NBER WORKING PAPER SERIES

GENDER, SELECTION INTO EMPLOYMENT, AND THE WAGE IMPACT OF IMMIGRATION

George J. Borjas
Anthony Edo

Working Paper 28682
http://www.nber.org/papers/w28682

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2021

We are grateful to Christoph Albert, Michael Amior, Axelle Arquié, Yvonne Giesing, Thomas Grjebine, Daniel Hamermesh, Gordon Hanson, Joan Llull, Joan Monras, Jacques Melitz, Valérie Mignon, Ariell Reshef, Camilo Umana Dajud, and Vincent Vicard for providing valuable comments on an earlier draft. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Gender, Selection into Employment, and the Wage Impact of Immigration
George J. Borjas and Anthony Edo
NBER Working Paper No. 28682
April 2021
JEL No. E24,F22,J21,J23

ABSTRACT

Immigrant supply shocks are typically expected to reduce the wage of comparable workers. Natives may respond to the lower wage by moving to markets that were not directly targeted by immigrants and where presumably the wage did not drop. This paper argues that the wage change observed in the targeted market depends not only on the *size* of the native response, but also on *which* natives choose to respond. A non-random response alters the composition of the sample of native workers, mechanically changing the average native wage in affected markets and biasing the estimated wage impact of immigration. We document the importance of this selection bias in the French labor market, where women accounted for a rapidly increasing share of the foreign-born workforce since 1976. The raw correlations suggest that the immigrant supply shock did not change the wage of French women, but led to a sizable decline in their employment rate. In contrast, immigration had little impact on the employment rate of men, but led to a sizable drop in the male wage. We show that the near-zero correlation between immigration and female wages arises partly because the native women who left the labor force had relatively low wages. Adjusting for the selection bias results in a similar wage elasticity for both French men and women (between -0.8 and -1.0).

George J. Borjas
Harvard Kennedy School
79 JFK Street
Cambridge, MA 02138
and NBER
gborjas@harvard.edu

Anthony Edo
CEPII
20 avenue de Ségur
75007 Paris, France
anthony.edo@cepii.fr

Gender, Selection into Employment, and the Wage Impact of Immigration

George J. Borjas and Anthony Edo*

1. Introduction

All other things equal, an immigration-induced increase in the size of the workforce is expected to reduce the wage of comparable workers. A voluminous literature attempts to estimate the impact of such supply shocks on the wage of native-born workers in the United States and other immigrant-receiving countries (see Blau and Mackie, 2016, for an exhaustive survey).

One key insight from this literature is that native workers may respond to the supply shock by moving to labor markets that were not directly affected by immigration and where presumably the wage did not drop. Some natives, for instance, might move to cities that received fewer immigrants and now pay relatively higher wages (Borjas, 2006; Amior, 2020; Monras, 2021); some natives might change their skill set to avoid the excess competition from equally skilled immigrants (Hunt, 2017; Llull, 2018); some natives might change their occupations (Card, 2001; Foged and Peri, 2016; Cortés and Pan, 2019); and some natives might leave the labor force altogether (Angrist and Kugler, 2003; Glitz, 2012; Dustmann, Schönberg and Stuhler, 2017). Regardless of the type of "switch," all of these responses help to attenuate the negative impact of immigration on the wage of comparable workers by effectively diffusing the shock across many other markets.

This diffusion implies that standard difference-in-differences comparisons of wages across markets may provide little information about how the wage in the market targeted by immigrants changed as a result of the supply shock. After all, the observed (relative) wage change in the targeted market will reflect not only the wage drop subsequent to the shock, but also the attenuation of that wage effect as some of the supply shock gets transmitted to other markets through the native response.

* We are grateful to Christoph Albert, Michael Amior, Axelle Arquié, Yvonne Giesing, Thomas Grjebine, Daniel Hamermesh, Gordon Hanson, Joan Llull, Joan Monras, Jacques Melitz, Valérie Mignon, Ariell Reshef, Camilo Umana Dajud, and Vincent Vicard for providing valuable comments on an earlier draft.

This paper argues that this approach to understanding how the native response biases the measured wage impact of immigration is incomplete. The wage change observed in a targeted market will depend not only on the *size* of the native response, but also on its *composition*. Put differently, the wage change observed in a labor market after an immigrant supply shock depends not only on the number of natives who "switched" markets, but also on which native workers switched. A non-random response changes the composition of the sample of native workers, and this compositional shift mechanically changes the average native wage in the affected markets. Depending on the context, the selection bias may exacerbate or further attenuate the measured wage impact of immigration.

We document the empirical relevance of this type of selection bias by examining how immigration differentially affected the employment and wages of men and women in the French labor market. The French context is particularly suitable for exploring the central hypothesis proposed in this paper for a simple reason: France experienced a remarkable "feminization" of its immigrant labor force in the past few decades, witnessing a very rapid rise in the female share of foreign workers. Because men and women could be imperfect substitutes, the rising number of immigrant women relative to immigrant men could affect the labor market outcomes of native men and women differentially (Acemoglu, Autor and Lyle, 2004; Edo and Toubal, 2017). Moreover, female labor supply tends to be more elastic than that of men at the extensive margin (Blau and Kahn, 2017; Evers, De Mooij and Van Vuuren, 2008). As a result, the supply shock may have had a considerable impact on the labor force participation rate of native women, potentially producing a sizable selection bias in the measurement of the wage impact of immigration.

In response to the economic crisis caused by the first oil shock of 1973, the French government stopped recruiting foreign labor in July 1974. In April 1976, however, France granted its pre-existing foreign-born population the right to family reunification, making it far easier for wives to join their husbands. A direct consequence of this policy shift was a rapid rise in the number of female immigrants entering the country. Between 1962 and 1975, the immigrant population (aged 18-64) grew by 620.8 thousand persons, and only 37.1 percent of this growth was due to female immigration. The immigrant population grew by another 1.1 million persons between 1975 and 2007, and women accounted for 75.6 percent of this increase.

Figure 1 illustrates key trends in the size and gender composition of the foreign-born labor force in France. The top panel of the figure shows how the policy shift led to an immediate decline in the immigrant share of the labor force. In 1975, 10.3 percent of labor force participants were foreign-born. By 1999, the immigrant share had dropped to 8.8 percent. This decline is entirely attributable to a large drop in the relative number of immigrant men. In contrast, the immigrant share in the female labor force rose steadily, almost doubling (from 5.7 to 9.2 percent) between 1968 and 2007.

The bottom panel of Figure 1 further demonstrates the feminization of French immigration by contrasting the French experience with that of the United States. In France, the share of women in the foreign-born labor force increased from 18.7 percent in 1962 to 22.8% in 1975, and then nearly doubled to 42.4 percent by 1999. In contrast, the share of women in the foreign-born labor force in the United States barely changed between 1970 and 2000, rising only from 39.8 to 41.1 percent over those three decades.

It is important to emphasize that family reunification was not the only factor changing the gender composition of foreign-born workers in France. Beauchemin, Borrel, and Régnard (2013, p. 4) note that "more and more of the women who arrive in France are single or 'pioneers' migrating ahead of their partner." In short, the historic feminization of the immigrant labor force was not only due to an increasing number of female migrants entering the country to reunite with their families, but also to a growing number of women who migrate on their own in search of better prospects.

The French context thus provides a unique opportunity for examining the link between gender and the labor market impact of immigration. The existing literature typically focuses on documenting how immigration affects the earnings of native men or sometimes pools all natives and disregards the gender composition of the labor force (some exceptions include Cortés and Tessada, 2011; Cortés and Pan, 2019; Farré, González and Ortega, 2011; and Llull, 2021).[1] However, the impact of immigration on the labor

[1] Furtado and Hock (2010) and Seah (2018) document that the fertility of native women might be influenced by immigrant supply shocks, further complicating their labor supply decision.

market outcomes of women may differ from that of men because, as noted earlier, men and women may not be perfect substitutes and female labor supply may be more elastic than male labor supply. When analyzing the impact of immigration on women's labor market outcomes, one important margin of adjustment would then obviously be the female labor force participation or employment rate. We argue that a change in the labor supply of native women, and the selection biases generated by such a supply response, may contaminate the wage change observed in a labor market hit by a supply shock.

Our study is guided by a theoretical framework that isolates the three key channels through which an immigrant supply shock changes the observed mean wage in a labor market. The first is the wage decline produced by the direct effect of immigration—the downward movement along the labor market's short-run labor demand curve. The second is the attenuation produced by the native response. Some natives will move out of the labor market targeted by immigrants, partially reversing the initial shift in the supply curve and attenuating the initial wage drop. The third is the selection bias. Because the natives who respond to the supply shock are unlikely to be a random sample of all native workers, the composition of the native labor force changes in the post-immigration regime, producing a spurious change in the wage that further complicates the interpretation of the observed wage change after a supply shock.[2]

Using standard results from the selection bias literature (Heckman, 1979), we show how the generic regression model estimated in the immigration literature (relating wages to the size of the immigrant supply shock) can be easily expanded to adjust for selection bias. The selection bias correction generates valuable information not only about which natives respond to the supply shock, but also produces an estimate of the wage elasticity that is not contaminated by the selection effect.

Our empirical study uses data from various French population censuses merged with information on labor market outcomes from the Labour Force Surveys (LFS) in the 1982-2016 period. The "raw" data reveal a striking gender asymmetry. The correlation

[2] There is, of course, an additional channel of adjustment as firms expand to take advantage of the lower price of labor. We abstract from this potential adjustment mechanism throughout the paper. Because the empirical analysis reported below does not formally account for this potential response by holding capital constant, our estimates of the wage impact are unlikely to measure the short-run effect.

between immigration and wages (across regions and over time) is negative for native men. At the same time, however, there is no correlation between immigration and the employment rate of native men. In contrast, the correlation between immigration and wages for native women is zero or even weakly positive. But the correlation between immigration and employment rates for native women is strongly negative. In short, the raw data seem to suggest that immigration had an adverse impact on French men along the wage margin, while immigration had an adverse impact on French women along the employment margin.

We show that the "zero wage elasticity" implied by the raw data for French women is partly an artifact of selection bias. The native women who left the labor market after the supply shock tended to be low-wage women, automatically increasing the average wage in the regions targeted by immigrants simply because the composition of the sample of working native women had changed. After taking account of the selection bias, our evidence indicates that the adjusted wage elasticity for native women is also negative and roughly the same size (between -0.8 and -1.0) as that found for native men. These results are robust to using alternative sample periods, alternative definitions of the immigrant supply shock, and to implementing an instrumental variable strategy that exploits the historical distribution of immigrants by country of origin across French regions (drawn from the 1968 Census) as an instrument for current immigrant penetration.

Although our analysis is the first to precisely delineate and empirically document how selection bias contaminates estimates of the wage impact of immigration, it is closely related to several recent studies that jointly consider the wage and labor supply responses to immigration in various European contexts: Bratsberg and Raaum (2012) for Norway; Dustmann, Schönberg and Stuhler (2017) for Germany; and Ortega and Verdugo (2016) for France. All of these studies find that low-wage workers are more likely to respond to immigration by leaving or not entering the workforce in the cities or industries targeted by immigrants. The studies then exploit the panel structure of their data and "track" the earnings of individual natives who are continuously employed, thus holding constant the composition of the sample of native workers over the period. This tracking approach produces a far more adverse wage effect than the raw correlations between immigration and wages would suggest, underlining the important role played by selective attrition in underestimating (or masking) the true effect of immigration.

It is crucial to note, however, that the approach of measuring the impact of immigration by tracking the wage of labor force "survivors" does not solve the selection problem. The subsample of survivors, after all, is self-selected from the at-risk population that was initially exposed to the immigrant supply shock, and hence their experience does not correctly measure the wage impact that would have been observed had the workers who left the labor force remained at work.[3] In other words, the wage trends observed in the subsample of workers who survived are contaminated by a classic example of selection bias, and do not necessarily represent the wage trends that would have been observed in the population at risk.[4]

Our conceptual framework and empirical analysis have implications that extend beyond the French context. Although we focus on the employment margin, the selection bias problem taints most existing estimates of the wage impact of immigration if natives respond along any margin. Conceptually, it does not matter if the native response is from employment to household production, or if the move is from one geographic area to another, or from one type of job to another. All of these native flows will be endogenous, are unlikely to be random, and generate selection biases that contaminate the observed change in the market wage after a supply shock.

The paper is organized as follows. Section 2 describes the data used in the analysis and documents the striking gender asymmetry in the wage and employment impact of immigration in France. Section 3 derives a theoretical framework that combines a model of

---

[3] For example, Ortega and Verdugo (2016) track the earnings of male native workers who are continually in the labor force even if they moved to other cities. The observed wage change is then affected by the selection bias produced by transitions in and out of the labor market (which is the focus of this paper), and by the wage growth resulting from endogenous internal migration decisions. In short, the tracking exercise does not produce the critical counterfactual of the wage change that provoked the native response.

[4] Although selection bias corrections are rare in the immigration literature, an empirical exercise reported in Card (2001) hints at their potential importance. In particular, Card uses the Gronau (1974) model of selection bias in grouped data to perform a back-of-the-envelope calculation that illustrates how the bias affects occupational wage differences created by differential supply shocks across occupation groups. The sample selection issue is also noted by Winter-Ebmer and Zweimüller (1996) who use a two-step Heckman selection model to estimate the probability of being a blue- versus white-collar worker, and then analyze the impact of immigration in the subsample of blue-collar workers.

labor demand with a model of an individual's decision to participate in the labor force, delineating the distinct channels through which immigration affects wages. Section 4 introduces the econometric framework, and describes the instruments used to control for the endogeneity of key variables as well as the selection correction. Section 5 reports the main empirical results and evaluates the sensitivity of our findings to a number of alternative model specifications. Section 6 expands the empirical analysis by documenting the wage impact of immigration across and within different skill groups. Finally, Section 7 summarizes the results and draws some conclusions.

2. Data and Descriptive Evidence

Our analysis of the French labor market uses data drawn from population censuses and the Labour Force Surveys (LFS) conducted by the French National Institute for Statistics and Economic Studies (INSEE). Specifically, we use the French censuses from 1968, 1975, 1982, 1990, 1999, 2007, and 2016 to calculate the size of the population and labor force in each census year (by gender and immigration status). The pre-2000 census extracts consist of a random sample of 25 percent of the French population, while the post-2000 censuses consist of a random sample of 14 percent of the population. The high sampling rates allow us to precisely estimate the number of immigrants in different French regions, helping to reduce the role of sampling error in the analysis (Aydemir and Borjas, 2011). We define an immigrant as a person born outside France who is either a noncitizen or a naturalized citizen. All other persons are classified as natives.

The annual LFS reports wages at the individual level beginning in 1982. Our empirical analysis of the wage impact of immigration thus covers the 1982-2016 period.[5] The LFS provides information on each person's labor force and employment status during the reference week, as well as detailed information on many demographic, social, and economic characteristics (including age, gender, nationality, education, marital status, and

[5] Between 1982 and 2002, the LFS surveyed a random sample of the French population, with a sampling rate equal to 0.3%. Since 2003, the annual sampling rate varies between 0.7% and 1.0%. Unless otherwise noted, we use the personal weight computed by INSEE throughout the analysis in order to make our sample representative of the French population.

number of children).[6] The LFS also reports the worker's monthly wage net of employee payroll tax contributions.[7] Wages are reported in nominal terms, and we deflate the data using the French Consumer Price Index produced by the INSEE. Since the LFS is designed to be representative of the population at the regional level (there are 22 regions in European France), we follow INSEE's advice and conduct our empirical analysis mainly at this geographic level.[8]

Our empirical analysis focuses on the monthly wage of full-time native workers to have a more precise measure of the "price of labor." The emphasis on full-time workers is standard in the literature (see Katz and Murphy, 1992; and Mulligan and Rubinstein 2008) and allows us to control for any change in the sample composition of part-time workers (given the presumably low labor market attachment of these workers). However, we will test the sensitivity of our results by also using the reported average number of hours worked per week to compute the hourly wage for all native workers.

Our sample is restricted to persons aged 18-64 living in European France. We exclude all persons who are self-employed, are in military occupations, are enrolled in school, or do not report their educational attainment. In our wage analysis, we will exclude observations that have extreme values of the hourly wage. Specifically, we exclude workers who are either in the top 0.5% or bottom 0.5% of the hourly wage distribution.

Table 1 summarizes some of the key characteristics of our data. Perhaps the most striking trend is the rapid increase in the employment rate of native women between 1962 and 2016, almost doubling from 37.2 to 70.1 percent. At the same time, the employment rate of French men declined noticeably, from 89.4 to 73.6 percent.

---

[6] The definition of employment status differs between the population census and the LFS. In the census, a person is defined as "employed" if he/she has a job at the time of the census. The LFS uses the International Labour Organization's definition, where a person is employed if he/she works for any amount of time, even if only for one hour, during a reference week.

[7] The monthly wage in the 1982 LFS is reported as a categorical variable with 19 bands. We impute the person's monthly wage by assigning the midpoint of each closed interval, 1000 francs for the "less than 1000 francs" band, and 45,000 francs for the "30,000 or more" band. The INSEE performs outlier corrections and imputes the wage data for surveyed individuals who do not report that information using statistical methods that rely on Mincerian wage equations.

[8] Moreover, the waves of the LFS between 2002 and 2012 only report geographic information at the regional level.

The data also indicate that the size of the supply shock was roughly similar for low- and high-educated native women. The immigrant share rose from 3.2 to 9.2 percent for women with a baccalaureate degree and from 5.7 to 14.1 percent for women without the degree. In contrast, immigration had a larger impact on the number of high-educated men, where the immigrant share rose from 5.3 to 10.0 percent while the immigrant share among the low-educated men was roughly constant (between 12 and 14 percent).

