5 \times 2 \times 2 = 20$ observations.

[45]The F statistics for the difference in the coefficients from the regressions with and without the selection correction are obtained from a pooled regression where both models are estimated simultaneously.

Discussion of main results I will now quickly attempt to provide some context for these results. As noted by Dahl (2002), in a model such as this where both wage and nonwage factors influence the latent utility associated with locations and where the latent utility function that determines locational choice remains unspecified, the interpretation of the reasons for coefficients changing after correction for selectivity bias is difficult. In Dahl's case, the coefficient of interest is the state-specific returns to education at a single point in time. In this case, the coefficient will be biased to the extent that the *levels* of general and state-specific ability in (3.17) vary across states. This interpretation would remain the same if there was only one year in my regressions; in this case, if correcting for selection resulted in the estimates of the immigration wage impact γ becoming more negative, it would imply that the mean selectivity bias of native workers, $\mathbb{E}\left[u_{ikt}|d_{ijkt}=1\right]$, in state-education-experience cells with higher levels of immigration is higher than those in cells that have experienced less immigration.

However, my regression equation (4.1) includes not only variation over time, but a large collection of fixed effects. As I mentioned in Section 4, identification of γ comes solely off of variation over time in the wages of state-education-experience cells. The fixed effects thus control for differences in the levels of unobserved ability across state-education-experience cells; that is, to the extent that the mean selectivity bias of native workers in high-immigration state-skill cells is higher than low-immigration cells, this will *not* result in bias if this difference remains constant over time. Bias will exist only if changes in the mean selectivity bias level of state-skill cells is correlated with changes in the immigrant density in that cell.

It follows from this that the fact that the estimate of γ becomes more negative after correcting for selection implies that changes in unobserved ability in state-skill cells is *positively* correlated with changes in cell immigrant density, so that the ability levels within immigrant-heavy state-skill cells is increasing over time. In the selection correction framework used in this paper, changes in the mean selectivity bias in a state-education-experience cell can occur for three reasons: changes in the average unobserved ability of movers, changes in the average unobserved ability of stayers, and – to the extent that the mean selectivity bias of movers and stayers within a cell are different – the proportion of the cell that were born in the state. Since determining the exact reasons why selectivity bias exists in this framework is difficult, I briefly return to descriptive analysis similar to that in Section 5.2.

Figure 7 plots the change from 1970 to 2010 in the mean selectivity bias among native movers in state-education-experience cells, where education and experience are again divided into two groupings each, against the change in immigrant density in the cell during the same time frame. There is little evidence that there is any relationship between changes in the mean selectivity bias of movers within cells is related to increases in immigration.

Figure 8 plots the change in the mean selectivity bias among native stayers from 1970-2010 against the change in immigrant density in the same time period, again using two groupings each for education and experience.[46] This figure suggests that the selection bias of γ is driven by *increases* in the selectivity bias of workers who choose to live in their state of birth. To examine this more clearly, Figure 10 plots the control function for stayers as estimated in Column 3 of Table 2. The relative increase the value of the control function for stayers with mean values of staying probabilities in "immigrant-intensive" states (the 15 states listed in Table 1) from 1970 to 2010 was 6.22 log points across all skill groups.[47] For some states this is even higher: the gap in the value of the control function between California and non-immigrant intensive states decreased by 13.02 log points during the same period.[48]

Figure 9 plots the difference between the mean selectivity bias of stayers and the mean selectivity bias of movers within cells in 2010 against the immigrant density in 2010. A value above zero implies that stayers are estimated to have higher unobserved ability; a value below zero implies the reverse. There appears to be little relationship between the two variables. Across the U.S., the mean value for this difference, weighted by the population of each state-education-experience-year cell, is -2.0 log points in 2010 and -2.3 when using all data from 1970-2010. This implies that reductions in the "mover density" as described in Section 5.2 can only explain a small portion of the increase in the estimated wage impact of immigration after correction for selection, since non-negligible increases in the mean selectivity bias within cells could only be accomplished by very large *increases* in the mover density.

Thus, the selection bias appears to be driven by decreases in the probability that native workers born in immigrant-intensive states choose to remain in their state of birth. The intuition behind this is that, as the propensity for natives born in these states to stay in their state of birth decreases to around that of natives born in non-immigrant intensive states, the mean value of the control function increases, perhaps because those whose skills are relatively least valued in their state of birth are the first to leave. To the extent that this then results in increases in the magnitude of the impact of immigration, we might say that it is likely that the skills that are valued relatively less in immigrant-intensive states are those that immigrants provide disproportionately relative to natives.[49] However, since the regression sample is restricted to natives who work full time, this phenomenon could also be caused by natives with relatively high estimated staying probabilities choosing to exit the labour market rather than face

[46] In the scatter plots, I regress wages on all independent variables *except* the immigrant density when determining the mean selectivity bias, $\mathbb{E}\left[u_{ikt}|d_{ijkt}=1\right]$.

[47] This is based on mean staying probabilities of 76.0 percent in 1970 and 65.9 percent in 2010 for intensive states, and 65.9 percent and 66.4 percent in 2010 for non-intensive states, calculated using Census person weights.

[48] This is based on staying probabilities across all skill groups in California decreasing from 84.7 percent in 1970 to 68.6 percent in 2010.

[49] Of course, many natives will move for wage-related reasons other than immigration competition or for non-wage amenity reasons. Further, many internal migrants move from one immigrant intensive state to another. 52.6 percent of native movers from intensive states moved to another intensive state in 2010; for comparison, 51.3 percent of native movers from non-intensive states moved to an intensive state.

greater levels of competition for jobs.[50]

The results in Table 2 also suggest that the selection correction has less of an impact when using finer education groupings. Figure 11 plots the change in the mean selectivity bias among native stayers from 1970-2010 against the change in immigrant density in the same time period, again using four education and two experience groups; the relationship between changes in mean selectivity bias and immigrant density is no longer positive, so the selection correction has little impact when using four education groups. One possible reason for this is that the relative size of the inflow is much higher for dropouts than any other group: the average state increase in the immigrant density over the period, weighted by state populations, was 0.4824 for dropouts but only 0.1219 for high school graduates. If native dropouts are not commensurately more responsive in terms of changes in their staying probability relative to high school graduates, changes in the mean selectivity bias will have relatively small effect on estimates of the impact of immigration. This suggests that the selection response to low-skill immigration is largely driven by native high school graduates even though the low-skill inflow consists disproportionately of dropouts.

IV estimates I now proceed to the IV estimates, using the enclave shift-share instrument described in Section 4. The instrument does well in the first stage in predicting actual immigrant densities, though it is more predictive when there are only two education groups. To the extent that immigrants tend to self-select into states with favourable economic conditions, we would expect to see the magnitude of estimates increase after instrumenting the immigrant density. This is the case for the regressions with two education groups, though not as much for the four education group regressions. The relative increase in the magnitude of the estimates after correcting for selection is similar in the IV regressions as in OLS, which suggests that the effects of the selection correction are not driven by high-ability natives choosing to live in states with favourable economic conditions.

The two-education group IV regressions increase in magnitude by a similar proportion after correcting for selection: 28.9 percent in the two experience group model and 40.5 percent with four experience groups. However, the four experience group regressions with and without selection correction are quite similar in the 2SLS results, although the selection terms remain significant.

6.2 Extensions

In this section I briefly discuss two possibilities for extensions to the control function included in the regression model: adding additional selection probabilities to the

[50] To the extent this was the case, this would also bias studies that use national level data. In a recent paper, Llull (2018) finds that natives choosing to drop out of the labour force can result in significant biases in national level estimates of the effect of immigration on wages at the lower ends of the wage distribution.

correction function, and allowing the control function to vary over time; both of these extensions only strengthen my results.

6.2.1 Expanding the set of selection probabilities

As I noted in Section 3.2, the selection correction procedure in Dahl (2002) relies on what is referred to as an "index sufficiency assumption" for feasibility, where only a subset of the selection probabilities are assumed to contain all information relevant to the selection decision. Further, if – as I hypothesize – immigration results in changes in native locational decisions, then the index sufficiency assumption will fail, since there will be correlations between the value of workers' subutility functions across states and the immigrant density.

To see this, consider a worker born in a non-immigrant intensive state who chooses to stay in his state of birth even though workers of similar characteristics are relatively likely to move to immigrant-intensive states. To the extent that native workers in immigrant-intensive cells are *positively* selected, the fact that this worker chose to stay in his state of birth might suggest he has lower unobserved ability. However, in the procedure in Dahl (2002), only the workers' state of birth for stayers is included in the correction function.

The only way to test this assumption is to add more moving probabilities and determine if this changes the results, but it is important to select relevant probabilities. One suggestion made by Dahl is the highest estimated probability outside of the chosen location p_{ijk} and the staying probability p_{ijj}. As I noted in Table 1, the states with the highest concentration of immigrants has remained remarkably consistent over the length of my sample. Since I am particularly interested in the relationship between unobserved ability and the probability of choosing to live in immigrant-intensive states, I add quadratics in the highest five estimated probabilities of moving to one of the 15 immigrant-intensive states listed in Table 1, excluding those where this probability is either the state of birth or residence.[51]

The results are presented in Table 5. Columns 1-4 represent the results for two education groups and two or four experience groups; Columns 1 and 3 present the results for two education groups and two experience groups when estimated using OLS and 2SLS, respectively, while 2 and 4 do the same but with four experience groups. Columns 5-8 are analogous but use four education groups. Each regression includes a Wald test for the core selection terms (the quadratics in the staying and moving probabilities) as well as a separate Wald test for the additional terms; both are jointly significant in all eight regressions.

Adding the additional selection terms results in increases in the magnitude of the estimates, which is consistent with correlation of individual earnings errors across states with the immigrant density. For instance, the OLS estimate when using two education

[51] This allows for the selectivity bias of workers who are born and live inside and outside immigrant-intensive states to vary based on their observed probability of moving to immigrant-intensive states.

and two experience groups is -0.371, as opposed to -0.353 in the core selection OLS model and -0.270 in OLS without correction, while the IV estimate is -0.519, compared to -0.473 in the core selection IV model and -0.367 in the IV model without selection correction.

It would be undesirable if these results were particularly dependent on the set of states considered to be "immigrant intensive". In Table 4, I estimate alternate regressions where quadratics in the top five of the *ten* most immigrant intensive states are added instead. The results are very similar to those in Table 5.

