

# Who Competes with Whom? Using Occupation Characteristics to Estimate the Impact of Immigration on Native Wages

Christopher Bollinger<sup>†</sup>

James Sharpe<sup>‡</sup>

## Abstract

Prior studies have examined the impact of immigration on native-born wages. These studies have relied upon education-experience groups to define labor markets and identify the wage elasticity of supply of immigrants. However, evidence suggests that immigrants' educational attainment is treated differently in the labor market and constructing labor markets based upon this characteristic leads to potentially biased conclusions. We utilize O\*NET occupational characteristics to form a different set of labor markets. Our analysis finds higher partial equilibrium effects on native born wages than prior work, as expected. These larger effects, however, are shown to be concentrated on the least skilled natives. Estimates of the total wage effect along the distribution of occupational skills confirm that the negative wage effect is concentrated on native workers in the bottom tail of the distribution. Natives in the upper tail of the distribution experience wage gains as a result of immigration.

**JEL Classification:** J24; J31; J61

**Keywords:** Immigration; Occupation-specific Skills; Native-Immigrant Complementarities

---

<sup>†</sup> University of Kentucky, Department of Economics, Email: [crboll@email.uky.edu](mailto:crboll@email.uky.edu)

<sup>‡</sup> Berry College, Department of Economics, Email: [jsharpe@berry.edu](mailto:jsharpe@berry.edu)

## 1. Introduction

A simple labor market model of supply and demand implies that immigrant (or any migrant) entrance to a local labor market will result in falling wage, *ceteris paribus*. Examining the implications of immigration on local labor markets has been an important topic in recent years, both within the economic literature and through the popular press. However, as recent surveys suggest (Kerr and Kerr, 2011; Dustmann et al., 2016), the results are far from uniform. A difficulty in estimation of the impact of immigration on earnings is identifying and isolating labor market cohorts. It has become standard in the literature to analyze the impact of immigration on similarly skilled natives within cohorts defined by education and work experience. This approach, pioneered in the immigration literature by Borjas (2003), implicitly assumes that within these cohorts, immigrants and natives are perfect substitutes. Recently, however, the assumption of perfect substitutability has been challenged, and estimates suggest that a degree of *imperfect* substitutability exists between immigrants and natives within these cohorts (Card, 2009; Ottaviano and Peri, 2012; Dustmann and Preston, 2012; Dustman et al, 2012; Manacorda et. al, 2012). As pointed out by Ottaviano and Peri (2012), this fact is nontrivial. If immigrants and natives are imperfect substitutes, then any wage effect of immigration would be concentrated on existing immigrants, not natives.

The result that education is an imperfect proxy for overall skill level, and especially so for immigrants, is well-documented in the literature. Immigrants earn less than similarly educated natives (Bratsberg and Terrell, 2002; Bratsberg and Ragan, 2002; Ferrer and Riddell, 2008; Friedberg, 2000). This wage difference is attributed to differing employment distributions across occupations and a lower return to education for immigrant workers. More significantly, immigrants face differential returns to education due to downgrading upon arrival to the US. (Dustmann et al., 2016, Mattoo et al., 2008; Neagu, 2009; Sharpe, 2015). More recently, Sharpe (2015) uses O\*NET data for the required level of education needed to adequately perform the job and finds that immigrants are twice as likely to be overeducated as native-born for the positions they hold. The difference in the rates increases with the amount of schooling and is most profound for highly-educated newly arrived immigrants, who are as much as four times more likely to be overeducated for their positions. That is, immigrants enter the US and are pushed toward jobs in which they possess too much education compared to the average worker.

Prior studies have used several methods to account for the imperfect substitutability between immigrants and natives. Some have used a structural approach to estimate the general equilibrium effects of immigration on native wages (Llull, 2010; Mancorda et. al, 2012; Ottaviano and Peri, 2012; Peri and Sparber, 2009; among others). The validity of this approach, however, has been questioned as the

estimated elasticities of substitution are not robust to changes in key assumptions (Borjas et al., 2012). Others have moved away from the education-experience cohort analysis altogether and analyzed varying definitions of skill cohorts: along the native wage distribution (Dustmann et. al, 2012), industry cohorts (Dustmann and Glitz, 2015), and occupation cohorts (Camarota, 1997; Card, 2001; Orrenius and Zavodny, 2007; Steinhardt, 2011).

In this paper, we propose that an improved measure of labor market competition is to stratify the labor market by occupation. Existing studies incorporating occupations as a proxy for skill are relatively sparse. To our knowledge, only four such studies exist. Camarota (1997) uses one CPS cross-section to estimate the impact of immigration on wages within occupations and finds that a 1% increase in immigration will decrease the wages of the average native worker by 0.5%. However, the use of a single cross-section and small within-occupation sample sizes, make causal inference difficult. Card (2001) estimates city-specific impacts of immigration on occupational wages for 175 cities using 1990 US Census data and finds that the immigration inflows of the 1980's decreased wages in low-skilled occupations in high-immigration cities by no more than 3%. Orrenius and Zavodny (2007) use CPS data from 1994 – 2000 and INS immigration data to estimate the impact of immigration on native wages in three broad occupation categories. The authors estimate that the change in immigrants over the data period decreased wages in low-skilled, manual occupations 0.8% and had no impact for medium-skilled and high-skilled occupations. In the case of the German labor market, Steinhardt (2011) shows that the wage effect of immigration is significantly larger when stratifying the labor market by occupation (-0.394) compared to stratifying the labor market by education (-0.0478).

To improve upon the prior studies, we propose constructing labor market cohorts by using occupational groups with similar *skill* requirements along with experience. Whereas prior studies using occupations have relied on broad Census-defined occupation groups, we construct occupation groups using skill data from the O\*NET. The advantage of using the O\*NET data is that we are able to construct occupation groups with a greater degree of homogeneity in overall skill level, regardless of nationality and citizenship status, than those using either education groups or broad occupation classifications. Section two presents the details of our data and the methodology to construct occupation-experience groups.

Our initial result, presented in section 3, incorporates our occupation classification to estimate the impact of immigrants on wages in the national labor market. This approach is compared to estimates obtained by (Borjas, 2003). As expected, if occupational characteristics are an improvement in constructing labor markets, we find that the partial equilibrium effect of immigration on native wages is 4

-5 times larger than the effect estimated utilizing cohorts stratified by education group. This result is robust to several different definitions of occupation groups defined on the basis of occupational skill.

Several potential concerns arise from the initial analysis. First, stratifying labor markets by occupations may introduce bias, as occupational choice is likely endogenous. Second, the national labor market approach identifies the partial equilibrium effect of immigration on native wages. The effect of immigration on the wages of experienced native workers relative to inexperienced native workers in the same occupation (or education) group (Dustmann et al., 2016). Third, the consensus in the literature is that the effect of immigration on natives would be concentrated on the least-skilled native workers who are most substitutable to immigrant labor.

In section 3 we address the first concern by augmenting the usual education-experience skill cohorts with demographic data to identify native-born who are comparable to immigrants. The resulting estimated wage elasticity is comparable to the wage elasticity we estimated using the occupation-experience skill cohorts. This supports our analysis above and allows other researchers an alternative methodology.

In section 4, we incorporate our occupation classification into the total wage effect framework and estimate the total wage effect of immigration along the distribution of occupational skills. This addresses both the second and third concern, as it estimates a general equilibrium analysis over the distribution, allowing for significant heterogeneity. While the estimated wage effect is not statistically distinguishable from zero in the middle of the skill distribution, we estimate large and statistically significant effects in each tail of the distribution. Specifically, we estimate a wage elasticity of -0.6 in the bottom tail where natives are substitutable to immigrant labor and a wage elasticity of 0.7 in the upper tail where natives are complementary to immigrant labor.

## **2. Data**

We draw from several data sources in this paper. Labor supply and wage data derive from the 1960, 1970, 1980, 1990, and 2000 Public Use Micro Samples (PUMS) of the U.S. Census, and the 2009, 2010, and 2011 Public Use American Community Survey (ACS). The ACS data are pooled together to form a single 2010 cross-section. Following the work of Borjas (2003), we restrict our sample to men, aged 18-64, who earned positive wage income. A full description of both the employment and wage samples can be found in the Data Appendix.

We sort workers into skill groups based on potential experience and occupation. As is customary in this literature, we calculate potential experience based on educational attainment. It is assumed that

workers with less than a high school diploma enter the labor market at 17 years old, workers with a high school diploma or GED enter the labor market at 19, workers with some college enter the labor market at 21, and those with a college degree enter the labor market at 23. Following Borjas (2003), we limit the sample to men who have 1-40 years of potential experience and group workers into 5-year potential experience groups (i.e. 1-5 years of potential experience, 6-10 years, etc.).