Finally, note that the number of observations available in the LFS is relatively small (between 32,000 and 83,000 per cross-section). The sample size limits the possibility of examining the wage impact of immigration in very narrowly defined markets.

We begin the empirical analysis by merging the employment rates and the data on the relative number of immigrants reported in the 1982, 1990, 1999, 2007, and 2016 censuses with the concurrent LFS wage data for native workers. The merged data helps illustrate the "raw" relationship between immigration and labor market outcomes of native men and women across French regions over the 1982-2016 period.

In this descriptive analysis, the unit of observation is a region-year cell. For each cell, we estimated the mean log monthly wage of full-time workers (separately by gender) as well as the immigrant share defined by $m_{rt} = \log(1 + M_{rt}/N_{rt})$, where M_{rt} gives the total number of (male and female) immigrants in the labor force in region r at time t and N_{rt} gives the corresponding number of natives.[9] We then calculated the adjusted mean wage as the residual from a regression (estimated separately by gender) of the mean log monthly wage on vectors of region and year fixed effects. We also calculated the adjusted supply shock by obtaining the residuals from a regression of m_{rt} on vectors of region and year fixed effects. The adjusted wage and immigrant share variables measure deviations in the log wage and in the size of the supply shock from the region's mean after netting out period effects that affect all regions equally.

Figures 2A and 2B document the asymmetric relationship between immigration and wages for native men and women in France. The scatter diagrams show a weak positive

[9] Bratsberg and Raaum (2012) also use this algebraic definition of the immigrant share to measure the supply shock. Most studies in the literature, however, define the supply shock as either M/N, or as the fraction of immigrants in the labor force, $M/(M+N)$. Both of these choices approximate the measure of the supply shock implied by a labor demand model, which as we show below, is our definition of m_{rt}.

correlation between immigration and the wage of native women (the estimated coefficient of the regression line is 0.11, with a standard error of 0.07). At the same time, however, there is a strong negative correlation between immigration and the wage of native men (the estimated coefficient and standard error are -0.42 and 0.16, respectively).

Using a similar approach, we also calculated the gender-specific adjusted employment rates for each region-year cell. These data, also illustrated in Figure 2, provide further evidence of the gender asymmetry in the labor market impact of immigration.[10] Figures 2C and 2D show a strong negative correlation between employment rates and immigration for native women, but essentially a zero correlation between employment and immigration for native men.[11]

The raw data suggest an important interaction between gender and the observed labor market impact of immigration on employment and wages. In the case of French men, a group with inelastic labor supply, immigration affected their labor market opportunities along the wage margin. In the case of women, a group with much more elastic labor supply, immigration affected their labor market opportunities by reducing the number of native women employed. These correlations hint at the possibility that immigration may have had a crowd-out effect on the employment of native women. This crowd-out could have attenuated any (initial) downward wage pressure generated by the immigrant supply shock on the female wage. The attenuation effect would be strengthened if the women who left the labor market had relatively low wages. In other words, the finding of a zero correlation between wages and immigration for native women may simply be a consequence of elastic female labor supply—and the ensuing selection bias—and does not necessarily reflect the wage response to immigration.[12]

[10] The regression coefficients (and standard errors) in the employment regressions are -0.98 (0.10) for women and -0.02 (0.10) for men. We find a similar gender asymmetry when correlating immigration and the labor force participation rates of natives.

[11] The asymmetric impact of immigration on the employment of native men and women is also found by Angrist and Kugler (2003) in Europe, Edo (2020) in France, and Smith (2012) in the United States.

[12] The empirical work reported below shows that the descriptive correlations summarized in the various panels of Figure 2 are robust to using an instrumental variable strategy, alternative sample periods, using departments (instead of regions) to define the local labor market, using unweighted regressions, and using alternative measures of the supply shock.

In sum, the descriptive evidence reveals a conspicuous gender asymmetry in the labor market impact of immigration on French workers. We now turn to a description of how these striking differences can be explained by a model that links the determination of the market wage and individual labor force participation decisions, and that also incorporates the selection biases created by the elastic labor supply response of native women.

3. Theoretical Framework

We illustrate the conceptual importance of the selection bias that contaminates the wage effect of an immigrant supply shock by focusing on a two-period model where natives choose whether or not to participate in the labor force.[13] To fix ideas, and without loss of generality, we assume that all natives in labor market k are working in the pre-migration period ($t = 0$). A standard labor demand framework that assumes a Cobb-Douglas aggregate production function with labor and capital as inputs implies that the *market* wage in the pre-migration period is given by:

$$\log w_{k0} = \varphi_{k0} + \eta \log N_{k0}\,, \tag{1}$$

where w_{k0} is the wage in market k at time 0; φ_{k0} is a market-specific parameter; and N_{k0} gives the total number of native "efficiency units" employed in the market. We will often refer to N_{k0} as "total native employment" in the pre-migration period. The parameter η is the wage elasticity.

There will inevitably be some wage dispersion in market k because the native workers in that market (though they share some key characteristics that help define the market, such as location, occupation, education, or work experience) will also exhibit some differences. For example, some natives may have higher quality education, or may differ in their drive or motivation, or may have a racial or ethnic background that is favored or penalized by employers. As a result, the *individual* earnings function giving the wage for

[13] For expositional convenience, we use the terms labor force participation and employment interchangeably.

person i in the pre-migration period depends not only on market conditions, but also allows for variation in individual wages because of differences in (observable and unobservable) characteristics captured by a variable ϵ_{i0}:

$$\log w_{ik0} = \varphi_{k0} + \eta \log N_{k0} + \epsilon_{i0}\,, \tag{2}$$

We assume that the individual variation in earnings $\epsilon_{i0} \backsim N(0, \sigma_{\epsilon 0}^2)$.

An influx of M_k immigrants enters the market, and we assume that their labor supply is perfectly inelastic. Native labor supply, however, might respond to the supply shock, so that the total number of native efficiency units offered to the market in the post-migration period ($t = 1$) is N_{k1}, changing total employment from N_{k0} to $M_k + N_{k1}$. The individual earnings function for native workers in market k in the post-migration period (*after* natives have responded) can then be written as:

$$\log w_{ik1} = \varphi_{k1} + \eta \log(M_k + N_{k1}) + \epsilon_{i1}\,, \tag{3a}$$

$$\log w_{ik1} = \varphi_{k1} + \eta\, m_k + \eta \log N_{k1} + \epsilon_{i1}\,, \tag{3b}$$

where the immigrant share $m_k = \log(1 + M_k/ N_{k1})$ measures the supply shock (as a fraction of native employment in the post-migration period) and $\epsilon_{i1} \backsim N(0, \sigma_{\varepsilon 1}^2)$.

The change in the wage of individual i is then given by:

$$\Delta \log w_{ik} = \varphi_k + \eta\, m_k + \eta \log \frac{N_{k1}}{N_{k0}} + \upsilon_i, \tag{4}$$

where $\varphi_k = \varphi_{k1} - \varphi_{k0}$; $\upsilon_i = \epsilon_{i1} - \epsilon_{i0}$; and $\upsilon_i \backsim N(0, \sigma_\upsilon^2)$.

Equation (4) shows that the wage change experienced by an individual worker in market k is determined by both the size of the immigrant supply shock (m_k) and by the change in the size of the native workforce (i.e., the induced native supply shift). If there were near-complete employment crowd-out as immigrants entered the market, the immediate drop in wages associated with the influx of foreign-born workers would be mostly offset by a corresponding decline in the number of native workers. It is instructive to formally represent this attenuation of the wage effect of immigration by assuming that the percent change in the number of native workers depends on the size of the supply shock (Borjas and Monras, 2017):

$$\frac{N_{k1} - N_{k0}}{N_{k1}} \approx \log \frac{N_{k1}}{N_{k0}} = \gamma \log\left(1 + \frac{M_k}{N_{k1}}\right) \approx \gamma \frac{M_k}{N_{k1}}, \tag{5}$$

where the parameter γ $(-1 \leq \gamma \leq 0)$ measures the crowd-out effect, the number of native workers who leave the labor market for every immigrant who enters. By substituting equation (5) into (4), we obtain a type of "reduced-form" equation relating the wage change to the immigrant supply shock, or:

$$\Delta \log w_{ik} = \varphi_k + \eta(1 + \gamma)\, m_k + \upsilon_i. \tag{6}$$

The reduced form in (6) represents the generic regression model estimated in the immigration literature, which produces an estimate of the "reduced-form wage elasticity" $\eta(1 + \gamma)$, the wage elasticity that incorporates the native supply shift. The potential crowd-out response has typically been examined in the context of internal migration: If native workers are mobile after the supply shock, the detrimental impact of immigration on local wages will be dispersed throughout the national economy and the (relative) short-run wage drop in a particular locality may not be detectable.

Our theoretical framework assumes that immigration is exogenous to economic conditions in market k at time 0. In practice, however, immigrants are not randomly distributed across labor markets and likely prefer to settle in markets that offer thriving economic opportunities. This endogeneity of immigration produces a bias in the estimate of the reduced-form wage elasticity $\eta(1 + \gamma)$. To address this identification problem, many recent studies use a variant of the shift-share instrument to isolate the variation in immigrant inflows across markets that is not determined by wages or other factors that influence wages (Altonji and Card, 1991; Card, 2001). Putting aside the well-known limitations of using such an instrument (Jaeger, Stuhler and Ruist, 2018), it is worth emphasizing that the IV estimate of the wage effect of immigration using equation (6) still only identifies the amalgam of parameters $\eta(1+\gamma)$, and does not provide any information about the true value of the wage elasticity to immigration (except when $\gamma=0$).

It turns out, however, that even if we could account for the native response to immigration along the lines of equation (4) or simply attempt to estimate the reduced-form wage impact of immigration in equation (6), these IV regressions would *still* yield inconsistent estimates of the corresponding wage impact of immigration. The wage change actually observed in labor market k depends crucially on *which* native workers chose to withdraw from the labor force. Put simply, the observed wage change is contaminated by selection bias.

For instance, the exit of low-skill natives from the labor force would artificially increase the market's average wage because of sample composition effects, so that the generic regression model could end up suggesting that immigration increased wages. The exit of high-skill workers, in contrast, would produce a decrease in the observed average wage, and might end up suggesting a very large (negative) wage elasticity. Even in the extreme case when $\gamma=-1$, so that the size of the total workforce remains constant in the post-migration period, the change in sample composition created by the self-selection of the "surviving" sample of native workers contaminates the empirical analysis and biases empirical estimates of the wage impact of immigration. In general, the non-random selection of the native workforce in the post-migration period implies that we cannot use the wage change observed in the sample of "survivors" to infer what would have happened to the wage of the natives who left the labor market.

To derive equations (4) and (6), we simplified the model by assuming that all natives work in the pre-migration period. The participation decision at $t=1$, however, is based on a comparison of an individual's reservation wage and the wage that natives can now earn given the more crowded labor market. We write the distribution of the reservation wage in the native population as:

$$\log \mathcal{R}_{ik} = \bar{\mathcal{R}}_k + u_{i,} \tag{7}$$

where $\mathcal{R}_{ik}$ gives the reservation wage for native worker i in market k; $\bar{\mathcal{R}}_k$ gives the mean reservation wage in the market; and $u_i \backsim N(0, \sigma_u^2)$.

Each native worker makes a labor supply decision for period 1 by comparing the wage that firms offer in the aftermath of the supply shock to the reservation wage. More precisely, the wage offered to native worker i *immediately* after the supply shock is given by:

$$\log w'_{ik} = \varphi_{k0} + \eta \log(M_k + N_{k0}) + \epsilon_{i0}\,, \tag{8a}$$

$$\log w'_{ik} = \varphi_{k0} + \eta\, m'_k + \eta \log N_{k0} + \epsilon_{i0}\,, \tag{8b}$$

where $m'_k = \log\,(1 + M_k/\, N_{k0})$. The wage offer w'_{ik} defined by equation (8b) reflects the wage adjustment made by firms as they move down the short-run labor demand curve in market k. It defines the counterfactual wage that would have been observed after the supply shock in a hypothetical scenario without a native supply response. The sample selection rule that determines if a particular native worker stays in the labor force *and* that we can observe the wage in both the pre- and post-migration periods is given by:

$$\Delta \log w_{ik} \text{ is observed if } Z_i^* = \log w'_{ik} - \log \mathcal{R}_{ik} > 0, \tag{9}$$

where Z_i^* is the continuous latent variable that generates the sample of workers at $t = 1$. This latent variable can be rewritten as:

$$Z_i^* = \varphi_{k0} - \bar{\mathcal{R}}_k + \eta\, m'_k + \eta \log\ N_{k0} + \epsilon_{i0} - u_i\,, \tag{10a}$$

$$Z_i^* = C_k + \tau_i, \tag{10b}$$

where $C_k = \varphi_{k1} - \bar{\mathcal{R}}_k + \eta\, m'_k + \eta \log\ N_{k0}$; $\tau_i = \epsilon_{i0} - u_i$; and $\tau_i \sim N(0, \sigma_\tau^2)$. The labor force participation rate of natives in the post-migration period is given by $\pi_k = 1 - \Phi(\alpha_k)$, where $\alpha_k = -C_k/\sigma_\tau$, and Φ denotes the standard normal distribution function.

Let the indicator variable $Z_i = 1$ if $Z_i^* > 0$, and $Z_i = 0$ otherwise. The change in the market wage from one period to the next is only observed for the subsample of workers who remained in the workforce in the post-migration period, or when $Z_i = 1$. Using standard results from the sample selection literature in grouped data (Gronau, 1974), the *observed* wage change in market *k* between the pre- and post-migration period is given by:

$$E(\Delta \log w_k \mid Z_i = 1) = \varphi_k + \eta\, m_k + \eta \log\left(\frac{N_{k1}}{N_{k0}}\right) + E(\upsilon_i \mid Z_i = 1), \tag{11a}$$

$$E(\Delta \log w_k \mid Z_i = 1) = \varphi_k + \eta\, m_k + \eta \log\left(\frac{N_{k1}}{N_{k0}}\right) + \rho_{\upsilon\tau}\, \sigma_\upsilon\, \lambda(\pi_k), \tag{11b}$$

where $E(\upsilon_i \mid Z_i = 1) = \rho_{\upsilon\tau}\, \sigma_\upsilon\, \lambda(\pi_k)$; and the inverse Mills ratio $\lambda(\pi_k) = \phi(\alpha_k)/\pi_k$, with $\phi(\cdot)$ representing the standard normal density. Note that the sign of $\rho_{\upsilon\tau}$ reveals which subsample of the native workforce responded to the immigrant supply shock and chose to withdraw from the labor market. The correlation $\rho_{\upsilon\tau}$ would be positive if the natives who left the labor market tend to be low-wage workers and would be negative if the natives who left tend to be high-wage workers.

Equation $(11b)$ illustrates the insight that both the *size* of the native response and the *skill composition* of the response help determine the observed wage impact of immigration. Specifically, note that the supply shock m_k has *three* distinct effects on the observed wage change in market *k*. The first is simply the direct short-run effect of the supply shock, captured by the (negative) wage elasticity η. This is the downward movement along the short-run labor demand curve in the absence of any native response.

The second term captures the possibility that immigrants crowd out the supply of natives. The percent change in the number of natives working, measured by $\log(N_{k1}/N_{k0})$, generates its own wage effect, as that supply response helps the labor market move back up the labor demand curve (and the elasticity η again comes into play). If the negative wage impact observed immediately after the immigrant supply shock reduces the size of the native workforce, the second term in equation (11b) is positive and helps attenuate the short-run wage impact.

Finally, the third term gives the selection bias resulting from the fact that the natives who drop out of the labor market are not randomly selected from the initial sample of native workers. The inverse Mills ratio λ is positive as long as some natives do leave the labor force (and $\lambda = 0$ otherwise). The sign of the third term, therefore, is determined by the sign of the correlation $\rho_{\upsilon\tau}$, which is positive if the natives who remain employed have above-average earnings. In this case, the positive wage growth produced by selection bias helps to further attenuate, and perhaps even reverse, the negative wage impact of the immigrant supply shock.[14]

Following Heckman (1979), our framework suggests that the selection bias can be viewed as a specification error in the generic wage regression that is commonly estimated in the immigration literature (and omits the λ term). In particular, substitute equation (6) into ($11b$) to obtain the reduced form of the regression model:

$$\mathrm{E}(\Delta \log w_k \mid Z_i = 1) = \varphi_k + \eta(1+\gamma)\, m_k + \rho_{\upsilon\tau}\, \sigma_\upsilon\, \lambda(\pi_k). \tag{12}$$

Suppose that the sorting of immigrants across markets was exogenous. Suppose further that we had aggregate-level data on wages, immigration, and native labor force participation rates across many markets and estimated a variant of equation (11b) that excluded the selection term. The omitted variable formula implies that:

$$\mathrm{plim}\, \eta\widehat{(1+\gamma)} = \eta(1+\gamma) + \rho_{\upsilon\tau}\, \sigma_\upsilon\, \delta_{\lambda m}, \tag{13}$$

where $\delta_{\lambda m}$ is the coefficient from a regression of the excluded inverse Mills ratio (λ) on the immigrant share m_k.

