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, Revised June 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, Revised June 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 should reduce the wage of comparable workers. A voluminous literature attempts to estimate the impact of such supply shocks on the wage of native workers (see Blau and Mackie, 2016, for a survey). One key insight is that natives may respond by moving to labor markets not directly affected by immigration and where presumably the wage did not drop. Some natives 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 competition (Hunt, 2017; Llull, 2018); some natives might change their occupations (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," these responses help to attenuate the negative wage impact of immigration by effectively diffusing the shock across many other markets.

This diffusion implies that difference-in-differences comparisons of wages across markets may not identify the wage impact in the market targeted by immigrants. The observed (relative) wage change in the targeted market will reflect not only the immediate wage drop after the shock, but also the attenuation of that wage effect as some of the shock gets transmitted elsewhere through the native response.

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 a supply

* 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, Jan Stuhler, Stephen Trejo, Camilo Umana Dajud, and Vincent Vicard for providing valuable comments on an earlier draft.

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 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 differently (Acemoglu, Autor and Lyle, 2004; Edo and Toubal, 2017). Moreover, female labor supply tends to be more elastic 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 foreign-born population the right to family reunification, making it far easier for wives to join their husbands.[1] A direct consequence of this policy shift was a rapid rise in the number of female immigrants. 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.[2]

[1] Family reunification was not the only factor changing the gender composition of foreign-born workers in France. As Beauchemin, Borrel, and Régnard (2013, p. 4) note, “more and more of the women who arrive in France are single or ‘pioneers’ migrating ahead of their partner.”

[2] The French context provides a unique opportunity for studying the link between gender and the impact of immigration. Most studies typically examine how immigration affects the earnings of native men or pool all

Figure 1 documents key trends in the size and gender composition of the foreign-born labor force in France. The top panel shows how the policy shift led to an immediate drop 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 fallen to 8.8 percent. This decline is entirely attributable to a 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 contrasts the French experience with that of the United States. In France, the female share of the foreign-born labor force rose from 18.7 percent in 1962 to 22.8% in 1975, and then nearly doubled to 42.4 percent by 1999. In the United States, the female share barely changed between 1970 and 2000, rising only from 39.8 to 41.1 percent over those three decades.

Our analysis is guided by a theoretical framework that isolates the three key channels through which an immigrant supply shock changes the mean wage in a labor market.[3] 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 due to the native response. Some natives may move out of the labor market targeted by immigrants, partially reversing the initial shift in the supply curve. The third is the selection bias. Because native workers are not randomly selected from the population, the composition of the native workforce may change after the supply shock, producing a spurious change in the wage. Using standard results from the selection bias literature (Heckman, 1979), we show how the generic regression model relating the wage to the size of the immigrant supply shock in repeated cross sections can be easily modified to incorporate a selection bias correction and identify the wage impact of immigration.

Our empirical study uses data from 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 between immigration and wages (across regions and over time) is negative for native men, yet immigration and

---

natives and ignore the gender composition of the labor force. Some exceptions include Cortés and Tessada (2011), Cortés and Pan (2019), Edo and Toubal (2017), Farré, González and Ortega (2011), and Llull (2021).

[3] There is an additional channel of adjustment as firms expand to take advantage of the lower price of labor. We abstract from this adjustment mechanism throughout the paper.

the male employment rate are uncorrelated. In contrast, the correlation between immigration and female wages is zero (or even weakly positive), but the correlation between immigration and female employment is strongly negative.

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 correcting for selection bias, the adjusted wage elasticity for native women is negative and roughly the same size (between -0.8 and -1.0) as that found for native men.[4]

Although our analysis is the first to delineate and 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. 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 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. The panel analysis produces a more adverse wage effect than the raw correlations between immigration and wages would suggest.

However, our theoretical framework shows that measuring the impact of immigration by tracking the wage of labor force "survivors" does not solve the selection problem. The reason is that the survivors are self-selected from the at-risk population that was initially exposed to the supply shock, and their experience does not correctly measure the wage impact that would have been observed had the workers who left the labor force

[4] Although selection bias corrections are rare in the immigration literature, an empirical exercise in Card (2001) hints at their potential importance. Card performs 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.

remained at work. In other words, the wage change observed in the subsample of survivors is contaminated by a classic case of selection bias and may not represent the wage change that would have been observed in the population at risk. In fact, our framework shows that the tracking of continuously employed workers may produce a more biased estimate of the wage impact than simply comparing mean wages across cross-sections. Regardless of whether the analysis uses cross-section or panel data, the measurement of the wage impact of immigration requires the explicit modeling of the selection bias diagnosed in this paper.

Our analysis has 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 these flows are endogenous and will generate selection biases that contaminate the observed change in the market wage after a supply shock.

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). 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, reducing 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.[5]

[5] The 1982 LFS does not report information on nationality at birth. We define a native in that survey as someone born in France. This definition implies that the native sample in 1982 excludes persons born outside France with French nationality at birth.

The annual LFS reports wages at the individual level beginning in 1982. Our empirical analysis covers the 1982-2016 period.[6] The LFS also reports each person's labor force and employment status during the reference week, and many demographic and economic characteristics (including age, gender, nationality, education, marital status, and number of children).[7] The LFS reports the worker's monthly wage net of employee payroll tax contributions.[8] Our analysis focuses on the monthly wage of full-time native workers to have a more precise measure of the "price of labor." 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.

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 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 key characteristics of our data. Perhaps the most striking trend is the 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. The data also indicate that the size of the immigrant 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

[6] 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.

[7] The definition of employment status differs between the census and the LFS. A person in the census is "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 during a reference week.

[8] Wages are reported in nominal terms, and we deflate using the Consumer Price Index produced by the INSEE. The reported monthly wage in the 1982 LFS is a categorical variable with 19 bands. We impute a monthly wage for workers in that survey 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.

5.3 to 10.0 percent while the immigrant share among the low-educated men was roughly constant (between 12 and 14 percent).

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 native labor market outcomes 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 relation between immigration and wages in France. The scatters show a weak positive correlation between immigration and female wages (the coefficient of the regression line is 0.11, with a standard error of 0.07), but a strong negative correlation between immigration and male wages (the coefficient and standard error are -0.42 and 0.16, respectively). Using a similar approach, we calculated the gender-specific adjusted employment rates for each region-year cell. These data, also illustrated in Figure 2, further document the gender asymmetry. Figures 2C and 2D show a strong negative correlation between employment and immigration for native women (the

[9] Bratsberg and Raaum (2012) also use this definition of the immigrant share. Most studies in the literature define the supply shock as either M/N or as $M/(M+N)$. Either variable approximates the measure of the supply shock implied by a labor demand model, which as we show below, is our definition of m_{rt}. Our results would be very similar if we used the approximations in the literature.

coefficient and standard error are -0.98 and 0.10), and a zero correlation for native men (the coefficient and standard error are -0.02 and 0.10).[10]

The raw data suggest an important interaction between gender and the observed 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 French women, a group with more elastic labor supply, immigration affected their opportunities by reducing the number of women employed. These correlations suggest that immigration may have had a crowd-out effect on female employment. This crowd-out would have attenuated the (initial) wage reduction produced by the supply shock. The attenuation effect would be magnified if the women who left the labor market had relatively low wages. In other words, the 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 initial wage impact of the supply shock.

3. Theoretical Framework

It is instructive to begin with a graphical representation of the link between the self-selection of the native workforce and the observed wage impact of immigration. Consider the stylized two-period model summarized by:[11]

Wage offer at t = 0:	$\log w_{i0} = \mu + \epsilon_{i0}$,	$(1a)$
Wage offer at t = 1:	$\log w_{i1} = \mu + \delta + \epsilon_{i1}$,	$(1b)$
Reservation wage:	$\log \mathcal{R}_i = \bar{\mathcal{R}} + h_i$.	$(1c)$

where w_{it} gives the wage of person *i* at time *t*; μ is the initial mean of the population wage distribution; $\mathcal{R}_i$ is the reservation wage; and $\bar{\mathcal{R}}$ is the mean (log) reservation wage. The ϵ's and h capture (unobserved) individual variation in wage offers and reservation wages.

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

[11] For expositional convenience, we use the terms labor force participation and employment interchangeably throughout this section.

The parameter δ measures the wage impact of an immigrant supply shock that occurs between the two periods. We focus on the case where immigrants and natives are substitutes, so that $\delta < 0$ (although it will be evident that the selection problem also arises when immigrants and natives are complements). By assumption, the shock only shifts the mean of the population wage distribution. To simplify further, suppose that a single (unobserved) "skills" factor, ω, determines both the wage offer and the reservation wage (i.e., $\epsilon_{i0} = \beta_w\ \omega_i$; $\epsilon_{1i} = \beta_w\ \omega_i$; and $h_i = \beta_h\ \omega_i$). This assumption implies that $Corr(\epsilon_0, h) = Corr(\epsilon_1, h) = Corr(\epsilon_0, \epsilon_1) = 1$. Figure 3 illustrates the model with $\delta < 0$. As drawn, the wage curves indicate that the returns to skills are greater in the labor market than in household production, leading to a positively selected workforce.

All persons with skills above threshold θ_0 work at t = 0 and the supply shock increases this threshold from θ_0 to θ_1. Suppose that the distribution of skills ω is uniformly distributed over the interval depicted in the figure. The mean wage of *workers* in the initial period is given by point A, the midpoint of the wage curve between θ_0 and the maximum wage. Similarly, the mean wage of workers after the supply shock is given by point B (the midpoint between θ_1 and the maximum wage). The vertical difference between A and B (which is very small) does not identify the wage impact δ (i.e., the vertical difference between the two wage curves).[12]

We can retrieve the correct wage effect by applying a Heckman selection correction to the *cross-section* data. If we estimate a selection-corrected wage equation in the initial cross-section, we identify the mean wage in the population at t = 0 (or point A_S, the midpoint of the wage curve). Similarly, a selection-corrected wage equation in the second cross-section estimates B_S, the mean wage in the population after the shock. The vertical difference between A_S and B_S identifies δ. The uncorrected cross-section wage growth ($A - B$) underestimates the true impact ($A_S - B_S$) because the workforce is positively selected, and the least skilled natives exited the labor market after the immigrants arrived.

In short, the change in the average wage earned by native workers depends crucially on *how many* and *which* native workers choose to respond to the shock. The exit of low-

[12] The exit of some natives from the labor force implies that part of the wage drop observed immediately after the supply shock is attenuated. The parameter δ then measures the "net" impact of the shock. The attenuation effect is discussed in greater detail below.

skill natives from the labor force artificially increases the market's average wage because of composition effects, so that the typical repeated cross-sections comparison could end up suggesting that immigration had little impact or even increased wages. The exit of high-skill workers reduces the observed average wage and might end up suggesting a large adverse wage effect. The non-random selection of the native workforce *and* the fact that immigration influences the participation decision imply that we cannot use the wage change observed between cross-sections (i.e., the classic identification strategy in the literature) to infer how supply shocks shifted the mean of the wage distribution.

We now generalize the model to document both the source of the wage impact δ and to determine exactly which data comparisons, including the potential availability of panel data, correctly identify the impact. A labor demand framework that assumes a Cobb-Douglas aggregate production function with labor and capital as inputs implies that the *market* wage w_{kt} in market k and period t (t = 0, 1) is:

$$\log w_{kt} = \varphi_{kt} + \eta \log L_{kt}\,, \tag{2}$$

where φ_{kt} is a parameter specific to cell (k, t); L_{kt} gives the total number of workers employed in the market; and η is the wage elasticity.

This market has received immigrant supply shocks in the past, and the workforce has N_{k0} natives and M_{k0} immigrants in the initial period. We assume the two groups are perfect substitutes. It is convenient to rewrite equation (2) as:

$$\log w_{k0} = \varphi_{k0} + \eta \log(M_{k0} + N_{k0}) = \varphi_{k0} + \eta\, m_{k0} + \eta \log N_{k0}\,, \tag{3}$$

where the immigrant share $m_{kt} = \log(1 + M_{kt}/\, N_{kt})$ approximates the fraction of the workforce that is foreign-born (as a fraction of native workers in the same period).

A new influx of immigrants enters the market, increasing the total number of foreign workers to M_{k1}. We again assume that immigrants and natives are perfect substitutes. We also assume that immigrant labor supply is perfectly inelastic. Native labor supply, however, might respond to the supply shock, so that the total number of native workers changes to N_{k1}. The native response consists of movements in and out of the labor force (and not migration to labor market k'). The market wage (*after* natives have responded) is:

$$\log w_{k1} = \varphi_{k1} + \eta \log(M_{k1} + N_{k1}) = \varphi_{k1} + \eta\, m_{k1} + \eta \log N_{k1}\,. \tag{4}$$

The change in the market wage is given by:

$$\Delta \log w_k = \varphi_k + \eta\,(m_{k1} - m_{k0}) + \eta \log \frac{N_{k1}}{N_{k0}} = \varphi_k + \eta\, \Delta m_k + \eta\, \Delta \log N_k\,, \tag{5}$$

where $\varphi_k = \varphi_{k1} - \varphi_{k0}$; $\Delta m_k = m_{k1} - m_{k0}$; and $\Delta \log N_k = \log N_{k1} - \log N_{k0}$. Equation (5) shows that the wage change in market k is determined by both the size of the supply shock and 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 supply shock Δm_k would be mostly offset by a corresponding decline in the number of native workers. Following Borjas and Monras (2017), it is convenient to write the change in the number of native workers as a function of the supply shock:

$$\Delta \log N_k = \gamma\, \Delta m_k, \tag{6}$$

where γ ($-1 \leq \gamma \leq 0$) captures the crowd-out effect, approximating the number of native workers who leave the labor market for every immigrant who enters. By substituting equation (6) into (5), we obtain a type of "reduced-form" equation:[13]

$$\Delta \log w_k = \varphi_k + \eta(1+\gamma)\, \Delta m_k. \tag{7}$$

[13] Of course, the repercussions of the attenuation effect noted in equation (5) would continue over time. The wage increase induced by the exit of $\Delta \log N_k$ natives immediately after the supply shock encourages some natives to enter the labor force in the next period. The market wage observed τ periods after the initial (one-time) shock can then be written as:

$$\Delta \log w_{k\tau} = \varphi_k + \eta\, \Delta m_k + \eta\, \Delta \log N_{k1} + \eta\, \Delta \log N_{k2} + \cdots + \eta\, \Delta \log N_{k\tau}.$$

Suppose that the native supply response for periods $t \geq 2$ can be modeled as in equation (6), so that $\Delta \log N_{kt} = \gamma\, \Delta \log N_{kt-1}$. It then follows that:

$$\Delta \log w_{k\tau} = \varphi_k + \eta(1 + \gamma + \gamma^2 + \cdots + \gamma^\tau)\, \Delta m_k \approx \varphi_k + \frac{\eta}{1-\gamma} \Delta m_k.$$

Accounting for all feedback effects still leads to a "reduced-form" relating the wage change in market k to the change in the immigrant share, with the total attenuation effect measured by $1/(1-\gamma)$. We abstract from these details to simplify the presentation and to focus on the crucial role of the selection issue.