### 2.1. Occupation Groups

The occupation groups constructed in this paper follow from Peri and Sparber (2009). We assume that occupations are distinguished by two occupation-specific indices of task intensity: manual task intensity and communicative task intensity. Individual occupations are then grouped based on their relative communicative-to-manual task intensity.

Occupation-specific task indices are constructed using the Department of Labor’s O\*NET survey, which provides comprehensive data on characteristics of occupations. The O\*NET content model is partitioned into several different domains, each providing different worker-specific and occupation-specific data. Unlike Peri and Sparber (2009), we make use of *both* worker-specific data on abilities, knowledge, and skills *and* occupation-specific data on work activities to generate these task intensity indices (throughout the rest of the paper, we will refer to all four of these measures as “skill groups”).<sup>3</sup> Table A1 of the Data Appendix lists each skill used in constructing the task intensity indices.

One challenge when working with occupations over this many Census years is that occupation classifications change over time. Additionally, O\*NET data are assigned to 2000 SOC (standard occupation classification) occupations. To remedy this problem, we use a modified occupation classification developed by Autor and Dorn (2013) (AD classification, hereafter). This occupation classification system creates a consistent, balanced panel of occupations across all years. To construct the occupation groups used in this paper, we merge skill data from the O\*NET survey to the AD classification and group occupations on the basis of their occupation-specific skills.

The O\*NET data assigns each skill a score for importance ( $I$ ) with a range of 0-5 and a score for level ( $L$ ) with a range of 0-7 for each occupation.<sup>4</sup> To create the occupation-specific skill index, we first standardize the importance and level scores such that each has a range of 0-100. Then, we create a normalized “task-intensity score” ( $TS$ ) for each skill by multiplying the standardized importance score

---

<sup>3</sup> Peri and Sparber (2009) rely solely on “abilities” from the O\*NET survey.

<sup>4</sup> Importance and Level scores measure different aspects. There are occupations in which a given skill is equally important; however, one occupation needs to use the skill at a much higher level. An example from the O\*NET is speaking ability for lawyers and paralegals. Speaking is important in both occupations; however, lawyers need a high level of speaking skills to argue cases, while paralegals simply need an average level of speaking skill (<https://www.onetonline.org/help/online/scales>).

and standardized level score – a higher task-intensity score suggests a given task is more important to performing a given occupation. We then calculate the average manual and communicative task-intensity score for each skill group and occupation. For example, within the worker ability domain, both physical abilities and psychomotor abilities are classified as manual abilities. Thus, for each occupation, we calculate the average manual task-intensity score by averaging the task-intensity of physical and psychomotor abilities. Lastly, the final manual (communicative) task-intensity score is the average of all skill-group specific manual (communicative) task-intensity scores. Analytically, the manual task intensity index for each occupation ( $j$ ) is calculated as<sup>5</sup>:

$$(1) \quad M_j = \frac{1}{n} \sum_i (\overline{TS}_{ij}) \quad \forall i = (Ability, Knowledge, Skill, Work Activity).$$

For each occupation in the AD classification, we create the ratio of communicative task intensity to manual task intensity, which is the basis for defining our occupation groups. From this ratio, we construct three occupation classifications based on the distribution of this skill ratio across occupations: 1) a four occupation group classification where each group is a quartile of the distribution, 2) a five occupation group classification where each group is a quintile of the distribution, and 3) a six occupation group classification where each group is a sextile of the distribution.

### 3. National Labor Market Approach

To be responsive to existing literature, we start by following Borjas et al. (1997) and Borjas (2003) by moving away from the area studies of Card (2001) and Orrenius and Zavodny (2007) and treat the U.S. as one national labor market. Concern arises over endogeneity in immigration to area labor markets. This endogeneity may take several forms. Immigrants may choose to locate in high wage cities, natives may respond to immigrant inflows by relocating, or firms may reallocate capital to high-immigrant cities in order to take advantage of the abundance of cheaper labor. To alleviate this concern, Borjas et al. (1997) suggested that the analysis move away from analyzing local labor markets; rather, researchers should use national-level data and treat the entire US as one labor market.

The empirical model is derived from a theoretical framework of the demand side of the labor market. Assuming output is produced using a CES production function where labor and capital are separable, the relative wage of a given skill group is a function of 1) the population share within the group

---

<sup>5</sup> The communicative task intensity index is constructed analogously.

and 2) a group specific productivity component.<sup>6</sup> Following Borjas (2003), this group-specific productivity component is absorbed by a collection of fixed effects:

$$(2) \quad w_{ijt} = \beta s_{ijt} + \theta_i + \varphi_j + \tau_t + (\theta_i * \tau_t) + (\varphi_j * \tau_t) + (\theta_i * \varphi_j) + \varepsilon_{ijt}.$$

Here,  $w_{ijt}$  is the mean of the log weekly wage of natives in occupation group  $i$  and experience group  $j$  at time  $t$ .  $s_{ijt}$  is the share of immigrants in occupation group  $i$ , experience group  $j$  at time  $t$ , making  $\beta$  the coefficient of interest. The share of immigrants in a skill group ( $i, j$ ) is represented as the percent of total hours worked by immigrants. The remaining controls are vectors of linear fixed effects for occupation group ( $\theta_i$ ), experience group ( $\varphi_j$ ) and year ( $\tau_t$ ) to control for differences in average wages across occupation groups, experience groups, and over time. The interaction of occupation fixed effects with time ( $\theta_i * \tau_t$ ) and experience group fixed effects with time ( $\varphi_j * \tau_t$ ) control for the fact that the impact of occupation or experience on average wages may change over time. Lastly, the interaction of occupation fixed effects and experience group fixed effects ( $\theta_i * \varphi_j$ ) controls for any differences in the impact of experience on average wages across occupation groups. Thus, the impact of immigration on native wages is identified by variation in immigrant shares within occupation groups and experience groups over time.

---

<sup>6</sup> For derivation of the model in the context of immigration, see Card (2001) or Borjas (2003).

**Table 2: Reduced Form Estimates of  $s_{ijt}$** 

VARIABLES	(1)	(2)	(3)	(4)	(5)
	Educ-Exp $w_{ijt}$	Occ-Exp (Quartile) $w_{ijt}$	Occ-Exp (Quintile) $w_{ijt}$	Occ-Exp (Sextile) $w_{ijt}$	Occ-Exp (Dorn) $w_{ijt}$
<b>Panel A: Preferred Model</b>					
$s_{ijt}$	-0.171* (0.099) [-0.126]	-0.925*** (0.215) [-0.680]	-0.674** (0.271) [-0.496]	-0.769*** (0.199) [-0.565]	-0.356** (0.158) [-0.262]
<b>Panel B: Includes Native Labor Force as Control Variable</b>					
$s_{ijt}$	-0.194* (0.108) [-0.142]	-0.794*** (0.236) [-0.584]	-0.607** (0.302) [-0.446]	-0.821*** (0.198) [-0.603]	-0.353 (0.214) [-0.260]
<b>Panel C: Dependent Variable is Residualized Log Wages</b>					
$s_{ijt}$	-0.134 (0.105) [-0.099]	-0.593*** (0.197) [-0.436]	-0.529** (0.208) [-0.389]	-0.527*** (0.136) [-0.387]	-0.224* (0.128) [-0.165]

- 1) Each column and row represent a unique specification estimated using data from the 1960-2000 Census and the pooled 2007-2010 ACS. Each column differs based on the definition of skill (education or one of the occupation groups), while each row differs based on the label – estimates of equation (3) in Panel A, estimates of equation (3) including the native labor force as a control variable in Panel B, and estimates of equation (3) using residualized log wages as the dependent variable. The dependent variable is mean of the log weekly wage of natives in each skill group (except the final specification where the dependent variable is residualized log wages). The independent variable of interest is the share of total hours worked by immigrants in a given skill group. All specifications include year fixed effects, occupation (or education in column 1) fixed effects, experience group fixed effects, and interactions of all fixed effects. Regressions are weighted by the number of native persons used to calculate the average wage in each cohort.
- 2) Robust standard errors clustered by skill group are reported in parentheses and the corresponding elasticity estimates are listed in brackets.