It is well known (Heckman, 1979, p. 156) that the inverse Mills ratio is a monotonically decreasing function of the labor force participation rate π_k, so that the

[14] Equation ($11b$) makes it easy to see how relaxing the assumption that all natives work at $t = 0$ would change the analysis. Assuming that the fundamentals that determine the type of selection (and are captured by the coefficient of the inverse Mills ratio) do not change over time, the selection term in ($11b$) would now be written as $\rho_{\upsilon\tau}\, \sigma_\upsilon\, [\lambda(\pi_{k1}) - \lambda(\pi_{k0})]$, where π_{kt} gives the labor force participation rate in market k at time t. Equation ($11b$) represents the special case where $\lambda(\pi_{k0}) = \lambda(1) = 0$.

coefficient $\delta_{\lambda m}$ would be positive if the supply shock reduced the employment rate of the native population. Equation (13) shows that the failure to take account of the selection bias produced by the supply response of native workers systematically biases the estimate of the wage impact of immigration as long as $\rho_{\upsilon\tau} \neq 0$. The bias is positive if $\rho_{\upsilon\tau} > 0$ and negative if $\rho_{\upsilon\tau} < 0$. In other words, the wage elasticity estimated by exploiting data on the wage growth observed in the subsample of persons who worked both before and after the supply is "too positive" if the workforce is positively selected and "too negative" if the workforce is negatively selected.

As we noted earlier, Bratsberg and Raaum (2012), Dustmann, Schönberg and Stuhler (2017), and Ortega and Verdugo (2016) find that the labor force participation decision of low-wage workers is particularly sensitive to immigrant supply shocks. They then track the wage of the subsample of "survivors" (i.e., those persons who remained in the workforce) to quantify the wage impact of immigration. Our theoretical framework indicates precisely why this tracking approach may not correctly measure the true wage impact. In particular, equation (8b) delineates how immigration changed the wage of the native population at risk, measured *after* immigrants have entered the labor market but *before* any native labor supply response. Equation (12) describes the evolution of wages in the subsample of survivors, and this wage growth depends not only on the supply shock itself but also on the fact that the survivors may not be a random sample of the population at risk. The application of the omitted-variable formula in equation (13) then shows the bias created by estimating the wage impact of immigration using only the wage trends observed in the self-selected subsample of survivors.

In short, the discussion in this section implies that correctly estimating the wage impact of immigration requires that we take account of the selection bias generated by the native response to immigrant supply shocks. As documented earlier, the raw data suggests a strong negative correlation between immigration and the employment rate of French native women, but a weak positive correlation between immigration and female wages. We now proceed to use the insights provided by our conceptual framework to estimate a regression model that incorporates a Heckman sample selection correction. This approach allows us to retrieve the true wage impact of the immigrant supply shock.

4. Econometric Framework

4.1. The Wage Equation

We estimate the selection-adjusted wage impact of immigration by turning to individual-level data and applying the Heckman selection correction. Based on the theoretical discussion, the general form of the individual earnings function that forms the basis for our empirical analysis is:

$$\log w_{irt} = \theta_a + \theta_e + \theta_r + \theta_t + \alpha_P P_{it} + \beta_1 m_{rt} + \beta_2 \log N_{rt} + \beta_3 \lambda_{it} + \mu_{it}, \qquad (14)$$

where $\log w_{irt}$ gives the log monthly wage of native worker i in region r at time t; θ_a, θ_e, θ_r, and θ_t are vectors of age, education, region and time fixed effects, respectively; and P_{it} is a vector of personal characteristics (discussed in detail below).[15] The vector of regional fixed effects in equation (14) eliminates all time-invariant regional characteristics that may affect the level of wages and the spatial distribution of immigrants, and the vector of time fixed effects adjusts for common factors specific to the survey year. We estimate equation (14) separately in the samples of working men and women.[16]

Following our theoretical framework, we measure the immigrant supply shock as $m_{rt} = \log(1 + M_{rt}/N_{rt})$. In principle, the coefficient β_1 in equation (14) would then measure the wage elasticity—the downward movement along the short-run labor demand

---

[15] The age fixed effects consist of six age categories (18-24, 25-32, 33-39, 40-47, 48-55, 56-64) and the education fixed effects consist of four education categories (college graduates, persons with some college, high school graduates, and persons with less than a high school diploma).

[16] Other studies that use individual-level (rather than cell-level) regressions to estimate the wage impact of immigration include Friedberg (2001) and Bratsberg and Raaum (2012). The individual-level approach provides the simplest way of correcting for selection bias. Gronau (1974) suggests how cell-level regressions could be used to adjust for selection bias in the absence of endogenous regressors. Specifically, each region-education-age-year cell would be partitioned into smaller cells defined by the exogenous variables that are assumed to affect labor supply but not wages. In our context, for example, a cell would consist of female workers in region r, year t, school level e, age group a, who are married, have three young children, and own a home. For each of these cells, we would calculate the employment rate of the corresponding group, which would produce an estimate of the inverse Mills ratio for that cell. This grouped regression may be problematic for two reasons. First, the breakdown of the data into a very large number of cells would lead to estimates of the inverse Mills ratio that likely contain substantial measurement error. Second, the labor demand function includes an endogenous variable giving the size of the native workforce at the region-year level, and the instrument for this variable will depend on some of the family variables used to define the cells in the analysis.

curve after immigrants enter the local labor market. The regression specification implies that the wage elasticity β_1 is estimated from within-region changes in the wage and in the size of immigrant shocks. The estimated elasticity gives the percent change in the wage of natives induced by a one percent immigration-induced increase in the size of the region's labor force.

The immediate wage drop that presumably follows the supply shock might encourage some natives to withdraw from the labor force, and the regression includes the (log) number of native workers N_{rt} to adjust for this reverse movement of the supply curve in labor market (r, t).

Finally, it is unlikely that such native flows in and out of the labor market are random. As a result, the regression model must also adjust for the resulting selection bias by including the inverse Mills ratio λ_{it}. The specification of the employment probit regression that generates our estimate of the inverse Mills ratio will be discussed shortly.

Following the work of Dustmann, Schönberg, and Stuhler (2016, 2017) and Jaeger, Ruist, and Stuhler (2018), we define the immigrant share m_{rt} at the region-year level (instead of assigning workers to different skill groups and then calculating a supply shock specific to the region-skill-year cell). This estimation strategy has the advantage that it accounts for all channels through which an immigrant supply shock in region r can affect the wage of workers in that region. Put differently, the estimate of β_1 does not only capture the "own" effect of a particular supply shift on the wage of competing workers. It also captures the complementary effects of the supply shock on the wage of workers with different skills as well as the wage adjustments produced by post-immigration changes in capital accumulation. This approach also does not rely on the pre-assignment of workers to particular skill groups, helping to avoid the potential mismeasurement of the immigrant supply shock in a given skill cell because of the possibility that employers might downgrade the skills that immigrants offer to the labor market (Dustmann, Frattini, and Preston, 2013).[17]

[17] Section 6 documents the robustness of our findings by showing that the key results are unchanged when we shift the unit of analysis to a region-education-year cell or to a region-education-age-year cell.

The three key variables in the expanded regression model in equation (14) are either endogenous (m_{rt} and $\log N_{rt}$) or need to be estimated (the inverse Mills ratio λ_{it}). We now turn to a discussion of the instruments and the estimation of the first stage probit model examining an individual's labor force participation decision.

4.2. Endogeneity of the Immigrant Supply Shock

It is well known that estimating the regression model in (14) using OLS will not provide consistent values of the wage elasticity because of the non-random sorting of immigrants across regions (with income-maximizing immigrants more likely to settle in regions that offer the best job opportunities). To address this issue, we follow the existing literature in using an instrumental variable approach, with the instrument based on past immigration patterns. This approach was pioneered by Altonji and Card (1991) and then used in many other studies (Jaeger, Ruist and Stuhler, 2018).

To build our instrument, we follow the procedure implemented in the study by Edo, Giesing, Poutvaara and Öztunc (2019) that investigates the political consequences of immigration in France over the 1988-2015 period. Specifically, we use the 1968 spatial distribution of immigrants from a given nationality for a given education group to predict the sorting of immigrants in subsequent periods. We use 11 nationality groups and four education groups.[18] We predict the number of immigrants for each region-time cell at time t ($t > 1968$) by multiplying the 1968 spatial distribution of immigrants in each origin-education group by the total number of immigrants from that group at time t, as follows:

$$\widehat{M}_r(t) = \sum_n \sum_e \frac{M_r^{ne}(1968)}{M^{ne}(1968)} \cdot M^{ne}(t), \tag{15}$$

where $M_r^{ne}(t)$ gives the number of immigrants in year t in national origin group n, education group e, and region r; and $M^{ne}(t) = \sum_r M_r^{ne}(t)$. We use an analogous approach

[18] The nationality groups are: Italian, Portuguese, Spanish, other European, Algerian, Moroccan, Tunisian, other African, Turkish, the rest of the world, and French for those immigrants who acquired the French citizenship. The education groups are college graduates, persons with some college, high school graduates, and persons with less than a high school diploma.

to predict the number of natives in the region because the actual number of natives is unlikely to be independent from regional conditions:

$$\hat{N}_r(t) = \sum_e \frac{N_r^e(1968)}{N^e(1968)} \cdot N^{\mathrm{e}}(t). \tag{16}$$

The shift-share instrument is then defined by:

$$\hat{m}_{rt} = \log\left(1 + \frac{\hat{M}_r(t)}{\hat{N}_r(t)}\right). \tag{17}$$

Despite its widespread use, it is important to emphasize that the shift-share instrument in (17) does not satisfy the exclusion restriction imposed by the IV strategy if: *(a)* the 1968 spatial distributions of immigrants and natives are correlated with persistent local factors that affect labor market outcomes; and/or *(b)* current economic outcomes are still adjusting to past immigration (Jaeger, Ruist and Stuhler, 2018).[19]

4.3. Endogeneity of Native Labor Supply

Although the generic regression model used in the immigration literature simply relates the wage in a particular market to the immigrant share in that market, the labor demand framework implies that a fully specified regression model should also include the size of the native labor force. Few studies, however, pursue this implication of the theory (exceptions include Borjas, 2003; and Bratsberg, Raaum, Røed and Schøne, 2014). As shown in Section 3, the exclusion of this variable identifies a reduced-form estimate of the wage elasticity that is contaminated by the size of the crowd-out effect, or $\eta(1 + \gamma)$.

[19] As shown by Jaeger, Ruist and Stuhler (2018), an important criterion to satisfy the exclusion restriction of shift-share instruments is to exploit periods with substantial changes in the national origin mix of immigrants. Edo, Giesing, Poutvaara and Öztunc (2019) and Ortega and Verdugo (2016) demonstrate that the serial correlation in the distribution of immigrants by country of origin is much lower in France than in the United States as the French immigration patterns changed quite drastically since 1968. The Jaeger, Ruist and Stuhler (2018) critique, therefore, is less relevant in our setting.

The size of the native labor force is endogenous to local economic conditions. We construct an instrument by combining the shift-share projection of the native population with information on gender and such (presumed) exogenous variables as the presence of young children in the household. The summary statistics reported in Table 1 suggest that a major determinant of changes in the size of the native workforce was the increase in the employment rate of women. As in many other countries, the presence of young children is a key deterrent to female labor supply in France (Piketty, 1998; Gurgand and Margolis, 2008). Let $\theta_r(t)$ be the fraction of the native population in region r at time t that is female and that does *not* have children under the age of 6.[20] Our instrument for the (log) size of the native workforce is given by:

$$\log \hat{F}_r(t) = \log \left[\theta_r(t) \cdot \widehat{\mathcal{N}}_r(t)\right], \tag{18}$$

where $\widehat{\mathcal{N}}_r(t)$ is an adjusted measure of the shift-share prediction $\widehat{N}_r(t)$ of the native population. The variable $\hat{F}_r(t)$ thus gives the predicted female native labor force in region r at time t.

The construction of $\widehat{N}_r(t)$ in equation (16) only took into account the geographic allocation of natives at the time of the 1968 cross-section, and ignored that there were region-specific long-run trends that were systematically changing that allocation prior to our sample period. Unlike changes in the population of immigrants, where sudden and sizable shocks can occur due to exogenous policy shifts (such as granting foreigners the right to family reunification) or economic and political shocks in source countries, future projections of the native population are much more dependent on pre-existing trends.

To construct the instrument in equation (18), we adjust the shift-share projection $\widehat{N}_r(t)$ for the long-term regional differences in population growth rates. In particular, we calculate the (baseline) annual growth rate of the native population in region r between 1968 and 1982, g_r, as well as the growth rate of the shift-share projection over the same period, $\hat{g}_r$, and define $\Delta g_r = g_r - \hat{g}_r$. The adjusted shift-share projection is then given by:

[20] The variable $\theta_r(t)$ equals the share of the population that is female (drawn from the census) times the share of the female population that does not have young children (drawn from the LFS).

$$\widehat{\mathcal{N}}_r(t) = \widehat{N}_r(t)(1 + \Delta g_r)^{t-1968}. \tag{19}$$

The adjusted projection $\widehat{\mathcal{N}}_r(t)$ equals the "cross-section" projection $\widehat{N}_r(t)$ if the geographic allocation of natives is constant prior to the sample period (i.e., $\Delta g_r = 0$).[21]

The omission of the log N_{rt} variable from the typical labor demand regressions estimated in the immigration literature is likely due to the difficulty in finding good instruments for native labor supply. Our extension of the shift-share approach to create the adjusted projection in equation (19) relies on the same types of assumptions typically made to justify the validity of shift-share instruments in other contexts. Specifically, both the geographic allocation of natives at a point in time and the pre-existing trends in this allocation are assumed to be independent of current wages.[22] However, the validity of our instrument also hinges crucially on the assumption that "shocks" in the presence of young children in the household affect female labor supply decisions but do not affect wages. Although this assumption is common in the literature (Blau and Kahn, 2017), it is unlikely to be strictly true (and particularly in a context that will also examine male labor supply and wages). Our empirical analysis, however, will demonstrate that the key insight of our framework—i.e., that selection biases matter when estimating the wage impact of immigration—is valid even when we simply attempt to estimate the reduced form wage elasticity identified by the generic equation in the literature.

4.4. The Inverse Mills Ratio

[21] The population data for the Île-de-France region, which includes Paris, illustrates the importance of this type of adjustment. This region's population grew by only 0.5 percent per year between 1968 and 1982, as compared to a national growth rate of 1.3 percent. The 2016 shift-share prediction $\widehat{N}_r(t)$ for Île-de-France is 8.5 million persons, as compared to an actual native population of only 5.0 million. The adjustment in equation (19) produces a prediction of 4.5 million. The adjustment is empirically less important when projecting the growth of the immigrant population, as levels of the future immigrant population are often determined by policy shocks or exogenous events in source countries.

[22] Ideally, we would use information on the cross-section geographic allocation of natives and the secular change in that allocation many years before the 1982-2016 sample period. Our results are robust if we start the sample period in 1990 or if we use the period 1968-1975 to measure the pre-existing growth rate in population. To ensure compatibility with existing studies, we ignored the adjustment in equation (19) when we constructed the instrument for the immigrant share. Our estimates of the wage impact of immigration are not sensitive to this additional correction.

Finally, the regression model in equation (14) includes the inverse Mills ratio to adjust for the selection bias produced by the non-random labor force participation decision of natives after the immigrant supply shock. We construct the inverse Mills ratio by first estimating a probit model that relates a native person's decision to work to the various regressors in the model, including a vector of characteristics Z that, by assumption, do not enter the labor demand function:

$$\mathrm{P}(EMP_{irt} = 1) = \Phi(\theta_a + \theta_e + \theta_r + \theta_t + \alpha_P P_{it} + \alpha_Z Z_{it} + \gamma_1 m'_{rt} + \gamma_2 \log N_{rt-1} + v_{it}), \quad (20)$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution. We estimate the probit model in equation (20) separately for men and women.

We initially categorize the population into working or not working based on person i's employment status in the reference week of the LFS data. In other words, EMP_{irt} is a binary variable indicating whether native person i in region r at time t is employed. Our empirical analysis will test the sensitivity of our evidence to alternative definitions of "work." In particular, we will also use full-time employment status to classify each person into the two employment categories, as in Mulligan and Rubinstein's (2008) related analysis of selection bias in the measurement of the gender wage gap.

The definition of the latent variable that determines the participation decision in equation (10a) implies that the employment probit regression should use the immigrant supply shock $m'_{rt} = \log\,(1 + M_{rt}/N_{rt-1})$ and the lagged size of the native labor force (N_{rt-1}). The use of these variables to measure labor market conditions in region r incorporates the idea that the labor force participation response of native workers is driven by the shift in the wage observed *immediately* after immigrants enter the labor market and prior to any native response.[23]

Because there are gender differences in the determinants of labor supply and wages, we use slightly different baseline specifications of the probit and wage regressions in equations (20) and (14) for the two groups. Our approach for the analysis of female

[23] The empirical results are nearly identical if we instead used the contemporaneous measure of the supply shock and of the native labor force in the probit regressions.

outcomes follows that used in the literature (Mulligan and Rubinstein, 2008; Blau and Kahn, 2017, p. 810). In particular, the probit regression includes variables that adjust for individual differences in the reservation wage, and the variables that are often used in the female labor supply context are marital status and the presence of young children (under age 6) in the household.[24] It is typically assumed that these family characteristics affect the reservation wage of women, but do not affect their wage.

The LFS data allow us to expand this generic specification as it contains an approximate measure of household wealth (so that we can also control for income effects on labor force participation). The available measure of household wealth indicates if the person owns their home free of any debt.[25] As long as leisure is a normal good, the neoclassical labor-leisure model implies that higher levels of household wealth increase the reservation wage and would have a negative effect on the probability of participating in the labor force.

In short, the regression specification for the joint study of female employment and wages can be summarized as follows: The wage regression includes vectors of age, education, region, and time fixed effects. The probit regression includes all of these variables *plus* the family characteristics (marital status and the presence of young children) and household wealth. Note that the independent variation in the inverse Mills ratio in the female wage regression is generated by the presence of both family characteristics and household wealth in the first-stage probit.