The reduced form in (7) is the typical regression model estimated in the immigration literature. It produces an estimate of the "reduced-form wage elasticity" $\eta(1+\gamma)$, the elasticity that incorporates the native supply shift. The crowd-out response has typically been examined in the context of internal migration: If natives are mobile, the detrimental impact of immigration on local wages is diffused throughout the national economy and the (relative) short-run wage drop in market k may not be detectable.

At any point in time, there will be some within-market wage dispersion because persons in market k (though they share characteristics that help define the market, such as location or education) also exhibit some differences. Some natives have higher quality schooling, or differ in their drive or motivation, or have a racial or ethnic background that is favored or penalized by employers. The wage offer made by firms to potential native workers then depends not only on market conditions, but also allows for individual variation because of differences in (unobserved) characteristics captured by ϵ_{it}:

$$\log w_{ikt} = \varphi_{kt} + \eta\, m_{kt} + \eta \log N_{kt} + \epsilon_{it}, \tag{8}$$

where $\epsilon_{it} \backsim N(0, \sigma_\epsilon^2)$. We assume supply shocks only shift the mean of the population wage distribution, implying that the variance σ_ϵ^2 is constant over time.

The distribution of the reservation wage in the native population is:

$$\log \mathcal{R}_{ik} = \bar{\mathcal{R}}_k + h_i, \tag{9}$$

where $\mathcal{R}_{ik}$ gives the reservation wage for worker i in market k; $\bar{\mathcal{R}}_k$ gives the mean (log) reservation wage; and $h_i \backsim N(0, \sigma_h^2)$. We assume the distribution of reservation wages is constant over time, and that supply shocks do not change the correlation between the reservation wage and the wage offer, so that $\rho_h = Corr(h_i, \epsilon_{it})\ \forall\ t$.

At each point in time, the labor force participation decision is based on a comparison of the reservation wage to the wage offer made by firms in the aftermath of the supply shock between periods t – 1 and t, but prior to any native supply response. For example, the labor force participation decision at t = 1 is based on the wage:

$$\log w'_{ik1} = \varphi_{k1} + \eta \log(M_{k1} + N_{k0}) + \epsilon_{i1} = \varphi_{k1} + \eta\, m'_k + \eta \log N_{k0} + \epsilon_{i1}\,, \tag{10}$$

where $m'_{k1} = \log(1 + M_{k1}/\ N_{k0})$. The wage offer w'_{ik1} defined by equation (10) reflects the initial wage adjustment observed as firms moved down the short-run labor demand curve in market *k* after the number of immigrant workers increased to M_{k1}. Put differently, it defines the counterfactual wage that would have been observed after the entry of the new immigrants in a hypothetical scenario without a native supply response. The selection rule that determines if native *i* in market *k* works in year *t* is then given by:

$$\log w_{ikt} \text{ is observed if } Z^*_{ikt} = \log w'_{ikt} - \log \mathcal{R}_{ik} > 0, \tag{11}$$

where Z^*_{ikt} is the continuous latent variable that generates the sample of workers at time *t*. This latent variable can be written as:

$$Z^*_{ikt} = \varphi_{kt} - \bar{\mathcal{R}}_k + \eta\, m'_{kt} + \eta \log\ N_{kt-1} + \epsilon_{it} - h_i = C_{kt} + v_{it}\,, \tag{12}$$

where $C_{kt} = \varphi_{kt} - \bar{\mathcal{R}}_k + \eta\, m'_{kt} + \eta \log\ N_{kt-1}$; $v_{it} = \epsilon_{it} - h_i$; and $v_{it} \sim N(0, \sigma_v^2)$. The assumptions that the distribution of the reservation wage is stable and that the supply shock only changes the mean of the wage distribution implies that the variance σ_v^2 is time-invariant. The labor force participation rate of natives in cell (k, t) is $\pi_{kt} = 1 - \Phi(\alpha_{kt})$, where $\alpha_{kt} = -C_{kt}/\sigma_v$, and Φ denotes the standard normal distribution function.

Let I_{ikt} represent the event that person *i* in market *k* at time *t* is employed (i.e., $Z^*_{ikt} > 0$). Using standard results from the sample selection literature (Gronau, 1974; Heckman, 1979), the average wage observed in market *k* at time *t* is:

$$\begin{aligned} \mathrm{E}(\log w_{ikt} \mid I_{ikt}) &= \varphi_{kt} + \eta\, m_{kt} + \eta \log N_{kt} + E(\epsilon_{it} \mid I_{ikt}) \\ &= \varphi_{kt} + \eta\, m_{kt} + \eta \log N_{kt} + \sigma_\epsilon\, \rho_{v\epsilon}\, \lambda(\pi_{kt}), \end{aligned} \tag{13}$$

where $\rho_{v\epsilon} = Corr(v_t, \epsilon_t)$ and is time-invariant because $\sigma_\epsilon^2, \sigma_h^2$, and ρ_h are constant and the inverse Mills ratio $\lambda(\pi_{kt}) = \phi(\alpha_{kt})/[1 - \Phi(\alpha_{kt})]$, with ϕ denoting the standard normal density. The sign of $\rho_{v\epsilon}$ reveals which subsample of the native population is employed: $\rho_{v\epsilon}$ is positive if the natives who work tend to be high-wage persons and is negative if they tend to be low-wage persons. The observed change in the mean native wage in market *k* across cross-sections is:

$$\Delta \log w_k|_{CS} = E[\log w_{ik1} \mid I_{ik1}) - E[\log w_{ik0} \mid I_{ik0})$$
$$= \varphi_k + \eta\, \Delta m_k + \eta\, \Delta \log N_k + \sigma_\epsilon\, \rho_{v\epsilon}\, [\lambda(\pi_{k1}) - \lambda(\pi_{k0})]. \quad (14)$$

Equation (14) shows that both the *size* of the native response and the *skill composition* of the response determine the observed wage impact. Specifically, immigration has three distinct effects on the wage change in market k. The first is the direct short-run effect of the shock Δm_k, 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 $\Delta \log N_k$, generates its own attenuating 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).

Finally, the third term gives the selection bias resulting from the fact that native workers are not randomly selected. The inverse Mills ratio is a monotonically decreasing function of the labor force participation rate π_k (Heckman, 1979, p. 156). As a result, the difference $[\lambda(\pi_{k1}) - \lambda(\pi_{k0})]$ is positive if the supply shock lowers the native labor force participation rate. The direction of the selection bias is then determined by the sign of $\rho_{v\epsilon}$, which is positive if the workforce is positively selected. 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.

Following Heckman (1979), the selection bias can be viewed as a specification error in the repeated cross-sections wage regression commonly estimated in the immigration literature. Substitute equation (6) into (14) to obtain the reduced form:

$$\Delta \log w_k|_{CS} = \varphi_k + \eta(1+\gamma)\, \Delta m_k + \rho_{v\epsilon}\, \sigma_v\, \Delta\lambda(\pi_k). \quad (15)$$

Suppose we have aggregate data on wages and immigration across many markets and estimate the variant of equation (15) that excludes the selection term $\Delta\lambda(\pi_k)$.[14] The omitted variable formula implies:

[14] The discussion abstracts from the endogeneity of the supply shock. Immigrants are not randomly distributed across markets and likely prefer to settle in markets that offer thriving economic opportunities.

$$\text{plim}\, \eta\widehat{(1+\gamma)} = \eta(1+\gamma) + \rho_{\upsilon\epsilon}\, \sigma_\upsilon\, \beta_{\lambda m}, \tag{16}$$

where $\beta_{\lambda m}$ is the coefficient from a regression of the excluded inverse Mills ratio variable ($\Delta\lambda$) on the supply shock Δm_k. The coefficient $\beta_{\lambda m}$ is positive if the supply shock reduced the employment rate of the native population. The failure to account for the selection bias produced by the non-random selection of native workers biases the estimate of the wage impact of immigration if $\rho_{\upsilon\epsilon} \neq 0$. The bias is positive if $\rho_{\upsilon\epsilon} > 0$ and negative if $\rho_{\upsilon\epsilon} < 0$. In other words, the wage elasticity estimated by correlating observed wage changes and supply shocks across markets is "too positive" if the workforce is positively selected and "too negative" if the workforce is negatively selected. Of course, this bias can be entirely avoided by including the inverse Mills ratio in the cross-section earnings regressions. As in Figure 3, the difference in the selectivity-corrected mean earnings across cross-sections produces an unbiased estimate of the wage impact of immigration.

A few recent papers track the panel of persons who worked both before and after a supply shock to identify the wage impact of immigration. This tracking tends to produce more adverse wage effects than those implied by comparing mean wages across cross-sections. This finding suggests that the panel approach, by holding constant the composition of the workforce, perhaps addresses the identification problem introduced by the self-selection of workers.

Moreover, the special case of the model in Figure 3 indeed shows that the panel strategy identifies the wage impact *without* using any selection correction. The observed wage in the panel of workers who are continuously employed is given by the midpoint of the wage curves after the threshold θ_1 in each cross-section, or points A_P and B. The difference $(A_P - B)$ exactly equals δ. In short, a selection correction is not needed if both wage offers and reservation wages are determined by a single skill factor ω and if immigration only changes the mean of the wage distribution. We can then identify the wage impact by tracking earnings in the subsample of continuously employed workers.

Even in the absence of selection bias, this endogeneity biases the estimate of the reduced-form elasticity $\eta(1+\gamma)$. We discuss this problem in detail below.

It is important to note, however, that this property of panel data does not generalize. In fact, depending on the joint distribution of the unobservables, the use of panel data could produce an even larger bias than comparing (uncorrected) means across cross-sections. We illustrate this result by returning to the stylized model in equations ($1a$) - ($1c$), where immigration shifts the mean of the wage distribution by δ_k, but now assume $\epsilon_{it} \backsim N(0, \sigma_\epsilon^2)$ and $h_i \backsim N(0, \sigma_h^2)$. A person works at t = 0 if $v_{i0} = (\epsilon_{i0} - h_i) > (\bar{\mathcal{R}}_k - \mu_k)$, and works at t = 1 if $v_{i1} = (\epsilon_{i1} - h_i) > (\bar{\mathcal{R}}_k - \mu_k - \delta_k)$.[15] Let $\rho_{01} = Corr(\epsilon_{i0}, \epsilon_{i1})$; $\rho_{v_0 v_1} = Corr(v_{i0}, v_{i1})$; and v_{it}^* be the standard normal transformation of v_{it}. The Appendix shows that the average wage growth observed in the panel of persons continuously employed in market k is:

$$\Delta \log w_k|_P = E[\log w_{ik1} \mid I_{ik0} \cap I_{ik1}] - E[\log w_{ik0} \mid I_{ik0} \cap I_{ik1}]$$

$$= \delta_k + \left[\frac{1 - \rho_{01}}{1 - \rho_{v_0 v_1}}\right] \frac{\sigma_\epsilon^2}{\sigma_v} \{E[v_{i1}^* \mid I_{ik0} \cap I_{ik1}] - E[v_{i0}^* \mid I_{ik0} \cap I_{ik1}]\}. \tag{17}$$

The conditional expectations in (17) are defined by:

$$E[v_{it}^* \mid I_{ik0} \cap I_{ik1}] = \frac{\int_{\bar{\mathcal{R}}_k - \mu_k}^{\infty} \int_{\bar{\mathcal{R}}_k - \mu_k - \delta_k}^{\infty} v_t^* \, f(v_0, v_1) dv_1 dv_0}{\int_{\bar{\mathcal{R}}_k - \mu_k}^{\infty} \int_{\bar{\mathcal{R}}_k - \mu_k - \delta_k}^{\infty} f(v_0, v_1) dv_1 dv_0}, \tag{18}$$

where $f(v_0, v_1)$ is the bivariate normal distribution. As in Figure 3, the panel wage growth in equation (17) identifies δ_k if earnings are perfectly correlated over time ($\rho_{01} = 1$). In general, however, $\Delta \log w_k|_P$ is a biased estimate of δ_k. The wage growth in the subsample of workers who worked continuously does not represent what would have been observed in the population.