Equation (3) is estimated via OLS and the estimated coefficients are reported in Table 2. Each column/row of Table 2 represents a different specification of (3). The columns differ by skill group classification (i.e. Education-Experience, Occupation (4 group)-Experience, etc.). Panel A reports our preferred specification where the regression is weighted by the number of observations used to calculate the average wage within a cell. We also present several robustness checks. Panel B presents estimates when we include native labor force as an explanatory variable. Because the key explanatory variable is simply the immigrant share of total hours worked within a skill group, an increase in  $s_{ijt}$  would occur from either an increase in immigrant labor supply *or* a decrease in native labor supply. As such, these estimates report the impact of  $s_{ijt}$  holding native labor supply constant. Panel C reports estimates where

the dependent variable is the residualized log wages in each cell (Card, 2001; Jaeger et al., 2018).<sup>7</sup> In each panel, we also reports the corresponding elasticities from the estimated coefficients in brackets.<sup>8</sup>

We start by discussing our preferred estimates in Panel A. Column (1) reports estimates using the traditional education-experience classification found in the existing literature.<sup>9</sup> The baseline results are slightly lower than those found by Borjas (2003).<sup>10</sup> Focusing on the estimated elasticity in brackets, the results suggest that a 1% supply shock (an inflow of immigrants that increases total hours worked within an education-experience cohort by 1%) will reduce native wages by a modest 0.13%. Columns (2) – (5) use different occupation classifications in the estimation of equation (3). Columns (2) – (4) use occupation groups defined by the distribution of the communicative-to-manual task intensity ratio. When we group workers based on occupation-specific skills, the estimated impact of immigration is much larger. Again, focusing on the elasticities in brackets, the results suggest a 1% supply shock within a given occupation-experience cohort will decrease native wages by 0.68%, 0.50%, and 0.57%, respectively. The results support the hypothesis that defining skill groups on the basis of education may attenuate the effects of immigration. By grouping workers into skill groups defined by occupation, the estimated impact on native wages is 4-5 times larger depending on specification.

Concern arises that our results are not driven by a careful construction of markets, but rather by other factors associated with defining labor markets by occupational groups. To test this, we estimate the model using the occupation classification system developed by Autor and Dorn (2013). Results are reported in column (5). Recall that these occupation groups are similar to the typical occupation classifications used in the U.S. Census and are not defined based on occupation-specific skills.<sup>11</sup> If the results are driven simply because we use occupations to define skill groups, we would expect the impact of immigration to be similar to columns (2)-(4). When using this occupational group classification however, the impact of immigration is significantly lower and similar in magnitude to the estimates when using education-based skill groups. This is unsurprising as AD rely on average educational attainment when constructing these groups, not occupation-specific skills.<sup>12</sup> To see this in the data, Figure 1 plots

---

<sup>7</sup> To residualize wage, we fit an OLS regression of log wages on observable individual characteristics (education, potential experience and its square, indicator variables for marital status, full-time worker status, and race) and Census region fixed effects

<sup>8</sup> The share of immigrants within a skill group ( $s_{ijt}$ ) in Eq. 3 is not in log form rather an approximation. As such, we calculated the corresponding elasticities as in Borjas (2003).

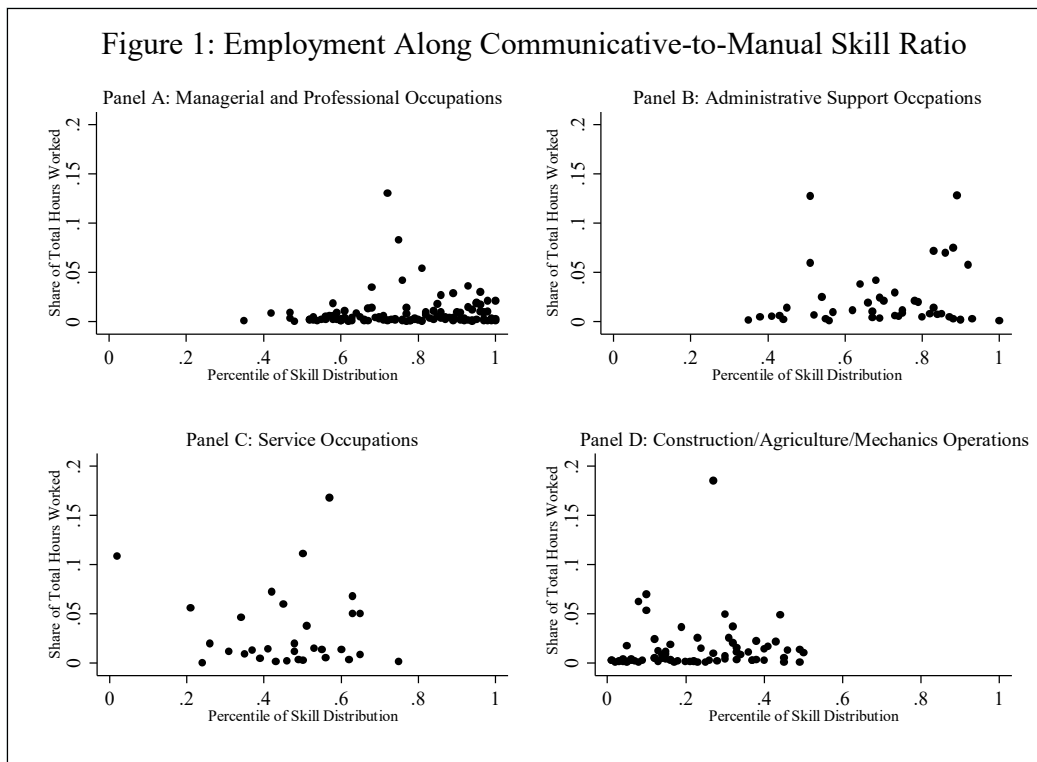
<sup>9</sup> In this specification, we use the four-group classification described above (Less than HS, HS grad, some college, college grad).

<sup>10</sup> Borjas (2003) estimates a point estimate of -0.572; however, this estimate does not use data from 2010 and uses CPS data for 2000. We provide replication results and sensitivity tests in the Appendix. When using the methodology above and the same data described in Borjas (2003), produce very similar results. Thus, the methodology used above is consistent with the past literature.

<sup>11</sup> The occupation groups are as follows: 1) Management/Professional/Technical/Financial/Public Security, 2) Administrative Support and Retail Sales, 3) Low-Skill Services, 4) Precision Production and Craft Occupations, 5) Machine Operators, Assemblers, and Inspectors, and 6) Transportation/Construction/Mechanics/Mining/Agricultural.

<sup>12</sup> In describing one of the occupation groups, the authors claim: “Technical, sales, and administrative support occupations cover a workforce that is on average better educated than any other occupation group apart from managers and professionals”.

the share of total hours worked along the distribution of our skill ratio within AD occupation groups. Panels A and B are white-collar jobs (i.e. management occupations, etc.) and panels C and D are low-wage blue-collar jobs (i.e. construction, services, etc.). Though labor supply is skewed in the expected direction for each occupation group (white-collar occupations are skewed toward relatively more communicative skills, and vice versa), the variance is quite high. Because of this variability, it is reasonable to assume that, similar to skill groups defined by educational attainment, not all workers will directly compete in the labor market. Thus, we take the result in column (5) as support for the claim that occupation-specific *skills*, not occupations themselves, are the important component in constructing skill groups for which labor market competition is identified.



Lastly, several robustness checks lend credence to the preferred specification. The results in Panel B suggest the results are not sensitive to controlling for native labor supply. As before, when defining labor market cohorts on the basis of occupational skill, the results suggest a 1% supply shock would lower native wages by roughly 0.45 – 0.60%. The estimates in Panel DC, however, suggest slightly lower wage elasticities, regardless of specification. When we purge the effects of demographic characteristics from average wages, the results suggest a 1% supply shock would lower wages by roughly 0.4%. Although the estimated elasticities are lower in absolute value, the general pattern is the same. When we

stratify labor markets by occupational skill, the partial equilibrium (relative) wage effect is 4-5 times larger compared to the results obtained when stratifying labor markets by education.

### 3.1. *Who Competes With Whom?*

The question of “who competes with whom?” in the labor market is the motivation for this paper. The motivation for stratifying the labor market into skill cohorts is to estimate the impact of immigration on the wages of demographically comparable natives. To this point, we have argued that occupation-experience cohorts are superior to education-experience cohorts because we define skill groups for which immigrants and natives directly compete in the labor market. That is, immigrants and natives with similar work experience are perfect substitutes within occupations while imperfect substitutes within education groups. While the above results suggest this to be the case, two concerns arise from the above methodology.<sup>13</sup> First, there may be some concern regarding the seeming arbitrariness with which we define the number occupation groups.<sup>14</sup> Second, occupational choice of immigrants is potentially endogenous in that immigrants may choose occupations based on favorable labor market conditions. If so, the estimates in Table 2 would be biased upward. However, if immigrants are systemically under-placed in the labor market and forced into lower wage jobs, then the estimates in Table 2 would be biased downward.<sup>15</sup> It is the latter that influenced the use of education-experience cohorts in the early literature.