6.2.2 Time-varying control function

Since patterns of internal migration have changed over time and immigrant inflows into the U.S. have also been increasing during the time period under consideration, it is important to verify that the increase in the estimated effect of immigration after correcting for selection is actually a result of selection pressures being different in immigrant intensive states. In Table 6 I estimate the OLS and IV models, correcting for selection and only using the core probabilities, but allowing the selection correction to vary over time by interacting each control function term with the Census year. The ordering of the columns is the same as it was in Tables 5 and 4.

The magnitude of the estimates increase by a larger amount when allowing the control function to vary over time than adding additional probabilities. The OLS estimate for two education and two experience groups is -0.474, while the IV estimate is -0.622; the wage elasticities implied by these estimates are 0.387 and 0.508 respectively. An exception is the 16-skill IV model, which is around the same as the original uncorrected estimate. Of course, since the 16-skill model has much lower estimates of the wage elasticity, the impact of the selection correction procedure is less precise.

6.3 Summary of results

Table 7 shows the change in the magnitude of the selection-corrected OLS and 2SLS estimates after including the various control functions proposed in this paper. Row 1 shows the change in the magnitude of the coefficient after adding the base control function to the specficication; Row 2 shows the change in magnitude compared to the uncorrected estimates after adding the additional probabilities to the base correction function; Row 3 shows the change in magnitude compared to the uncorrected estimates after allowing the base control function to vary by year.

Across all three alternate selection correction specifications, the two education group models see much larger increases in the estimates of the wage impact of immigration, both in relative and absolute terms. For instance, the OLS estimate of the wage impact of immigration when using two education and experience groups is 1.9 times that of the equivalent estimate using four education and two experience groups (Table 2), but 2.5 times when using time-varying control functions (Table 6). This suggests that the native self-selection may be occuring within broadly-defined (high or low skill) but between finer education groups. This could, for instance, be the result of low-ability high school graduates selecting out of states with high immigrant dropout populations. However, any further analysis of the exact causes of the selectivity bias is beyond the scope of this paper.

7 Conclusion

Most of the attention given to the effect that native internal migration may have on biasing estimates of the wage impact of immigration obtained through area studies have focused solely on the effect that decreases in net native inflows may have on the effective size of the immigrant supply shock. In this paper, I have argued that native internal migration may result in bias of the labour market impact of immigration beyond simply reducing the size of the increase in labour supply, since if immigrant competition tends to be higher in the lower parts of the wage distribution within skill cells, native displacement that results from this competition will lead to increases in mean levels of unobserved ability within skill cells in immigrant-intensive states. Internal migration is self-selected, so to the extent that native-born workers change their choice of location based on levels of immigration it is natural that those workers who face the highest degree of immigrant competition are the most likely to choose to live elsewhere. I apply a selection correction procedure proposed by Dahl (2002) to correct for this potential bias.

My results suggest that selectivity bias due to native locational decisions result in substantial underestimation of the wage effects of immigration, though the magnitude of the bias greatly depends on how skill groups are defined; estimates of the impact of immigration increase by around 30 to 40 percent when using four and eight skill models which include only two education groups after including the core correction function, but the increase is smaller – both in absolute and relative terms – when four education groups are used. This suggests that selection is occurring within broadly defined but between more fine definitions of education. Simple extensions to the control function – such as expanding the number of selection probabilities allowed to enter the function or allowing the function to vary over time – result in increases in the estimated wage elasticity of up to 90 percent. However, the estimated impact of immigration is much lower in the finest skill grouping I use – 16 cells per state, consisting of four education and experience groups – and the impact of the selection correction is also inconsistent across specifications of the control function for the 16 skill group model, which perhaps suggests that measurement error in the immigrant densities is effecting estimates in state level regressions that include this number of skill groups.

Overall, the results suggest that state level estimations of the wage impact of immigration that use coarse skill groupings, a commonly used instrument for immigrant densities using information on past immigrant settlement, and a selection correction procedure can provide similar estimates than national level studies. The actual estimated elasticities have limited real-world significance due to the fact that they are only partial effects, which ignore the complementarities that arise from immigration of workers outside of natives' own skill groups. However, the same selection pressures which result in bias in these estimates are likely to effect area studies that examine the total wage effect of immigration using *relative* labour supplies as well. However, it is very difficult to interpret the reasons why bias is present using this selection correction procedure, so more research into the types of occupations worked by individuals who move in and out of immigrant-intensive states would be useful, as well as the effect in general that immigration has on natives' occupational, locational, and labour force participation choice.

8 References

Ahn, Hyungtaik, and James L Powell. 1993. "Semiparametric estimation of censored selection models with a nonparametric selection mechanism." *Journal of Econometrics* 58 (1-2): 3–29.

Altonji, Joseph G, and David Card. 1991. "The effects of immigration on the labor market outcomes of less-skilled natives." In *Immigration, trade, and the labor market*, 201–234. University of Chicago Press.

Amior, Michael. 2018. "The Contribution of Foreign Migration to Local Labor Market Adjustment." Working paper.

Amuedo-Dorantes, Catalina, and Sara De La Rica. 2011. "Complements or substitutes? Task specialization by gender and nativity in Spain." *Labour Economics* 18 (5): 697–707.

Aydemir, Abdurrahman, and George J Borjas. 2007. "Cross-country variation in the impact of international migration: Canada, Mexico, and the United States." *Journal of the European Economic Association* 5 (4): 663–708.

— . 2011. "Attenuation bias in measuring the wage impact of immigration." *Journal of Labor Economics* 29 (1): 69–112.

Bartik, Timothy J. 1991. "Who benefits from state and local economic development policies?" *WE Upjohn Institute for Employment Research*.

Beaudry, Paul, David A Green, and Benjamin Sand. 2012. "Does industrial composition matter for wages? A test of search and bargaining theory." *Econometrica* 80 (3): 1063–1104.

Blanchard, Olivier Jean, and Lawrence F Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity* 23 (1): 1–76.

Borjas, George J. 2001. "Does immigration grease the wheels of the labor market?" *Brookings Papers on Economic Activity* 2001 (1): 69–119.

— . 2003. "The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market." *Quarterly Journal of Economics* 118 (4): 1335–1374.

— . 2006. "Native internal migration and the labor market impact of immigration." *Journal of Human Resources* 41 (2): 221–258.

— . 2013. "The analytics of the wage effect of immigration." *IZA Journal of Migration* 2 (1): 22.

Borjas, George J., Richard B. Freeman, and Lawrence F. Katz. 1996. "Searching for the Effect of Immigration on the Labor Market." *The American Economic Review* 86 (2): 246–251.

Borjas, George J., Jeffrey Grogger, and Gordon H. Hanson. 2008. "Imperfect Substitution between Immigrants and Natives: A Reappraisal." Working Paper, Working Paper Series 13887. National Bureau of Economic Research.

— . 2012. "Comment: On Estimating Elasticities of Substitution." *Journal of the European Economic Association* 10 (1): 198–210.

Borjas, George J, Richard B Freeman, and Lawrence F Katz. 1992. "On the labor market effects of immigration and trade." In *Immigration and the workforce: economic consequences for the United States and source areas*, 213–244. University of Chicago Press.

Borjas, George J, and Lawrence F Katz. 2007. "The evolution of the Mexican-born workforce in the United States." In *Mexican immigration to the United States*, 13–56. University of Chicago Press.

Bound, John, and George Johnson. 1992. "Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations." *The American Economic Review* 82 (3): 371–392.

Bourguignon, François, Martin Fournier, and Marc Gurgand. 2007. "Selection bias corrections based on the multinomial logit model: Monte Carlo comparisons." *Journal of Economic Surveys* 21 (1): 174–205.

Cadena, Brian C. 2014. "Recent immigrants as labor market arbitrageurs: Evidence from the minimum wage." *Journal of Urban Economics* 80:1–12.

Cadena, Brian C, and Brian K Kovak. 2016. "Immigrants equilibrate local labor markets: Evidence from the Great Recession." *American Economic Journal: Applied Economics* 8 (1): 257–90.

Card, David. 2001. "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration." *Journal of Labor Economics* 19 (1): 22–64.

— . 2005. "Is the new immigration really so bad?" *The Economic Journal* 115 (507).

— . 2009. "Immigration and inequality." *American Economic Review* 99 (2): 1–21.

Card, David, and John DiNardo. 2000. "Do immigrant inflows lead to native outflows?" *American Economic Review* 90 (2): 360–367.

Card, David, and Thomas Lemieux. 2001. "Can falling supply explain the rising return to college for younger men? A cohort-based analysis." *The 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 to the United States*, 193–228. University of Chicago Press.

Correia, Sergio. 2016. "Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator." Tech. rep. Working Paper.

Cortes, Patricia. 2008. "The effect of low-skilled immigration on US prices: evidence from CPI data." *Journal of Political Economy* 116 (3): 381–422.

Dahl, Gordon B. 2002. "Mobility and the return to education: Testing a Roy model with multiple markets." *Econometrica* 70 (6): 2367–2420.

Dustmann, Christian, Tommaso Frattini, and Ian P Preston. 2012. "The effect of immigration along the distribution of wages." *Review of Economic Studies* 80 (1): 145–173.

Dustmann, Christian, Uta Schönberg, and Jan Stuhler. 2017. "Labor supply shocks, native wages, and the adjustment of local employment." *The Quarterly Journal of Economics* 132 (1): 435–483.

Filer, Randall. 1992. "The effect of immigrant arrivals on migratory patterns of native workers." In *Immigration and the workforce: Economic consequences for the United States and source areas*, 245–270. University of Chicago Press.

Friedberg, Rachel M, and Jennifer Hunt. 1995. "The impact of immigrants on host country wages, employment and growth." *Journal of Economic Perspectives* 9 (2): 23–44.

Glitz, Albrecht. 2012. "The labor market impact of immigration: A quasi-experiment exploiting immigrant location rules in Germany." *Journal of Labor Economics* 30 (1): 175–213.

Heckman, James J. 1979. "Sample selection bias as a specification error." *Econometrica* 47 (1): 153–161.

Hunt, Jennifer. 2017. "The impact of immigration on the educational attainment of natives." *Journal of Human Resources* 52 (4): 1060–1118.