There is no simple expression for the bias in (17) because the expectations involve integrals of the bivariate normal. Nevertheless, we show in the Appendix that a sufficient condition for the bias to be positive is $\pi_{k1} \geq 0.5$ (i.e., a majority of natives work even after

---

[15] These thresholds simplify the model of the participation decision in equation (11), where the reservation wage is compared to w'_{ikt}, the wage offer that would have been observed after the supply shock in a scenario without a native supply response. We can adopt the approach in (11) and reframe the analysis of the stylized model so that δ'_k would be the corresponding wage effect that determines labor force participation. This generalization complicates the presentation without changing any of the implications.

the supply shock). The uncorrected (for selection bias) wage growth in a panel, just like the uncorrected wage growth between cross-sections, will then produce estimates of the wage impact that understate the adverse effect.

The bias is easy to quantify in one special case, allowing for a direct comparison of the cross-section and panel estimators. Suppose $\rho_{v_0 v_1} = 0$, so that employment outcomes are uncorrelated over time. The panel wage growth collapses to:[16]

$$\Delta \widetilde{\log w}_k|_P = \delta_k + (1 - \rho_{01}) \frac{\sigma_\epsilon^2}{\sigma_v} [\lambda(\pi_{k1}) - \lambda(\pi_{k0})]. \tag{19}$$

The participation rate drops after the supply shock if $\delta_k < 0$, implying $\lambda(\pi_{k1}) > \lambda(\pi_{k0})$. Equation (19) trivially shows that the average wage growth observed among continuously employed workers imparts a positive bias on the estimate of δ_k if $\rho_{01} \neq 1$. Remarkably, despite the intuitive appeal of the conjecture that a panel would lead to more accurate results, the bias produced by the panel might exceed the cross-section bias. The observed wage growth between cross-sections in the stylized model is:[17]

$$\Delta \log w_k|_{CS} = \delta_k + \frac{\sigma_\epsilon \, \sigma_h}{\sigma_v} \left[\frac{\sigma_\epsilon}{\sigma_h} - \rho_h\right] [\lambda(\pi_{k1}) - \lambda(\pi_{k0})]. \tag{20}$$

The excess bias of the panel approach (in the special case of employment independence) is:

$$\Delta \widetilde{\log w}_k|_P - \Delta \log w_k|_{CS} = -\frac{\sigma_\epsilon \, \sigma_h}{\sigma_v} \left[\frac{\sigma_\epsilon}{\sigma_h} \rho_{01} - \rho_h\right] [\lambda(\pi_{k1}) - \lambda(\pi_{k0})]. \tag{21}$$

Equations (20) and (21) imply that workers will be positively selected *and* the excess bias will be positive if:

$$\frac{\sigma_\epsilon}{\sigma_h} > \rho_h > \rho_{01} \frac{\sigma_\epsilon}{\sigma_h}. \tag{22}$$

---

[16] Equation (19) follows from (17) because employment independence implies $E[v_{it}^* | \, I_{ik0} \cap I_{ik1}] = E[v_{it}^* | \, I_{ikt}] = \lambda(\pi_{kt})$; see the Appendix for details.

[17] The cross-section wage growth in equation (20) is identical to that derived in equation (15). The wage impact of immigration $\delta_k = \eta(1+\gamma)\,\Delta m_k$, and $\rho_{v\epsilon} = \frac{\sigma_h}{\sigma_v}\left[\frac{\sigma_\epsilon}{\sigma_h} - \rho_h\right]$.

In other words, the panel wage growth will impart a larger positive bias on the observed wage effect if the correlation between market and reservation wages is "sufficiently large" or if the serial correlation in earnings is "sufficiently small." Of course, it would be incorrect to infer that the panel approach typically produces a more biased estimate of δ_k than the comparison of cross-sections. What the analysis does indicate, however, is that there are regions of the bivariate normal distribution where such an outcome is possible.

We have shown that the self-selection of the native workforce biases standard estimates of the wage impact of immigration if two conditions hold: (a) native workers are not randomly selected from the population; and (b) supply shocks influence the labor force participation decision of natives. The identification of the wage impact of immigration will then require either a *selectivity-corrected* analysis of the mean wage of workers across repeated cross-sections, or a *selectivity-corrected* analysis of the wage growth observed in a panel of persons who worked continuously through the sample period.

The cross-section approach, which we pursue below, has two advantages.[18] First, panel data suitable for analysis in the immigration context are rarely available (and are very limited for France). Second, the selection correction required to analyze repeated cross-sections is far simpler than the correction required by the panel approach. The cross-section correction is a straightforward application of the Heckman two-step procedure, applied to each cross-section to retrieve the population mean wage in a particular market at a point in time. The correction required to purge the panel of selection bias is far more complex because the complement of the sample of continuously employed workers contains three distinct groups with three different truncations: persons who worked before the supply shock, but not after; persons who did not work before the shock, but worked after; and persons who never worked.

[18] Heckman and Robb (1985, p. 240) write: "Although longitudinal data are widely regarded in the social science and statistical communities as a panacea for selection and simultaneity problems, there is no need to use longitudinal data to identify the impact of training on earnings if conventional specifications of earnings functions are adopted. Estimators based on repeated cross-section data for unrelated persons identify the same parameter." Our analysis echoes the Heckman-Robb conclusion, replacing the word "training" by "immigration." We have also shown that a longitudinal analysis that ignores the selection problem can potentially produce a larger bias than a cross-section analysis that ignores the selection problem.

4. Econometric Framework

4.1. The Wage Equation

We estimate the selection-adjusted wage impact of immigration by turning to individual-level data in a pooled sample of cross-sections and applying the Heckman selection correction. The individual-level earnings function is:

$$\log w_{irt} = \theta_a + \theta_e + \alpha P_{it} + \theta_{rt} + \varphi \lambda_{it} + \mu_{it}, \tag{23a}$$

where $\log w_{irt}$ gives the log monthly wage of native worker i in region r at time t; θ_a and θ_e are vectors of age and education fixed effects, respectively; P_{it} is a vector of personal characteristics (discussed below); θ_{rt} is a vector of interacted region-time fixed effects; and λ_{it} is the inverse Mills ratio calculated from a first-stage probit on the probability that the individual is employed.[19] The specification of the probit regression will be discussed shortly. Crucially, the self-selection correction implies that the fixed effects θ_{rt} measure (age- and education-adjusted) mean earnings in the *population* of cell (r, t). We estimate equation ($23a$) separately in the samples of working men and women.

We have shown that the wage impact of immigration can be identified by examining how supply shocks change the mean of the population earnings distribution across repeated cross-sections. The estimates of θ_{rt} from equation (23a) then allow us to run the following regression using cell-level data:

$$\theta_{rt} = \theta_r + \theta_t + \beta_1 m_{rt} + \beta_2 \log N_{rt} + \xi_{it}, \tag{23b}$$

where θ_r and θ_t are vectors of region and time fixed effects, respectively. We measure the immigrant supply shock as $m_{rt} = \log{(1 + M_{rt}/N_{rt})}$. The coefficient β_1 in equation ($23b$) measures the wage elasticity—the downward movement along the short-run labor demand curve after immigrants enter the local labor market. The elasticity β_1 is estimated from within-region changes in the (selection-corrected) wage and immigrant shocks. The immediate wage drop that presumably follows the supply shock might encourage some

[19] 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).

natives to withdraw from the labor force, and the regression also includes the (log) number of native workers N_{rt} to adjust for this reverse shift of the supply curve in cell (r, t).[20]

Following Dustmann, Schönberg, and Stuhler (2016, 2017) and Jaeger, Ruist, and Stuhler (2018), we initially define the immigrant share m_{rt} at the region-year level (instead of assigning workers to different skill groups and calculating a supply shock specific to a region-skill-year cell). This strategy 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 specific supply shock on the wage of competing workers. It also captures the complementary effects 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 need to pre-assign workers to specific skill groups, avoiding the potential mismeasurement of the supply shock in a skill cell because of the possibility that employers might downgrade the skills that immigrants offer to the labor market (Dustmann, Frattini, and Preston, 2013).

The three key variables in the regression model in equations $(23a)$ and $(23b)$ need to be estimated (the inverse Mills ratio λ_{it}) or are endogenous (m_{rt} and $\log N_{rt}$). We now turn to a discussion of the first stage probit depicting an individual's labor force participation decision and of the instruments used to correct for the endogeneity.

4.2. The Inverse Mills Ratio

The regression model in equation $(23a)$ includes the inverse Mills ratio to adjust for the selection bias induced by the endogenous labor force participation decision of natives. 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 wage equation:

---

[20] By combining equations $(23a)$ and $(23b)$, the two-stage model collapses into a single regression:

$$\log w_{irt} = \theta_a + \theta_e + \theta_r + \theta_t + \alpha_P P_{it} + \beta_1 m_{rt} + \beta_2 \log N_{rt} + \beta_3 \lambda_{it} + \mu'_{it}.$$

The estimates of the coefficients β_1 and β_2 are nearly identical regardless of whether we estimate the model in one pass of the data or as the two-stage process described in the text. We prefer the two-equation framework as it more clearly shows the source of the identification of the wage impact and makes the analysis comparable to the cell-level studies that dominate the immigration literature.

$$\mathrm{P}(EMP_{irt} = 1) = \Phi(\theta_a + \theta_e + \alpha_P P_{it} + \alpha_Z Z_{it} + \theta_{rt} + v_{it}), \tag{24}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution. Note that the probit equation does not include any immigration-related variables, but instead includes the vector of region-time fixed effects θ_{rt} (which subsumes all potential measures of the shock and native supply response). 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. We estimate the probit model in equation (24) separately for men and women.

Because there are gender differences in the determinants of labor supply and wages, we use slightly different baseline specifications of the wage and probit regressions in equations (23a) and (24) for the two groups. Our approach for the analysis of female 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.[21] 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 a rough 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.[22] As long as leisure is a normal good, higher levels of household wealth increase the reservation wage and would have a negative effect on the probability of participating in the labor force.

---

[21] 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 the "unmarried" group.

[22] The fraction of persons who own their home without a mortgage 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 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.

In short, the regression specification for the joint study of female employment and wages can be summarized as follows: The individual-level wage regression will include vectors of age, education, and region-time fixed effects. The probit regression includes all these variables *plus* the family characteristics (marital status and the presence of young children) and household wealth. 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 Rubinstein, 2008). This assumption, however, may not be applicable in France (or other European contexts), where the unemployment rate among prime-age men is high, and 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 individual-level 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.

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

4.3. Endogeneity of the Immigrant Supply Shock

It is well known that estimating the cell-level regression model in $(23b)$ using OLS produces inconsistent estimates of the wage impact of immigration because of the non-random sorting of immigrants across regions (i.e., income-maximizing immigrants are more likely to settle in regions that offer the best job opportunities). To address this issue, we use 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.[23] 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:

$$\widehat{M}_r(t) = \sum_n \sum_e \frac{M_r^{ne}(1968)}{M^{ne}(1968)} \cdot M^{ne}(t), \tag{25}$$

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

$$\widehat{N}_r(t) = \sum_e \frac{N_r^e(1968)}{N^e(1968)} \cdot N^{\mathrm{e}}(t). \tag{26}$$

The shift-share instrument is then defined by:

[23] 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.

$$\widehat{m}_{rt} = \log\left(1 + \frac{\widehat{M}_r(t)}{\widehat{N}_r(t)}\right). \tag{27}$$

Despite its widespread use, it is important to emphasize that the shift-share instrument in (27) 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).[24]

4.4. 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 $\eta(1 + \gamma)$.

The size of the native labor force is endogenous to local economic conditions. Our instrument combines 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 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 other countries, the presence of young children deters female labor supply in France (Piketty, 1998; Gurgand and Margolis, 2008). Let $\psi_r(t)$ 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.[25] Our instrument for the (log) size of the native workforce is given by:

[24] As noted by Jaeger, Ruist and Stuhler (2018), an important criterion required 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 French immigration patterns changed substantially after 1968.

[25] The variable $\psi_r(t)$ 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).

$$\log \hat{F}_r(t) = \log\left[\psi_r(t) \cdot \hat{\mathcal{N}}_r(t)\right], \tag{28}$$

where $\hat{\mathcal{N}}_r(t)$ is an adjusted measure of the shift-share prediction $\hat{N}_r(t)$ 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 $\hat{N}_r(t)$ in equation (26) only took into account the geographic allocation of natives at the time of the 1968 cross-section, and ignored 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 or economic and political shocks in source countries, future projections of the native population are more dependent on pre-existing trends.

To construct the instrument in equation (28), we adjust the shift-share projection $\hat{N}_r(t)$ for the long-term regional differences in population growth rates. 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:

$$\hat{\mathcal{N}}_r(t) = \hat{N}_r(t)(1 + \Delta g_r)^{t-1968}. \tag{29}$$

The adjusted projection $\hat{\mathcal{N}}_r(t)$ equals the "cross-section" projection $\hat{N}_r(t)$ if the geographic allocation of natives is constant prior to the sample period (i.e., $\Delta g_r = 0$).[26]

The exclusion of the $\log N_{rt}$ variable from the typical regression model 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 instrument in (28) relies on the same types of assumptions used to justify the validity of shift-share instruments. Specifically, both the geographic allocation of natives at a point in time and

[26] 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 $\hat{N}_r(t)$ 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 (29) produces a prediction of 4.5 million.

the pre-existing trends in this allocation are assumed to be independent of current wages.[27] The validity of our instrument also hinges on the assumption that "shocks" in the presence of young children affect female labor supply decisions but do not affect wages. Although this is a common assumption (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.

Because of the absence of compelling exogenous shocks in native labor supply, our empirical analysis will report estimates of the wage impact of immigration both excluding and including the native labor supply variable. The evidence will demonstrate that the key insight of our framework—i.e., that selection biases matter when estimating the wage impact—is valid even when we restrict our attention to the reduced form wage elasticity identified by the generic equation in the literature.