An alternate way to approach the question of “who competes with whom?” is to let the data determine which native workers are demographically comparable to immigrants. In this section, we return to the standard education-experience skill cohort. The use of education-based skill cohorts in this section is advantageous for two reasons. First, switching occupations is significantly easier than switching education groups. As discussed above, there may be doubt as to whether the estimates in Table 2 result from defining more homogeneous skill groups or bias introduced by using occupations. Second, this analysis provides a test to the claim that imperfect substitutability within education groups is the primary force behind the counterintuitive results seen in the previous literature.

---

<sup>13</sup> Another potential concern is the use of occupation defined skill groups while treating the U.S. as a single labor market. This national labor market approach has been widely adopted in the literature when assessing the impact of immigration on native wages. However, previous studies using occupation groups have used the area approach – treating metropolitan areas as their own distinct labor market. We have estimated the *local* labor market effect of immigration, treating individual metropolitan areas as local labor markets. The results are similar (an elasticity of -0.4) and available upon request.

<sup>14</sup> While this is a legitimate concern, we have estimated the above model using occupation classifications with as many as 10 occupation groups (dividing the skill distribution by centiles) and the underlying result does not change. These results are available upon request.

<sup>15</sup> We tested for endogeneity of occupational choice across skill groups. When regressing immigrant penetration in a particular skill group ( $s_{ijt}$ ) on lagged native wages ( $w_{ijt-10}$ ), the resulting coefficient is not statistically significant and essentially zero.

To identify demographically comparable natives, we begin by modeling the relationship between observable characteristics and the nativity of the worker. We first estimate, using the same data as above, the following flexible probit model on male workers for each Census year separately:

$$(3) \quad \Pr(I_i = 1) = \Phi(\beta X_i + \gamma OCC_i + \delta GEOG_i).$$

$I_i$  is a dummy variable equal to 1 if the worker is an immigrant;  $X_i$  is a vector of worker demographics including education, marital status, race, a quartic in potential experience, and a full set of education-by-demographic interactions;  $OCC_i$  is a vector of occupation-specific controls including AD occupation group fixed effects and the communicative-to-manual skill ratio;  $GEOG_i$  is a vector of geographic location controls including metropolitan status, state fixed effects, and a state-by-metro interaction.<sup>16</sup> We use the estimated coefficients to predict the probability of being an immigrant for all natives in the sample. We assume that native workers who more closely resemble immigrants in the data are also more likely to compete with immigrants in the labor market.

Table 3 reports the average labor market and demographic characteristics of native workers in four quartiles that reflect the intensity with which they will compete with immigrants in the labor market (i.e. Quartile 1 are the native workers least like immigrants in the data). Hours worked, weeks worked, potential experience, and the percentage of workers who are part-time are all fairly constant across quartiles. Perhaps counterintuitively, average weekly wages are *higher* among natives that are *more* likely to compete with immigrants in the data. However, this confounding result can be explained by the fact that those in quartiles 3 and 4 are much more likely to reside in metropolitan areas where wages are higher. In addition, native minorities are much more likely to compete with immigrants—the proportion of white workers decreases uniformly across the quartiles. Lastly, the differences across education and occupation groups are as expected. Native workers who are more likely to compete with immigrants are those with less education and work in low-skill services and transportation, construction, mining, and agricultural occupations.

To estimate the impact of immigration on the native wages, we estimate the same reduced-form model in equation (2). The lone difference is the dependent variable is now the average log weekly wage of demographically comparable natives in a given competition quartile within a given education-experience cohort. The results are presented in Table 4 below.

---

<sup>16</sup>We estimated the initial probit models without occupation fixed effects and the skill ratio and the results are not sensitive to their exclusion. These results are available upon request.

**Table 3: Native Worker Characteristics by Intensity of Competition with Immigrants (2000)**

	<b>Low Competition</b>	<b>Medium Competition</b>	<b>High Competition</b>	<b>Very High Competition</b>
<b>Observations (N)</b>	<b>1,329,187</b>	<b>1,329,188</b>	<b>1,329,189</b>	<b>1,329,182</b>
<i>Weekly Wage</i>	\$638.06	\$714.06	\$796.91	\$742.65
<i>Hours Worked per Week</i>	39.67	40.14	40.36	40.15
<i>Weeks Worked per Year</i>	45.05	45.83	45.95	45.46
<i>Potential Experience</i>	17.92	18.74	18.82	18.69
<i>White</i>	93.17%	89.75%	86.77%	64.12%
<i>African-American</i>	6.83%	10.20%	12.59%	13.30%
<i>Full-Time</i>	69.54%	72.76%	73.74%	71.20%
<i>Live in Metropolitan Area</i>	27.58%	69.03%	87.63%	93.13%
<b>Education Groups</b>				
<i>Less Than High School</i>	5.20%	7.10%	7.76%	11.92%
<i>High School Graduate (or GED)</i>	39.13%	41.71%	40.21%	46.41%
<i>Some College</i>	25.13%	24.69%	24.58%	25.69%
<i>College Graduate</i>	30.53%	26.50%	27.45%	15.99%
<b>Occupation Groups (AD)</b>				
<i>Management &amp; Professional</i>	37.39%	36.70%	37.93%	32.56%
<i>Administrative Support &amp; Retail Sales</i>	31.17%	25.79%	24.52%	19.77%
<i>Low-Skill Services</i>	8.48%	11.36%	12.12%	17.94%
<i>Precision Production &amp; Craft</i>	3.11%	3.63%	3.61%	3.88%
<i>Machine Operators &amp; Assemblers</i>	4.32%	6.28%	6.59%	6.84%
<i>Transportation, Construction, Mining, Agricultural</i>	15.54%	16.24%	15.24%	19.02%
<b>Occupation Groups (Skill Based)</b>				
<i>Quintile 1</i>	11.15%	12.92%	12.40%	14.33%
<i>Quintile 2</i>	11.65%	13.12%	12.94%	16.08%
<i>Quintile 3</i>	22.03%	21.30%	20.82%	22.95%
<i>Quintile 4</i>	29.34%	27.00%	26.81%	24.28%
<i>Quintile 5</i>	25.83%	25.66%	27.03%	22.35%

Column (1) of table 4 presents the estimates from the full model (as in Table 2). Then, in columns (2) – (4), we report the estimates impact on the wages in each intensity quartile. For example, the dependent variable in column (2) is the average log weekly wage of natives in the lowest competition intensity quartile. Recall that by modeling skill groups on the basis of education and experience, the implicit assumption is that all workers within these skill groups are perfect substitutes. In theory, we would expect the impact of immigration on the wages to be the same across all columns because all natives should compete equally with immigrants in the labor market. From the estimates in Table 4, we

see that the theory does not hold. For those natives that least resemble immigrants in the data (column 2), the estimated effect of immigration on wages is *positive*, reflecting complementarities between those natives and immigrant labor. As the intensity of competition increases, the impact of immigration on native wages becomes more negative and statistically significant. For natives that most close resemble immigrants in the data (Very High Competition), the effect of immigration is highly statistically significant with an implied elasticity of -0.33.

The same general pattern is seen in columns (6) and (7). Column (6) presents results where the dependent variable is average log weekly wage of demographically comparable natives in the high and very high competition groups, while column 7 presents results for natives in the low and medium competition groups. From column (6), the wage elasticity is similar to those in Table 2 and suggests that a 1% immigration shock would decrease the wages of these natives by 0.37%. As a result, we can conclude that the results in Table 2 are not the result of the endogeneity of occupational choice; rather, it is the construction of a more homogeneous group of perfectly substitutable workers that directly compete in the labor market.