Imai, Susumu, Derek Stacey, and Casey Warman. 2018. "From engineer to taxi driver? Language proficiency and the occupational skills of immigrants." Tech. rep. GLO Discussion Paper.

Jaeger, David A. 2007. "Skill Differences and the Effect of Immigrants on the Wages of Natives." *US Bureau of Labor Statistics Working Paper* 273.

Johnson, George E. 1980. "The labor market effects of immigration." *ILR Review* 33 (3): 331–341.

Kaplan, Greg, and Sam Schulhofer-Wohl. 2017. "Understanding the long-run decline in interstate migration." *International Economic Review* 58 (1): 57–94.

Katz, Lawrence F, and Kevin M Murphy. 1992. "Changes in relative wages, 1963–1987: supply and demand factors." *The Quarterly Journal of Economics* 107 (1): 35–78.

Kennan, John, and James R Walker. 2011. "The effect of expected income on individual migration decisions." *Econometrica* 79 (1): 211–251.

Kerr, Sari Pekkala, and William R Kerr. 2011. "Economic impacts of immigration: A survey." Tech. rep. National Bureau of Economic Research.

Kerr, Sari Pekkala, William R Kerr, and William F Lincoln. 2015. "Skilled immigration and the employment structures of US firms." *Journal of Labor Economics* 33 (S1): S147–S186.

Kerr, Sari Pekkala, et al. 2017. "High-skilled migration and agglomeration." *Annual Review of Economics* 9:201–234.

LaLonde, Robert J, and Robert H Topel. 1991. "Labor market adjustments to increased immigration." In *Immigration, trade, and the labor market*, 167–199. University of Chicago Press.

Lee, Lung-Fei. 1982. "Some approaches to the correction of selectivity bias." *The Review of Economic Studies* 49 (3): 355–372.

— . 1983. "Generalized econometric models with selectivity." *Econometrica*: 507–512.

Lewis, Ethan, et al. 2004. "How did the Miami labor market absorb the Mariel immigrants?" Tech. rep. Federal Reserve Bank of Philadelphia Philadelphia, PA.

— . 2005. "How do local labor markets in the US adjust to immigration?" *Philadelphia Fed Business Review Q* 1:16–25.

Llull, Joan. 2018. "Immigration, Wages, and Education: A Labour Market Equilibrium Structural Model." *The Review of Economic Studies* 85 (3): 1852–1896.

Longhi, Simonetta, Peter Nijkamp, and Jacques Poot. 2005. "A meta-analytic assessment of the effect of immigration on wages." *Journal of Economic Surveys* 19 (3): 451–477.

— . 2008. "Meta-Analysis of Empirical Evidence on the Labour Market Impacts of Immigration." Population Studies Centre Discussion Papers dp-67. University of Waikato, Population Studies Centre.

Manacorda, Marco, Alan Manning, and Jonathan Wadsworth. 2012. "The impact of immigration on the structure of wages: theory and evidence from Britain." *Journal of the European Economic Association* 10 (1): 120–151.

Molloy, Raven, Christopher L Smith, and Abigail Wozniak. 2011. "Internal migration in the United States." *Journal of Economic Perspectives* 25 (3): 173–96.

Monras, Joan. 2015. "Immigration and wage dynamics: Evidence from the mexican peso crisis." Working paper.

Moser, Petra, Alessandra Voena, and Fabian Waldinger. 2014. "German Jewish émigrés and US invention." *American Economic Review* 104 (10): 3222–55.

Okkerse, Liesbet. 2008. "How to measure labour market effects of immigration: A review." *Journal of Economic Surveys* 22 (1): 1–30.

Ottaviano, Gianmarco I. P., and Giovanni Peri. 2012. "Rethinking the effect of immigration on wages." *Journal of the European Economic Association* 10 (1): 152–197.

Peri, Giovanni, Kevin Shih, and Chad Sparber. 2015. "STEM workers, H-1B visas, and productivity in US cities." *Journal of Labor Economics* 33 (S1): S225–S255.

Peri, Giovanni, and Chad Sparber. 2009. "Task specialization, immigration, and wages." *American Economic Journal: Applied Economics* 1 (3): 135–69.

— . 2011. "Assessing inherent model bias: An application to native displacement in response to immigration." *Journal of Urban Economics* 69 (1): 82–91.

Røed, Marianne, and Pål Schøne. 2012. "Does immigration increase labour market flexibility?" *Labour Economics* 19 (4): 527–540.

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

Ruggles, Steven, et al. 2017. *Integrated Public Use Microdata Series: Version 7.0. [dataset]. Minneapolis, MN.*

Rybczynski, Tadeusz M. 1955. "Factor endowment and relative commodity prices." *Economica* 22 (88): 336–341.

Sand, Benjamin M. 2013. "A re-examination of the social returns to education: Evidence from US cities." *Labour Economics* 24:97–106.

Welch, Finis. 1979. "Effects of cohort size on earnings: the baby boom babies' financial bust." *Journal of Political Economy* 87 (5, Part 2): S65–S97.

Wozniak, Abigail, and Thomas J Murray. 2012. "Timing is everything: Short-run population impacts of immigration in US cities." *Journal of Urban Economics* 72 (1): 60–78.

Appendix A Wage determination in the canonical model

Since the wage for each type of worker is equal to their marginal productivity, the impact of immigration can be determined by the impact on the marginal productivity of each worker type. Firms choose a level of capital and labour such that the wage and rental rates are equal to the marginal productivity of labour and capital, respectively.

Therefore, in equilibrium, the wages of workers is given by

$$
\begin{aligned}
w_{srt} &= \frac{\partial Q_t}{\partial L_{srt}} = \frac{\partial Q_t}{\partial L_t}\frac{dL_t}{dL_{srt}} = \frac{\partial Q_t}{\partial L_t}\frac{\partial L_t}{\partial L_{st}}\frac{\partial L_{st}}{\partial L_{srt}} \\
w_{srt} &= (\alpha_t)(A_t)\left(\frac{K_t}{L_t}\right)^{1-\alpha} \times (L_t)^{1-\rho}\theta_{st}L_{st}^{\rho-1} \times (L_{st})^{1-\eta}\lambda_{sr}L_{srt}^{\eta-1} \\
w_{srt} &= (\alpha_t)(A_t)(K_t)^{1-\alpha}(L_t)^{\alpha_t-\rho}\theta_{st}L_{st}^{\rho-\eta}\lambda_{sr}L_{srt}^{\eta-1}, \qquad \text{(A.1)}
\end{aligned}
$$

where w_{srt} is the wage of a worker with education s and experience r at time t. Equation (A.1) can be simplified by taking logs:

$$
\begin{aligned}
\ln w_{srt} = \ln(\alpha_t) + \ln(A_t) + (1-\alpha)(\ln K_t) + \ln\theta_{st} + (\alpha_t - \rho)(\ln L_t) \\
+ (\rho-\eta)(\ln L_{st}) + log\lambda_{ij} + (\eta - 1)(\ln L_{srt}). \qquad \text{(A.2)}
\end{aligned}
$$

We can now examine the response of native wages to immigration. First, note that equation (A.2) implies that the change in the wages of a worker with education s and experience r due to input changes from $t-1$ to t is

$$
\begin{aligned}
\Delta \ln w_{srt} = (1-\alpha)(\Delta\ln K_t) + (\alpha_t-\rho)(\Delta\ln L_t) + (\rho-\eta)(\Delta\ln L_{st}) \\
+ (\eta-1)(\Delta\ln L_{srt}). \qquad \text{(A.3)}
\end{aligned}
$$

I assume that the native labour supply is time invariant and that the initial stock of immigrants in the labour force is zero.[52] Define M_{srt} as total labour supplied by immigrant workers with education s and experience r at time t, and N_{srt} as the corresponding total labour supplied by native workers. Then $m_{srt} = M_{srt}/N_{srt}$, the ratio of immigrant to native labour supply within labour cell (i, r, t), represents the percentage increase in labour supply in the cell relative to the counterfactual of zero immigration; for relatively small increases from zero in the immigrant-native ratio $\Delta\ln L_{srt} \approx \Delta m_{srt}$.[53] This implies that the change in wages due to immigration can be approximated by

$$
\begin{aligned}
\Delta \ln w_{srt} \approx (1-\alpha)(\Delta\ln K_t) + (\alpha_t-\rho)(\Delta m_t) + (\rho-\eta)(\Delta m_{st}) \\
+ (\eta-1)(\Delta m_{srt}), \qquad \text{(A.4)}
\end{aligned}
$$

where Δm_t and Δm_{st} represent, respectively, the change in the ratio of immigrant to native supply, in efficiency units, of the labour composites L_t and L_{st}. Note that the first term equals zero in the short run under the assumption of no capital adjustment, so the short run wage adjustment in response to immigration is approximately

$$
\Delta \ln w_{srt} \approx (\alpha_t-\rho)(\Delta m_t) + (\rho-\eta)(\Delta m_{st}) + (\eta-1)(\Delta m_{srt}). \qquad \text{(A.5)}
$$

[52] Dustmann et al. (2017) allows native labour supply to respond to immigration.

[53] More generally, $\Delta\ln L_{srt} \approx (\Delta M_{i,r,t})/(L_{i,r,t-1}) = (\Delta M_{i,r,t})/(M_{i,r,t-1} + N_{i,r,t-1})$. If labour supply from natives is constant, $(\Delta M_{i,r,t})/(M_{i,r,t-1} + N_{i,r,t-1}) = \Delta m_{srt}$ when immigrant labour supply at $t-1$ is zero. It is convenient to express wage changes in terms of m_{srt}, as in the empirical estimation that follows.

In the long run, capital adjusts such as to maintain a constant rental rate. Taking the log of the first order condition for capital, we see that

$$\begin{aligned}\ln r &= \ln(1-\alpha_t) + \ln(A_t) + (\alpha_t)(\ln L_t - \ln K_t) \\ \ln K_t &= \ln L_t + \frac{1}{\alpha_t}\left(\ln(1-\alpha_t) + \ln(A_t) - \ln r\right)\end{aligned} \tag{A.6}$$

in the long run, where r is the long-run rental rate of capital. Since a_t, A_t, and r are invariant to increases to labour supply, equation (A.6) implies that the long run wage adjustment of an education-experience cell to sufficiently small levels of immigration is approximately

$$\Delta \ln w_{srt} \approx (1-\rho)(\Delta m_t) + (\rho-\eta)(\Delta m_{st}) + (\eta-1)(\Delta m_{srt}), \tag{A.7}$$

where equation (A.7) is strictly greater than equation (A.5) as $\Delta m_t > 0, 1 > \alpha_t$.[54] The "total effect" of immigration is the change in the wages of an education-experience cell due to the increase in aggregate labour supply of the cell (Δm_{srt}), the cell's education group that result from immigration (Δm_{st}), and the increase in aggregate labour supply (Δm_t).