5. Empirical Results

5.1. First-Stage IV Estimates

Table 2 presents the first stage of our baseline IV wage regressions for both native women and native men. Our simplest regression specification relates the wage to the immigrant share. Panel A of the table presents the first-stage regression associated with this 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)), region, and time fixed effects.

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.

---

[27] Ideally, we would use information on the geographic sorting of natives and the change in that sorting 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. To ensure compatibility with existing studies, we ignored the adjustment in equation (29) when we constructed the instrument for the immigrant share. Our estimates of the wage impact of immigration are not sensitive to this additional correction.

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 ($\log N_{rt}$). The instruments are the predicted population of immigrants (i.e., $\widehat{M}_r(t)$ defined in equation (25)) and the predicted female native labor force (i.e., $\hat{F}_r(t)$ defined in equation (28)).[28] All regressions again include region and time fixed effects.

The results indicate that the immigrant share is positively correlated with $\widehat{M}_r(t)$ and negatively correlated with $\hat{F}_r(t)$. The positive correlation is in line with the literature on the immigrant shift-share instrument, while the negative correlation probably arises because a rise in the predicted number of working women would mechanically reduce the ratio of immigrant to native workers. There is also a very strong positive correlation between the predicted number of working women and the size of the native 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 12.9 and 16.5, indicating that our instruments are reasonably strong.

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. Columns 1 and 3 estimate equation (24) separately for native women and native men. These probits are used to compute the inverse Mills ratio included in the individual-level wage regressions discussed below.

The table reports the estimated coefficients on the variables that adjust for differences in reservation wages, such as marital status, presence of young children, and home ownership. We find that marriage lowers the probability of employment for women (by 1 percentage point), and increases it for men (by 10 percentage points). The presence of young children in the household also predicts employment, and the sign of the correlation again differs between men and women. In particular, the presence of young

---

[28] 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 $\hat{F}_r(t)$ are both functions of the shift-share prediction of the native population. In fact, using $\widehat{m}_{rt}$ and $\hat{F}_r(t)$ as instruments leads to weaker first-stage estimates and less significant estimated coefficients in the second-stage IV regressions.

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 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.

For illustrative purposes, columns 2 and 4 estimate the direct impact of the immigrant share on the native employment probability using an IV strategy. As suggested by our modeling of the participation decision in equation (12), we measure the immigrant supply shock by $m'_{rt} = \log\left(1 + M_{rt}/N_{rt-1}\right)$ and include in the probit regression the log predicted native population in the prior census (as measured by the shift share projection $\widehat{N}_{rt-1}$).[29] We instrument m'_{rt} by using $\log\left(1 + \widehat{M}_{rt}/\widehat{N}_{rt-1}\right)$, where $\widehat{M}_{rt}$ and $\widehat{N}_{rt-1}$ are the corresponding shift share predictions.

The regression results in columns 2 and 4 reveal a negative correlation between the immigrant share and the probability of employment for native women, while the same correlation is 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 resembles the descriptive evidence in Section 2. Table 3 also shows that the lagged size of the predicted native population tends to be negatively correlated with the employment probability of natives.

5.3. The Wage of Native Workers

We used the probit regressions reported in columns 1 and 3 of Table 3 to calculate the inverse Mills ratio for each person, and then estimated (separately by gender) the individual-level earnings regressions in equation (23a).[30] This exercise produces the selectivity-corrected mean wage of the population in cell (r, t). The cell means become the dependent variable in equation (23b) that examines how immigration affects the wage

[29] In particular, equation (12) 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.

[30] The individual-level regressions in the female (male) sample have 71,326 (103,704) observations.

distribution. Table 4 reports the (OLS and IV) estimated impact of the immigrant supply shock on the adjusted log wage of native women (Panel A) and men (Panel B) at the regional level between 1982 and 2016.

The cell-level regressions are weighted by cell size (i.e., the sum of the individual weights in the cell), and we cluster the standard errors at the region level to account for the possibility of within-group error correlation. Because the number of regions may be too small to estimate the correct cluster-robust standard errors, we implement the wild cluster bootstrap method (Cameron, Gelbach, and Miller, 2008, p. 427) using 1,000 replications and report the corresponding *p*-values.[31] We will show that the evidence is robust when we estimate the impact of immigration (a) at the departmental level (using 94 French departments, instead of 22 regions); and (b) at the region-education-age level (using two education groups and two age groups, and exploiting variation across 88 clusters).

Consider initially the results in Table 4 for native women. Column 1 presents the simplest specification, where the (adjusted) mean wage in the cell is calculated from an individual-level regression that does not correct for selection bias and the mean wage in the cell is related only to the immigrant share (plus region and year fixed effects). The OLS coefficient of the immigrant share is insignificant and numerically close to zero, reproducing the descriptive evidence in Figure 2 (which did not adjust for individual differences in education and age).

Column 2 adjusts for the selection bias created by the non-random selection of the workforce (i.e., the dependent variables is the region-time fixed effect from an individual-level wage equation that includes the inverse Mills ratio as a regressor). As shown in Table 4, the estimated coefficient of the inverse Mills ratio from the individual-level wage regression is strongly positive, suggesting that female workers are positively selected from the female population.[32] The mean value of the inverse Mills ratio for women is 0.48, so that the self-selection of female workers increases the mean of the observed wage

---

[31] Cameron, Gelbach and Miller (2008) show 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 analysis of the wage impact of immigration.

[32] Positive selection of women into employment is also found 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.

distribution by about 10.1 percent (or the product of the coefficient of the inverse Mills ratio and its mean) relative to the population mean.

Note that the OLS estimate of the coefficient of the immigrant share becomes significantly negative, with a value of -0.44 (0.08). The change in the impact of immigration between columns 1 and 2 is predicted by our theoretical framework if the women who exit the labor force in the post-migration period have relatively low wages. In other words, ignoring the positive selection of the sample of working women produces an estimate of the wage impact of immigration that is positively biased.[33]

Columns 5 and 6 present the analogous IV regressions when the immigrant share is instrumented using the 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 negative and significant (with a value of -0.43, and a standard error of 0.10) when the regression adjusts for selection.

The remaining columns of Table 4 expand the basic model. Columns 3 and 7 do not adjust for selection but add the variable measuring the (log) size of the native labor force. As noted earlier, although 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 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).[34] It is 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 wage effect suggests the

---

[33] The results reported in Table 4 would be nearly identical if we estimated the model in one pass of the data (as discussed in footnote 20). In the specification in column 2, for example, the one-pass approach replaces the interacted region-year fixed effects with the immigrant share variable in both the probit and earnings regressions. The coefficient of the immigrant share is -0.42 (0.08), and the coefficient of the inverse Mills ratio is 0.21 (0.02). This similarity extends to all other columns of the table.

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

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 increases to -0.95 (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.

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 -0.95, while the corresponding estimate of the reduced-form elasticity $\eta(1+\gamma)$ in column 6 is -0.43. The implied estimate of the crowd-out parameter γ is 0.5, so that about half of the initial impact of the supply shock is eroded by the native employment response. This result is consistent with other estimates. 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 reports the regressions using the sample of native men. There are several interesting gender differences and one important similarity. First, the individual-level estimates from equation (23b) suggest weaker selection for men. The estimated coefficient of the inverse Mills ratio for men is insignificant and much smaller than the corresponding coefficient for women (0.05 as compared to 0.21). 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 increases the mean of the observed wage distribution for men by only about 1.7 percent, as compared to 10.1 percent for women.

Second, the coefficient of the native labor supply variable is positive, but close to zero and insignificant. As we noted earlier, the instrument for the supply variable log N_{rt} (based on shift-share projections of the female native population and the presence of small children in the household) may not fully resolve the endogeneity of male labor supply. The regression coefficient may also be reflecting factors specific to the French context, where the growth of the native workforce in recent decades was driven by the rise in labor force participation among native women. If men and women are not perfect substitutes, the increased number of native workers need not lead to lower wages for native men.

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

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 value 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 0.49).[35]

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 4.2, our baseline model employed slightly different specifications of the selection model by gender. The probit regressions included the family variables and household wealth for both men and women. However, the individual-level 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 estimated wage elasticity when using alternative specifications of the selection model. The baseline specification reproduces the estimates from Table 4. Specification 2 includes the family variables in the female individual-level wage regression,

---

[35] Although many area studies find negligible wage effects from immigration (Blau and Mackie, 2016; Edo, 2019), our estimates resemble those reported in several recent studies. Ortega and Verdugo (2016) estimate a wage elasticity between -0.2 and -1.0 in France and Jaeger, Ruist, and Stuhler (2018) report a short-run elasticity between -0.9 and -1.6 in the United States. Several studies of massive and unexpected supply shocks document sizable adverse wage effects. The short-run wage elasticities reported in Borjas (2017), Edo (2020), and Monras (2020) are between [-0.5; -1.5], [-1.0; -2.0] and [-0.7; -1.4], respectively.

but excludes them from the male individual-level wage regression. In specifications 3 and 4, we only use the family variables (specification 3) or the home ownership indicator (specification 4) to generate independent variation in the inverse Mills ratio. Finally, the last row in each panel follows Olsen (1980) by assuming that the error term in the selection equation is uniformly distributed, thereby replacing the inverse Mills ratio in equation (23a) with the predicted employment probability calculated from a linear probability model. This specification helps show that the impact of our sample selection correction is not driven by the nonlinearity of the probit model.

While the OLS and IV estimates reported in columns 1 and 4 of Panel A are virtually zero for women, adjusting for selection in columns 2 and 5 *always* results in a negative and significant elasticity—regardless of the specification of the selection model. This pattern documents the bias in the estimated wage impact of immigration produced by the self-selection of working women. Columns 3 and 6 add the size of the native labor force to the regression and again 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 five specifications). Panel B shows that the estimated wage elasticities for men are stable across specifications and columns.[36] In sum, the robustness of the estimated elasticity suggests that changes in how the model corrects for sample selection does not change the estimated wage impact of immigration.[37]

Tables 6 to 10 provide additional sets of robustness tests. These 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.

---

[36] Table 4 reports the estimated wage elasticities if we simply asserted that the selection problem is not relevant for the study of male earnings. The no-selection parameter estimates are represented by the regressions in 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.

[37] 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. This coefficient turns negative when 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 the male employment probability *and* on male earnings. Excluding the family vector from the male wage regression imparts a negative bias on the coefficient of the inverse Mills ratio (which is negatively correlated with the probability of employment).

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 analysis 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 reduce the potential correlation between the shift-share instrument (based on the 1968 census) and current labor market outcomes.

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 (0.12) in column 3 when the regression adjusts for sample selection, and more than doubles to -1.07 (0.30) in column 4 when the regression holds constant the size of the native labor force. In contrast, 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 used cross-sections that were spaced apart in roughly equal intervals, and skipped over the 2012 data. Columns 5-8 of Table 6 reproduce the regressions using all the available census data since 1982, expanding the study to six separate cross-sections. It is evident that including the additional 2012 cross-section barely affects our results.

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. Columns 1-4 of Table 7 use the alternative measure given by log $(1 + M_{rt}/ N_{rt-1})$. In other words, we use the size of the native labor force in the prior census as the base that defines the immigrant share.[38] This alternative measure addresses the 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 use gender-specific immigrant shares to measure the supply shock, only using women to compute the immigrant share in Panel A and men in Panel B.[39] All the coefficients are similar to the baseline results in Table 4. The estimated effect of immigration on the female wage is insignificant in the simplest model (columns 1-2 and 5-6), and 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 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 weight the cell-level regressions. The last four columns expand the individual-level wage equation (23a) and probit equation (24) 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.

Each of the specifications confirms that the selection-corrected wage impact of immigration for women is larger than the corresponding uncorrected estimate. For example, adjusting for selection in the female wage equation changes the IV wage elasticity from 0.01 to -0.58 in the unweighted regression model and from -0.35 to -0.56 in the full interaction model.[40] In contrast, within each of the two alternative specifications, the estimated wage impact is relatively stable for native men (the wage elasticity is about -0.7 in columns 2-4 and ranges between -0.9 and -1.2 in columns 6-8).

The dependent variable in the baseline probit specification in Table 3 indicated if a native person was employed and we then examined earnings in the subsample of full-time

[38] 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.

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

[40] The loss of precision in the unweighted estimates as compared to Table 4 is consistent with the fact that weighted least squares estimation corrects for heteroskedastic error terms and thereby achieves more precisely estimated coefficients than unweighted estimation.

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 indicates if the person is employed full-time (with the alternative outcome including both those not employed and those employed part-time). This alternative approach does not change any of our baseline results. The wage elasticity for men is negative and between -0.8 and -0.9, 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.[41] The individual-level hourly wage regressions are then estimated using the entire sample of both full- and part-time workers. As with our baseline estimates, the selectivity-corrected estimates in the female wage regression leads to a far more negative wage elasticity; it doubles from -0.47 (0.09) to -0.93 (0.10) 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.[42]

Table 10 performs a final robustness check by using an alternative definition of a labor market. Instead of defining the market in terms of the 22 regions in European France, we use the geographically smaller definition of a department (of which there are 94).[43] This sampling framework significantly increases the number of cells and introduces much more variation in immigration and wages into the analysis.[44]

[41] 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). The reported weekly hours likely contains substantial measurement errors, which may affect the estimated wage elasticities (Barrett and Hamermesh, 2019; Laroque and Salanié, 2002; and Ortega and Verdugo, 2016).

[42] The individual-level hourly wage regressions used to predict the selectivity-corrected hourly wage in the cell has 98,451 (108,198) observations in the female (male) sample. 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.

[43] 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.