**Table 4: Impact on Demographically Comparable Natives**

	(1) All	(2) Low Competition	(3) Medium Competition	(4) High Competition	(5) Very High Competition	(6) High and Very High Competition	(7) Low and Medium Competition
VARIABLES	$w_{ijt}$	$w_{ijt}^{Q1}$	$w_{ijt}^{Q2}$	$w_{ijt}^{Q3}$	$w_{ijt}^{Q4}$	$w_{ijt}^{Top50}$	$w_{ijt}^{Bot50}$
<i>Immigrant Share</i> ( $s_{ijt}$ )	-0.171* (0.0988) [-0.126]	0.265* (0.143) [0.195]	-0.217 (0.176) [-0.159]	-0.365** (0.171) [-0.268]	-0.447** (0.186) [-0.329]	-0.508*** (0.139) [-0.373]	0.063 (0.137) [0.046]
Observations	192	192	192	192	192	192	192
R-squared	0.999	0.998	0.998	0.998	0.998	0.999	0.999

1. Each column represents a different specification. The dependent variable in column (1) is the mean log native wage in a given education-experience group. The dependent variables in columns (2) – (5) are the mean log wages of natives in competition intensity quartile  $j$  in each education-experience group. In column (6), the dependent variable is the mean log wages of natives in competition intensity quartiles 3 and 4 (thus, the top 50% most likely to compete with immigrants). Column (7) reports estimates for the bottom 50%. The independent variable of interest is the share of total hours worked by immigrants in each education-experience group. Robust standard errors clustered by skill group are reported in parentheses. Elasticities are reported in brackets.
2. All regressions are weighted. The weights are the sample size used to create the average log weekly wage in a given cohort.

Overall, these results are consistent with the previous literature in one important way. Recent research suggests that the effect of immigration on native wages, if there is one at all, is concentrated on the least educated natives (Cortes, 2008; Ottaviano and Peri, 2012; Dustman et al., 2013). So, while the estimated elasticity is larger (in absolute value), the results of this paper fit nicely with this interpretation. Immigrant inflows have no effect on the wages for those native workers who are least similar to

immigrants (columns 2, 3, and 7); however, the effect increases as the similarity between natives and immigrants increases. This may also explain how different researchers obtain different estimates. Depending on the mix of occupation and the associated skills, the average response over the entire labor market – typically measured by researchers-- would be a mix of these different coefficients, and hence would vary.

#### 4. Effect Along the Occupational Skill Distribution

While the national labor market approach is appealing for the reasons discussed above, past studies suggest this approach may overstate the negative effect of immigration on wages. Dustmann et al. (2016) suggest that the national labor market approach identifies the *relative* wage effect of immigration (within a given skill cell), not the *total* wage effect that accounts for skill complementarities. In order to uncover the total effect of immigration on native wages, researchers have turned to a spatial correlation approach that relies on variation in immigrant inflows across regions (i.e. metropolitan areas, commuting zones, or states) to identify the effect on native wages.

Regardless of the definition of the labor market (i.e. local vs. national), however, empirical methods estimating the average effect likely masks significant heterogeneity along the skill distribution. In fact, the general consensus in the literature is that the effect of immigration, to the extent that one exists, is concentrated on the least skilled natives. The two most salient examples in the existing literature are Altonji and Card (1991), who report a wage elasticity of -1.1 for white male high school dropouts, and Dustman et al. (2012), who find large negative effects (-0.5) at the bottom of the income distribution and large positive wage effects (0.4) near the top of the wage distribution. Given that immigrant inflows have become less skilled, on average, over time and immigrants typically downgrade upon arrival, one may also expect that the wage effect should be concentrated at the bottom tail of the occupational skill distribution. As such, in this section, we incorporate our occupation group classification into the spatial correlation approach to estimate the *total* wage effect of immigration along the occupational skill distribution.

In the spirit of Dustmann et al. (2012), we estimate the effect of immigration along the distribution of occupational skills at 10 percentage point intervals. We start by defining the occupational skill distribution based on the skill index defined in Section 2. Occupations are then placed into one of 10 groups ( $j$ ) according to their position on the skill distribution. For example, occupations with a communicative-to-manual skill index below the 10<sup>th</sup> percentile are grouped in  $j=1$ , while occupations with a communicative-to-manual skill index above the 90<sup>th</sup> percentile are grouped in  $j=10$ . We then estimate the effect of immigrant inflows, separately, for each group. Specifically, we regress changes in the

average log wage of native workers at a given point in the skill distribution ( $j$ ) across local labor markets ( $k$ ) on changes in the immigrant ratio in local labor market ( $k$ ), local labor market-specific controls for changes in demographics ( $\Delta X_{kt}$ ), and year fixed effects ( $\theta_t$ ). Because measurement error is a concern for cells with small sample sizes, labor markets are not stratified by potential experience in this analysis.<sup>17</sup>

The formal model is written as:

$$(4) \quad \Delta \log w_{jkt} = \beta \Delta s_{kt} + \gamma \Delta X_{kt} + \theta_t + \Delta \epsilon_{jkt}$$

We estimate (4) using the sample of working-age noninstitutionalized male population from IPUMS data from the 1990 Census, 2000 Census, and 2010 American Community Survey. We follow the existing literature (Card, 2001; Jaeger et al., 2018) by constructing the dependent variable using residualized log wages.<sup>18</sup>  $\Delta X_{kt}$  includes controls for changes in the average age of the population, changes in average educational attainment, changes in the percentage minority, and the changes in the average occupational skill ratio. As is common in the literature, we define local labor markets ( $k$ ) as metropolitan areas (MSAs). One potential concern with this approach, however, is a lack of observations for certain MSAs. Aydemir and Borjas (2011) show that measurement error from small sample sizes leads to significant attenuation bias in the estimated effect of immigration on native wages. As such, we limit the analysis to the 150 metropolitan areas with the largest number of immigrant observations.<sup>19</sup>

Because the location choices of immigrants across local labor markets are endogenous, equation (4) is estimated via 2SLS using a variant of the shift-share instrumental variable (e.g. Antonji and Card, 1991; Card, 2001). Specifically, we predict an immigrant inflow to each metropolitan area based on the country of origin-specific historical migration patterns and occupational distributions of immigrants. The idea is that while current immigrant location choices are endogenous with respect to local labor market conditions, historical migration patterns are not. This identifying assumption relies on the common result that the most important determinant of location choice is the share of the existing population that is foreign born.

Our version of the shift-share instrument follows most closely to that in Card (2001). The predicted immigrant inflows assume that each metropolitan area will receive the same share of

---

<sup>17</sup> The more traditional approach to estimating the total wage effect would also stratify labor markets by potential experience (as suggested by Dustmann et al, 2016). We have estimated such a specification for several different subgroups of the population and report the estimates in Table A1 of the Appendix. The results complement those in the paper.

<sup>18</sup> To residualize wage, we fit an OLS regression of log wages on observable individual characteristics (education, potential experience and its square, indicator variables for marital status, full-time worker status, and race) and Census region fixed effects.

<sup>19</sup> Though not reported here, we also estimated the model using different samples of MSAs. The results are quantitatively and qualitatively similar when we restrict the analysis to the top-100 CBSAs or expand the sample to include all CBSAs (available upon request).

immigrants from country  $c$  that worked in occupation  $o$ . In this analysis, the initial shares are based on the immigrant distribution in 1980. Specifically, we predict immigrant inflows by:

$$(5) \quad \hat{I}_{kt} = \sum_{k=1}^n \sum_{c=1}^n \gamma_{c,o,1980} * \alpha_{c,k,1980} * I_{c,US,t}.$$

The first term on the right-hand side,  $\gamma_{c,j,1980}$ , is the share of newly arriving immigrants employed in occupation  $o$  (in their home country) that arrived in the US from 1972-79. The second term,  $\alpha_{c,k,1980}$ , is the share of immigrants from country  $c$  that migrated to metropolitan area  $k$  in 1980. While the INS data in 1972-1979 provide detailed information on occupation, these data do not identify metropolitan area of intended residence. As such, we follow the existing literature and make use of US Census data to construct  $\alpha_{c,k,1980}$ . Specifically, we use the 1980 Census and calculate the share of newly arriving immigrants (defined as those who arrived at the US in the last 10 years) from country  $c$  that reside in metropolitan area  $k$ . The third term,  $I_{c,US,t}$ , is the total number of immigrants from country  $c$  that arrived in the US in each subsequent decade (1990, 2000, and 2010). We utilize annual data from the INS (from the Department of Homeland Security in more recent years) to construct the decadal inflows of immigrants from 1980-2010 ( $I_{c,US,t}$ ).