The change in the wage of an education-experience cell due to an increase in the labour supply of workers with the same level of education and experience as a result of immigration Δm_{srt} is captured by the last term on the right hand side of equation (A.7) but also the effect that the increase in L_{srt} has on the labour aggregates L_{st} and L_t.[55] This is referred to in the literature as the "partial effect" of immigration.

Recalling α_t is the labour share of income, define α_{st} as the share of total income accruing to workers of education s, and α_{srt} as the share of total income accruing to workers of education s and experience r. The change in the value of a CES aggregate can be expressed in terms of the income shares of its inputs (Borjas 2013); in particular, it can be shown that $\Delta \ln L_t \approx \left[\sum_{s=1}^{S}(\alpha_{st}/\alpha)\Delta m_{st}\right]$ and $\Delta \ln L_{st} \approx \left[\sum_{s=1}^{S}(\alpha_{srt}/\alpha_{st})\Delta m_{st}\right]$. This implies that the partial effect - the change in wages for a worker in skill cell (i, r, t) to an immigrant inflow of the same skill - can be given approximately by

$$\begin{aligned}\frac{dW_{srt}}{dm_{srt}} &\approx \left[(\alpha_t-\rho)\frac{\alpha_{srt}}{\alpha} + (\rho-\eta)\frac{\alpha_{srt}}{\alpha_{st}} + (\eta-1)\right] \\ \frac{dW_{srt}}{dm_{srt}} &\approx \left[(1-\rho)\frac{\alpha_{srt}}{\alpha} + (\rho-\eta)\frac{\alpha_{srt}}{\alpha_{st}} + (\eta-1)\right]\end{aligned} \tag{A.8}$$

in the short run and long run respectively.

[54]Similarly, it can be shown that the long-run wage changes from immigration when workers of similar education but differing experience are perfect substitutes is approximated by

$$\Delta \ln w_{st} \approx (1-\rho)(\Delta m_t) + (\rho-1)(\Delta m_{st}).$$

[55]$(\eta-1) = -(1/\sigma_x)$ where σ_x is the elasticity of substitution between experience groups.

Appendix B Figures

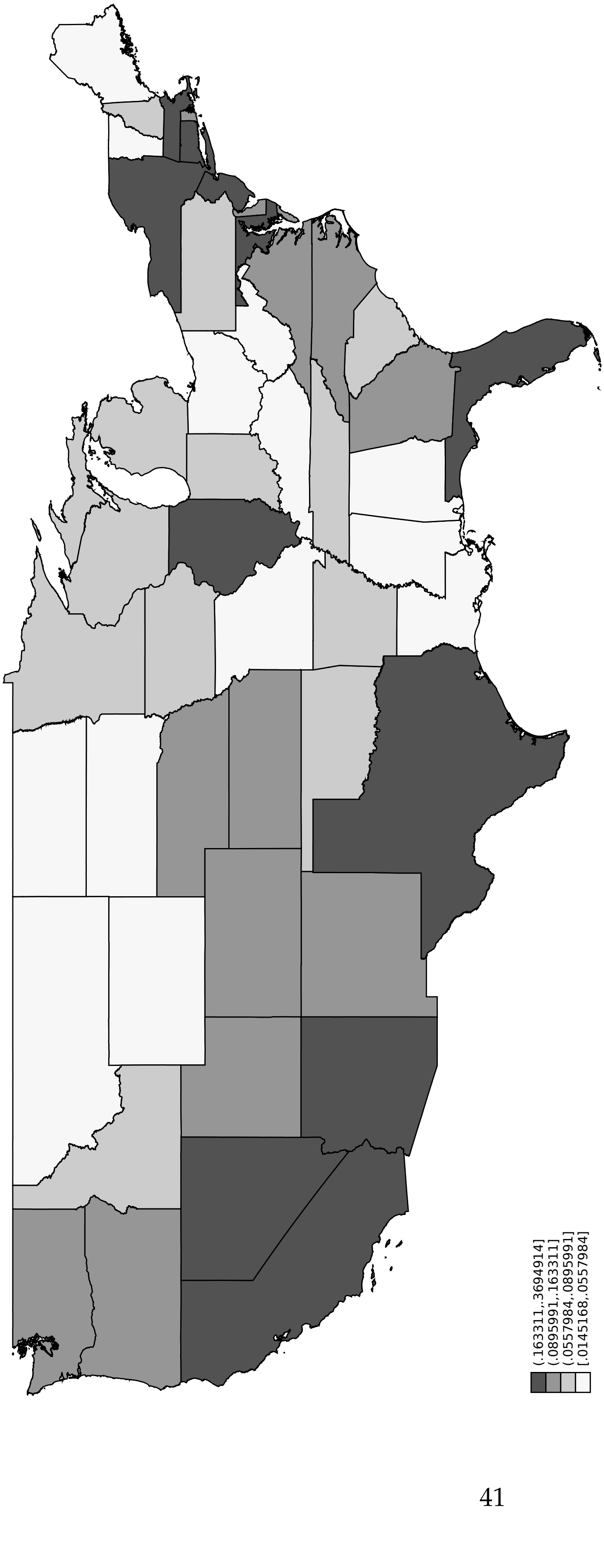
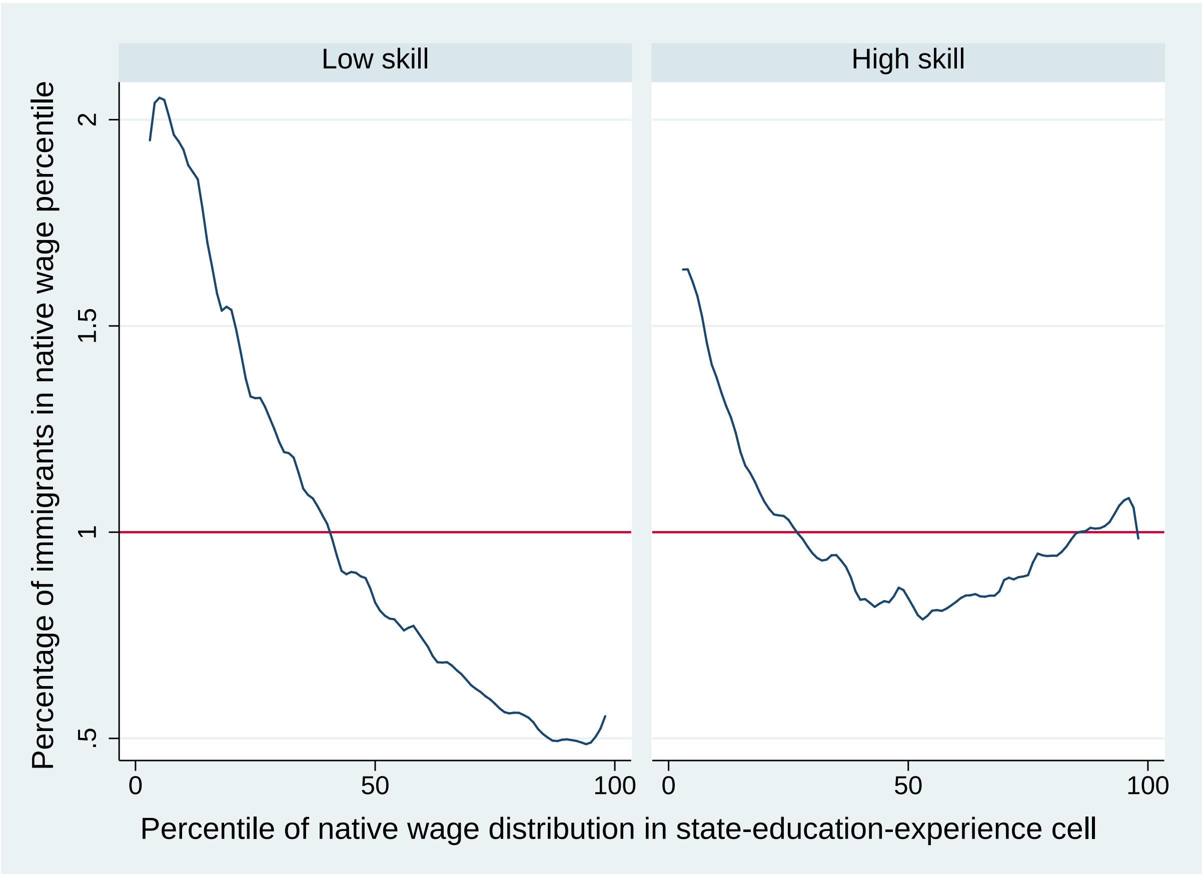


Figure 1: Immigrant density across the continental United States, 2010

Notes: Each data point corresponds to a state-education-experience cell, where there are two experience groups and education is defined as "high" or "low" skill.