It is not uncommon in the U.S. labor supply literature to simply assert that the selection problem is not empirically relevant for men (Pencavel, 1986, p. 55; Mulligan and

---

[24] The marital status variable in the LFS classifies individuals into one of four groups: single, widowed, divorced, or married. We pool all single, divorced, or widowed natives into an "unmarried" group, as opposed to married natives.

[25] This variable is set to unity if the person owns their home without a mortgage, and zero otherwise. The mean of this variable rose from 22.0 to 32.2 percent between 1982 and 2016. The homeownership information was not collected for a random half of the sample in the 2016 LFS cross-section. We impute the missing values by running a probit regression in the pooled 1982-2016 data that relates the homeownership indicator (if available) to age, education, interacted region-time fixed effects, and a full set of interactions between gender, marital status, presence of young children, and region fixed effects. We then impute a value of 1 or 0 to the missing observations based on whether the predicted probability of home ownership was above or below 0.6. Our results are similar if we simply excluded the 2016 observations that had missing information on homeownership assets.

Rubinstein, 2008). This assumption, however, may not be applicable in France (or other European contexts), where the unemployment rate among prime-age men is sufficiently large that the assumption that male workers are randomly selected from the population becomes less plausible. During our sample period, for example, the average unemployment rate of native men aged 25-59 was 8.1 percent.

Our specification of the regression models for the joint study of male labor supply and earnings differs slightly from what is typically used in the female context because marriage may have a productivity-related positive effect on male earnings (Choi, Joesch, and Lundberg, 2008; and McDonald, 2020), and fatherhood may also increase male earnings (Lundberg and Rose, 2000). In other words, even if these family characteristics did not affect the male reservation wage, they would need to enter *both* the probit and the wage regressions because they affect the male wage directly. As a result, the family variables do not produce independent variation for the inverse Mills ratio in the male wage regression. This independent variation is instead produced by the measure of household wealth that we assume only affects reservation wages (for both men and women).

The baseline regression specification for the joint study of male employment and wages can be summarized as follows: The wage regression includes vectors of age, education, region, and time fixed effects, as well as marital status and presence of young children. The probit regression includes all of these variables *plus* the measure of household wealth. We will show below that our estimates of the wage elasticity are very robust to alternative modeling strategies (including the assumption of no selection for men).

For any particular specification of the probit model, the estimated coefficients are then used to calculate the value of the inverse Mills ratio for each person in our sample. The OLS regressions that relate a worker's wage to the immigrant share will use an inverse Mills ratio computed directly from the probit model in (20). The calculation of the inverse Mills ratio for the IV regression models uses an expanded probit specification that will be discussed below in Section 5.2.

5. Empirical Results

5.1. First-Stage IV Estimates

Table 2 presents the first stage of our baseline IV individual-level wage regressions for both samples of native women and native men. Our simplest regression specification will simply relate the wage to the immigrant share. Panel A of the table presents the first-stage regression associated with this generic model, where we regress m_{rt} (i.e., the single endogenous regressor) on $\widehat{m}_{rt}$ (i.e., the shift-share instrument defined in equation (17)) and age, education, region, and time fixed effects (as well as any other individual control variables that may enter the wage equation).[26]

Not surprisingly, the first stage shows a strong positive and significant correlation between the instrument and the endogenous variable. We also report the Kleibergen-Paap rk Wald F statistics as this test accounts for the non-i.i.d. structure of the residual (Kleibergen and Paap, 2006). They are larger than the lower bound of 10 suggested by the literature on weak instruments (Stock, Wright and Yogo, 2002), indicating that our IV estimates are unlikely to suffer from a weak instrument problem.

Panel B reports the first-stage estimates for the expanded specification that has two endogenous variables, the immigrant share m_{rt} and native labor supply ($\log N_{rt}$). As instruments, we use the predicted population of immigrants (i.e., $\widehat{M}_r(t)$ defined in equation (15)) and the predicted female native labor force (i.e., $\widehat{F}_r(t)$ defined in equation (18)).[27] All regressions again include the age, education, region, and time fixed effects and the additional individual control variables included in the wage regression.

The results indicate that the immigrant share is positively correlated with $\widehat{M}_r(t)$ and negatively correlated with $\widehat{F}_r(t)$. The positive correlation is in line with the literature on the immigrant shift-share instrument, while the negative correlation is likely produced by the fact that a rise in the predicted number of working women would mechanically reduce the

---

[26] The first-stage regressions estimated in the male sample also include variables indicating marital status and the presence of young children.

[27] We use $\widehat{M}_r(t)$ instead of $\widehat{m}_{rt}$ as an instrument to avoid potential collinearity issues arising from the fact that $\widehat{m}_{rt}$ and $\widehat{F}_r(t)$ are both functions of the shift-share prediction of the native population. In fact, using $\widehat{m}_{rt}$ and $\widehat{F}_r(t)$ as instruments leads to weaker first-stage estimates and less significant estimated coefficients in the second-stage IV regressions.

ratio of immigrant to native workers. There is also a very strong positive correlation between the predicted number of working women and the actual number of natives in the labor force.

To evaluate the strength of our two instruments, we use the IV first-stage F-statistics for the case of multiple endogenous variables proposed by Sanderson and Windmeijer (2016). The first-stage F-tests of excluded instruments are between 17 and 22, indicating that our instruments are reasonably strong.[28]

5.2. The Probability of Employment

Table 3 reports the estimates of the probit regression on whether native person i in region r at time t is employed in the reference week. The inverse Mills ratio that we include in our baseline IV wage regressions (using one single endogenous regressor) is calculated from the coefficient estimates reported in column 1 for native women and column 3 for native men.

We follow the estimation approach proposed by Wooldridge (2002, pp. 567-570) for statistical models with sample selection and endogenous regressors. Specifically, the probit regression includes our instrument for the immigrant share (i.e., the predicted immigrant share in the population). As suggested by equation (10a) of our theoretical framework, the regression also includes the log predicted native population in the prior census (as measured by the shift share projection $\widehat{N}_{rt-1}$).[29] The table also reports the coefficients of the additional variables that adjust for family characteristics and other differences in reservation wages, such as marital status, presence of young children, and home ownership.

For illustrative purposes, columns 2 and 4 of Table 3 estimate the direct impact of the immigrant share m'_{rt} on the native employment probability using an IV strategy. We

---

[28] The first-stage IV regressions in models that adjust for sample selection will also include the inverse Mills ratio as an additional regressor. Adding that regressor does not change any of the first-stage results summarized in Table 2.

[29] In particular, equation (10a) suggests that the probit model should include the log native labor force in census year t-1 (i.e., prior to the immigrant supply shock) as a regressor. Because this variable is likely to be endogenous, we instead use the lag of the shift-share prediction of the native population.

instrument m'_{rt} by using log $(1 + \widehat{M}_{rt}/\widehat{N}_{rt-1})$, where $\widehat{M}_{rt}$ and $\widehat{N}_{rt-1}$ are the corresponding shift share predictions.

The probit regressions reveal a strong and negative relationship between the immigration variables and the probability of employment for native women, while the same correlations are close to zero for native men. The marginal effect from column 2 implies that a 10 percent immigration-induced increase in the size of the labor force reduces the probability of working by 11.6 percentage points for native women. The asymmetric impact of immigration on native employment by gender is consistent with the descriptive evidence summarized in Section 2, which suggested that immigration reduced the employment rate of native women but was not correlated with the employment rates of native men.

The other coefficients reported in Table 3 show that the lagged size of the predicted native population is negatively correlated with the employment probability of natives. We find that marriage lowers the probability of employment for women (by 1 percentage point), and substantially increases it for men (by 10 percentage points). The presence of young children in the household is found to be an important predictor of employment, and the sign of the correlation again differs between men and women. In particular, the presence of young children lowers the probability of employment by 10 percentage points for women, and increases it by 7 percentage points for men. Finally, the probit regressions reveal that household wealth (as proxied by the homeownership variable) has a strong negative effect on the employment probability for both men and women. Persons who own their home free of debt have a 2 to 4 percentage point lower probability of working.[30]

5.3. The Wage of Native Workers

Table 4 reports the OLS and IV estimated impact of the immigrant supply shock on the log monthly wage of full-time native women (Panel A) or men (Panel B) between 1982 and 2016.

[30] We will also use the Wooldridge (2002) method to estimate the selection models when the wage regressions have two endogenous variables. The probit regressions used to generate the inverse Mills ratio then includes the instruments for both endogenous variables (i.e., the log predicted immigrant population and the log predicted number of working women).

Consider initially the results for native women. Column 1 presents the simplest model, where the log wage is related only to the immigrant share (and age, education, region, and year fixed effects). The OLS coefficient of the immigrant share is insignificant and numerically close to zero. This finding reproduces the descriptive evidence illustrated in Figure 2 (which did not adjust for individual differences in education and age).

Column 2 adds the inverse Mills ratio to adjust for the selection bias resulting from the change in the sample composition of female workers following the supply shock. The coefficient of the inverse Mills ratio is positive, suggesting that female workers are positively selected from the female population (i.e., the unobserved factors that make labor force participation more likely tend to be associated with higher wages).[31] The mean value of the inverse Mills ratio for women is 0.48, so that the self-selection of the sample of female workers increases the mean of the observed wage distribution by about 10.6 percent (or the product of the coefficient of the inverse Mills ratio and its mean) relative to the population mean.

Note also that the OLS estimate of the wage elasticity implied by the coefficient of the immigrant share becomes significantly negative with a value of -0.42 (and a standard error of 0.08). The change in the measured impact of immigration between columns 1 and 2 is precisely the prediction made by the model in equation (13) if the women who exit (or do not enter) the labor force in the post-migration period have relatively low wages. In other words, ignoring the self-selection of the sample of working women produces an estimate of the wage impact of immigration that is positively biased.

Columns 5 and 6 of Panel A present the analogous IV regression results where the immigrant share is instrumented using the corresponding shift-share prediction. The OLS and IV coefficients for the simplest model are quite similar. The IV coefficient of the immigrant share in column 5 is essentially zero, and the coefficient becomes much more negative and significant (with a value of -0.44, and a standard error of 0.10) when the regression adjusts for selection. Finally, note that the coefficient of the inverse Mills ratio remains positive, confirming the positive selection found in the OLS regressions.

[31] Positive selection of women into employment is also reported by Mulligan and Rubinstein (2008) for the United States, Olivetti and Petronglo (2008) for a panel of OECD countries, and Dolado, Garcia-Penalosa and Tarasonis (2020) for Europe.

The remaining columns of Table 4 expand the basic model relating wages to immigration. In particular, columns 3 and 7 exclude the inverse Mills ratio but instead add the variable measuring the size of the native labor force, or log N_{rt}. As noted earlier, even though the presence of this variable in the equation is implied by the simplest labor demand framework, it has typically been excluded from the regressions estimated in the immigration literature. Because of the classic "supply-demand" endogeneity problem introduced by this variable, our discussion focuses on the IV results.

The log N_{rt} variable has a negative and significant impact on female wages (as predicted by theory).[32] It is also worth noting that the wage impact of immigration, as measured by the coefficient of the immigrant share variable, also becomes negative and significant (compared to the simplest model in column 5). The fact that holding constant the size of the native labor force results in a more negative immigration effect suggests the existence of a crowd-out effect. In terms of our theoretical framework, the coefficient of the immigrant share variable in a model that does not control for the size of the native labor force is contaminated by the crowd-out effect and equals $\eta(1 + \gamma)$.

Finally, columns 4 and 8 of Table 4 report the estimates from the full regression specification that controls for both the size of the native labor force and for sample selection. The estimated wage elasticity in the IV regression has now increased to -0.98 (with a standard error of 0.30). In other words, an immigration-induced 10 percent increase in the size of the labor force is predicted to lower the wage of native women by nearly 10 percent. Note also that the impact of the log N_{rt} variable remains negative and significant in the fully specified model, and that the coefficient of the inverse Mills ratio still suggests significant positive selection.[33]

---

[32] The coefficient of the native labor supply variable should equal the coefficient of the immigrant share, as both variables measure supply shocks in a labor demand framework. The sizable numerical difference between the two coefficients can probably be attributed to the fact that the instrument for the native labor supply variable may not fully resolve the endogeneity problems created when higher wages induce more natives to work.

[33] It may seem that the inverse Mills ratio, which is inversely related to the probability that a particular individual is employed, may be highly correlated with the variable giving the size of the native labor force, or log N_{rt}. In fact, the correlation between the two variables is very small, -0.12 for women and -0.05 for men. As a result, it seems that the two variables seem to be capturing different aspects of the economic problem examined in this paper. In particular, the size of the native labor force is controlling for the magnitude of the

Our theoretical framework implies that we can recover the crowd-out parameter γ from the coefficients of the immigrant share variable in columns 6 and 8. Specifically, the selection-adjusted estimate of the wage elasticity η in column 8 is -0.98, while the corresponding estimate of the reduced-form elasticity $\eta(1+\gamma)$ in column 6 is -0.44.[34] The implied estimate of the crowd-out parameter γ is 0.55, so that about 6 native women leave (or do not enter) the workforce for every 10 new immigrants. This result is consistent with other estimates in the European context. Using a panel of European countries, Angrist and Kugler (2003) find that 4 to 8 natives lose their jobs for every 10 immigrants in the labor force, while Glitz (2012) reports 3 native job losses for every 10 immigrants in Germany.

Panel B replicates the entire set of regressions using the sample of working native men. There are a number of interesting differences in the results between men and women. First, our estimates suggest weaker selection for men. In particular, the estimated coefficient of the inverse Mills ratio for men is insignificant and about a quarter the size as the corresponding coefficient for women (0.05 as compared to 0.22). Moreover, the mean of the inverse Mills ratio is smaller because men are more likely to work (the male mean is 0.34 as compared to 0.48 for women). As a result, selection would increase the mean of the observed wage distribution for men by only about 1.7 percent, as compared to 10.6 percent for women.

Second, the coefficient of the native labor supply variable is positive, but it is numerically close to zero (and insignificant in the IV regressions). As we noted earlier, the instrument for the supply variable $\log N_{rt}$ (based on shift-share predictions of the female native population and the presence of small children in the household) may not fully resolve the endogeneity issues created by the fact that although native men are more likely to work when the wage is high, male labor supply is generally inelastic. The non-negative male wage response to changes in the size of the native labor force may also be reflecting factors that are specific to the French context, where the growth of the native workforce in recent decades has been mostly driven by the rise in the labor force participation rate of

---

native supply response to the immigrant shock, while the inverse Mills ratio is adjusting for the selection bias produced by such a response.

[34] The difference between the two coefficients is marginally significant, with a *t*-statistic of 1.71.

native women.[35] If men and women are not perfect substitutes, the increased number of native workers need not lead to lower wages for native men. The male wage could even increase if men and women are complements in production.

Finally, regardless of the specification of the regression model, the estimated coefficient of the immigrant share variable for men is negative, significant, and hovers between -0.7 and -0.9. As our descriptive analysis in Figure 2 showed, there is a strong negative correlation between immigration and the wage of French native men. This correlation persists regardless of the econometric model used to measure the link between immigration and male wages.

It is also worth emphasizing that the simplest (generic) model in column 5 linking immigration and wages suggest a zero correlation between the two variables for women and a negative correlation for men. However, the correction of the biases introduced by the crowd-out effect and the self-selection of workers results in a wage elasticity that has roughly the same magnitude for the two groups. In fact, the difference between the -0.8 elasticity for men and the -1.0 elasticity for women reported in column 8 is not statistically significant (the *t*-statistic is 0.57).[36] Therefore, the magnitude of the short-run wage impact of immigration on French workers is roughly the same for men and women.[37]

[35] The labor force participation rate of native women increased from 58.9 percent in 1982 to 80.5 percent in 2016, while that of native men declined slightly from 87.2 to 84.6 percent.

[36] Our estimate of the wage elasticity (between -0.8 and -1.0) is larger than the estimates reported in several cross-area studies, which often find negligible wage effects due to immigration (Blau and Mackie, 2016; Edo, 2019). Nevertheless, our finding of a negative wage response is consistent with Ortega and Verdugo's (2016) examination of immigration and the earnings of prime-age men in France over the 1976-2007 period (they report a wage elasticity between -0.2 and -1.0 when estimating the local impact of immigration on the wage of unskilled blue-collar native workers). It is also consistent with some of the U.S. evidence based on cross-area variations. For example, Altonji and Card (1991, pp. 222) report that a 1 percentage point increase in the fraction of immigrants in a metropolitan area reduces the wage of white (black) male high school dropouts by around –1.1 (-1.9) percent, while Jaeger, Ruist and Stuhler (2018) report a short-run wage elasticity for all native workers between -0.9 and -1.6 for the 1970-1980 period. Some recent studies of massive and unexpected inflows of immigrants also find sizable adverse wage effects. In particular, the estimated short-run wage elasticities in Borjas (2017), Edo (2020) and Monras (2020) are between [-0.5; -1.5], [-1.0; -2.0] and [-0.7; -1.4], respectively.

[37] Although not statistically significant, the larger point estimate for women suggests that the two groups may be imperfect substitutes. Edo and Toubal (2017) report that French immigration contributed to the reduction of the relative wage of native women between 1990 and 2010, thereby contributing to a widening gender wage gap.

The standard errors reported in Table 4 have been adjusted for clustering at the regional level to account for the possibility of within-group error correlation (Moulton, 1990). However, with only 22 clusters, our cluster-robust standard errors may be downward biased (Cameron, Gelbach and Miller, 2008). We will show below, however, that the evidence summarized in Table 4 is robust when we estimate the impact of immigration *(a)* at the departmental level (thereby using 94 French departments, instead of 22 regions); and *(b)* at the skill-region level (using two education groups and two age groups, hence exploiting variation across 88 clusters).