Before proceeding, it is worth noting how this instrument differs from the one in Card (2001). To construct the initial occupation shares,  $\gamma_{c,o,1980}$ , we use data on newly arriving *legal* immigrants to the US using the “Immigrants Admitted to the US” data from the Immigration and Naturalization Service (INS). Using the 1972-1979 segments of the INS data, we are able to calculate the annual total number of immigrants from country  $c$  in occupation  $o$ .<sup>20</sup> The advantage of using these data is that both country of origin *and* occupation (in country of origin) are observed in the data. In Card (2001), the occupation distribution is constructed using observed occupations (reported in the US Census) after immigrating to the US. This is problematic as occupational choice is endogenous. Specifically, if the instrument is constructed based on the observed occupational choice upon arrival, the instrument is likely to suffer from the same endogenous variation as the observed inflows. So, while the traditional shift-share instrument may alleviate the concerns of endogenous sorting across metropolitan areas, it will not correct for the endogeneity of immigrant occupation choices. Because our instrument is based on the occupational choice in the immigrant’s country of origin, variation in predicted inflows will be based on occupational skills rather than endogenous occupational choice.

---

<sup>20</sup> While 1970-1979 would be ideal for constructing decadal shares, we do not have access to the data from 1970 or 1971.

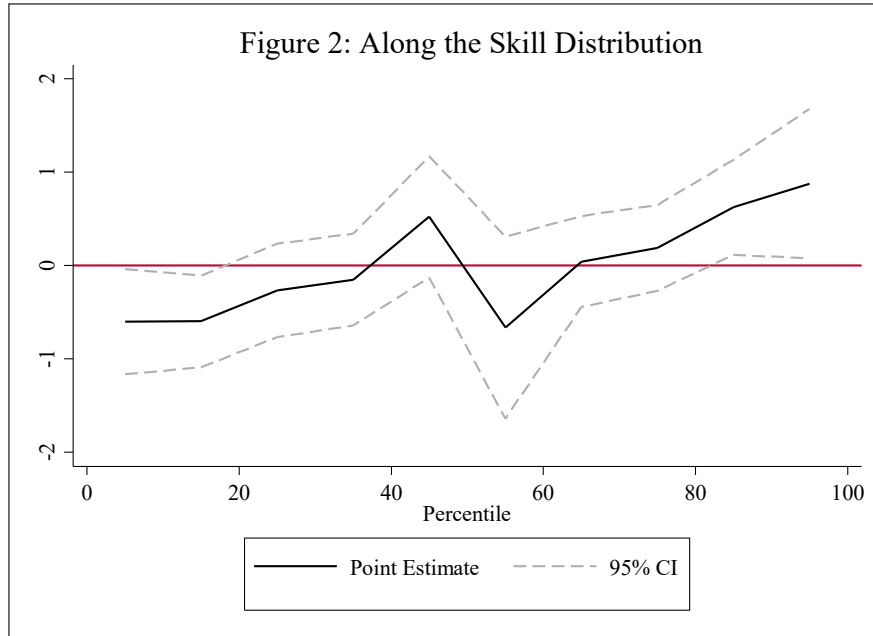
#### 4.1. Results

The estimated total wage effect at different points of the distribution are presented in Table 5. The OLS estimates are presented in column (1) and the 2SLS estimates in column (2). First stage estimates are reported in Table A2 of the Appendix. Consistent with the existing literature and the results in section 3, the effect of immigration is shown concentrated on the least skilled native workers. From the 2SLS estimates in column (2), the total wage effect is negative and highly statistically significant in the bottom tail (the 10<sup>th</sup> and 20<sup>th</sup> percentile) of the skill distribution. The results suggest that a 1% increase in the share of immigrants to a local labor market leads to a 0.6% decrease in the wages of natives in the bottom tail of the skill distribution. While the estimates are less precise in the middle of the distribution, immigrant inflows are shown to have a significant positive effect at the upper tail of the distribution: 1% increase in immigrants to a local labor market leads to a 0.6 – 0.8% *increase* in the wages of natives. Again, the result is consistent with the fact that higher-skilled native workers are largely complementary to immigrant labor.

<b>Table 5: Effects Along the Skill Distribution</b>		
VARIABLES	(1) OLS $\Delta \log \tilde{w}_{jkt}$	(2) 2SLS $\Delta \log \tilde{w}_{jkt}$
<i>10<sup>th</sup> percentile</i>	-0.121 (0.091)	-0.602** (0.288)
<i>20<sup>th</sup> percentile</i>	-0.034 (0.106)	-0.598** (0.250)
<i>30<sup>th</sup> percentile</i>	-0.097 (0.114)	-0.265 (0.255)
<i>40<sup>th</sup> percentile</i>	0.054 (0.122)	-0.151 (0.251)
<i>50<sup>th</sup> percentile</i>	-0.004 (0.117)	0.521 (0.331)
<i>60<sup>th</sup> percentile</i>	0.369*** (0.130)	-0.666 (0.497)
<i>70<sup>th</sup> percentile</i>	0.298** (0.135)	0.043 (0.249)
<i>80<sup>th</sup> percentile</i>	0.068 (0.133)	0.188 (0.234)
<i>90<sup>th</sup> percentile</i>	0.505*** (0.145)	0.624** (0.260)
<i>100<sup>th</sup> percentile</i>	0.523*** (0.166)	0.878** (0.408)

1. Each cell presents the estimated total wage effect at a given point along the skill distribution. In all cases, the dependent variable is the change in the mean log residualized wage of native men in skill group  $j$  in MSA  $k$ . Column (1) presents the OLS estimates, while column (2) presents the 2SLS estimates. All regressions are weighted. The weights are the sample size used to create the average log weekly wage in a given cohort. Robust standard errors clustered by MSA are reported in parentheses.

To see the effects along the full distribution, we also plot the 2SLS estimates in Figure 2 below. The solid black line represents the estimates from Column (2) of Table 6 and the dashed grey lines are the 95% confidence interval. Figure 2 tells a similar story. The negative wage effect identified in previous sections are concentrated on the least skilled native workers employed in occupations that require relatively more manual tasks. As skill level increases, the effect of immigration on native wages disappears. While imprecisely estimated at some points, the effect is positive for natives in the upper-tail of the distribution.



## 6. Conclusion

“Who competes with whom?” is an important question when trying to uncover the impact of immigration on native wages. The existing literature assessing the impact of immigration on native wages has yielded contradictory results. In this paper, we attribute these counterintuitive results to the fact that previous attempts have failed to compare immigrants and (demographically comparable) natives who directly compete in the labor market. Because immigrants and natives specialize in different skills and immigrants are often under placed in the labor market, immigrants and natives tend to cluster in different occupations.

Using the national labor market approach, we show that when labor markets are defined based on occupation-specific skills, the estimated (relative) impact of immigration is significantly more negative compared to defining labor markets on the basis of education. Specifically, the estimated impact of immigration on native wages is 2-3 times larger than those using education-experience cohorts. The results are robust to changes in occupation classification and suggest a 1% immigrant labor supply shock will decrease native wages by about 0.4-0.6%.

Because occupational choice is endogenous, we test the robustness of the main results by analyzing the impact of immigration on the wages of demographically comparable natives *within* education-experience groups. While a bit smaller in magnitude, the results are similar to those found when using occupation-experience groups. Findings suggest a 1% immigrant labor supply shock will

decrease the wages of natives most likely to compete with immigrants in the labor market by roughly 0.3-0.4%. For those least likely to compete with natives, however, the wage effect is zero.

Lastly, the national labor market approach identifies the partial equilibrium effect of immigrant inflows on native wages. To be responsive to the existing literature, we estimate the total wage effect of immigrant inflows along the distribution of occupation-specific skill. The analysis yields a familiar result: the wage impact of an immigrant inflow is concentrated on the least-skilled native workers. For those in the bottom two deciles of the skill distribution, immigration has large negative effects (elasticity estimate around -0.6). As we move up the skill distribution, this effect quickly disappears. For workers in the upper-tail of the distribution, the effect of immigration is positive (elasticity around 0.6 – 0.8).