[44] The information on a person's department of residence is not available in the LFS between 2002 and 2012. Our department-level analysis 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. The 2013 and 2016 LFS

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 department d at time t (constructed along the lines implied by equation (27)) *and* the predicted number of immigrants in a given region as defined in equation (25). This extension of the shift-share approach helps capture potential network effects 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 defined in equation (28)) and the analogously constructed log predicted female native labor force at the departmental level.

Regardless of the specification, the results 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 -0.94 (0.21) for women and -1.02 (0.25) for men.

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

do not report any natives living in the Lozère department, so we exclude it from the analysis. Lozère is the smallest department in France, containing only 0.12 percent of the native population in 2016.

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.[45]

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 education group helps measure 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 impact of immigration on the low education group. Correcting for sample selection increases the estimated elasticity for this skill group from -0.78 (0.18) in column 2 to -1.03 (0.20) in column 3. The inclusion of the native labor supply variable in column 4 increases the negative wage response even more. 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 low-skill workers.

We conclude our empirical exploration by changing the unit of analysis from the region-year cell to a region-skill-year cell. Specifically, we divide each regional market into four skill groups. We use the two education groups introduced above (those who have the Baccalaureate degree v. those who do not) and two age groups (18-40 years old v. 41-64 years old). The key difference between this empirical strategy and the baseline specification is that we will now measure the mean wage, the immigrant share, and the size of the native labor force at the region-skill-year level rather than at the region-year level.

We first estimate the mean wage for the region-skill cell from the individual-level regression estimated separately by gender:

[45] The share of immigrants in the low (high) educated segment of the labor force increased from 10.8 percent (4.1 percent) in 1982 to 13.9 percent (9.5 percent) in 2016.

$$\log w_{irst} = \alpha\, P_{it} + \theta_{rst} + \varphi\, \lambda_{it} + \mu_{it}, \tag{30a}$$

where $\log w_{irst}$ gives the log monthly wage of native worker i, in region r, skill group s, at time t; P_{it} is a vector of personal characteristics; θ_{rst} is a vector of fully interacted region-skill-time fixed effects; and λ_{it} is the inverse Mills ratio for each native worker calculated from a first-stage probit regression on the probability of employment. The regressors in the probit regression (also estimated separately by gender) include marital status, the presence of young children in the household, the home ownership variable, and the vector of region-skill-time fixed effects θ_{rst}.

The specification of the regression model at the cell level is:

$$\theta_{rst} = \theta_r + \theta_s + (\theta_r \times \theta_s) + \theta_t + \beta_1\, m_{rst} + \beta_2 \log N_{rst} + \nu_{srt}, \tag{30b}$$

where the dependent variable is the mean (adjusted) wage of natives in cell (r, s, t), and is estimated from equation $(30a)$.

Note that equation $(30b)$ 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. This estimation strategy implies that the wage impact of immigration is identified from changes that occur within region-skill groups over time.

Table 12 reproduces the structure of our baseline Table 4 by showing the OLS and IV regression coefficients from equation $(30b)$. We address the endogeneity of the immigrant share at the region-skill level by exploiting the same strategy introduced in Section 4.3, thereby instrumenting the immigrant share m_{rst} by using the corresponding shift-share prediction in a given region-skill group at time t.[46] In columns 7-8, we again account for the endogeneity of the log native labor force by using the log predicted female native labor force at the region-skill 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

[46] In columns 5-6, our instrument for m_{rst} is $\widehat{m}_{rst} = \log\left(1 + \widehat{M}_{rs}(t)/\widehat{N}_{rs}(t)\right)$. We predict $\widehat{M}_{rs}(t)$ and $\widehat{N}_{rs}(t)$ by multiplying the 1968 distribution of immigrants (natives) across region-skill cells for each country group n by the total number of immigrants (natives) from that group in subsequent years. In columns 7-8, our instrument for m_{rst} is $\log\left(\widehat{M}_{rs}(t)\right)$.

makes the OLS and IV estimates of the impact of immigration on female wages more negative. The IV wage elasticity jumps from -0.04 (0.10) to -0.38 (0.15). 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 from the individual-level wage regression is always significantly positive for women, and weaker and insignificant for men. Finally, holding native labor supply constant produces a more negative wage elasticity for women, suggesting a crowd-out effect at the region-education-age level.

Relative to the baseline estimates in Table 4, the wage elasticities reported in Table 12 are somewhat smaller for women and somewhat larger for men. The intuition behind the different approaches (i.e., the unit of analysis being the region-year cell or the region-skill-year cell) suggests that the skill-cell approach is more likely to isolate the "own" effect of immigration and may miss the complementary cross-cell effects. As a result, the estimated wage elasticity would be expected to be more negative when using the region-skill-year breakdown.

However, 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, perhaps resulting in attenuated estimates of the wage elasticity.

Further, if immigrants are placed in jobs that require less education than they have, assignment to their nominal education groups may produce an inaccurate measure of the supply shock in a particular skill group (Dustmann, Frattini and Preston, 2013). In the same vein, immigrants may not necessarily compete with natives in the same age group, especially if firms value the prior work experience of immigrants and natives differently. The measurement error might generate additional biases in estimating the wage effect of immigration using a skill-cell approach.

Nevertheless, the lessons provided by exploiting information on supply shocks within specific skill cells confirm our key hypothesis: The measurement of the wage impact of immigration requires an analysis that pays careful attention to the self-selection of the native workforce and to the labor supply response induced by the supply shock.

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 implication of the textbook supply-demand model, the evidence is mixed, and there is still disagreement on even the direction of the wage impact of immigration 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. The wage change observed in a market targeted by immigrants depends 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 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 workforce.

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 self-selection of the native workforce, and the native response to the immigrant supply shock, contaminates estimates of the key parameters of the labor demand function. Our empirical application of this framework 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, 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).

It is important to emphasize that the selection bias identified and explored 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.

Appendix: Selection Bias and Identifying the Wage Impact of Immigration

Consider the stylized model:

Wage offer at $t = 0$:	$\log w_0 = \mu + \epsilon_0,$
Wage offer at $t = 1$:	$\log w_1 = \mu + \delta + \epsilon_1,$
Reservation wage:	$\log \mathcal{R} = \bar{\mathcal{R}} + h,$

where $\epsilon_t \sim N(0, \sigma_t^2)$ and $h \sim N(0, \sigma_h^2)$. We omit subscripts denoting market and individual variation. Immigration affects the wage structure by shifting the mean μ of the wage distribution by δ. A person works if the wage offer exceeds the reservation wage. The decision to work in each period is given by:

$$I_0: \quad v_0 = \epsilon_0 - h > \bar{\mathcal{R}} - \mu = \kappa.$$
$$I_1: \quad v_1 = \epsilon_1 - h > \bar{\mathcal{R}} - \mu - \delta = \kappa - \delta.$$

Let π_t be the participation rate at time t. It follows that $\pi_0 = 1 - \Phi(\kappa/\sigma_{v_0})$ and $\pi_1 = 1 - \Phi[(\kappa - \delta)/\sigma_{v_1}]$. Henceforth, we simplify notation by interpreting (when necessary) the variables κ and δ as being in their standardized form.

A1. Cross-Section Wage Growth

The average wage change observed among workers across cross-sections is:

$$\Delta_{CS} = E[\log w_1 \mid I_1] - E[\log w_0 \mid I_0] = \delta + E[\epsilon_1 \mid I_1] - E[\epsilon_0 \mid I_0].$$

Using standard results from the selection literature, we can write:

$$E[\epsilon_t \mid I_t] = \sigma_t \, \rho_{tv_t} \, \lambda(\pi_t),$$

where $\rho_{tv_t} = Corr(\epsilon_t, v_t)$; and $\lambda(z)$ is the inverse Mills ratio evaluated at z. The cross-section wage growth is:

$$\Delta_{CS} = \delta + \sigma_1 \, \rho_{1v_1} \, \lambda(\pi_1) - \sigma_0 \, \rho_{0v_0} \, \lambda(\pi_0).$$

The correlation ρ_{tv_t} is given by:

$$\rho_{tv_t} = Corr(\epsilon_t, v_t) = \frac{\sigma_h}{\sigma_{v_t}} \left(\frac{\sigma_t}{\sigma_h} - \rho_{ht} \right),$$

where $\rho_{ht} = Corr(h, \epsilon_t)$. The assumption that immigration only changes the mean of the wage distribution implies $\sigma_0^2 = \sigma_1^2 = \sigma_\epsilon^2$ and $\rho_{h0} = \rho_{h1} = \rho_h$. It follows that $\sigma_{v_0}^2 = \sigma_{v_1}^2 = \sigma_v^2$ and $\sigma_0 \, \rho_{0v_0} = \sigma_1 \, \rho_{1v_1} = \sigma_\epsilon \, \rho_{\epsilon v}$. The cross-section wage growth can be written as:

$$\Delta_{CS} = \delta + \frac{\sigma_\epsilon \sigma_h}{\sigma_v}\left(\frac{\sigma_\epsilon}{\sigma_h} - \rho_h\right)[\lambda(\pi_1) - \lambda(\pi_0)]. \qquad (A1)$$

A2. Panel Wage Growth

To derive the analogous equation for the wage growth in a sample of persons continuously employed, we use a well-known property of conditional expectations for a multivariate normal. Consider a vector of random variables $\mathbf{X}$, where $\mathbf{X}$ has dimension $n \times 1$, $\mathbf{X} \sim N_n(\mu, \Sigma)$, and partition the vector $\mathbf{X}$ as:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{\mathrm{A}} \\ \mathbf{X}_{\mathrm{B}} \end{bmatrix},$$

where $\mathbf{X}_{\mathrm{A}}$ has dimension $p \times 1$ ($p < n$). It then follows that:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_{\mathrm{A}} \\ \mu_{\mathrm{B}} \end{bmatrix}, \qquad \text{and} \qquad \boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_{\mathrm{AA}} & \Sigma_{\mathrm{AB}} \\ \Sigma_{\mathrm{BA}} & \Sigma_{\mathrm{BB}} \end{bmatrix}.$$

The conditional distribution of $\mathbf{X}_{\mathrm{A}}|\mathbf{X}_{\mathrm{B}} \sim N_p(\mu_{\mathrm{A|B}}, \Sigma_{\mathrm{A|B}})$, where:

$$\mu_{\mathrm{A|B}} = \mu_{\mathrm{A}} + \Sigma_{\mathrm{AB}}\,\Sigma_{\mathrm{BB}}^{-1}\,(\mathbf{X}_{\mathrm{B}} - \mu_{\mathrm{B}}), \qquad (A2)$$
$$\Sigma_{\mathrm{A|B}} = \Sigma_{\mathrm{AA}} - \Sigma_{\mathrm{AB}}\,\Sigma_{\mathrm{BB}}^{-1}\,\Sigma_{\mathrm{BA}}.$$

The panel wage growth is defined by:

$$\Delta_P = E[\log w_1 \mid I_0 \cap I_1] - E(\log w_0 \mid I_0 \cap I_1] = \delta + E[\epsilon_1 \mid I_0 \cap I_1] - E[\epsilon_0 \mid I_0 \cap I_1].$$

Transform $(\epsilon_0, \epsilon_1, v_0, v_1)$ into standard normal random variables $(\epsilon_0^*, \epsilon_1^*, v_0^*, v_1^*)$. We can write the panel wage growth as:

$$\Delta_P = \delta + \sigma_\epsilon\, E[\epsilon_1^* \mid I_0 \cap I_1] - \sigma_\epsilon\, E[\epsilon_0^* \mid I_0 \cap I_1]. \qquad (A3)$$

To examine the properties of Δ_P, we use equation $(A2)$ and define:

$$\mathbf{X}_{\mathrm{A}} = \begin{bmatrix} \epsilon_0^* \\ \epsilon_1^* \end{bmatrix}, \qquad \text{and} \qquad \mathbf{X}_{\mathrm{B}} = \begin{bmatrix} v_0^* \\ v_1^* \end{bmatrix}.$$

Because the random variables are standard normal it follows that $\mu_A = \mu_B = 0$, and:

$$\Sigma_{\mathrm{AB}} = \begin{bmatrix} \rho_{0v_0} & \rho_{0v_1} \\ \rho_{1v_0} & \rho_{1v_1} \end{bmatrix}, \quad \Sigma_{\mathrm{BB}} = \begin{bmatrix} 1 & \rho_{v_0v_1} \\ \rho_{v_1v_0} & 1 \end{bmatrix}, \quad \Sigma_{\mathrm{BB}}^{-1} = \frac{1}{1 - \rho_{v_0v_1}^2}\begin{bmatrix} 1 & -\rho_{v_0v_1} \\ -\rho_{v_1v_0} & 1 \end{bmatrix},$$

where $\rho_{v_0v_1} = Corr(v_0, v_1)$. We can then write:

$$E[\mathbf{X}_{\mathrm{A}}|\mathbf{X}_{\mathrm{B}}] = \frac{1}{1-\rho_{v_0v_1}^2}\begin{bmatrix}\rho_{0v_0} - \rho_{0v_1}\,\rho_{v_0v_1} & \rho_{0v_1} - \rho_{0v_0}\rho_{v_0v_1} \\ \rho_{1v_0} - \rho_{1v_1}\,\rho_{v_0v_1} & \rho_{1v_1} - \rho_{1v_0}\,\rho_{v_0v_1}\end{bmatrix}\begin{bmatrix}v_0^* \\ v_1^*\end{bmatrix}.$$