We also examined the precision of the estimated coefficients reported in Table 4 by implementing the wild cluster bootstrap method of Cameron, Gelbach and Miller (2008, p. 427) using 1,000 replications.[38] When using this method for the sample of native female workers, the wage elasticity to immigration is significant at the 5% level in columns 2 and 4, and at the 1% level in column 8 (with a corresponding Wild cluster bootstrap p-value of 0.002). For the sample of native men, the wild cluster bootstrap p-values for the wage elasticity in the IV regressions are always lower than 0.01.[39]

5.4. Robustness Tests

This section implements several robustness tests to assess the sensitivity of the "baseline" results reported in Table 4.

As discussed in Section 4.4, our baseline model employed slightly different specifications of the selection model by gender. In particular, the probit regressions included the family variables and household wealth for both men and women. However, the wage regression for men (but not for women) included controls for marital status and the presence of young children as these family variables may affect male earnings.

Table 5 reports the wage elasticity estimated using alternative specifications of the selection model. The baseline specification reproduces the elasticity estimates from Table

---

[38] Cameron, Gelbach and Miller (2008) demonstrate that this resampling method provides the most accurate cluster-robust inference in the case of a small number of clusters. Dustmann, Schonberg and Stuhler (2017) and Edo (2020) use this bootstrapping technique in their empirical analysis on the labor market impact of immigration.

[39] The wild cluster bootstrap estimates for the inverse Mills ratio in Table 4 indicate that the selection coefficient is always significant at the 1% level for native women and insignificant for native men.

4. Specification 2 includes the family variables in the female wage regression, but excludes them from the male wage regression. In the two last specifications, we only use the family variables (specification 3) or the home ownership indicator (specification 4) to generate independent variation in the inverse Mills ratio.

While the OLS and IV estimates reported in columns 1 and 4 are virtually zero for women, the inclusion of the inverse Mills ratio in columns 2 and 5 *always* results in a negative and significant elasticity—regardless of the specification of the selection model. This shift documents the bias in the estimated wage impact of immigration produced by the self-selection of women into employment. Columns 3 and 6 also control for the size of the native labor force and show that the estimated wage elasticity for women is robust to the modeling assumptions used (e.g., the IV elasticities range between -0.9 and -1.1 across the four specifications).

Panel B of Table 5 shows that the estimated wage elasticities for men are stable across specifications and columns.[40] The robustness of the estimated elasticity suggests that changes in how the regression model adjusts for the employment response and sample selection does not change the estimated impact of immigration on the male wage.[41]

Tables 6 to 11 provide additional sets of robustness tests. These additional tables all have the same structure and reproduce (separately by gender) the regressions reported in columns 1, 5, 6 and 8 of our baseline Table 4 using alternative specifications, samples, variable definitions, and dependent variables.[42]

---

[40] Table 4 reports the estimated wage elasticities if we simply assumed that the selection problem is not relevant for the study of male earnings. The no-selection parameter estimates are represented by the models that exclude the inverse Mills ratio (i.e., columns 1 and 3 for OLS, and columns 5 and 7 for IV). The wage elasticities obtained in this polar case are similar to those reported in Table 5 using alternative specifications of the selection model.

[41] Although the coefficient of the immigrant share is not sensitive to the specification of the selection model for either men or women, the (unreported) coefficient of the inverse Mills ratio is sensitive to the variables included in the *male* wage regression. In particular, this coefficient turns negative when the family characteristics are included in the male probit but excluded from the male wage regression (specifications 2 and 3). Marriage and the presence of young children have a very strong positive effect on both the male employment probability and male earnings. Excluding the family vector from the male wage regression then imparts a negative bias on the coefficient of the inverse Mills ratio (which is negatively correlated with the predicted probability of employment).

[42] Our baseline analysis applied a filter to exclude outliers who reported wages in either the top 0.5% or the bottom 0.5% of the hourly wage distribution. We apply this filter independently to each of the different samples being analyzed in the various robustness checks reported in this section. When applicable, we also re-estimated the homeownership probit regression used to impute missing values in the relevant sample. The

Table 6 estimates the model using alternative sample periods. Our baseline regressions merged data from the census and the LFS for the 1982, 1990, 1999, 2007, and 2016 cross-sections. In columns 1-4, we restrict the regressions to the 1990-2016 sample period for two reasons. The LFS adopted different sampling methods over time, so that the number of observations is much larger in the post-1990 surveys, leading to more precise wage measures for region-year cells in the latter part of the sample period. Moreover, starting the empirical analysis in 1990 helps to reduce the potential correlation between the shift-share instrument (based on the 1968 census) and current labor market outcomes. In other words, using an instrument based on the settlement pattern of immigrants who arrived more than 20 years earlier is more likely to satisfy the exclusion restriction imposed by the IV strategy.

The results from columns 1-4 of Panel A again illustrate the importance of accounting for sample selection and the size of the native employment response when estimating the wage elasticity for native women. The estimated IV coefficient of the immigrant share is -0.04 (0.12) in the simplest IV model reported in column 2, increases to -0.45 (0.11) in column 3 when the regression adjusts for sample selection, and more than doubles to -1.10 (0.33) in column 4 when the regression holds constant the size of the native labor force. In striking contrast to the variation in the estimated wage impact of immigration for women across the various columns, the estimated wage elasticity in the sample of native men is stable across specifications, hovering between -0.9 and -1.0.

The baseline analysis reported in Table 4 used data from five different cross-sections: 1982, 1990, 1999, 2007, and 2016. Since 2004, however, the French population censuses have been conducted annually. They can only be exploited every five years, so that an additional census is available for 2012. In our baseline analysis, we opted to use cross-sections that were spaced apart in roughly equal intervals, so that we skipped over the 2012 data. Columns 5-8 of Table 6 reproduce the regressions using all the available census data since 1982, thus expanding the study to six separate cross-sections. It is evident that adding in the additional 2012 cross-section barely changes our results.

---

homeownership probit used to impute missing values includes department fixed effects (instead of region fixed effects) in the department-level analysis.

Table 4 used the measure of the immigrant share implied by the theoretical framework, or log $(1 + M_{rt}/ N_{rt})$. We now use two alternative measures of the supply shock to determine the sensitivity of our results. In particular, columns 1-4 of Table 7 use the alternative measure given by log $(1 + M_{rt}/ N_{rt-1})$. In short, we use the size of the native labor force in the prior census as the base that defines the immigrant share.[43] This alternative measure of the supply shock addresses the potential concern that using the current native labor force to define the immigrant share may create a spurious correlation between immigration and regional wages (Card and Peri, 2016). Columns 5-8 of Table 7 use gender-specific immigrant shares to measure the supply shock, only using women to compute the immigrant share variable in Panel A and men in Panel B.[44] All the regression coefficients reported in Table 7 are similar to the baseline results in Table 4. While the estimated effects of immigration on female native wages are insignificant in the simplest model (columns 1-2 and 5-6), the wage response becomes stronger and statistically significant when controlling for selection bias and native labor supply. The estimated wage elasticity for men is again roughly similar across the different specifications and in line with the baseline elasticity estimates.

Table 8 uses two different regression specifications to estimate the wage impact of immigration. The first four columns report the coefficients when we do not use sampling weights in either the probit or wage regressions. The last four columns expand the (baseline) wage model in equation (14) by adding the full set of all possible (two- and three-way) interactions between the age, education, and region fixed effects and the age, education, and time fixed effects.[45]

Each of the specifications confirms that the selection-corrected wage impact of immigration for women is larger than the corresponding uncorrected estimate. For example, adding the inverse Mills ratio to the female wage equation changes the IV wage

[43] We construct the instrument for log $(1 + M_{rt}/ N_{rt-1})$ by following the same strategy described in Section 4 to predict M_{rt} and N_{rt-1} based on shift-share projections from the 1968 census.

[44] Although we used the same instruments as in Table 4 to implement our IV strategy, our estimated IV coefficients are robust to using gender-specific instruments.

[45] The full set of interactions is, of course, also included in the probit model used to calculate the inverse Mills ratio. Note that the full interaction model cannot include interactions between region and time fixed effects as the immigrant supply shock is measured at the region-year level.

elasticity from -0.1 to -0.4 in the unweighted regression model and from -0.4 to -0.6 in the full interaction model. In contrast, within each of the two alternative specifications, the estimated wage impact of immigration is relatively stable in the sample of native men (the wage elasticity is about -0.6 in columns 1-4 and ranges between -0.9 and -1.2 in columns 5-8).

The dependent variable in the baseline probit specification in Table 3 indicated whether a native person was employed and then we examined earnings in the subsample of full-time workers. Following Mulligan and Rubinstein (2008), columns 1-4 of Table 9 use an alternative probit model to compute the inverse Mills ratio. Specifically, the dependent variable is now a binary variable indicating whether the person is employed full-time (with the alternative outcome including both those not employed and those employed part-time). This alternative calculation of the Mills ratio does not change any of our baseline results. The wage elasticity for men is negative and hovers between -0.8 and -0.9 across specifications, while the wage elasticity for women again becomes more negative as the regression adjusts for sample selection and native labor supply.

Columns 5-8 of Table 9 extend the analysis by calculating the hourly wage rate for each worker in the sample.[46] The second-stage hourly wage regressions are then estimated using the entire sample of both full- and part-time workers. As with our baseline estimates, the inclusion of the inverse Mills ratio in the female wage regression leads to a far more negative wage elasticity; it doubles from -0.47 (0.09) to -0.94 (0.09) when we use the entire sample of female workers. In contrast, the wage elasticity estimated in the male sample is roughly constant across columns. Note that we find positive selection for both men and women when we use the hourly wage as the dependent variable, although the intensity of selection is again stronger for women.[47]

The wage elasticities estimated in the pooled sample of full- and part-time female

[46] The hourly wage rate is calculated by using information on usual hours worked in a typical week (except for the 1990 and 1999 LFS, which only report hours worked during the reference week). Because the reported number of weekly hours may contain substantial measurement errors, the estimated elasticities reported in Table 10 should be interpreted with caution (see Borjas, 1980, and Barrett and Hamermesh, 2019, for the United States; Laroque and Salanié, 2002, and Ortega and Verdugo, 2016, for France).

[47] The product of the coefficient of the inverse Mills ratio times its mean is 0.115 (or 0.23×0.50) for women and 0.05 (or 0.13×0.35) for men.

workers are larger than our baseline estimates: the female wage elasticity reported in column 8 of Table 4 for full-time workers is -0.98 (0.30), while the elasticity estimated in the entire sample of female native workers is -1.32 (0.22). This difference suggests a greater degree of substitutability between part-time native workers (84.2 percent of whom are women in our sample period) and immigrants.

Table 10 performs a final robustness check by using an alternative definition of a geographic region. Instead of defining the labor market in terms of the 22 regions in European France, we use the geographically smaller definition of a department (of which there are 94).[48] This sampling framework significantly increases the number of cells and introduces much more geographic variation in immigration and wages into the analysis.

The information on a person's department of residence, however, is not available in the LFS data collected between 2002 and 2012. As a result, our department-level analysis instead uses the 2013 LFS to obtain the wage and employment status of natives and merges this information with the population data provided in the 2012 census.[49]

Our instrument for the immigrant share differs slightly from that used at the region level. In particular, we instrument the immigrant share in the department-level regressions by using both the predicted share of immigrants in the region as defined in equation (17) *and* the predicted number of immigrants in a given department d at time t (constructed along the lines implied by equation (15)). This extension of the shift-share approach attempts to capture the possibility that network effects can be effective outside departmental boundaries. In particular, the presence of immigrants in one given department could affect the locational decision of co-nationals in neighboring areas within the same region. This IV strategy also has the advantage that it is less subject to potential bias introduced by sampling error if we only employed department-level shift-share instruments. To account for the endogeneity of the log native labor force, we use two analogous instruments: the log predicted female native labor force at the regional level (as

---

[48] Before 2016, European France was officially divided into 22 administrative regions, which represent the largest geographical units in the country. Each region is then divided into several administrative sub-regions called departments.

[49] The 2013 and 2016 LFS do not report any natives living in the Lozère department, so we exclude that department from the analysis. Lozère is the smallest department in France, containing only 0.12 percent of the native population in 2016.

defined in equation (18)) and the analogously constructed log predicted female native labor force at the departmental level.

Regardless of the specification, all of the results for native women and men at the department level are consistent with our baseline estimates and conclusions. The most general specification reported in column 8 indicates that the wage elasticity is essentially identical to the baseline estimate, and equals -1.00 (0.26) for men and -0.94 (0.22) for women.[50]

6. Skills and the Wage Impact of Immigration

This section extends the analysis by examining how the wage impact of immigration differs across skill groups. It also further tests the robustness of our results by adopting a variation of the skill-cell strategy (Borjas, 2003), where the wages of specific skill groups are linked directly to the influx of immigrants into the particular skill group.

Table 11 reports the coefficients resulting from an extension of the baseline analysis where we divide the sample into two education groups, workers who have completed their high school (by passing a French exam named the "Baccalauréat" giving access to college or an equivalent diploma) and those who have not. In 1982, only 21.2 percent of native workers had a Baccalaureate degree; by 2016, this fraction had increased to 56.0 percent.[51]

As we noted earlier, the measure of the supply shock in the baseline specification of Table 4 gives the immigration-induced percent increase in the size of the (entire) native labor force. This approach permits the estimated wage elasticity to capture both the "own" and the "cross" effects of immigration. Estimating the regression model separately by

[50] Edo, Giesing, Poutvaara and Öztunc (2019) document that immigrant supply shocks tend to generate native migration flows across departments over the 1988-2012 period, but they do not find any evidence of an internal migration response at the regional level (consistent with the fact that regions encompass a much larger area than departments). As a result, measuring the impact of the supply shock at the regional level (rather than at the smaller departmental level) reduces the potential concern that our estimated wage effects are contaminated by native internal migration.

[51] This classification splits the sample of full-time native workers into two relatively equally sized groups: 56.4 percent of French natives in our sample have a baccalaureate degree. The share of immigrants in the high (low) educated segment of the labor force increased from 4.1 percent (10.8 percent) in 1982 to 9.5 percent (13.9 percent) in 2016.

education group should provide some information about the relative wage effect of the *same* supply shock across skill groups.

Panel A of Table 11 reveals that the negative wage elasticity for women tends to be driven by the adverse impact of immigration on the low education group. The inclusion of the inverse Mills ratio increases the estimated wage elasticity for this skill group from -0.78 (0.18) in column 2 to -1.03 (0.21) in column 3. The inclusion of the native labor supply variable in column 4 increases the negative wage response even more, producing a wage elasticity of -1.32 (0.33). In contrast, the estimated IV wage effects in columns 6-8 for highly educated native women, although negative after accounting for sample selection, are not statistically significant.

The wage elasticities for men also suggest a stronger negative response for the low education group. The wage elasticity for low educated men ranges between -1.1 and -1.5, while the wage elasticity for highly educated men is between -0.4 and -0.5. In short, the data clearly point to a stronger adverse effect of immigration on the low-skill segment of the labor market. This finding likely reflects different degrees of competition between a specific supply shock and natives within each education group.

We conclude our empirical exploration by changing the unit of analysis from the region-year cell to a region-skill-year cell. We use two alternative definitions of a "skill group." We first use the two education groups introduced above (those who completed high school v. those who did not) to define the skill groups. We will also decompose each of the two education categories into two age groups (between 18-40 years old or 41-64 years old), to create four education-age skill cells. The key difference between this empirical strategy and the baseline specification is that we will now measure immigrant penetration and the size of the native labor force at the region-skill-year level rather than at the region-year level.

The general specification of the regression model is now given by:

$$\log w_{irst} = \theta_a + \theta_r + \theta_s + \theta_t + (\theta_r \times \theta_s) + \alpha_P P_{it} + \beta_1 m_{rst} + \beta_2 \log N_{rst} + \beta_3 \lambda_{it} + \mu_{it}, \quad (21)$$

where the dependent variable gives the monthly wage of individual i in region r and skill group s at time t. Note that the regression includes vectors of interacted region-skill fixed effects ($\theta_r \times \theta_s$) to control for unobserved time-invariant characteristics that are region-skill specific.[52] This estimation strategy implies that the wage elasticity to immigration is identified from changes that occur within region-skill groups over time. This empirical approach neglects cross-group complementarities and attempts to capture the own-effect of immigration on the earnings of similarly skilled natives.

We address the endogeneity of the immigrant share at the region-skill level by adapting the strategy introduced in our analysis of the department-level data in Table 10. In particular, we instrument the immigrant share m_{rst} by using the shift-share prediction of the regional immigrant share (i.e., the $\widehat{m}_{rt}$ defined in equation (17)) *and* the log predicted number of immigrants in a given region-skill group at time t, or $\log\ \widehat{M}_{rs}(t)$. To construct this variable, we adapt equation (15) and multiply the 1968 distribution of immigrants across region-skill cells for each country group n by the total number of immigrants from that group in subsequent years:[53]

$$\widehat{M}_{rs}(t) = \sum_{n} \frac{M_{rs}^{n}(1968)}{M^{n}(1968)} \cdot M^{n}(t), \tag{22}$$

where $M^n(t) = \sum_r \sum_s M_{rs}^n(t)$ and $M_{rs}^n(t)$ is the number of immigrants in national origin group n, skill group s, and region r at time t.

This IV strategy allows for the possibility that networks can impact the locational choice of immigrants regardless of their education level. It also helps mitigate the concern that sampling error or educational downgrading among immigrants could introduce some bias in the construction of our shift-share instrument for a region-skill cell.

We then follow the same empirical strategy as with the baseline specification. In particular, we compute an inverse Mills ratio for each native worker by estimating probit regressions (separately by gender) on the probability of employment. The probit model

[52] Although interacted education-time and region-time fixed effects cannot be included in the IV estimation, we added them to the OLS regressions in columns 1 and 2. This exercise makes the estimated OLS wage elasticities for women more negative and leave the estimated effect for men unchanged.