## References

- Altonji, J. G., & Card, D. (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market* (pp. 201-234). University of Chicago Press.
- Autor, D. & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *The American Economic Review*, 103(5), 1553-1597.
- Aydemir, A. B., & Borjas, G. J. (2011). Attenuation bias in measuring the wage impact of immigration. *Journal of Labor Economics*, 29(1), 69-112.
- Borjas, G. J. (2003). The labor demand curve is downward sloping: reexamining the impact of immigration on the labor market. *The quarterly journal of economics*, 118(4), 1335-1374.
- Borjas, G. J., Freeman, R. B., Katz, L. F., DiNardo, J., & Abowd, J. M. (1997). How much do immigration and trade affect labor market outcomes? *Brookings papers on economic activity*, 1-90.
- Borjas, G. J., Grogger, J., & Hanson, G. H. (2012). Comment: On estimating elasticities of substitution. *Journal of the European Economic Association*, 10(1), 198-210.
- Bratsberg, B., & Terrell, D. (2002). School quality and returns to education of US immigrants. *Economic Inquiry*, 40(2), 177-198.
- Bratsberg, B., & Ragan Jr, J. F. (2002). The impact of host-country schooling on earnings: a study of male immigrants in the United States. *Journal of Human resources*, 63-105.
- Camarota, S. (2005). Immigrants at Mid-Decade: A Snapshot of America's Foreign-Born Population in 2005. *Center for Immigration Studies Background*.
- Card, D. (2001). Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics*, 19(1), 22-64.
- Card, D. (2009). *Immigration and inequality* (No. w14683). National Bureau of Economic Research.
- Cortes, P. (2008). The effect of low-skilled immigration on US prices: evidence from CPI data. *Journal of political Economy*, 116(3), 381-422.
- Dustmann, C., Frattini, T., & Preston, I. P. (2012). The effect of immigration along the distribution of wages. *Review of Economic Studies*, 80(1), 145-173.
- Dustmann, C., & Preston, I. (2012). Comment: Estimating the effect of immigration on wages. *Journal of the European Economic Association*, 10(1), 216-223.
- Dustmann, C., & Glitz, A. (2015). How do industries and firms respond to changes in local labor supply? *Journal of Labor Economics*, 33(3), 711-750.
- Dustmann, C., Schönberg, U., & Stuhler, J. (2016). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives*, 30(4), 31-56.
- Ferrer, A., & Riddell, W. C. (2008). Education, credentials, and immigrant earnings. *Canadian Journal of Economics/Revue canadienne d'économie*, 41(1), 186-216.
- Friedberg, R. M. (2000). You can't take it with you? Immigrant assimilation and the portability of human capital. *Journal of Labor Economics*, 18(2), 221-251.
- Jaeger, D. A., Ruist, J., & Stuhler, J. (2018). *Shift-share instruments and the impact of immigration* (No. w24285). National Bureau of Economic Research.
- Kerr, S. P., & Kerr, W. R. (2011). *Economic impacts of immigration: A survey* (No. w16736). National Bureau of Economic Research.
- Llull, J. (2010). Immigration, Wages, and Education: A Labour Market Equilibrium Structural Model. *The Review of Economic Studies*.
- Manacorda, M., Manning, A., & Wadsworth, J. (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.
- Mattoo, A., Neagu, I. C., & Özden, Ç. (2008). Brain waste? Educated immigrants in the US labor market. *Journal of Development Economics*, 87(2), 255-269.

- Neagu, I. C. (2009). Career Placement of Skilled Migrants in the US Labor Market. *Research Working papers*, 1(1), 1-50.
- Orrenius, P. M., & Zavodny, M. (2007). Does immigration affect wages? A look at occupation-level evidence. *Labour Economics*, 14(5), 757-773.
- Ottaviano, G. I., & Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, 10(1), 152-197.
- Peri, G., & Sparber, C. (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, 1(3), 135-169.
- Sharpe, J. (2015). Three Essays on the Economic Impact of Immigration (Unpublished doctoral dissertation). University of Kentucky, Lexington, KY.
- Steinhardt, M. F. (2011). The wage impact of immigration in Germany-new evidence for skill groups and occupations. *The BE Journal of Economic Analysis & Policy*, 11(1).

## *Appendix*

**Table A1: Area Analysis By Subgroup**

VARIABLES	(1) OLS	(2) IV
White	0.157*** (0.039)	0.123 (0.092)
Minority	0.018 (0.077)	-0.857*** (0.203)
Existing Immigrant	0.149* (0.077)	-0.720** (0.293)
HS Diploma or Less	-0.166*** (0.054)	-0.752*** (0.134)
Some College or More	0.333*** (0.049)	0.379*** (0.111)
1-5 Years of Experience	-0.279** (0.122)	-0.977*** (0.282)
6-10 Years of Experience	-0.122 (0.103)	-0.598** (0.248)
11-15 Years of Experience	0.113 (0.101)	-0.172 (0.228)
16-20 Years of Experience	0.380*** (0.098)	0.382** (0.195)
21-25 Years of Experience	0.436*** (0.095)	0.664*** (0.204)
26-30 Years of Experience	0.335*** (0.101)	0.569*** (0.214)
31-35 Years of Experience	0.249*** (0.088)	0.342 (0.231)
36-40 Years of Experience	0.187* (0.111)	0.389 (0.288)
Occupation Group 1 (Least Skilled)	-0.018 (0.093)	-0.647*** (0.215)
Occupation Group 2	-0.028 (0.078)	-0.409** (0.177)
Occupation Group 3	0.051 (0.077)	0.149 (0.193)
Occupation Group 4	0.233*** (0.077)	0.022 (0.182)
Occupation Group 5 (Most Skilled)	0.451*** (0.089)	0.788*** (0.206)

1. Each cell presents the results from a unique regression. Estimates come from estimating equation (4) for different subpopulations. In all cases, we follow the existing literature and stratify labor markets by occupation (quintile) and potential experience. Columns differ by estimation technique: column (1) uses OLS, while column (2) uses 2SLS and the shift-share instrument described in the paper. Robust standard errors clustered by MSA are reported in parentheses.  $p < 0.01$ \*\*\*,  $p < 0.05$ \*\* ,  $p < .10$ \*

---

**Table A2: First Stage Estimates**

---

VARIABLES	(1)
<i>10<sup>th</sup> percentile</i>	0.015*** (0.004) [20.23]
<i>20<sup>th</sup> percentile</i>	0.019*** (0.004) [22.96]
<i>30<sup>th</sup> percentile</i>	0.017*** (0.003) [18.85]
<i>40<sup>th</sup> percentile</i>	0.017*** (0.003) [18.21]
<i>50<sup>th</sup> percentile</i>	0.018*** (0.003) [21.09]
<i>60<sup>th</sup> percentile</i>	0.012*** (0.004) [11.44]
<i>70<sup>th</sup> percentile</i>	0.018*** (0.003) [30.11]
<i>80<sup>th</sup> percentile</i>	0.018*** (0.003) [23.80]
<i>90<sup>th</sup> percentile</i>	0.017*** (0.003) [20.80]
<i>100<sup>th</sup> percentile</i>	0.016*** (0.003) [15.47]

---

1. The table displays the first stage coefficient corresponding to Table 5. Standard errors are reported in parentheses and the first stage F-stat is reported in brackets.

---

## *Data Appendix*

### *A. Creating Manual-to-Communicative Task Index*

We use the O\*NET database (version 18) to construct the manual-to-communicative task ratio. First, we use select attributes from Ability, Work Activity, Skill, and Knowledge descriptors from the O\*NET to create a communicative task-intensity index and a manual task-intensity index. Abilities, Skills, and Knowledge data describe the attributes of workers, while Work Activity describes occupation attributes. For the communicative task-intensity index, we use worker and occupation attributes related to communicating information, social skills, and listening. The manual task-intensity index uses attributes related to basic strength and related characteristics. A full list of attributes for each descriptor used in these calculations (along with their manual/communicative designation) can be found in Table A1 below.

We first compute a measure of overall intensity for each attribute in a given O\*NET occupation. The O\*NET provides two ratings for the attributes: Importance and Level. The importance rating indicates the importance of a particular attribute to a given occupation, while the level rating indicates the degree to which an attribute is needed to perform a job. We create an overall intensity measure by multiplying Importance (scale 1-5) and Level (scale 1-7). We then normalize each intensity measure to be in the range of 0-1 by dividing by 35.

One limitation of the O\*NET is that occupations do not change over time. In order to use these data for my entire sample, we match the occupation groups defined in the O\*NET (i.e. 11-1011) to the occupation classification (occ1990dd) of Autor and Dorn (2013). The advantage of the occupation classification of Autor and Dorn (2013) is that occupations are a consistent panel from 1960-2010. To do this, we first match O\*NET occupations to occupations defined by the U.S. Census (using the standard crosswalk file and OCC codes from the 2000 Census), then we match the Census OCC codes to the occ1990dd codes (using the files provide by the authors on their website). It should be noted that there are significantly more O\*NET occupation groups than occ1990dd occupation groups (841 O\*NET vs. 330 occ199dd); thus, there are multiple O\*NET occupation groups for each occ1990dd code.