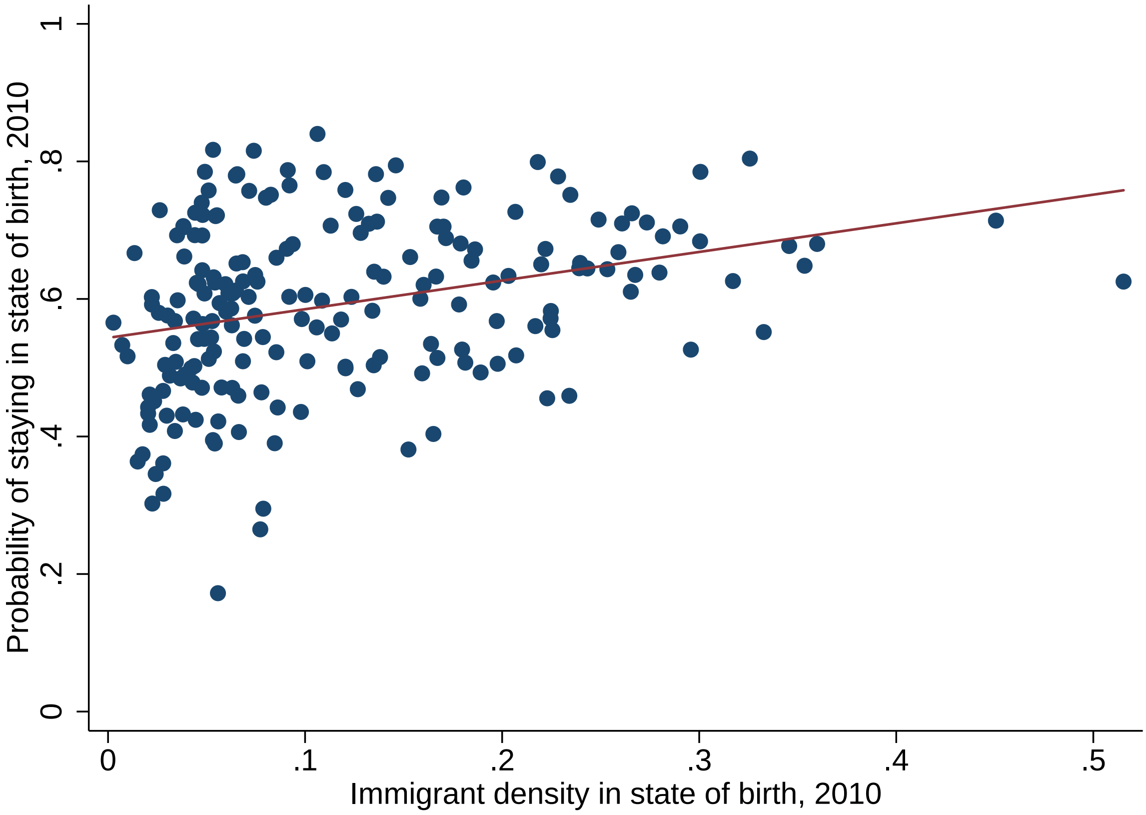
Source: Author's calculations from the 2008-2010 ACS. Sample selection criteria for construction of the immigrant density correspond to the description of the "employment sample" as described in text; for the construction of the native staying probability, the "wage sample" criteria are used

Figure 2: Position of immigrants in the native distribution of real weekly wages within state-education-experience-year cells



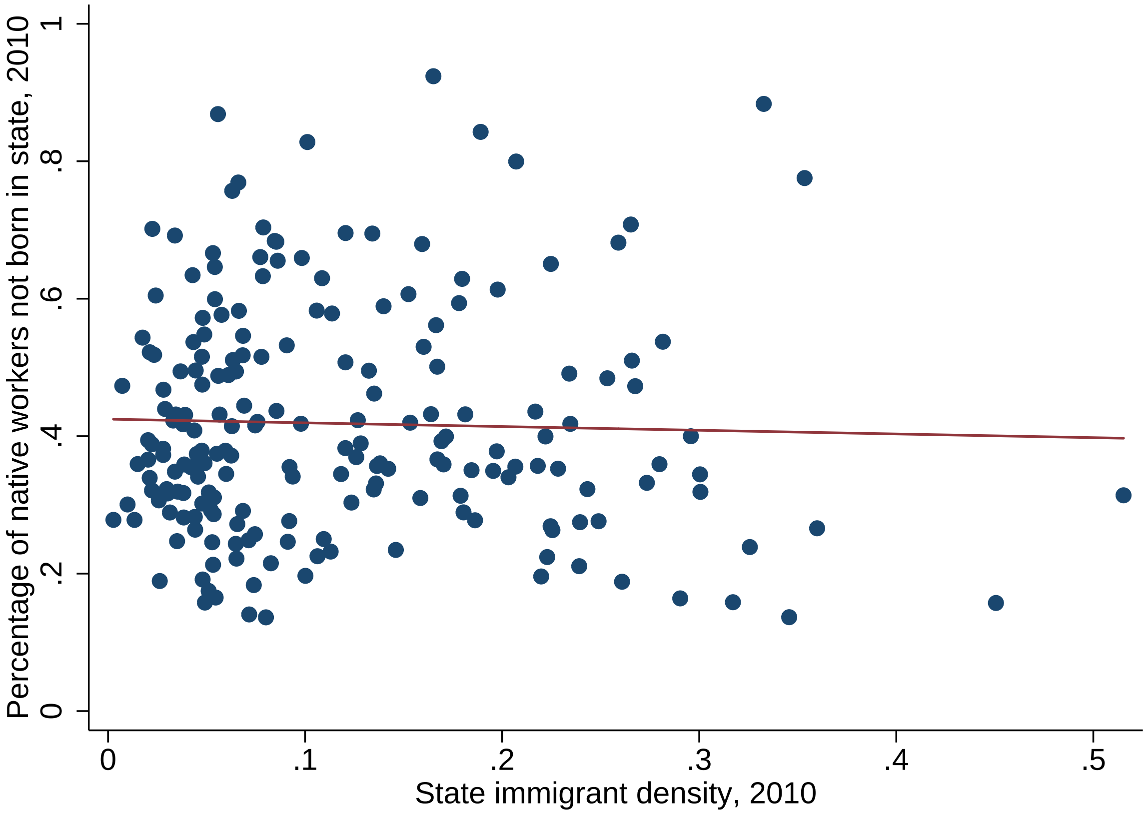
Notes: Each data point corresponds to a state-education-experience cell, where there are two experience groups and education is defined as "high" or "low" skill.
Source: Author's calculations from the 1970-2000 US Census and the 2008-2010 ACS. The "wage sample" selection criteria are used for the wages of both immigrants and natives for the construction of the wage percentiles for natives and immigrants.

Figure 3: Immigration within state-skill cells and native probability of staying in state of birth



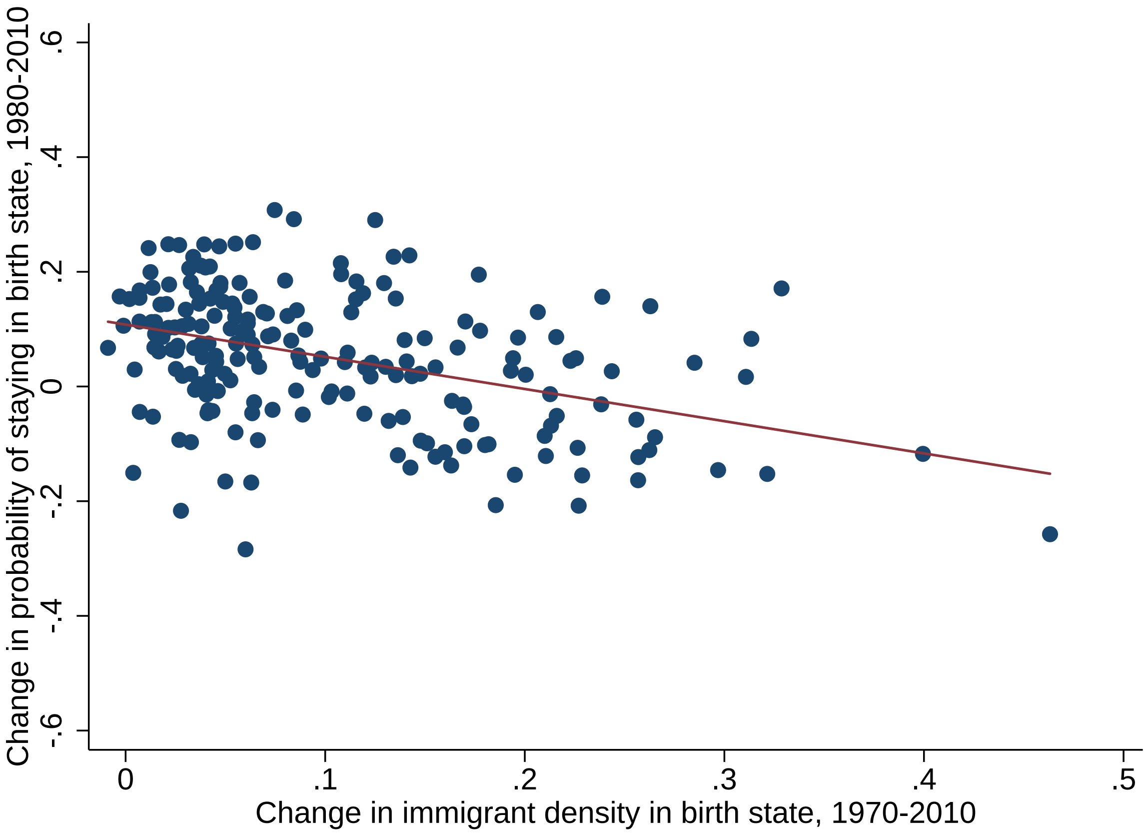
Notes: Each data point corresponds to a state-education-experience cell, where there are two experience groups and education is defined as "high" or "low" skill.
Source: Author's calculations from the 2008-2010 ACS. Sample selection criteria for construction of the immigrant density correspond to the description of the "employment sample" as described in text; for the construction of the native staying probability, the "wage sample" criteria are used. The probabilities of moving or staying in a state are constructed as described in the text.