Note that:

$$\rho_{1v_1} - \rho_{0v_1} = \rho_{0v_0} - \rho_{1v_0} = \frac{\sigma_\epsilon}{\sigma_v}(1-\rho_{01}),$$

where $\rho_{01} = Corr(\epsilon_0, \epsilon_1)$. The panel wage growth in $(A3)$ can be rewritten as:

$$\Delta_P = \delta + \left(\frac{1-\rho_{01}}{1-\rho_{v_0v_1}}\right)\frac{\sigma_\epsilon^2}{\sigma_v}\{E[v_1^* \mid I_0 \cap I_1] - E[v_0^* \mid I_0 \cap I_1]\}, \qquad (A4)$$

A3. Bias in Panel Wage Growth

It may be possible to sign the bias in equation (A4) by using the closed-form solutions for the moments of the truncated bivariate normal (Rosenbaum, 1961; and Muthén, 1990). In the context of our model, the conditional expectations in $(A4)$ are:

$$\Pi \cdot E[v_0^*| I_0 \cap I_1] = \phi(\kappa)\left\{1 - \Phi\left[\frac{(1-\rho_{v_0v_1})\kappa - \delta}{\left(1-\rho_{v_0v_1}^2\right)^{0.5}}\right]\right\} + \rho_{v_0v_1}\phi(\kappa-\delta)\left\{1 - \Phi\left[\frac{\kappa - \rho_{v_0v_1}(\kappa-\delta)}{\left(1-\rho_{v_0v_1}^2\right)^{0.5}}\right]\right\}$$

$$\Pi \cdot E[v_1^*| I_0 \cap I_1] = \phi(\kappa-\delta)\left\{1 - \Phi\left[\frac{\kappa - \rho_{v_0v_1}(\kappa-\delta)}{\left(1-\rho_{v_0v_1}^2\right)^{0.5}}\right]\right\} + \rho_{v_0v_1}\phi(\kappa)\left\{1 - \Phi\left[\frac{(1-\rho_{v_0v_1})\kappa - \delta}{\left(1-\rho_{v_0v_1}^2\right)^{0.5}}\right]\right\},$$

where $\Pi = \Pr(I_0 \cap I_1)$. Substituting into (A4) and combining terms yields:

$$\Delta_P = \delta + (1-\rho_{01})\cdot\frac{\sigma_\epsilon^2}{\sigma_v}\cdot\frac{\phi(\kappa-\delta)}{\Pi}\cdot\left\{1 - \Phi\left[\frac{(1-\rho_{v_0v_1})\kappa + \rho_{v_0v_1}\delta}{\left(1-\rho_{v_0v_1}^2\right)^{0.5}}\right]\right\}$$

$$-(1-\rho_{01})\cdot\frac{\sigma_\epsilon^2}{\sigma_v}\cdot\frac{\phi(\kappa)}{\Pi}\cdot\left\{1 - \Phi\left[\frac{(1-\rho_{v_0v_1})\kappa - \delta}{\left(1-\rho_{v_0v_1}^2\right)^{0.5}}\right]\right\}. \qquad (A5)$$

The assumption that $\delta < 0$ implies:

$$\Phi\left[\frac{(1-\rho_{v_0v_1})\kappa + \rho_{v_0v_1}\delta}{\left(1-\rho_{v_0v_1}^2\right)^{0.5}}\right] < \Phi\left[\frac{(1-\rho_{v_0v_1})\kappa - \delta}{\left(1-\rho_{v_0v_1}^2\right)^{0.5}}\right].$$

A sufficient condition for the bias in $(A5)$ to be positive is $\phi(\kappa-\delta) > \phi(\kappa)$. This restriction holds if the participation rate at $t = 1$ exceeds 0.5.

Analogous to equation (A1), equation (A5) shows the Heckman-type selection correction required for a panel regression to consistently estimate δ. The correction is far more complex than the one required in an analysis of repeated cross-sections.

A4. Panel Wage Growth with Independent Employment Outcomes

Suppose $\rho_{v_0 v_1} = 0$, so that employment outcomes I_0 and I_1 are independent. The conditional expectations in $(A4)$ can then be written as:

$$E(v_1^* | I_0 \cap I_1) = \frac{\int_\kappa^\infty \int_{\kappa-\delta}^\infty v_1^* f(v_0, v_1) dv_1 dv_0}{\int_\kappa^\infty \int_{\kappa-\delta}^\infty f(v_0, v_1) dv_1 dv_0} = \frac{\int_{\kappa-\delta}^\infty v_1^* \phi(v_1) \left[\int_\kappa^\infty \phi(v_0) dv_0\right] dv_1}{\int_{\kappa-\delta}^\infty \phi(v_1) dv_1 \int_\kappa^\infty \phi(v_0) dv_0} = E(v_1^* | I_1) = \lambda(\pi_1).$$

$$E(v_0^* | I_0 \cap I_1) = \frac{\int_\kappa^\infty \int_{\kappa-\delta}^\infty v_0^* f(v_0, v_1) dv_1 dv_0}{\int_\kappa^\infty \int_{\kappa-\delta}^\infty f(v_0, v_1) dv_1 dv_0} = \frac{\int_\kappa^\infty v_0^* \phi(v_0) \left[\int_{\kappa-\delta}^\infty \phi(v_1) dv_1\right] dv_0}{\int_{\kappa-\delta}^\infty \phi(v_1) dv_1 \int_\kappa^\infty \phi(v_0) dv_0} = E(v_0^* | I_0) = \lambda(\pi_0).$$

Substituting these expressions into equation $(A4)$ yields:

$$\tilde{\Delta}_P = \delta + (1 - \rho_{01}) \frac{\sigma_\epsilon^2}{\sigma_v} [\lambda(\pi_1) - \lambda(\pi_0)]. \qquad (A6)$$

Equation $(A6)$ also follows directly from $(A5)$ by setting $\rho_{v_0 v_1} = 0$, and using the implied property that $\Pi = [1 - \Phi(\kappa)] \cdot [1 - \Phi(\kappa - \delta)]$. Suppose $\delta < 0$, so that immigration reduces the native participation rate. The panel wage growth $\tilde{\Delta}_P$ in equation $(A6)$ understates the adverse wage impact of immigration if $\rho_{01} \neq 1$.

We can also show that the positive bias produced by the panel wage growth in this special case could exceed the positive bias from cross-section comparisons in equation $(A1)$. The difference in bias between the panel and cross-section estimators is given by:

$$\tilde{\Delta}_P - \Delta_{CS} = -\frac{\sigma_\epsilon \, \sigma_h}{\sigma_v} \left(\rho_{01} \frac{\sigma_\epsilon}{\sigma_h} - \rho_h\right) [\lambda(\pi_1) - \lambda(\pi_0)].$$

The panel estimator produces a larger positive bias if $\rho_{01}(\sigma_\epsilon / \sigma_h) < \rho_h$.

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. 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.

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.

Dolado, J. J., García-Peñalosa, C., & Tarasonis, L. 2020. The Changing Nature of Gender Selection into Employment over the Great Recession. Economic Policy.

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. 2017. Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment. Quarterly Journal of Economics, 132(1), 435-483.

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.

Gardner, J. 2020. Immigration Displaces Women. Working Paper, University of Mississippi.

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(6), 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.

Heckman, J. & Robb, R. (1985). Alternative Methods for Evaluating the Impact of Interventions: An Overview, Journal of Econometrics 30(1-2), 239-267.

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.

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 15(1), 174-203.

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 15(1), 204-35.

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

Muthén, Bengt. 1990. Moments of the Censored and Truncated Bivariate Normal Distribution, British Journal of Mathematical and Statistical Psychology 43(1), 131-143.

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.

Olsen, R.J. 1980. A least squares correction for selectivity bias. Econometrica 48(7), 1815-1820.

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.

Rosenbaum, S. 1961. Moments of a Truncated Bivariate Normal Distribution. Journal of the Royal Statistical Society, Series B (Methodological) 23(2), 405-408.

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.

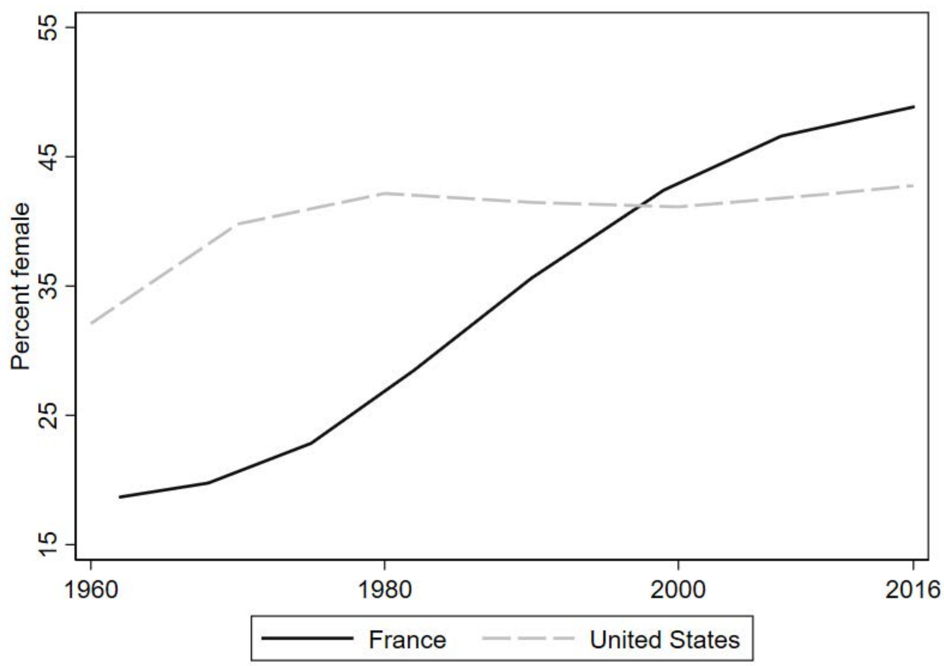
Stock, J., Yogo, M. & Wright, J. 2002. A Survey of Weak Instruments and Weak Identification in a Generalized Method of Moments,” Journal of Business and Economic Statistics 20(4), 518-529.

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

Figure 1. Immigration and gender

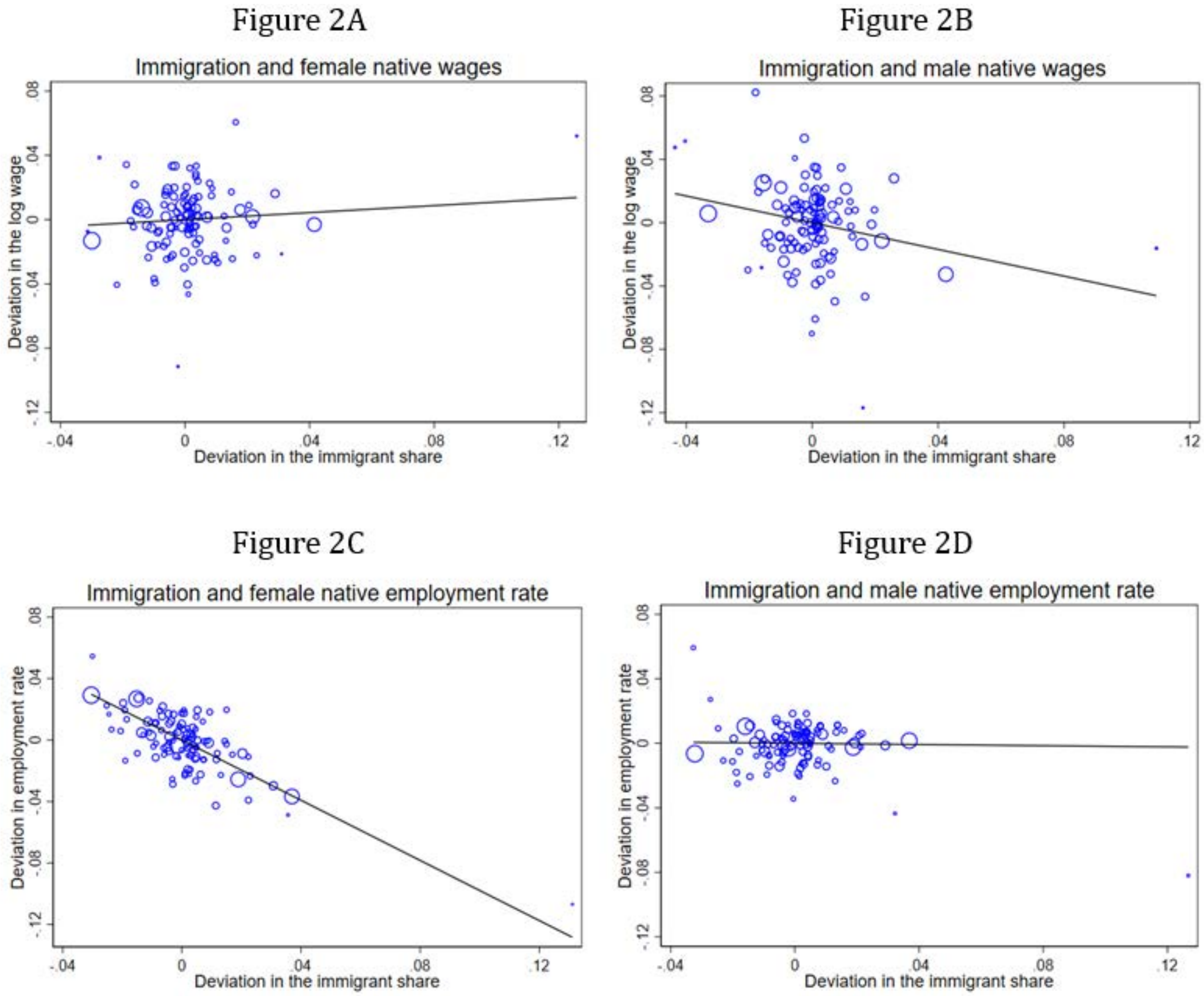
A. Trends in the immigrant share in the French labor force

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.