As such, the manual (communicative) task-intensity index for each occ1990dd code is simply the weighted average of all manual (communicative) attribute-specific intensity measures within a given occ1990dd code (weighted by total employment). Then, the manual-to-communicative ratio is calculated by dividing the manual task-intensity index by the communicative task-intensity index.

---

**Table A2: O\*NET Components Used in Manual-to-Communicative Ratio**

---

***Abilities***

Verbal (All)	Communicative
Idea Generation and Reasoning (Fluency of Ideas, Originality, Deductive Reasoning, Inductive Reasoning)	Communicative
Perceptual (Perceptual Speed)	Communicative
Sensory (Speech Recognition, Speech Clarity)	Communicative
Psychomotor (All)	Manual
Physical (All)	Manual

***Work Activities***

Interpreting the Meaning of Information for Others	Communicative
Communicating with Supervisors, Peers, or Subordinates	Communicative
Communicating with Persons Outside Organization	Communicative
Establishing and Maintaining Interpersonal Relationships	Communicative
Assisting and Caring for Others	Communicative
Selling or Influencing Others	Communicative
Resolving Conflicts and Negotiating with Others	Communicative
Performing for or Working Directly with the Public	Communicative
Performing General Physical Activities	Manual
Handling and Moving Objects	Manual
Controlling Machines and Processes	Manual
Operating Vehicles, Mechanized Devices, or Equipment	Manual

***Skills***

Reading Comprehension	Communicative
Active Listening	Communicative
Writing	Communicative
Speaking	Communicative
Installation	Manual
Operation Monitoring	Manual
Equipment Maintenance	Manual

***Knowledge***

English Language	Communicative
Communications	Communicative
Building and Construction	Manual
Mechanical	Manual

---

1. Abilities, Work Activities, Skills, and Knowledge are the descriptors
  2. Within each descriptor, we list all of the “attributes” used in the calculation of the task intensity indices.
-

## ***B. Sample Description***

### ***B.1 Wage Sample***

We calculate mean log wages for male workers in each year. Following Borjas (2003), we restrict the sample to include non-self-employed males, aged 18-64, who have positive weeks worked, valid earnings data, and that did not live in group quarters. Mean log wages are represented as constant 2010-dollars and we used hours worked ( $\text{perwt} \times \text{weeks} \times \text{hours} / 2000$ ) as weights in the calculation. As in Borjas (2003), we use potential experience as a proxy for actual experience. To calculate potential experience, we assume that workers with less than a high school diploma enter the labor force at 17; workers with a high school diploma or GED enter the labor force at 19; workers with some college (less than a bachelor's degree) enter the labor force at age 21; and workers with a college degree enter the labor force at 23. We drop those who report potential experience less than 0 or greater than 40.

### ***B.2 Employment Sample***

To calculate labor supply in each occupation-experience cohort, we limit the sample to males aged 18-64 who have positive weeks worked that did not reside in group quarters. Here, self-employed workers are included in the calculations. Labor supply in an occupation-experience cohort is the sum of all hours worked. Potential experience is defined as above.

## ***C. Borjas (2003) Replication***

One concern is that the results of our paper are driven by different sample selection criteria, variable construction, and/or weighting methods. To provide support for the analysis above, we present a replication of the work by Borjas (2003).

We start by noting the differences in our data and methodology from that of Borjas (2003). Borjas (2003) utilizes US Census data from 1960, 1970, 1980, and 1990; pooled CPS data from 1999, 2000, and 2001 to form 2000 data. The 1980 and 1990 data use the 5% extract for immigrant counts and 1% extract for native counts and native wage calculations. Regressions are weighted by the total sample size of the educ-exp-year cell (immigrants and natives). Immigrant share is defined in terms of the number of workers (not the share of hours worked by immigrants).

For the analysis in our paper, we make several changes to this original Borjas (2003) methodology and, instead, follow the sample selection from Borjas, Grogger, and Hanson (2012) (BGH, hereafter). For our analysis, we use US Census data from 1960 – 2010 with no differentiation between 1% and 5% extracts in 1980 or 1990. Though we use 2010 data in the main results of the paper, we omit 2010

in the analysis that follows. Regressions are weighted by the number of natives in each cell used to calculate the average wages. Immigrant share is defined as the percent of total hours worked.

Table A1 presents our replication of Borjas (2003). Each column represents a unique specification and all models include the full set of fixed effects described in the paper. Column (1) are the results reported by Borjas (2003). Column (2) presents our replication results. Columns (3) – (6) present the estimated impact of immigration on native wages using our occupation classifications when using these replication data. For comparison, the results in columns (2) – (6) should be compared to the estimates in Table 2 in the text.

Comparing columns (1) and (2) suggest a successful replication of Borjas (2003). Although not exact, the differences are trivial. There are, however, notable differences between the replication results and those in column (1) of Table 2. When using the methodology of BGH and including the 2010 data, we estimate an elasticity of -0.19 that is only marginally significant. We address this point below. Furthermore, columns (3) – (6) suggest that data and methodology are not driving the results in the paper. In fact, when using the methodology in Borjas (2003), the estimated impacts are even larger than reported in Table 2.

	(1) Educ-Exp (Borjas, 2003) $w_{ijt}$	(2) Educ-Exp (My Replication) $w_{ijt}$	(3) Occ-Exp (Quartile) $w_{ijt}$	(4) Occ-Exp (Quintile) $w_{ijt}$	(5) Occ-Exp (Sextile) $w_{ijt}$	(6) Occ-Exp (Dorn) $w_{ijt}$
<i>Weighted (All)</i>	-0.572*** (0.162)	-0.568*** (0.162)	-1.177*** (0.200)	-1.190*** (0.243)	-0.983*** (0.190)	-0.663*** (0.244)
<i>Elasticities</i>	-0.400	-0.398	-0.824	-0.833	-0.688	-0.464

1) Each column and row represent a unique specification. Each column differs based on the definition of skill (education or one of the occupation groups), while each row differs based on the label – weighted regression of (3) and the corresponding elasticities of the weighted regressions, unweighted regression of (3). The dependent variable is mean of the log weekly wage of natives in each skill group. The independent variable of interest is the share of total employment by immigrants in a given skill group. All specifications include year fixed effects, occupation (or education in column 1) fixed effects, experience group fixed effects, and interactions of all fixed effects. Robust standard errors clustered by skill group are reported in parentheses.

2) Weighted estimates in row 1 and 4 are weighted by the total number of natives used to calculate the average wage in each cohort.

Next, we report estimates when we alter one of the components of the original Borjas replication. Again, each column represents a different specification using a different skill classification (as above). Row 1 reports estimates when we use the exact Borjas (2003) data but weight the regressions by the number of natives used to calculate average wages (as in BGH). Row 2 reports the estimates when we use

the 2000 Census data (IPUMS) in place of the pooled cross-section from the CPS but utilize the weights in Borjas (2003). Row 3 reports estimate when we use the BGH weights *and* the 2000 Census data.

For each row, the reported estimates are smaller (in absolute value) when compared to Table A1. From row (1), however, the choice of weights has a relatively minor effect on the estimates. A more sizeable impact is seen in row (2). When we use the 2000 Census data in place of the CPS data, the estimates are much more in line with what we estimate in our paper. The differences (between row 2 of the table below and the table above) are fairly large; however, there is only a statistical difference (at the 5% level) in the Quintile occupation group (although it is very close for most of the others).

Taken together, Tables A1 and A2 suggest that variable construction, sample selection, and the choice of weight are not driving the results in the paper. While notable differences exist in some specifications, the results presented in the paper are smaller (in absolute value). Thus, any data or methodological issues are actually working against the narrative of the paper.

<b>Table A4: Changing Borjas Replication</b>					
VARIABLES	(1) Educ-Exp (My Replication) $w_{ijt}$	(2) Occ-Exp (Quartile) $w_{ijt}$	(3) Occ-Exp (Quintile) $w_{ijt}$	(4) Occ-Exp (Sextile) $w_{ijt}$	(5) Occ-Exp (Dorn) $w_{ijt}$
<i>Weighted (Natives)</i>	-0.481*** (0.104)	-1.217*** (0.209)	-1.26*** (0.249)	-1.054*** (0.189)	-0.717*** (0.245)
<i>Weighted (All; 2000 Census)</i>	-0.367*** (0.129)	-0.884*** (0.241)	-0.649** (0.275)	-0.721*** (0.223)	-0.496*** (0.183)
<i>Weighted (Natives; 2000 Census)</i>	-0.389*** (0.131)	-0.893*** (0.240)	-0.656** (0.275)	-0.729*** (0.224)	0.504*** (0.186)