Figure 4: Immigration within state-skill cells and percentage of native movers within cell



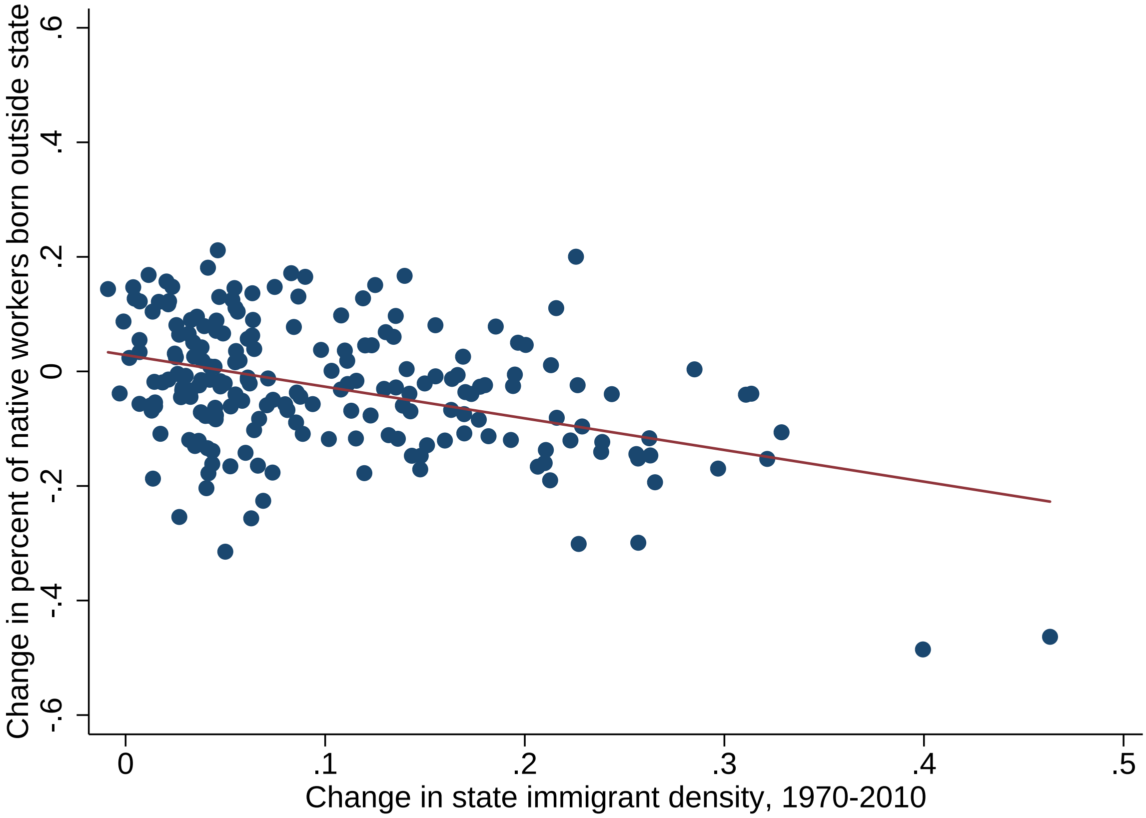
Notes: Each data point corresponds to a state-education-experience cell, where there are two experience groups and education is defined as "high" or "low" skill.
Source: Author's calculations from the 2008-2010 ACS. Sample selection criteria for construction of the immigrant density correspond to the description of the "employment sample" as described in text; for the construction of the mover density, the "wage sample" criteria are used.

Figure 5: Immigration within state-skill cells and change in the probability of staying in birth state



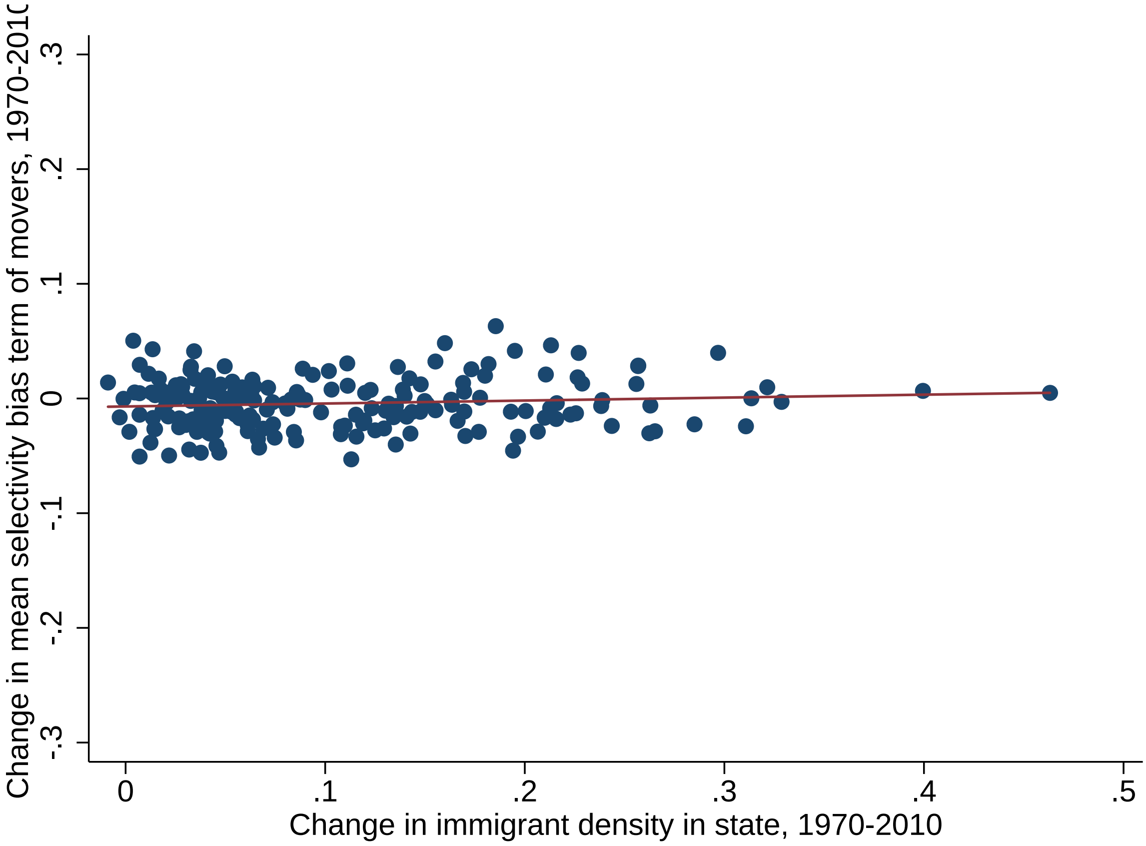
Notes: Each data point corresponds to a state-education-experience cell, where there are two experience groups and education is defined as "high" or "low" skill.
Source: Author's calculations from the 1970 U.S Census and 2008-2010 ACS. Sample selection criteria correspond to the description of the "employment sample" as described in text. Sample selection criteria for construction of the immigrant density correspond to the description of the "employment sample" as described in text; for the construction of the staying probabilities, the "wage sample" criteria are used. The probabilities of moving or staying in a state are constructed as described in the text.

Figure 6: Immigration within state-skill cells and change in the percentage of native movers within cell



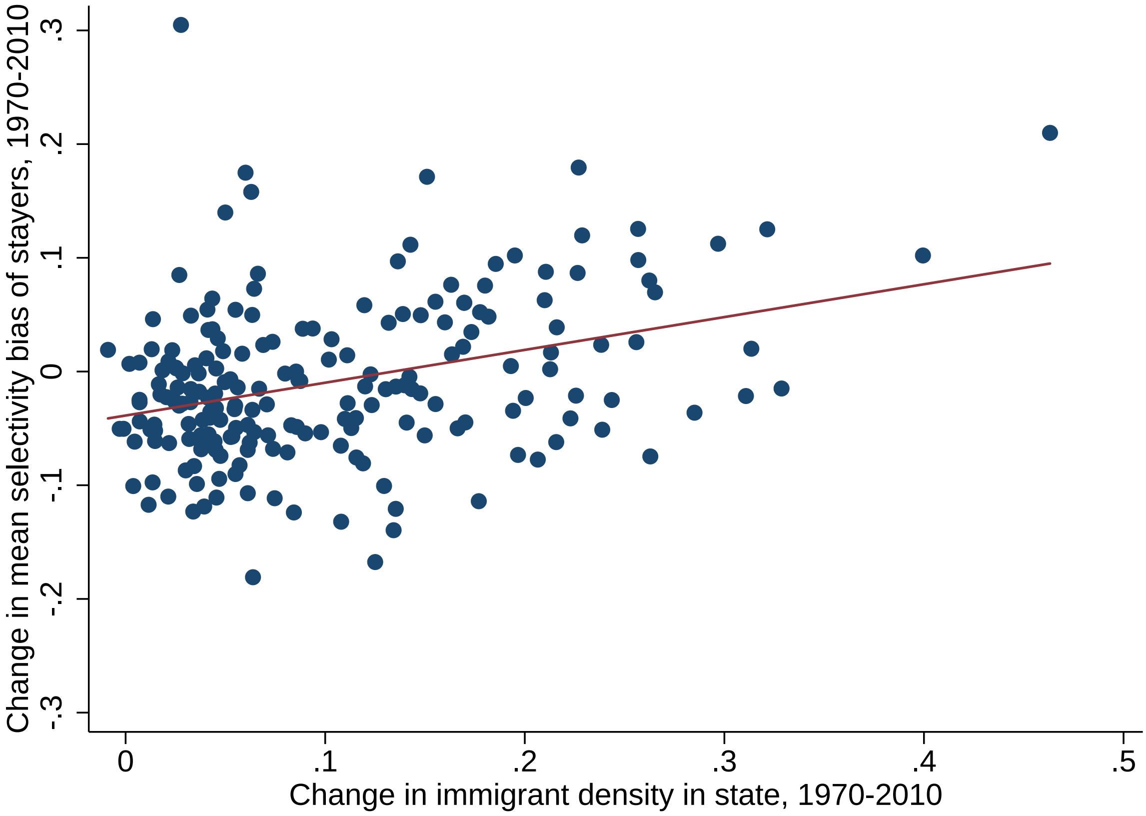
Notes: Each data point corresponds to a state-education-experience cell, where there are two experience groups and education is defined as "high" or "low" skill.
Source: Author's calculations from the 1970 U.S Census and 2008-2010 ACS. Sample selection criteria for construction of the immigrant density correspond to the description of the "employment sample" as described in text; for the construction of the mover density, the "wage sample" criteria are used.