Figure 3. Identifying the wage impact of immigration

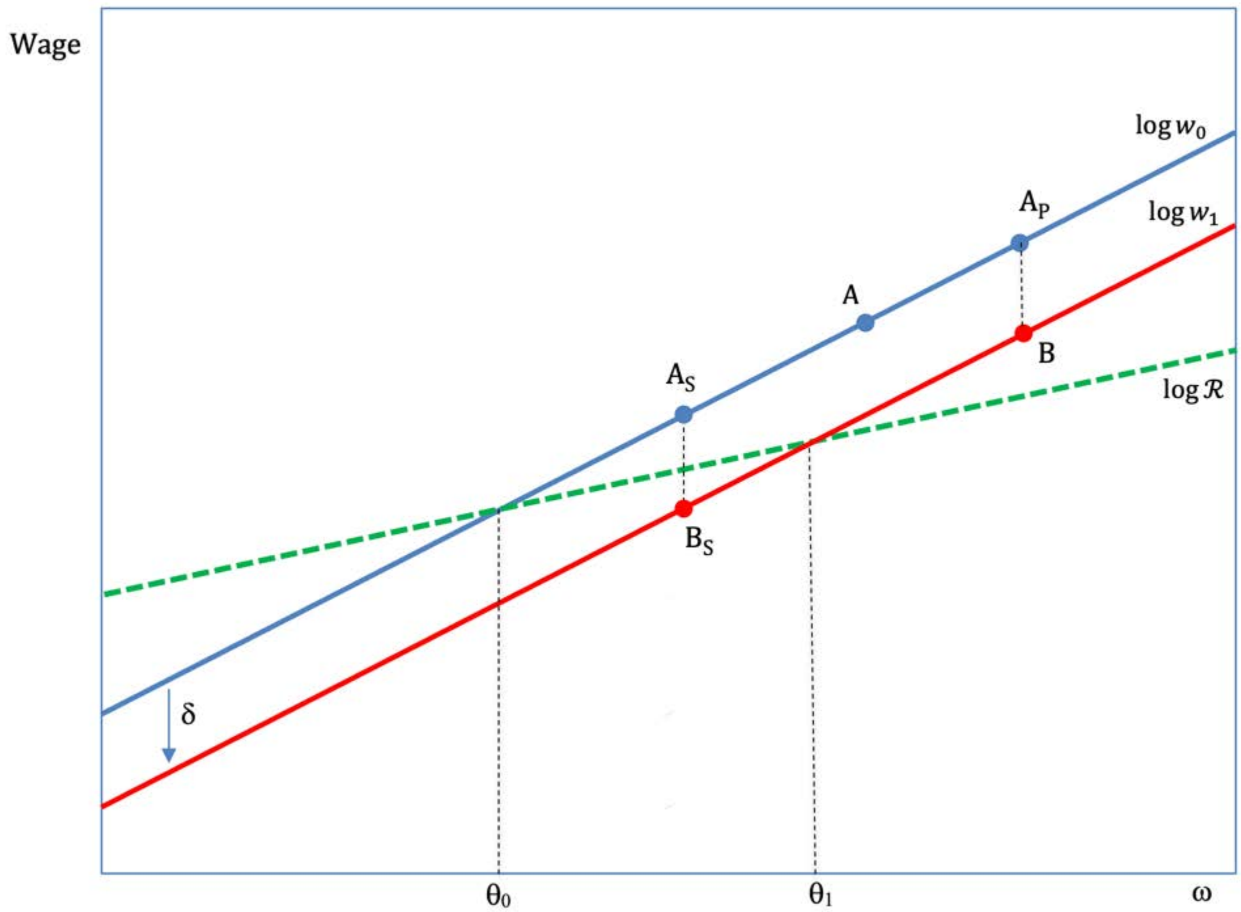


Table 1: Descriptive statistics

	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				
Predicted immigrant share in population	1.77***	-	1.71***	-
	(0.31)		(0.37)	
Kleibergen-Paap F-test of excluded instrument	32.06	-	21.00	-
	B. Two endogenous variables model			
Dependent variable: Immigrant share				
Log predicted immigrant population	-	0.12***	-	0.11***
		(0.02)		(0.03)
Log predicted female native labor force	-	-0.14***	-	-0.13***
		(0.04)		(0.04)
SW multivariate F-test of excluded instruments	-	12.91	-	14.09
Dependent variable: Log of native labor force				
Log predicted immigrant population	-	-0.09	-	-0.09
		(0.08)		(0.08)
Log predicted female native labor force	-	0.58***	-	0.58***
		(0.09)		(0.09)
SW multivariate F-test of excluded instruments	-	15.15	-	16.53

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 region and time fixed effects, and are weighted by cell size (i.e., the sum of individual weights in a cell). ***, **, * 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)
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	-0.11***	-0.12***	-0.09***	-0.09***
	(0.02)	(0.02)	(0.02)	(0.02)
Marginal effect	*-0.04*	*-0.04*	*-0.02*	*-0.02*
Immigrant share	-	-3.65***	-	0.21
		(1.06)		(0.95)
Marginal effect		*-1.16*		*0.05*
Log predicted native population in t-1	-	-0.19	-	-0.69***
		(0.26)		(0.24)
Marginal effect		*-0.06*		*-0.18*
Region f.e. x Time f.e.	Yes	-	Yes	-
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. 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 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	-0.44***	-0.00	-0.35**	-0.01	-0.43***	-0.61**	-0.95***
	(0.07)	(0.08)	(0.13)	(0.14)	(0.11)	(0.10)	(0.30)	(0.30)
Wild cluster bootstrap p-value	*0.71*	*0.02*	*0.98*	*0.03*	*0.95*	*0.05*	*0.13*	*0.00*
Log of native labor force	-	-	0.01	0.05	-	-	-0.25***	-0.21**
			(0.07)	(0.07)			(0.10)	(0.09)
Wild cluster bootstrap p-value			*0.85*	*0.43*			*0.11*	*0.10*
Selectivity-corrected estimates	-	Yes	-	Yes	-	Yes	-	Yes
Inverse Mills ratio	-	0.21***	-	0.21***	-	0.21***	-	0.21***
		(0.02)		(0.02)		(0.02)		(0.02)
Kleibergen-Paap F-test	-	-	-	-	32.06	32.06	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	12.91	12.91
SW multivariate F-test (log nat.)	-	-	-	-	-	-	15.15	15.15
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)	(0.13)	(0.17)	(0.17)	(0.09)	(0.09)	(0.18)	(0.18)
Wild cluster bootstrap p-value	*0.18*	*0.20*	*0.02*	*0.02*	*0.00*	*0.00*	*0.01*	*0.01*
Log of native labor force	-	-	0.08	0.09	-	-	0.02	0.04
			(0.06)	(0.06)			(0.07)	(0.07)
Wild cluster bootstrap p-value			*0.17*	*0.13*			*0.77*	*0.57*
Selectivity-corrected estimates	-	Yes	-	Yes	-	Yes	-	Yes
Inverse Mills ratio	-	0.05	-	0.05	-	0.05	-	0.05
		(0.07)		(0.07)		(0.07)		(0.07)
Kleibergen-Paap F-test	-	-	-	-	21.00	21.00	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	14.09	14.09
SW multivariate F-test (log nat.)	-	-	-	-	-	-	16.53	16.53

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The unit of observation is a region-year cell over the 1982-2016 period, and all regressions have 110 observations (22 regions and 5 years). The dependent variable is the age- and education-adjusted wage of native women (Panel A) or men (Panel B). Columns 3, 4, 6 and 8 further adjust wages for sample selection. 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 region and time fixed effects, and are weighted by cell size. Wild bootstrap p-values in italics are computed using 1,000 bootstrap replications. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

Table 5: Immigration and wages using alternative selection models

	Employment regression	Individual-level wage regression	OLS estimates (1)	(2)	(3)	IV estimates (4)	(5)	(6)
Regressors included: X = age, education, region, time f.e.; F = family characteristics; H = home ownership								
A. Impact on the wage of native women								
Baseline specification	(X,F,H)	(X)	-0.02	-0.44***	-0.35**	-0.01	-0.43***	-0.95***
			(0.07)	(0.08)	(0.14)	(0.11)	(0.10)	(0.30)
Wild cluster bootstrap p-value			*0.71*	*0.02*	*0.03*	*0.95*	*0.05*	*0.00*
Specification 2	(X,F,H)	(X,F)	-0.03	-0.40***	-0.32**	-0.01	-0.38***	-0.92***
			(0.07)	(0.08)	(0.14)	(0.11)	(0.10)	(0.30)
Wild cluster bootstrap p-value			*0.67*	*0.02*	*0.05*	*0.93*	*0.06*	*0.01*
Specification 3	(X,F)	(X)	-0.02	-0.40***	-0.32**	-0.01	-0.38***	-0.92***
			(0.07)	(0.08)	(0.14)	(0.11)	(0.10)	(0.30)
Wild cluster bootstrap p-value			*0.71*	*0.02*	*0.04*	*0.95*	*0.06*	*0.01*
Specification 4	(X,H)	(X)	-0.02	-0.52***	-0.41**	-0.01	-0.51***	-1.00***
			(0.07)	(0.08)	(0.15)	(0.11)	(0.11)	(0.30)
Wild cluster bootstrap p-value			*0.71*	*0.02*	*0.02*	*0.95*	*0.04*	*0.00*
Baseline specification using	(X,F,H)	(X)	-0.02	-0.52***	-0.42***	-0.01	-0.51***	-1.00***
linear probability model			(0.07)	(0.08)	(0.14)	(0.11)	(0.10)	(0.29)
Wild cluster bootstrap p-value			*0.71*	*0.01*	*0.02*	*0.95*	*0.04*	*0.00*
	Employment regression	Individual-level 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)	(0.13)	(0.17)	(0.09)	(0.09)	(0.18)
Wild cluster bootstrap p-value			*0.18*	*0.20*	*0.02*	*0.00*	*0.00*	*0.01*
Specification 2	(X,F,H)	(X,F)	-0.72***	-0.56***	-0.48**	-0.83***	-0.65***	-0.93***
			(0.12)	(0.10)	(0.19)	(0.09)	(0.10)	(0.25)
Wild cluster bootstrap p-value			*0.19*	*0.04*	*0.06*	*0.00*	*0.01*	*0.00*
Specification 3	(X,F,H)	(X)	-0.72***	-0.55***	-0.48**	-0.83***	-0.64***	-0.93***
			(0.12)	(0.11)	(0.19)	(0.09)	(0.10)	(0.25)
Wild cluster bootstrap p-value			*0.19*	*0.05*	*0.06*	*0.00*	*0.01*	*0.00*
Specification 4	(X,F)	(X)	-0.72***	-0.77***	-0.60***	-0.83***	-0.90***	-0.68***
			(0.12)	(0.13)	(0.17)	(0.09)	(0.10)	(0.20)
Wild cluster bootstrap p-value			*0.19*	*0.28*	*0.03*	*0.00*	*0.00*	*0.03*
Baseline specification using	(X,F,H)	(X,F)	-0.79***	-1.03***	-0.81***	-0.90***	-1.18***	-0.63**
linear probability model			(0.12)	(0.18)	(0.19)	(0.09)	(0.15)	(0.26)
Wild cluster bootstrap p-value			*0.18*	*0.47*	*0.02*	*0.00*	*0.00*	*0.06*
Selectivity-corrected estimates			-	Yes	Yes	-	Yes	Yes
Add log of native labor force as regressor			-	-	Yes	-	-	Yes

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The unit of observation is a region-year cell over the 1982-2016 period, and all regressions have 110 observations (22 regions and 5 years). The dependent variable is the age- and education-adjusted wage of native women (Panel A) or men (Panel B). Columns 2-3 and 5-6 further adjust wages for sample selection. Each row uses a specific set of variables to generate 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 include region and time fixed effects, and are weighted by cell size. Wild bootstrap p-values in italics are computed using 1,000 bootstrap replications. ***, **, * 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***	-1.07***	-0.03	0.00	-0.42***	-0.93***
	(0.07)	(0.12)	(0.12)	(0.30)	(0.07)	(0.11)	(0.10)	(0.28)
Wild cluster bootstrap p-value	*0.16*	*0.73*	*0.10*	*0.00*	*0.57*	*0.97*	*0.05*	*0.00*
Log of native labor force	-	-	-	-0.21**	-	-	-	-0.20**
				(0.10)				(0.09)
Wild cluster bootstrap p-value				*0.11*				*0.10*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.20***	0.20***	-	-	0.21***	0.21***
			(0.02)	(0.02)			(0.02)	(0.02)
Kleibergen-Paap F-test	-	27.12	27.12	-	-	33.36	33.36	-
SW multivariate F-test (imm. share)	-	-	-	11.32	-	-	-	13.78
SW multivariate F-test (log nat.)	-	-	-	11.92	-	-	-	16.30
B. Impact on the wage of native men								
Immigrant share	-0.91***	-0.99***	-1.02***	-0.88***	-0.77***	-0.87***	-0.89***	-0.75***
	(0.15)	(0.11)	(0.11)	(0.23)	(0.12)	(0.09)	(0.10)	(0.18)
Wild cluster bootstrap p-value	*0.26*	*0.01*	*0.00*	*0.02*	*0.23*	*0.00*	*0.00*	*0.01*
Log of native labor force	-	-	-	0.07	-	-	-	0.05
				(0.08)				(0.06)
Wild cluster bootstrap p-value				*0.43*				*0.46*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.04	0.04	-	-	0.04	0.04
			(0.06)	(0.06)			(0.07)	(0.07)
Kleibergen-Paap F-test	-	17.13	17.13	-	-	21.51	21.51	-
SW multivariate F-test (imm. share)	-	-	-	12.41	-	-	-	15.08
SW multivariate F-test (log nat.)	-	-	-	13.13	-	-	-	17.86