[53] We use the same nationality groups introduced in the construction of the shift-share instrument in Section 4.2.

again uses the Wooldridge method for models with endogenous regressors and selection bias. In the final step, we use the regression model in equation (21) using both OLS and IV to estimate the impact of the immigrant share on the wage of natives.

Table 12 reports the regression coefficients. While columns 1-6 exploit variations within region-education groups over time to identify the wage impact of immigration, columns 7-12 exploit variations within region-education-age groups. For each of these two alternative definitions of a region-skill cell, we report the main OLS and IV estimates. Note that columns 5 and 11 instrument the immigrant share only, though they control for the log native labor force. Columns 6 and 12 account for the endogeneity of this latter variable by using the log predicted female native labor force at the regional level and at the skill-region level.

The thrust of the evidence reported in Panel A (for native women) and Panel B (for native men) resembles our baseline findings. First, accounting for sample selection always makes the OLS and IV estimates of the impact of immigration on female wages much more negative. The IV wage elasticity jumps from -0.10 (0.11) to -0.31 (0.13) at the region-education level, and from -0.18 (0.10) to -0.43 (0.15) at the region-education-age level. In contrast, the male wage elasticity is much less responsive to the adjustment for selection bias. Second, the estimated coefficient of the inverse Mills ratio is always significantly positive for women, but is much less stable for men (and turns negative in the most general IV specification in the region-education-age analysis). Finally, holding native labor supply constant produces a more negative wage elasticity for women, suggesting a crowd-out effect at the region-skill level.

The wage elasticities reported in Table 12 are roughly of the same magnitude as our baseline estimates (particularly after taking standard errors into account). The intuition behind the different approaches (i.e., the unit of analysis being the region-year cell or the region-skill-year cell) would suggest a larger difference. In particular, the skill-cell approach summarized in equation (21) is more likely to isolate the "own" effect of immigration and may miss the complementary cross-cell effects (such as the impact of low-educated immigrants on the earnings of high-educated natives in the same region). As a result, the estimated wage elasticity would be expected to be more negative when using the region-skill-year breakdown.

However, the wage effects reported in Table 12 may be underestimated. First, there will likely be greater attenuation bias in an analysis that uses a "smaller" market (Aydemir and Borjas, 2011). The sample for estimating the immigrant share, the size of the native labor force, and the various instruments is far smaller when the analysis divides the regional labor market into distinct skill categories, resulting in attenuated estimates of the wage elasticity.

Second, if immigrants are placed in jobs that require less education than they actually have, their assignment to their nominal education groups produces an inaccurate measure of the size of the supply shock in a particular skill group (Dustmann, Frattini and Preston, 2013). Similarly, immigrants may not necessarily compete with natives in the same age group, especially if employers value the prior work experience of immigrants and natives differently. The measurement error might produce an additional attenuation bias in the estimated wage effect of immigration.

In sum, the lessons provided by exploiting information on supply shocks within specific skill cells confirm the key hypothesis proposed in this paper: A more complete evaluation of the wage impact of immigration requires the explicit modeling of the self-selected labor force produced by the native response.

7. Conclusion

The surge in international labor flows in the past few decades has inspired an equally large increase in the amount of economic research devoted to understanding and documenting the economic consequences of such flows. An important part of this rapidly expanding literature examines the impact that immigrants have on the labor market opportunities of native workers in the receiving countries. Much of this research is guided by an intuitive prediction of economic theory: An immigration-induced increase in the size of the labor force should reduce the wage of comparable workers, at least in the short run. Despite the intuitive appeal of this straightforward implication of the textbook supply-demand model, the evidence is mixed, and there is still substantial disagreement on even the direction of the wage impact of immigration on native workers despite three-decades worth of research on the subject.

Part of the difficulty in measuring the wage impact arises because native workers may respond to the supply shock by moving to labor markets that were not directly affected by immigration. This diffusion of the immigrant supply shock across markets attenuates the wage impact in the targeted market. As a result, standard comparisons of wages across markets may not truly measure the relative wage change experienced by the market targeted by immigrants.

This paper proposes and empirically explores a new hypothesis that provides a deeper understanding into how the diffusion might bias estimates of the wage impact of immigration. In particular, the wage change observed in a market targeted by immigrants will depend not only on the number of natives who respond by moving to other markets, but also on which native workers make the move. A non-random native response changes the composition of the sample of native workers, and this compositional shift artificially changes the average native wage in the affected markets. In the end, the selection bias may exacerbate, attenuate, or perhaps even reverse the sign of the measured wage impact of immigration.

We document the empirical relevance of this type of selection bias by examining how immigration differentially affected the employment and wages of men and women in France. Beginning with a policy shift in 1976, which gave foreign workers the right to family reunification and made it far easier for wives to join their husbands, France experienced a rapid "feminization" of its immigrant labor force.

The raw data in the French labor market reveals a striking gender asymmetry in how immigration correlates with wages and employment. The correlation between immigration and wages (across cities and over time) is negative for native men, but essentially zero for native women. At the same time, the correlation between immigration and employment rates is negative for native women, but essentially zero for native men.

Our theoretical framework combines a basic labor demand framework with the econometric model of selection to illustrate how the non-random response of native workers to the immigrant supply shock, and the subsequent self-selection of the native labor force, contaminates estimates of the key parameters of the labor demand function. Our empirical application of this framework in the context of the French labor market shows that the orthogonality between immigration and wages for French women is partly an artifact of selection bias. The native women who exited (or did not enter) the labor market after the supply shock tended to be low-wage women, mechanically increasing the average wage in those cities targeted by immigrants and making it seem as if immigration had no impact on the female wage. After adjusting for selection, our evidence indicates that the wage elasticity for native women is also negative and roughly the same size as that found for native men (where labor supply was much more inelastic). The gender asymmetry in the labor market impact of immigration can thus be reconciled by jointly examining how immigration affects wages and labor supply, and by accounting for the size and the inevitable selection bias in the native response.

It is important to emphasize that the selection bias problem we identify and explore in this paper probably contaminates many of the existing estimates of the wage impact of immigration. Immigrant supply shocks are likely to have an (immediate) effect on the labor market of receiving countries. Some native workers are likely to respond to these changes in economic opportunities. The native response is unlikely to be random, altering the composition of the native labor force after the supply shock. A valid assessment of the economic consequences of immigration inevitably requires a thorough examination of the direction and magnitude of the resulting selection bias.

References

Acemoglu, D., Autor, D., & Lyle, D. 2004. Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury. Journal of Political Economy 112(3): 497-555.

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, edited by J. Abowd and R. Freeman. University of Chicago Press, 201-234.

Amior, M. 2020. Immigration, Local Crowd-Out and Undercoverage Bias, Working Paper, Hebrew University of Jerusalem.

Angrist, J.A., & Kugler, A.D. 2003. Protective or Counter-Productive? Labor Market Institutions and the Effect of Immigration on EU Natives. Economic Journal 113 (488), F302-F331.

Aydemir, A., & Borjas, G.J. 2011. Attenuation Bias in Measuring the Wage Impact of Immigration. Journal of Labor Economics, 29(1), 69-112.

Barrett, G.F., & Hamermesh, D.S. 2019. Labor Supply Elasticities: Overcoming Nonclassical Measurement Error Using More Accurate Hours Data. Journal of Human Resources, 54(1), 255-265.

Beauchemin, C., Borrel, C., & Régnard, C. 2013. Immigrants in France: A Female Majority. Population Societies, (7), 1-4.

Blau, F. D., & Kahn, L.M. 2017. The Gender Wage Gap: Extent, Trends, and Explanations. Journal of Economic Literature, 55(3), 789-865.

Blau, F.D., & Mackie, C., eds. 2016. The Economic and Fiscal Consequences of Immigration. Washington, DC: National Academies Press.

Borjas, G.J. 1980. The Relationship Between Wages and Weekly Hours of Work: The Role of Division Bias. Journal of Human Resources 15(3): 409-423.

Borjas, G.J. 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.

Borjas, G.J. 2006. Native Internal Migration and the Labor Market Impact of Immigration. Journal of Human Resources 41(2): 221-258.

Borjas, G.J. 2017. "The Wage Impact of the Marielitos: A Reappraisal," Industrial and Labor Relations Review 70 (October): 1077-1110.

Borjas, G.J., & Monras, J. 2017. The Labour Market Consequences of Refugee Supply Shocks. Economic Policy, 32(91), 361-413.

Bratsberg, B., & Raaum, O. 2012. Immigration and wages: Evidence from construction. Economic Journal, 122(565), 1177-1205.

Bratsberg, B., Raaum, O., Røed, M., & Schøne, P. 2014. Immigration Wage Effects by Origin. The Scandinavian Journal of Economics, 116(2), 356-393.

Cameron, A.C., Gelbach, J., & Miller, D. 2008. Bootstrap-Based Improvements for Inference with Clustered Errors. Review of Economics and Statistics 90(3): 414-427.and Miller, 2008.

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., & Peri, G. 2016. Immigration Economics by George J. Borjas: A Review Essay. Journal of Economic Literature, 54(4), 1333-49.

Choi, H.J., Joesch, J.M., & Lundberg, S. 2008. Sons, Daughters, Wives, and the Labour Market Outcomes of West German Men. Labour Economics 15(5): 795-811.

Cortés, P., & Pan, J. 2019. When Time Binds: Substitutes for Household Production, Returns to Working Long Hours, and the Skilled Gender Wage Gap. *Journal of Labor Economics*, *37*(2), 351-398.

Cortés, P., & Tessada, J. 2011. Low-Skilled Immigration and the Labor Supply of Highly Skilled Women. *American Economic Journal: Applied Economics*, *3*(3), 88-123.

Dustmann, C., Frattini, T., & Preston, I.P. 2013. The Effect of Immigration along the Distribution of Wages. Review of Economic Studies, 80(1), 145-173.

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.

Dustmann, C., Schönberg, U., & Stuhler, J. 2017. Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment. Quarterly Journal of Economics, 132(1), 435-483.

Dolado, J. J., García-Peñalosa, C., & Tarasonis, L. 2019. The Changing Nature of Gender Selection into Employment over the Great Recession. Economic Policy.

Edo, A. 2019. The Impact of Immigration on the Labor Market. Journal of Economic Surveys, 33(3), 922-948.

Edo, A. 2020. The Impact of Immigration on Wage Dynamics: Evidence from the Algerian Independence War. Journal of the European Economic Association 18(6): 3210-3260.

Edo, A., Glesing, Y., Öztunc, J., & Poutvaara, P. 2019. Immigration and Electoral Support for the Far-Left and the Far-Right. European Economic Review 115: 99-143.

Edo, A., & Toubal, F. 2017. Immigration and the Gender Wage Gap. European Economic Review, 92, 196-214.

Evers, M., De Mooij, R., & Van Vuuren, D. 2008. The Wage Elasticity of Labour Supply: A Synthesis of Empirical Estimates. De Economist, 156(1), 25-43.

Farré, L., González, L., & Ortega, F. 2011. Immigration, Family Responsibilities and the Labor Supply of Skilled Native Women. The BE Journal of Economic Analysis & Policy, 11(1).

Foged, M. & Peri, G. 2016. Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data. American Economic Journal: Applied Economics 8(2): 1-34.

Friedberg, R. 2001. The Impact of Mass Migration on the Israeli Labor Market. Quarterly Journal of Economics 116(4): 1373-1408.

Furtado, D., & Hock, H. 2010. Low Skilled Immigration and Work-Fertility Tradeoffs Among High Skilled US Natives. American Economic Review 100(2): 224-28.

Glitz, A. 2012. The Labor Market Impact of Immigration: A Quasi-Experiment Exploiting Immigrant Location Rules in Germany, Journal of Labor Economics 30(1), 175–213.

Gronau, R. 1974. Wage Comparisons: A Selectivity Bias. Journal of Political Economy 82 (November/December 1974): 1119-43.

Gurgand, M., & Margolis, D.N. (2008). Does Work Pay in France? Monetary Incentives, Hours Constraints, and the Guaranteed Minimum Income. *Journal of Public Economics*, *92*(7), 1669-1697.

Heckman, J. 1979. Sample Selection Bias as a Specification Error, Econometrica 47(1): 153–161.

Hunt, J. 2017. The Impact of Immigration on the Educational Attainment of Natives. Journal of Human Resources, 52(4), 1060-1118.

Jaeger, D. A., Ruist, J., & Stuhler, J. 2018. Shift-Share Instruments and the Impact of Immigration (No. w24285). National Bureau of Economic Research. (updated version 2019)

Katz, L.F., & Murphy, K.M. 1992. Changes in Relative Wages, 1963-1987: Supply and Demand Factors. Quarterly Journal of Economics 107(1): 35-78.

Kleibergen, F., & Paap, R. 2006. Generalized Reduced Rank Tests Using the Singular Value Decomposition. Journal of Econometrics 133(1): 97-126.

Laroque, G., & Salanié, B. 2002. Labour Market Institutions and Employment in France. Journal of Applied Econometrics, 17(1), 25-48.

Llull, J. 2018. Immigration, Wages, and Education: A Labour Market Equilibrium Structural Model. Review of Economic Studies, 85(3), 1852-1896.

Llull, J. 2021. Immigration and Gender Differences in the Labor Market, Journal of Human Capital, forthcoming.

Lundberg, S., & Rose, E. 2000. Parenthood and the Earnings of Married Men and Women. Labour Economics 7(6): 689-710.

McDonald, P. 2020. The Male Marriage Premium: Selection, Productivity, or Employer Preferences? Journal of Marriage and Family, forthcoming.

Monras, J. 2020. "Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis," Journal of Political Economy 128 (August): 3017-3089.

Monras, J. 2021. Local Adjustment to Immigrant-driven Labor Supply Shocks, Journal of Human Capital, forthcoming.

Moulton, B.R. 1990. An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. Review of Economics and Statistics 72(2): 334-338.

Mulligan, C.B., & Rubinstein, Y. 2008. Selection, Investment, and Women's Relative Wages over Time, Quarterly Journal of Economics 123 (3): 1061-1110.

Ortega, J., & Verdugo, G. 2016. Moving up or down? Immigration and the selection of natives across occupations and locations. IZA Discussion Paper No. 10303. (updated version 2020)

Pencavel, J. 1986. Labor Supply of Men: A Survey. In Handbook of Labor Economics, Volume 1, edited by O. Ashenfelter and R. Layard. Elsevier Science Publishers, 3-102.

Piketty, T. 1998. L'impact des Incitations Financières au Travail sur les Comportements Individuels: Une Estimation pour le cas Français. Économie & prévision, 132(1), 1-35.

Olivetti, C., & Petrongolo, B. 2008. Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps. Journal of Labor Economics, 26(4), 621-654.

Sanderson E., & Windmeijer, F. 2016. A Weak Instrument F-Test in Linear IV Models with Multiple Endogenous Variables," Journal of Econometrics 190(2): 212-221.

Seah, K.K.C. 2018. The Effect of Immigration Shocks on Native Fertility Outcomes: Evidence from a Natural Experiment. IZA Journal of Development and Migration 8, 18.

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

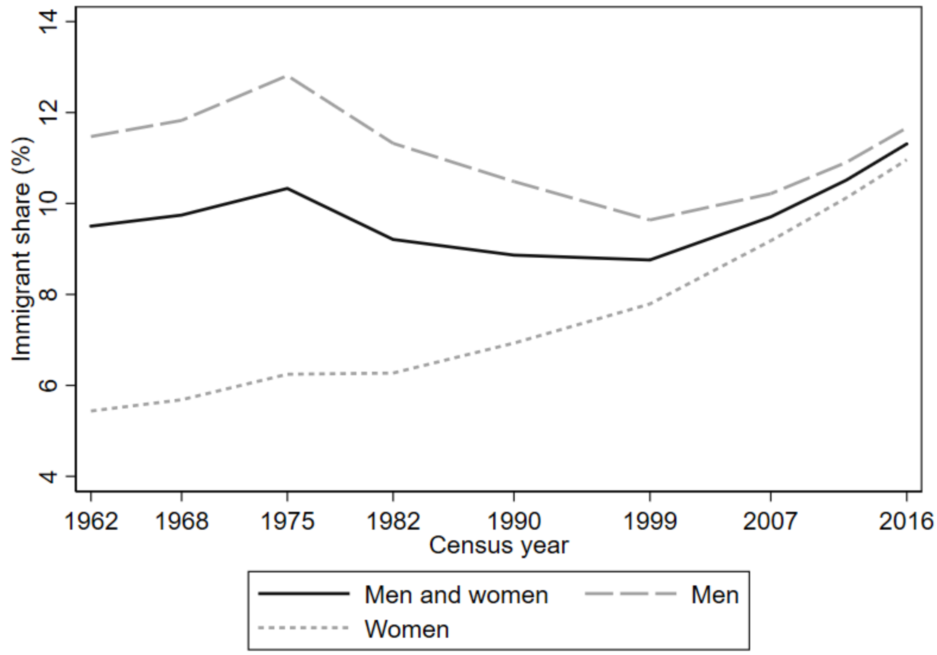
Stock, J., Yogo, M. & Wright, J. 2002. A Survey of Weak Instruments and Weak Identification in a Generalized Method of Moments," Journal of Business and Economic Statistics 20(4): 518-529.

Winter-Ebmer, R., & Zweimüller, J. 1996. Immigration and the Earnings of Young Native Workers. Oxford Economic Papers 48(3): 473-491.