Figure 7: Immigration within state-skill cells and change in the mean selectivity bias of native movers within cell



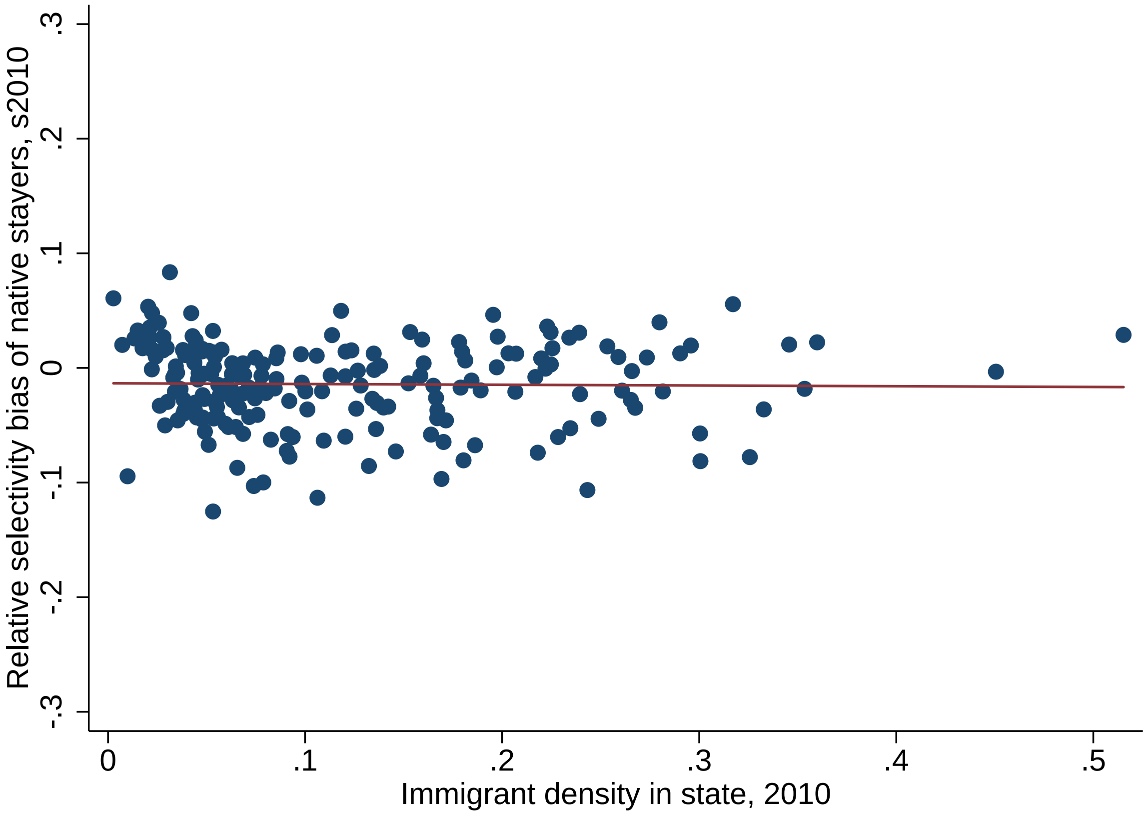
Notes: Each data point corresponds to a state-education-experience cell, where there are two experience groups and education is defined as "high" or "low" skill.
Source: Author's calculations from the 1970 U.S Census and 2008-2010 ACS. Sample selection criteria correspond to the description of the "employment sample" as described in text. Sample selection criteria for construction of the immigrant density correspond to the description of the "employment sample" as described in text; for the estimation of the error mean term for native movers, the "wage sample" criteria are used.

Figure 8: Immigration within state-skill cells and change in the mean selectivity bias of native stayers within cell (Two education groups)



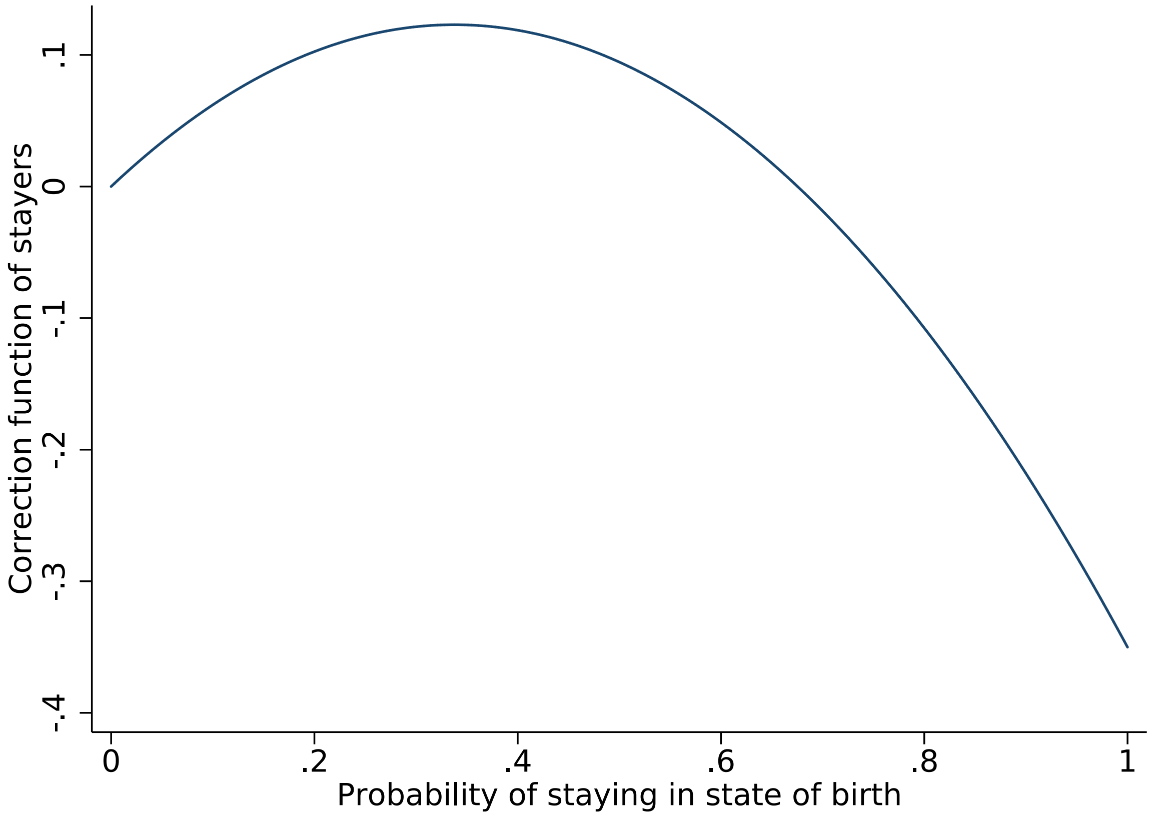
Notes: Each data point corresponds to a state-education-experience cell, where there are two experience groups and education is defined as "high" or "low" skill.
Source: Author's calculations from the 1970 U.S Census and 2008-2010 ACS. Sample selection criteria correspond to the description of the "employment sample" as described in text. Sample selection criteria for construction of the immigrant density correspond to the description of the "employment sample" as described in text; for the estimation of the error mean term for native stayers, the "wage sample" criteria are used.

Figure 9: Immigration within state-skill cells and selectivity bias of stayers relative to movers within cell



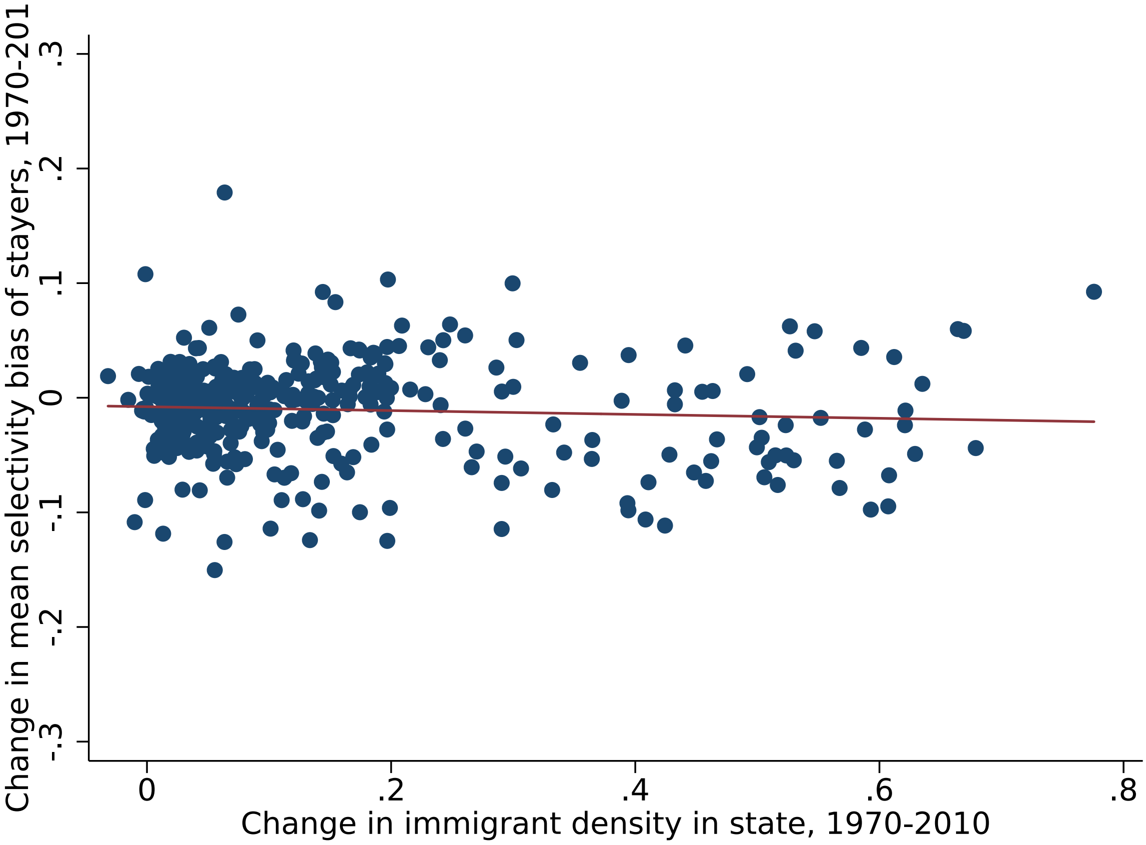
Notes: Each data point corresponds to a state-education-experience cell, where there are two experience groups and education is defined as "high" or "low" skill. As noted in the text, I calculate the "relative selectivity bias of stayers" by subtracting the mean selectivity bias of movers from the mean selectivity bias of stayers. Values above zero imply that stayers are more positively selected than movers and vice versa.
Source: Author's calculations from the 2008-2010 ACS. Sample selection criteria correspond to the description of the "employment sample" as described in text. Sample selection criteria for construction of the immigrant density correspond to the description of the "employment sample" as described in text; for the construction of the mean error terms of native movers and stayers, the "wage sample" criteria are used.

Figure 10: Control function for stayers



Notes: The estimates of the control function are obtained by OLS and correspond to the regression in Column 3 of Table 2. Workers are divided into 4 skill cells based on two categories each of education and experience. The control function is a quadratic in the staying probability; the coefficient on p_{ijj} is 0.728 (0.027) and the coefficient on p_{ijj}^2 is -1.079 (0.040). Since the intercept of the control function is unidentified, I set it to zero for simplicity.

Figure 11: Immigration within state-skill cells and change in the mean selectivity bias of native stayers within cell (Four education groups)



Notes: Each data point corresponds to a state-education-experience cell, where there are two experience groups and four education groups.
Source: Author's calculations from the 1970 U.S Census and 2008-2010 ACS. Sample selection criteria correspond to the description of the "employment sample" as described in text. Sample selection criteria for construction of the immigrant density correspond to the description of the "employment sample" as described in text; for the estimation of the error mean term for native stayers, the "wage sample" criteria are used.

Appendix C Tables

Table 1: List of immigrant intensive states

Rank	State	Mean immigrant density, 1970-2010
1	California	0.231
2	New York	0.173
3	New Jersey	0.163
4	Hawaii	0.138
5	Florida	0.127
6	Nevada	0.131
7	Texas	0.119
8	Illinois	0.117
9	Massachusetts	0.110
10	Conneticut	0.105
11	Rhode Island	0.105
12	Arizona	0.102
13	Maryland	0.089
–	U.S. *(national)*	0.097
14	Washington	0.082
15	Virginia	0.073

Notes: "U.S. (national)" represents the mean immigrant density of the U.S. as a whole over the timespan, not the mean of state densities.
Source: Author's calculations from 1970-2000 U.S. Census and 2008-2010 ACS. Sample selection criteria correspond to the description of the "employment sample" as described in text.

Table 2: Effect of immigration within skill groups on native wages, OLS

	Two education groups				Four education groups			
	OLS		OLS with selection		OLS		OLS with selection	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Immigrant share	-0.270**	-0.165**	-0.353**	-0.227**	-0.142**	-0.094**	-0.174**	-0.104**
	(0.074)	(0.054)	(0.081)	(0.058)	(0.027)	(0.019)	(0.030)	(0.020)
Selection correction terms (Birth and residence state)	No	No	Yes	Yes	No	No	Yes	Yes
Education groups	2	2	2	2	4	4	4	4
Experience groups	2	4	2	4	2	4	2	4
Wald test for selection terms	–	–	229.41	335.04	–	–	272.43	498.48
			(0.000)	(0.000)			(0.000)	(0.000)
F-stat for difference	–	–	17.47	24.80	–	–	22.32	15.21
			(0.000)	(0.000)			(0.000)	(0.000)
Clusters	1,000	2,000	1,000	2,000	2,000	4,000	2,000	4,000
N	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959

Notes: Each column represents a regression where the dependent variable is individual log real wages and the independent variable of interest is the immigrant share of the individual's skill group. Skill groups for each regression are defined in rows 5 and 6. All regressions are weighted by the individual's Census person weight and include education, experience, state, and time fixed effects, as well as state-year, education-year, experience-year, state-education, and state-education-experience interactions. Cluster-robust standard errors in parentheses; observations are clustered by state-education-experience-year. A single asterisk * represents significance at the 5 percent level and two asterisks ** represent significant at the 1 percent level.

Table 3: Effect of immigration within skill groups on native wages, IV (2SLS)

	Two education groups				Four education groups			
	IV		IV with selection		IV		IV with selection	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Immigrant share	-0.367**	-0.222**	-0.473**	-0.312**	-0.152**	-0.059	-0.164**	-0.070
	(0.085)	(0.062)	(0.093)	(0.068)	(0.043)	(0.036)	(0.046)	(0.036)
Selection correction terms (Birth and residence state)	No	No	Yes	Yes	No	No	Yes	Yes
Education groups	2	2	2	2	4	4	4	4
Experience groups	2	4	2	4	2	4	2	4
Wald test for selection terms	–	–	229.40	334.90	–	–	272.60	498.14
			(0.000)	(0.000)			(0.000)	(0.000)
F-stat for difference	–	–	47.42	33.09	–	–	1.42	8.18
			(0.000)	(0.000)			(0.233)	(0.004)
F-stat for first stage regression	632.96	657.26	633.54	657.37	49.86	68.54	49.93	68.63
Clusters	1,000	2,000	1,000	2,000	2,000	4,000	2,000	4,000
N	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959

Notes: Each column represents a regression where the dependent variable is individual log real wages and the independent variable of interest is the immigrant share of the individual's skill group. Skill groups for each regression are defined in rows 5 and 6. All regressions are weighted by the individual's Census person weight and include education, experience, state, and time fixed effects, as well as state-year, education-year, experience-year, state-education, and state-education-experience interactions. Cluster-robust standard errors in parentheses; observations are clustered by state-education-experience-year. A single asterisk * represents significance at the 5 percent level and two asterisks ** represent significant at the 1 percent level.

Table 4: Effect of immigration within skill groups on native wages, OLS and IV (Additional selection correction terms)

	Two education groups				Four education groups			
	OLS		IV		OLS		IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Immigrant share	-0.371**	-0.238**	-0.519**	-0.358**	-0.183**	-0.109**	-0.213**	-0.111**
	(0.082)	(0.020)	(0.094)	(0.071)	(0.030)	(0.020)	(0.045)	(0.036)
Selection correction terms								
(Birth and residence state)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional correction terms	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education groups	2	2	2	2	4	4	4	4
Experience groups	2	4	2	4	2	4	2	4
Wald test for core selection terms	155.56	255.84	155.40	225.63	161.22	196.37	161.02	196.24
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wald test for additional selection terms	104.42	196.58	104.52	196.64	69.51	131.67	69.74	131.94
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F-stat for first stage regression	–	–	633.80	657.41	–	–	50.10	68.81
Clusters	1,000	2,000	1,000	2,000	2,000	4,000	2,000	4,000
N	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959

Notes: Each column represents a regression where the dependent variable is individual log real wages and the independent variable of interest is the immigrant share of the individual's skill group. Skill groups for each regression are defined in rows 5 and 6. All regressions are weighted by the individual's Census person weight and include education, experience, state, and time fixed effects, as well as state-year, education-year, experience-year, state-education, and state-education-experience interactions. Cluster-robust standard errors in parentheses; observations are clustered by state-education-experience-year. "Additional correction terms" are quadratics of the 1^{st} to 5^{th}-best probabilities of the 10 most immigrant-intensive states. A single asterisk * represents significance at the 5 percent level and two asterisks ** represent significant at the 1 percent level.

Table 5: Effect of immigration within skill groups on native wages, OLS and IV (Additional selection correction terms)

	Two education groups				Four education groups			
	OLS		IV		OLS		IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Immigrant share	-0.378**	-0.245**	-0.528**	-0.364**	-0.173**	-0.100**	-0.186**	-0.092*
	(0.083)	(0.059)	(0.095)	(0.070)	(0.030)	(0.020)	(0.045)	(0.036)
Selection correction terms (Birth and residence state)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional correction terms	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education groups	2	2	2	2	4	4	4	4
Experience groups	2	4	2	4	2	4	2	4
Wald test for core selection terms	175.78	255.41	175.68	206.95	176.25	196.37	176.20	206.94
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wald test for additional terms	89.59	150.48	90.16	151.19	68.37	113.41	68.46	113.49
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F-stat for first stage regression	–	–	634.81	658.65	–	–	50.05	68.75
Clusters	1,000	2,000	1,000	2,000	2,000	4,000	2,000	4,000
N	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959

Notes: Each column represents a regression where the dependent variable is individual log real wages and the independent variable of interest is the immigrant share of the individual's skill group. Skill groups for each regression are defined in rows 5 and 6. All regressions are weighted by the individual's Census person weight and include education, experience, state, and time fixed effects, as well as state-year, education-year, experience-year, state-education, and state-education-experience interactions. Cluster-robust standard errors in parentheses; observations are clustered by state-education-experience-year. "Additional correction terms" are quadratics of the 1st to 5th-best probabilities of the 15 most immigrant-intensive states. A single asterisk * represents significance at the 5 percent level and two asterisks ** represent significant at the 1 percent level.

Table 6: Effect of immigration within skill groups on native wages, OLS and IV (Time-varying control functions)

	Two education groups				Four education groups			
	OLS		IV		OLS		IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Immigrant share	-0.474**	-0.320**	-0.622**	-0.424**	-0.187**	-0.113**	-0.197**	-0.090*
	(0.089)	(0.065)	(0.103)	(0.078)	(0.031)	(0.021)	(0.048)	(0.037)
Time-varying selection correction terms (Birth and residence state)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education groups	2	2	2	2	4	4	4	4
Experience groups	2	4	2	4	2	4	2	4
Wald test for selection terms	83.66	112.90	83.60	112.71	77.11	126.59	77.15	126.53
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F-stat for first stage regression	–	–	638.46	655.29	–	–	49.95	68.59
Clusters	1,000	2,000	1,000	2,000	2,000	4,000	2,000	4,000
N	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959	7,407,959

Notes: Each column represents a regression where the dependent variable is individual log real wages and the independent variable of interest is the immigrant share of the individual's skill group. Skill groups for each regression are defined in rows 5 and 6. All regressions are weighted by the individual's Census person weight and include education, experience, state, and time fixed effects, as well as state-year, education-year, experience-year, state-education, and state-education-experience interactions. Cluster-robust standard errors in parentheses; observations are clustered by state-education-experience-year. A single asterisk * represents significance at the 5 percent level and two asterisks ** represent significant at the 1 percent level.

Table 7: Effect of immigration within skill groups on native wages after selection correction relative to uncorrected estimates

	Two education groups				Four education groups			
	OLS		IV		OLS		IV	
Size of coefficients relative to uncorrected estimates	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Corrected (core terms) *(Tables 2 and 3)*	130.7%	137.8%	128.9%	140.5%	122.5%	110.6%	107.9%	118.6%
Corrected (core + additional terms) *(Table 4)*	137.4%	144.2%	141.4%	161.3%	128.9%	116.0%	140.1%	188.1%
Corrected (time varying control function) *(Table 6)*	175.6%	193.9%	169.4%	191.0%	131.7%	120.2%	138.7%	95.7%
Education groups	2	2	2	2	4	4	4	4
Experience groups	2	4	2	4	2	4	2	4

Notes: Calculated from the point estimates of the coefficient on state-education-experience immigrant density provided in Tables 2, 3, 4 and 6.