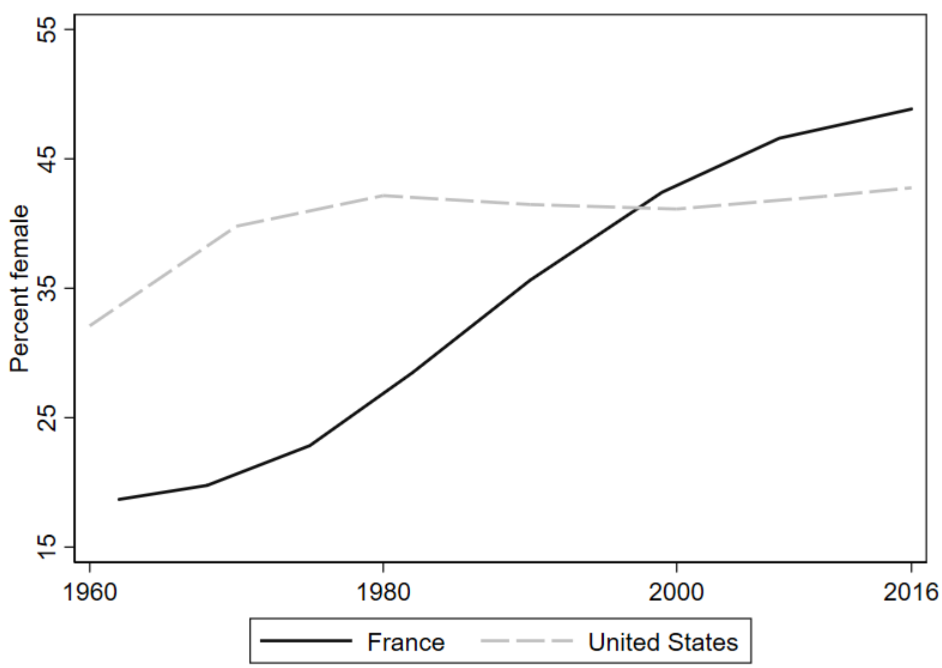
Wooldridge, J.M. 2002. Econometric Analysis of Cross Section and Panel Data, Cambridge, MA: MIT press.

Figure 1. Immigration and gender

A. Trends in the immigrant share in the French labor force

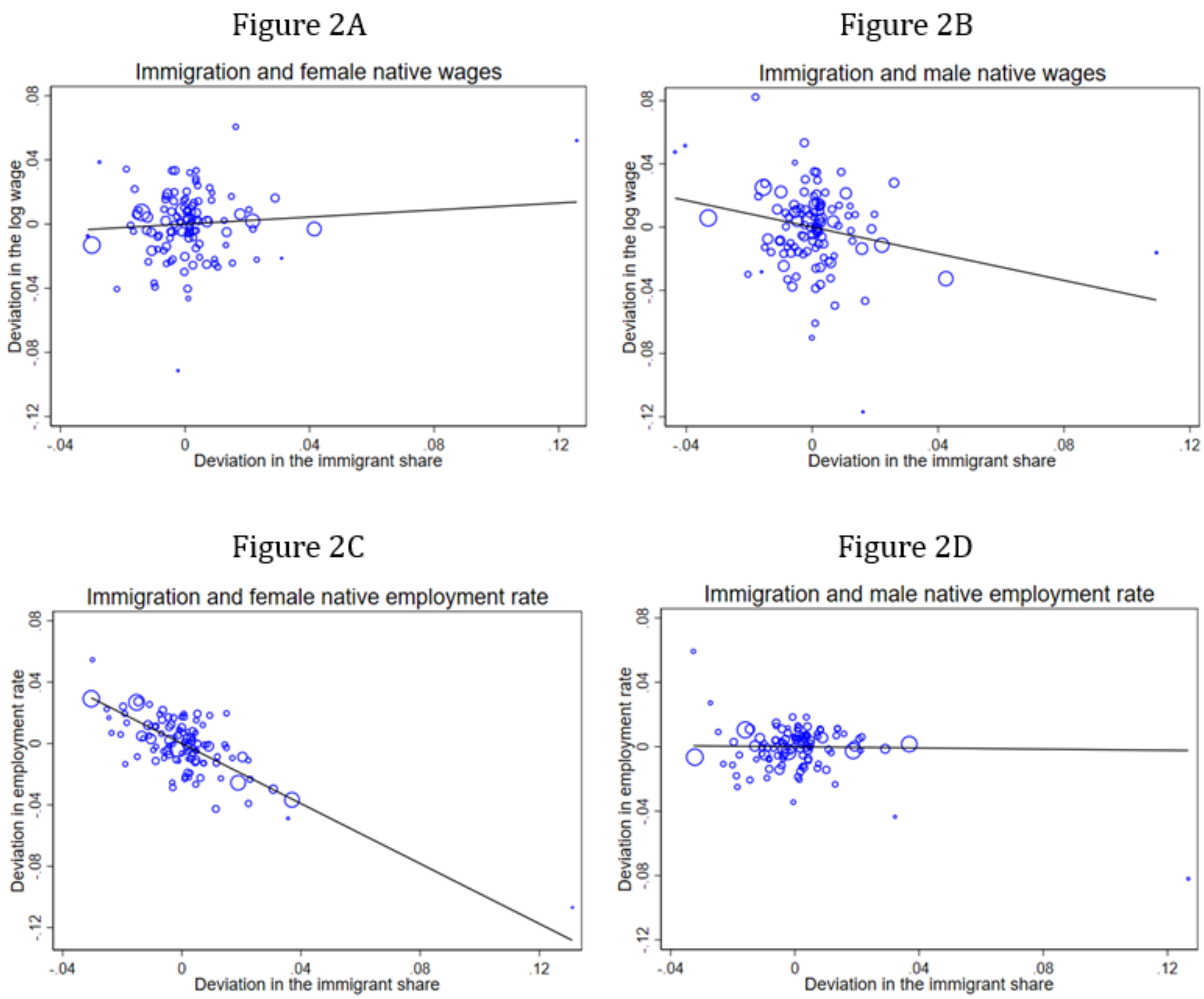


B. The feminization of the immigrant labor force, France v. USA



Source: INSEE, French censuses; IPUMS, USA decennial censuses and American Community Surveys.

Figure 2. Immigration, wages, and employment of native men and women



Notes: The unit of observation in the scatter diagrams is a region-year cell over the 1982-2016 period. We merged the following years of the LFS to create cross-section wage samples that correspond to the timing of the French censuses: 1982-1983, 1990-1991, 1998-1999-2000, 2006-2007-2008 and 2015-2016-2017. Figures 2A and 2B (Figures 2C and 2D) correlate the deviation in the log monthly wage (employment rate) of native women and men, respectively, to the deviation in the immigrant share after removing any year-specific effects that are common to all regions in a given census year. The deviations in the log wage, employment rate, or immigrant share are residuals from regressions of these variables on region fixed effects and census year fixed effects. The regression line in the figures weights the data by the number of observations used to compute the dependent variable.

Table 1: Descriptive statistics

	1962	1968	1975	1982	1990	1999	2007	2016
	A. French census data							
Immigrant share	9.50	9.75	10.33	9.21	8.86	8.76	9.71	11.31
Immigrant share, women	5.44	5.68	6.25	6.27	6.93	7.79	9.18	10.96
With a baccalaureate degree	3.16	2.94	3.05	3.25	4.58	5.82	7.38	9.19
With less than a baccalaureate degree	5.73	6.17	7.19	7.37	8.08	9.33	11.38	14.13
Immigrant share, men	11.47	11.83	12.81	11.32	10.48	9.64	10.22	11.66
With a baccalaureate degree	5.34	4.55	4.66	4.83	6.50	7.80	8.86	9.95
With less than a baccalaureate degree	12.18	12.96	14.63	13.17	11.99	10.69	11.40	13.78
Employment rate of female natives	37.15	40.08	47.55	51.51	56.40	62.19	67.76	70.06
Employment rate of male natives	89.43	87.36	86.62	81.05	77.34	75.47	75.45	73.64
	B. French labor force survey data							
Average wage of female natives	-	-	-	1626.8	1639.1	1746.2	1846.7	1896.3
Average wage of male natives	-	-	-	2049.8	2014.7	2047.9	2168.8	2213.6
Employment rate of female natives	-	-	-	55.18	56.60	61.73	63.57	65.83
Employment rate of male natives	-	-	-	83.41	78.69	76.28	71.55	70.42
Observations	-	-	-	32,446	78,531	83,311	59,414	75,446

Notes: The table uses data drawn from the French censuses (Panel A) and the French Labour Force Surveys (Panel B). The immigrant shares are computed using the sample of persons in the labor force and are defined as log $(1 + M/N)$, where M and N give the number of foreign-born and native labor force participants, respectively.

Table 2: Instrumental variables, first-stage regressions

	Sample of native women		Sample of native men	
	(1)	(2)	(3)	(4)
	A. Single endogenous variable model			
Dependent variable: Immigrant share in workforce				
Predicted immigrant share in population	1.77***	-	1.70***	-
	(0.27)		(0.32)	
Kleibergen-Paap F-test of excluded instrument	42.14	-	27.63	-
Observations	71,326	-	103,704	-
	B. Two-endogenous variables model			
Dependent variable: Immigrant share in workforce				
Log predicted immigrant population	-	0.12***	-	0.11***
		(0.02)		(0.02)
Log predicted female native labor force	-	-0.14***	-	-0.13***
		(0.03)		(0.04)
SW multivariate F-test of excluded instruments	-	17.17	-	18.76
Dependent variable: Log of native labor force				
Log predicted immigrant population	-	-0.09	-	-0.09
		(0.07)		(0.07)
Log predicted female native labor force	-	0.58***	-	0.58***
		(0.08)		(0.08)
SW multivariate F-test of excluded instruments	-	20.15	-	22.00
Observations	-	71,326	-	103,704

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The table reports the first-stage IV regressions in the estimation sample. In Panel A, the dependent variable is the immigrant share in the labor force. In Panel B, the dependent variables are the immigrant share and the log number of natives in the labor force. As instruments, we use the predicted immigrant share in the population based on the geographic settlement of immigrants and natives in the 1968 census and the predicted female native labor force based on the geographic settlement of natives in the 1968 census and the relative number of women with young children in subsequent years. As tests for weak instruments, Panel A reports the Kleibergen-Paap rk Wald F-test for the excluded instrument, while Panel B reports the Sanderson-Windmeijer (SW) F-tests of excluded instruments for each endogenous regressor. All regressions include age, education, region, and time fixed effects, and use the individual weight provided by INSEE. The regressions estimated in the male sample also include marital status and a variable indicating the presence of young children. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

Table 3: Probit regressions on the employment probability of natives

	Native women		Native men	
	Reduced form probit	Instrumental variable probit	Reduced form probit	Instrumental variable probit
	(1)	(2)	(3)	(4)
Predicted immigrant share in population	-4.44***	-	-0.15	-
	(1.10)		(1.27)	
Marginal effect	*-1.41*		*-0.04*	
Immigrant share	-	-3.65***	-	0.21
		(1.06)		(0.95)
Marginal effect		*-1.16*		*0.05*
Log predicted native population in t-1	-0.50***	-0.19	-0.64***	-0.69***
	(0.19)	(0.26)	(0.18)	(0.24)
Marginal effect	*-0.16*	*-0.06*	*-0.16*	*-0.18*
Married	-0.04*	-0.04*	0.39***	0.39***
	(0.02)	(0.02)	(0.02)	(0.02)
Marginal effect	*-0.01*	*-0.01*	*0.10*	*0.10*
Presence of children below 6	-0.32***	-0.32***	0.26***	0.26***
	(0.02)	(0.02)	(0.03)	(0.03)
Marginal effect	*-0.10*	*-0.10*	*0.07*	*0.07*
Home ownership	-0.12***	-0.12***	-0.09***	-0.09***
	(0.02)	(0.02)	(0.02)	(0.02)
Marginal effect	*-0.04*	*-0.04*	-0.02	-0.02
Observations	173,432	173,432	155,716	155,716

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. Below the standard errors, we report the marginal effect of each variable computed at the mean value of the sample. The dependent variable is a binary variable equal to one if the individual is employed and zero otherwise. We use the regressions estimated in columns 1 and 3 to predict the inverse Mills ratio that is included in the IV second-stage individual-level wage regressions with one endogenous variable. In columns 2 and 4, the immigrant supply shock is defined as the number of immigrants relative to the native labor force in census year $t-1$, and its instrument is the corresponding shift-share instrument based on the 1968 French census. All regressions include age, education, region, and time fixed effects, and use the individual weight provided by INSEE. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

Table 4: Impact of immigration on native individual wages

	OLS estimates				IV estimates			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. Impact on the wage of native women							
Immigrant share	-0.02	-0.42***	-0.00	-0.34***	-0.01	-0.44***	-0.61**	-0.98***
	(0.06)	(0.08)	(0.11)	(0.12)	(0.11)	(0.10)	(0.30)	(0.30)
Inverse Mills ratio	-	0.21***	-	0.21***	-	0.22***	-	0.22***
		(0.02)		(0.02)		(0.02)		(0.02)
Log of native labor force	-	-	0.01	0.05	-	-	-0.25**	-0.21**
			(0.06)	(0.06)			(0.10)	(0.09)
Kleibergen-Paap F-test	-	-	-	-	42.14	43.60	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	17.17	17.26
SW multivariate F-test (log nat.)	-	-	-	-	-	-	20.15	20.35
	B. Impact on the wage of native men							
Immigrant share	-0.79***	-0.81***	-0.65***	-0.66***	-0.90***	-0.93***	-0.80***	-0.78***
	(0.12)	(0.14)	(0.15)	(0.16)	(0.09)	(0.11)	(0.18)	(0.18)
Inverse Mills ratio	-	0.05	-	0.06	-	0.06	-	0.05
		(0.07)		(0.07)		(0.06)		(0.06)
Log of native labor force	-	-	0.08	0.09*	-	-	0.02	0.04
			(0.05)	(0.05)			(0.07)	(0.07)
Kleibergen-Paap F-test	-	-	-	-	27.63	27.89	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	18.76	19.17
SW multivariate F-test (log nat.)	-	-	-	-	-	-	22.00	22.48

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The dependent variable is the log monthly wage of full-time native women (Panel A) or men (Panel B) over the 1982-2016 period. Columns 5-6 instrument the share of immigrants with the shift-share instrument computed using the 1968 French census; columns 7-8 instrument both the share of immigrants and the log native labor force by using the shift-share instrument and the predicted (log) size of the female native labor force. All regressions include age, education, region, and time fixed effects, and use the individual weight provided by INSEE. The regressions in Panel A (B) have 71,326 (103,704) observations. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

Table 5: Immigration and wages using alternative selection models

	Regressors included: X = age, education, region, time f.e.; F = family characteristics; H = home ownership		OLS estimates			IV estimates		
			(1)	(2)	(3)	(4)	(5)	(6)
	Probit regression	Wage regression	**A. Impact on the wage of native women**					
Baseline specification	(X,F,H)	(X)	-0.02	-0.42***	-0.34***	-0.01	-0.44***	-0.98***
			(0.06)	(0.08)	(0.12)	(0.11)	(0.10)	(0.30)
Specification 2	(X,F,H)	(X,F)	-0.03	-0.38***	-0.31**	-0.01	-0.42***	-0.99***
			(0.06)	(0.11)	(0.13)	(0.11)	(0.13)	(0.31)
Specification 3	(X,F)	(X)	-0.02	-0.38***	-0.31**	-0.01	-0.39***	-0.95***
			(0.06)	(0.08)	(0.11)	(0.11)	(0.11)	(0.31)
Specification 4	(X,H)	(X)	-0.02	-0.49***	-0.40***	-0.01	-0.56***	-1.09***
			(0.06)	(0.13)	(0.14)	(0.11)	(0.15)	(0.32)
	Probit regression	Wage regression	**B. Impact on the wage of native men**					
Baseline specification	(X,F,H)	(X,F)	-0.79***	-0.81***	-0.66***	-0.90***	-0.93***	-0.78***
			(0.12)	(0.14)	(0.16)	(0.09)	(0.11)	(0.18)
Specification 2	(X,F,H)	(X)	-0.72***	-0.57***	-0.51***	-0.83***	-0.65***	-0.92***
			(0.11)	(0.11)	(0.14)	(0.09)	(0.09)	(0.22)
Specification 3	(X,F)	(X)	-0.72***	-0.56***	-0.51***	-0.83***	-0.64***	-0.92***
			(0.11)	(0.11)	(0.15)	(0.09)	(0.10)	(0.22)
Specification 4	(X,H)	(X)	-0.72***	-0.77***	-0.59***	-0.83***	-0.90***	-0.68***
			(0.11)	(0.13)	(0.15)	(0.09)	(0.10)	(0.20)
Inverse Mills ratio			-	Yes	Yes	-	Yes	Yes
Log of native labor force			-	-	Yes	-	-	Yes

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The dependent variable is the log monthly wage of full-time native women (Panel A) or men (Panel B). Each row uses a specific set of variables to compute the inverse Mills ratio and estimate the wage regressions. Columns 4-5 instrument the share of immigrants with the shift-share instrument computed using the 1968 French census; column 6 instruments both the share of immigrants and the log native labor force by using the shift-share instrument and the predicted (log) size of the female native labor force. All regressions use the individual weight provided by INSEE. The regressions in Panel A (B) have 71,326 (103,704) observations. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

Table 6: Immigration and wages using alternative sample periods

	Sample period: 1990-2016				Baseline period, adds 2012			
	OLS	IV estimates			OLS	IV estimates		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. Impact on the wage of native women							
Immigrant share	-0.11*	-0.04	-0.45***	-1.10***	-0.03	0.00	-0.43***	-0.96***
	(0.06)	(0.12)	(0.11)	(0.33)	(0.06)	(0.11)	(0.09)	(0.29)
Inverse Mills ratio	-	-	0.20***	0.21***	-	-	0.22***	0.23***
			(0.02)	(0.02)			(0.02)	(0.02)
Log of native labor force	-	-	-	-0.20*	-	-	-	-0.20**
				(0.11)				(0.09)
Kleibergen-Paap F-test	-	38.07	39.04	-	-	42.07	43.61	-
SW multivariate F-test (imm. share)	-	-	-	16.15	-	-	-	17.64
SW multivariate F-test (log nat.)	-	-	-	17.13	-	-	-	20.95
	B. Impact on the wage of native men							
Immigrant share	-0.91***	-0.99***	-1.02***	-0.88***	-0.77***	-0.87***	-0.90***	-0.75***
	(0.14)	(0.12)	(0.13)	(0.23)	(0.12)	(0.10)	(0.12)	(0.18)
Inverse Mills ratio	-	-	0.05	0.05	-	-	0.04	0.04
			(0.06)	(0.06)			(0.06)	(0.06)
Log of native labor force	-	-	-	0.07	-	-	-	0.05
				(0.09)				(0.07)
Kleibergen-Paap F-test	-	24.05	24.22	-	-	27.14	27.33	-
SW multivariate F-test (imm. share)	-	-	-	18.02	-	-	-	19.54
SW multivariate F-test (log nat.)	-	-	-	18.97	-	-	-	23.14

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The dependent variable is the log monthly wage of full-time native women (Panel A) or men (Panel B). The regressions in columns 1-4 use the 1990-2016 cross-sections; the regressions in columns 5-8 use the original 1982-2016 cross-sections and add the 2012 panel. Columns 2-3 and 6-7 instrument the share of immigrants with the shift-share instrument computed using the 1968 French census; columns 4 and 8 instrument both the share of immigrants and the log native labor force by using the shift-share instrument and the predicted (log) size of the female native labor force. All regressions include age, education, region, and time fixed effects, and use the individual weight provided by INSEE. The female (male) regressions in columns 1-4 have 64,015 (92,400) observations. The female (male) regressions in columns 5-8 have 89,934 (128,263) observations. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

Table 7: Immigration and wages using alternative measures of the supply shock

	Immigrants to pre-existing natives				Gender-specific supply shock			
	OLS	IV estimates			OLS	IV estimates		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Impact on the wage of native women								
Immigrant share	-0.05	-0.02	-0.47***	-0.94***	0.02	-0.01	-0.36***	-1.03**
	(0.08)	(0.12)	(0.11)	(0.25)	(0.05)	(0.09)	(0.10)	(0.46)
Inverse Mills ratio	-	-	0.22***	0.22***	-	-	0.22***	0.23***
			(0.02)	(0.02)			(0.02)	(0.02)
Log of native labor force	-	-	-	-0.18**	-	-	-	-0.32*
				(0.07)				(0.17)
Kleibergen-Paap F-test	-	17.90	19.17	-	-	27.61	27.78	-
SW multivariate F-test (imm. share)	-	-	-	22.51	-	-	-	7.30
SW multivariate F-test (log nat.)	-	-	-	21.24	-	-	-	10.32
B. Impact on the wage of native men								
Immigrant share	-0.73***	-0.87***	-0.90***	-0.74***	-0.81***	-1.16***	-1.19***	-0.79***
	(0.15)	(0.14)	(0.16)	(0.17)	(0.28)	(0.19)	(0.21)	(0.20)
Inverse Mills ratio	-	-	0.05	0.05	-	-	0.05	0.05
			(0.06)	(0.06)			(0.06)	(0.06)
Log of native labor force	-	-	-	0.07	-	-	-	0.11
				(0.07)				(0.07)
Kleibergen-Paap F-test	-	13.80	14.04	-	-	46.20	47.07	-
SW multivariate F-test (imm. share)	-	-	-	24.91	-	-	-	41.91
SW multivariate F-test (log nat.)	-	-	-	22.88	-	-	-	27.25

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The dependent variable is the log monthly wage of full-time native women (Panel A) or men (Panel B). The regressions in columns 1-4 define the immigrant share as the number of immigrants in census year t relative to the number of native workers in census year t-1; columns 5-8 use the gender-specific immigrant share in the labor force. Columns 2-3 and 6-7 instrument the share of immigrants with the shift-share instrument computed using the 1968 French census; columns 4 and 8 instrument both the share of immigrants and the log native labor force by using the shift-share instrument and the predicted (log) size of the female native labor force. All regressions include age, education, region, and time fixed effects, and use the individual weight provided by INSEE. The regressions in Panel A (B) have 71,326 (103,704) observations. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

Table 8: Immigration and wages using alternative specifications

	Unweighted regression model				Full interaction model			
	OLS	IV estimates			OLS	IV estimates		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Impact on the wage of native women								
Immigrant share	0.02	-0.08	-0.38***	-0.72***	-0.33***	-0.36***	-0.57***	-0.76***
	(0.06)	(0.11)	(0.11)	(0.23)	(0.08)	(0.11)	(0.11)	(0.27)
Inverse Mills ratio	-	-	0.16***	0.16***	-	-	0.15***	0.15***
			(0.01)	(0.01)			(0.02)	(0.02)
Log of native labor force	-	-	-	-0.17***	-	-	-	-0.06
				(0.07)				(0.10)
Kleibergen-Paap F-test	-	24.61	25.79	-	-	42.64	43.36	-
SW multivariate F-test (imm. share)	-	-	-	20.73	-	-	-	17.14
SW multivariate F-test (log nat.)	-	-	-	23.57	-	-	-	20.68
B. Impact on the wage of native men								
Immigrant share	-0.58***	-0.58***	-0.62***	-0.62***	-1.07***	-1.21***	-1.19***	-0.87***
	(0.15)	(0.19)	(0.21)	(0.21)	(0.19)	(0.15)	(0.18)	(0.22)
Inverse Mills ratio	-	-	0.07	0.07	-	-	-0.02	-0.03
			(0.04)	(0.04)			(0.08)	(0.08)
Log of native labor force	-	-	-	0.01	-	-	-	0.11
				(0.08)				(0.10)
Kleibergen-Paap F-test	-	18.56	18.88	-	-	28.16	28.95	-
SW multivariate F-test (imm. share)	-	-	-	23.07	-	-	-	20.15
SW multivariate F-test (log nat.)	-	-	-	26.05	-	-	-	24.10

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The dependent variable is the log monthly wage of full-time native women (Panel A) or men (Panel B). We do not weight the probit and wage regressions in columns 1-4; the regressions in columns 5-8 use the individual weights provided by INSEE, and also include as controls all interacted age-education-region fixed effects and all interacted age-education-time fixed effects. Columns 2-3 and 6-7 instrument the share of immigrants with the shift-share instrument computed using the 1968 French census; columns 4 and 8 instrument both the share of immigrants and the log native labor force by using the shift-share instrument and the predicted (log) size of the female native labor force. All regressions include age, education, region, and time fixed effects. The regressions in Panel A (B) have 71,326 (103,704) observations. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

Table 9: Immigration and wages using alternative samples of native workers

	Probit on full-time employment				Full- and part-time workers			
	OLS	IV estimates			OLS	IV estimates		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. Impact on the wage of native women							
Immigrant share	-0.02	-0.01	-0.24**	-0.72***	-0.49***	-0.47***	-0.94***	-1.32***
	(0.06)	(0.11)	(0.10)	(0.28)	(0.04)	(0.09)	(0.09)	(0.22)
Inverse Mills ratio	-	-	0.14***	0.14***	-	-	0.23***	0.23***
			(0.01)	(0.01)			(0.03)	(0.03)
Log of native labor force	-	-	-	-0.20**	-	-	-	-0.16**
				(0.09)				(0.07)
Kleibergen-Paap F-test	-	42.14	42.60	-	-	33.73	34.78	-
SW multivariate F-test (imm. share)	-	-	-	17.30	-	-	-	17.18
SW multivariate F-test (log nat.)	-	-	-	20.35	-	-	-	20.24
	B. Impact on the wage of native men							
Immigrant share	-0.79***	-0.90***	-0.93***	-0.77***	-0.96***	-1.00***	-1.07***	-0.85***
	(0.12)	(0.09)	(0.11)	(0.18)	(0.11)	(0.12)	(0.14)	(0.20)
Inverse Mills ratio	-	-	0.06	0.06	-	-	0.13**	0.13**
			(0.07)	(0.07)			(0.05)	(0.05)
Log of native labor force	-	-	-	0.05	-	-	-	0.09
				(0.07)				(0.08)
Kleibergen-Paap F-test	-	27.63	28.05	-	-	27.41	27.65	-
SW multivariate F-test (imm. share)	-	-	-	19.61	-	-	-	19.03
SW multivariate F-test (log nat.)	-	-	-	22.73	-	-	-	22.31

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The dependent variable in columns 1-4 is the log monthly wage of full-time native women (Panel A) or men (Panel B). The inverse Mills ratio in columns 3-4 is derived from probit regressions where the dependent variable is a full-time indicator (instead of an employment indicator as in our baseline regressions or in columns 5-8). The dependent variable in columns 5-8 is the log hourly wage of both full- and part-time native workers. Columns 2-3 and 6-7 instrument the share of immigrants with the shift-share instrument computed using the 1968 French census; columns 4 and 8 instrument both the share of immigrants and the log native labor force by using the shift-share instrument and the predicted (log) size of the female native labor force. All regressions include age, education, region, and time fixed effects, and use the individual weight provided by INSEE. The female (male) regressions in columns 1-4 have 71,326 (103,704) observations. The female (male) regressions in columns 5-8 have 98,451 (108,198) observations. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

Table 10: Immigration and wages using geographic variation across departments

	OLS estimates				IV estimates			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Impact on the wage of native women								
Immigrant share	-0.11	-0.42***	-0.13	-0.41***	-0.05	-0.51***	-0.63***	-0.94***
	(0.07)	(0.09)	(0.10)	(0.12)	(0.20)	(0.19)	(0.19)	(0.22)
Inverse Mills ratio	-	0.24***	-	0.24***	-	0.24***	-	0.24***
		(0.05)		(0.05)		(0.04)		(0.04)
Log of native labor force	-	-	-0.02	0.01	-	-	-0.25***	-0.19***
			(0.04)	(0.04)			(0.08)	(0.07)
Kleibergen-Paap F-test	-	-	-	-	9.17	8.98	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	15.48	15.55
SW multivariate F-test (log nat.)	-	-	-	-	-	-	16.58	16.86
B. Impact on the wage of native men								
Immigrant share	-0.56***	-0.59***	-0.49***	-0.51***	-0.81***	-0.85***	-1.00***	-1.00***
	(0.12)	(0.12)	(0.14)	(0.13)	(0.20)	(0.20)	(0.25)	(0.26)
Inverse Mills ratio	-	0.08**	-	0.08**	-	0.08**	-	0.08**
		(0.04)		(0.04)		(0.04)		(0.04)
Log of native labor force	-	-	0.06	0.06	-	-	-0.05	-0.04
			(0.04)	(0.04)			(0.09)	(0.09)
Kleibergen-Paap F-test	-	-	-	-	11.33	11.41	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	17.99	18.05
SW multivariate F-test (log nat.)	-	-	-	-	-	-	16.35	16.44

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by department. The dependent variable is the log monthly wage of full-time native women (Panel A) or men (Panel B). Columns 5-6 instrument the share of immigrants with two shift-share instruments constructed using the 1968 French census, giving the predicted immigrant share for the region and the predicted (log) number of immigrants in the department; columns 7- 8 instrument both the share of immigrants and the log native labor force by using the shift-share instruments and the predicted (log) size of the female native labor force at the region and department levels. All regressions include department and time fixed effects, and use the individual weight provided by INSEE. The female (male) regressions have 75,492 (108,135) observations. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

Table 11: Immigration and wages, by education group

	Less than a baccalaureate degree				Baccalaureate degree			
	OLS	IV estimates			OLS	IV estimates		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Impact on the wage of native women								
Immigrant share	-0.65***	-0.78***	-1.03***	-1.32***	0.24**	0.28	-0.11	-0.62
	(0.09)	(0.18)	(0.21)	(0.33)	(0.10)	(0.18)	(0.15)	(0.47)
Inverse Mills ratio	-	-	0.12***	0.12***	-	-	0.23***	0.24***
			(0.02)	(0.02)			(0.04)	(0.04)
Log of native labor force	-	-	-	-0.18*	-	-	-	-0.16
				(0.11)				(0.17)
Kleibergen-Paap F-test	-	27.64	29.12	-	-	59.52	60.36	-
SW multivariate F-test (imm. share)	-	-	-	18.74	-	-	-	16.24
SW multivariate F-test (log nat.)	-	-	-	20.05	-	-	-	20.75
B. Impact on the wage of native men								
Immigrant share	-1.22***	-1.45***	-1.38***	-1.07***	-0.45***	-0.48***	-0.54***	-0.40*
	(0.21)	(0.16)	(0.17)	(0.24)	(0.12)	(0.12)	(0.13)	(0.23)
Inverse Mills ratio	-	-	-0.10*	-0.11**	-	-	0.12	0.12
			(0.06)	(0.06)			(0.09)	(0.09)
Log of native labor force	-	-	-	0.00	-	-	-	0.13
				(0.12)				(0.11)
Kleibergen-Paap F-test	-	17.52	18.11	-	-	52.08	52.19	-
SW multivariate F-test (imm. share)	-	-	-	21.34	-	-	-	17.26
SW multivariate F-test (log nat.)	-	-	-	23.07	-	-	-	22.61

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The dependent variable is the log monthly wage of full-time native women (Panel A) or men (Panel B). Columns 1-4 use the sample of native workers with less than a baccalaureate degree, while columns 5-8 use the sample of native workers with a baccalaureate degree. Columns 2-3 and 6-7 instrument the share of immigrants with the shift-share instrument computed using the 1968 French census; columns 4 and 8 instrument both the share of immigrants and the log native labor force by using the shift-share instrument and the predicted (log) size of the female native labor force. All regressions include age, education, region, and time fixed effects, and use the individual weight provided by INSEE. The female (male) regressions using the sample of natives without a baccalaureate degree have 36,409 (65,847) observations; the female (male) regressions using the sample of natives with a baccalaureate degree have 34,917 (37,857) observations. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

Table 12: Immigration and wages using the skill-cell approach

	Regressions at the region-education level						Regressions at the region-education-age level					
			IV estimates						IV estimates			
	OLS estimates		Single endogenous variable model			Two-endogenous variable model	OLS estimates		Single endogenous variable model			Two-endogenous variable model
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Impact on the wage of native women												
Immigrant share	-0.17***	-0.33***	-0.10	-0.31**	-0.54***	-0.56***	-0.18**	-0.43***	-0.18*	-0.43***	-0.73***	-0.62***
	(0.06)	(0.06)	(0.11)	(0.13)	(0.12)	(0.13)	(0.08)	(0.07)	(0.10)	(0.15)	(0.17)	(0.19)
Inverse Mills ratio	-	0.23***	-	0.22***	0.29***	0.29***	-	0.27***	-	0.21***	0.22***	0.15***
		(0.02)		(0.02)	(0.02)	(0.02)		(0.02)		(0.02)	(0.02)	(0.03)
Log of native labor force	-	-	-	-	-0.05***	-0.06***	-	-	-	-	-0.04***	-0.04***
					(0.01)	(0.01)					(0.01)	(0.01)
Kleibergen-Paap F-test	-	-	11.85	12.51	9.81	-	-	-	14.50	12.33	34.24	-
SW F-test (imm. share)	-	-	-	-	-	13.92	-	-	-	-	-	15.63
SW F-test (log nat.)	-	-	-	-	-	197.43	-	-	-	-	-	394.96
B. Impact on the wage of native men												
Immigrant share	-0.57***	-0.59***	-0.71**	-0.74**	-0.80***	-0.76***	-0.58***	-0.65***	-0.84***	-0.95***	-1.04***	-0.88***
	(0.14)	(0.14)	(0.29)	(0.30)	(0.05)	(0.06)	(0.17)	(0.16)	(0.30)	(0.31)	(0.19)	(0.20)
Inverse Mills ratio	-	0.12*	-	0.10*	0.08	0.08	-	0.14***	-	0.12***	-0.01	-0.11***
		(0.06)		(0.06)	(0.05)	(0.06)		(0.05)		(0.04)	(0.04)	(0.04)
Log of native labor force	-	-	-	-	-0.07***	-0.07***	-	-	-	-	-0.07***	-0.08***
					(0.01)	(0.01)					(0.01)	(0.01)
Kleibergen-Paap F-test	-	-	12.51	12.66	7.38	-	-	-	13.06	15.00	36.68	-
SW F-test (imm. share)	-	-	-	-	-	10.55	-	-	-	-	-	18.99
SW F-test (log nat.)	-	-	-	-	-	753.55	-	-	-	-	-	820.48

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered at the region-skill level. The dependent variable is the log monthly wage of full-time native women (Panel A) or men (Panel B). Columns 1-6 run the regressions at the region-education level, while columns 7-12 run the regressions at the region-education-age level. We use two education groups (baccalaureate v. non-baccalaureate) and two age groups (18-40 v. 41-64) to build the region-skill-year cells. Columns 3-4-5 and 9-10-11 instrument the share of immigrants with two shift-share instruments computed using the 1968 French census, giving the predicted immigrant share for the region and the predicted (log) number of immigrants in the skill-region group. Columns 6 and 12 instrument both the share of immigrants and the log native labor force by using the shift-share instruments and the predicted (log) size of the female native labor force in the region and in the region-skill group. All regressions include age, time, and interacted region-skill fixed effects. Each regression uses the individual weight provided by INSEE. The regressions in Panel A (B) have 71,326 (103,704) observations. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.