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The unit of observation is a region-year cell. The regressions in columns 1-4 use the 1990-2016 cross-sections and have 88 observations (22 regions and 4 years); the regressions in columns 5-8 use the original 1982-2016 cross-sections and add the 2012 panel, thus having 132 observations (22 regions and 6 years). The dependent variable is the age- and education-adjusted wage of native women (Panel A) or men (Panel B). Columns 3-4 and 7-8 further adjust wages for sample selection. 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 region and time fixed effects, and are weighted by cell size. Wild bootstrap p-values in italics are computed using 1,000 bootstrap replications. ***, **, * 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	-0.44***	-0.91***	0.02	-0.01	-0.35***	-0.98**
	(0.09)	(0.12)	(0.11)	(0.25)	(0.06)	(0.09)	(0.09)	(0.43)
Wild cluster bootstrap p-value	*0.47*	*0.89*	*0.02*	*0.00*	*0.70*	*0.95*	*0.07*	*0.02*
Log of native labor force	-	-	-	-0.18**	-	-	-	-0.31*
				(0.07)				(0.17)
Wild cluster bootstrap p-value				*0.06*				*0.19*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.21***	0.21***	-	-	0.21***	0.21***
			(0.02)	(0.02)			(0.02)	(0.02)
Kleibergen-Paap F-test	-	13.61	13.61	-	-	21.01	21.01	-
SW multivariate F-test (imm. share)	-	-	-	16.44	-	-	-	5.70
SW multivariate F-test (log nat.)	-	-	-	15.87	-	-	-	7.89
B. Impact on the wage of native men								
Immigrant share	-0.73***	-0.87***	-0.89***	-0.74***	-0.81**	-1.16***	-1.19***	-0.79***
	(0.16)	(0.13)	(0.14)	(0.17)	(0.30)	(0.17)	(0.18)	(0.20)
Wild cluster bootstrap p-value	*0.28*	*0.01*	*0.01*	*0.01*	*0.34*	*0.01*	*0.02*	*0.01*
Log of native labor force	-	-	-	0.06	-	-	-	0.11*
				(0.06)				(0.06)
Wild cluster bootstrap p-value				*0.34*				*0.15*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.05	0.05	-	-	0.05	0.05
			(0.07)	(0.07)			(0.07)	(0.07)
Kleibergen-Paap F-test	-	10.48	10.48	-	-	35.13	35.13	-
SW multivariate F-test (imm. share)	-	-	-	18.49	-	-	-	31.36
SW multivariate F-test (log nat.)	-	-	-	17.05	-	-	-	20.47

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The unit of observation is a region-year cell over the 1982-2016 period, and all regressions have 110 observations (22 regions and 5 years). The dependent variable is the age- and education-adjusted wage of native women (Panel A) or men (Panel B). Columns 3-4 and 7-8 further adjust wages for sample selection. 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 region and time fixed effects, and are weighted by cell size. Wild bootstrap p-values in italics are computed using 1,000 bootstrap replications. ***, **, * 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	0.46	0.01	-0.58**	-0.75	-0.32***	-0.35***	-0.56***	-0.73**
	(0.28)	(0.30)	(0.30)	(0.51)	(0.09)	(0.10)	(0.10)	(0.29)
Wild cluster bootstrap p-value	*0.59*	*0.97*	*0.04*	*0.17*	*0.08*	*0.09*	*0.02*	*0.02*
Log of native labor force	-	-	-	-0.09	-	-	-	-0.05
				(0.20)				(0.11)
Wild cluster bootstrap p-value				*0.71*				*0.64*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.21***	0.21***	-	-	0.14***	0.14***
			(0.02)	(0.02)			(0.02)	(0.02)
Kleibergen-Paap F-test	-	23.41	23.41	-	-	32.06	32.06	-
SW multivariate F-test (imm. share)	-	-	-	21.69	-	-	-	12.91
SW multivariate F-test (log nat.)	-	-	-	7.96	-	-	-	15.15
	B. Impact on the wage of native men							
Immigrant share	-0.35**	-0.66***	-0.70***	-0.71***	-1.09***	-1.22***	-1.21***	-0.88***
	(0.15)	(0.17)	(0.17)	(0.24)	(0.18)	(0.12)	(0.12)	(0.22)
Wild cluster bootstrap p-value	*0.20*	*0.02*	*0.03*	*0.01*	*0.30*	*0.00*	*0.00*	*0.02*
Log of native labor force	-	-	-	-0.05	-	-	-	0.11
				(0.11)				(0.08)
Wild cluster bootstrap p-value				*0.65*				*0.24*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.05	0.05	-	-	-0.02	-0.02
			(0.07)	(0.07)			(0.08)	(0.08)
Kleibergen-Paap F-test	-	23.41	23.41	-	-	21.15	21.15	-
SW multivariate F-test (imm. share)	-	-	-	21.69	-	-	-	14.24
SW multivariate F-test (log nat.)	-	-	-	7.96	-	-	-	16.85

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The unit of observation is a region-year cell over the 1982-2016 period, and all regressions have 110 observations (22 regions and 5 years). The dependent variable is the age- and education-adjusted wage of native women (Panel A) or men (Panel B). Columns 3-4 and 7-8 further adjust wages for sample selection. We do not weight the regressions in columns 1-4. The regressions in columns 5-8 are weighted by cell size, and include all interacted age-education-region fixed effects and all interacted age-education-time fixed effects to generate the inverse Mills ratio. 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 region and time fixed effects. Wild bootstrap p-values in italics are computed using 1,000 bootstrap replications. ***, **, * 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				Hourly wage of 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**	-0.71**	-0.49***	-0.47***	-0.93***	-1.30***
	(0.07)	(0.11)	(0.10)	(0.29)	(0.04)	(0.09)	(0.10)	(0.23)
Wild cluster bootstrap p-value	*0.71*	*0.95*	*0.14*	*0.05*	*0.00*	*0.02*	*0.01*	*0.01*
Log of native labor force	-	-	-	-0.20**	-	-	-	-0.16**
				(0.09)				(0.08)
Wild cluster bootstrap p-value				*0.71*				*0.19*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.14***	0.14***	-	-	0.23***	0.23***
			(0.01)	(0.01)			(0.03)	(0.03)
Kleibergen-Paap F-test	-	32.06	32.06	-	-	25.65	25.65	-
SW multivariate F-test (imm. share)	-	-	-	12.91	-	-	-	12.86
SW multivariate F-test (log nat.)	-	-	-	15.15	-	-	-	15.07
	B. Impact on the wage of native men							
Immigrant share	-0.79***	-0.90***	-0.92***	-0.76***	-0.96***	-1.00***	-1.07***	-0.85***
	(0.12)	(0.09)	(0.09)	(0.18)	(0.12)	(0.12)	(0.13)	(0.20)
Wild cluster bootstrap p-value	*0.18*	*0.00*	*0.00*	*0.01*	*0.09*	*0.01*	*0.02*	*0.02*
Log of native labor force	-	-	-	0.04	-	-	-	0.09
				(0.07)				(0.07)
Wild cluster bootstrap p-value				*0.54*				*0.25*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.05	0.05	-	-	0.13**	0.13**
			(0.07)	(0.07)			(0.06)	(0.06)
Kleibergen-Paap F-test	-	21.00	21.00	-	-	20.83	20.83	-
SW multivariate F-test (imm. share)	-	-	-	14.09	-	-	-	13.99
SW multivariate F-test (log nat.)	-	-	-	16.53	-	-	-	16.41

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The unit of observation is a region-year cell over the 1982-2016 period, and all regressions have 110 observations (22 regions and 5 years). The dependent variable is the age- and education-adjusted wage of native women (Panel A) or men (Panel B). Columns 3-4 and 7-8 further adjust wages for sample selection. 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 adjusted measure of the mean wage in the cell in columns 5-8 is based on 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 region and time fixed effects, and are weighted by cell size. Wild bootstrap p-values in italics are computed using 1,000 bootstrap replications. ***, **, * 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	-0.14	-0.49***	-0.16	-0.46***	-0.22	-0.67***	-0.61***	-0.94***
	(0.10)	(0.08)	(0.13)	(0.10)	(0.23)	(0.19)	(0.20)	(0.21)
Wild cluster bootstrap p-value	*0.42*	*0.00*	*0.30*	*0.00*	*0.44*	*0.02*	*0.01*	*0.00*
Log of native labor force	-	-	-0.02	0.02	-	-	-0.26***	-0.19***
			(0.05)	(0.04)			(0.08)	(0.06)
Wild cluster bootstrap p-value			*0.76*	*0.54*			*0.02*	*0.02*
Selectivity-corrected estimates	-	Yes	-	Yes	-	Yes	-	Yes
Inverse Mills ratio	-	0.25***	-	0.25***	-	0.25***	-	0.25***
		(0.04)		(0.04)		(0.04)		(0.04)
Kleibergen-Paap F-test	-	-	-	-	7.39	7.39	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	14.75	14.75
SW multivariate F-test (log nat.)	-	-	-	-	-	-	11.87	11.87
	B. Impact on the wage of native men							
Immigrant share	-0.62***	-0.64***	-0.54***	-0.56***	-0.80***	-0.80***	-1.02***	-1.02***
	(0.12)	(0.12)	(0.15)	(0.15)	(0.18)	(0.18)	(0.24)	(0.25)
Wild cluster bootstrap p-value	*0.00*	*0.01*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
Log of native labor force	-	-	0.06	0.06	-	-	-0.04	-0.03
			(0.05)	(0.05)			(0.08)	(0.08)
Wild cluster bootstrap p-value			*0.18*	*0.15*			*0.63*	*0.71*
Selectivity-corrected estimates	-	Yes	-	Yes	-	Yes	-	Yes
Inverse Mills ratio	-	0.06*	-	0.06*	-	0.06*	-	0.06*
		(0.03)		(0.03)		(0.03)		(0.03)
Kleibergen-Paap F-test	-	-	-	-	7.92	7.92	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	17.02	17.02
SW multivariate F-test (log nat.)	-	-	-	-	-	-	11.99	11.99

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by department. The unit of observation is a department-year cell over the 1982-2016 period, and all regressions have 470 observations (94 regions and 5 years). The dependent variable is the age- and education-adjusted wage of native women (Panel A) or men (Panel B). Columns 3, 4, 6 and 8 further adjust wages for sample selection. 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 department and the predicted (log) number of immigrants in the region; 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 are weighted by cell size. Wild bootstrap p-values in italics are computed using 1,000 bootstrap replications. ***, **, * 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***	-1.29***	0.24*	0.28	-0.10	-0.62
	(0.11)	(0.18)	(0.20)	(0.35)	(0.12)	(0.18)	(0.16)	(0.43)
Wild cluster bootstrap p-value	*0.02*	*0.04*	*0.03*	*0.00*	*0.25*	*0.24*	*0.61*	*0.24*
Log of native labor force	-	-	-	-0.18	-	-	-	-0.16
				(0.12)				(0.16)
Wild cluster bootstrap p-value				*0.08*				*0.38*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.11***	0.11***	-	-	0.22***	0.22***
			(0.02)	(0.02)			(0.04)	(0.04)
Kleibergen-Paap F-test	-	21.03	21.03	-	-	45.34	45.34	-
SW multivariate F-test (imm. share)	-	-	-	13.84	-	-	-	12.27
SW multivariate F-test (log nat.)	-	-	-	14.83	-	-	-	15.55
B. Impact on the wage of native men								
Immigrant share	-1.22***	-1.45***	-1.40***	-1.06***	-0.45***	-0.48***	-0.53***	-0.40*
	(0.25)	(0.16)	(0.14)	(0.24)	(0.13)	(0.12)	(0.12)	(0.22)
Wild cluster bootstrap p-value	*0.36*	*0.00*	*0.00*	*0.02*	*0.09*	*0.07*	*0.05*	*0.02*
Log of native labor force	-	-	-	0.01	-	-	-	0.12
				(0.11)				(0.10)
Wild cluster bootstrap p-value				*0.91*				*0.27*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	-0.09	-0.09	-	-	0.09	0.09
			(0.06)	(0.06)			(0.10)	(0.10)
Kleibergen-Paap F-test	-	13.33	13.33	-	-	39.69	39.69	-
SW multivariate F-test (imm. share)	-	-	-	15.21	-	-	-	12.75
SW multivariate F-test (log nat.)	-	-	-	16.51	-	-	-	16.66

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The unit of observation is a region-year cell over the 1982-2016 period, and all regressions have 110 observations (22 regions and 5 years). The dependent variable is the age- and education-adjusted wage of native women (Panel A) or men (Panel B). Columns 3, 4, 6 and 8 further adjust wages for sample selection. 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 region and time fixed effects, and are weighted by cell size. Wild bootstrap p-values in italics are computed using 1,000 bootstrap replications. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

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

	OLS estimates				IV estimates			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. Impact on the wage of native women							
Immigrant share	-0.21**	-0.37*	-0.32***	-0.47**	-0.04	-0.38**	-0.26*	-0.60***
	(0.08)	(0.20)	(0.10)	(0.21)	(0.10)	(0.15)	(0.15)	(0.21)
Wild cluster bootstrap p-value	0.01	0.01	0.01	0.00	0.68	0.12	0.24	0.08
Log of native labor force	-	-	-0.04***	-0.03**	-	-	-0.04***	-0.04***
			(0.01)	(0.01)			(0.01)	(0.01)
Wild cluster bootstrap p-value			0.00	0.01			0.00	0.01
Selectivity-corrected estimates		Yes		Yes		Yes		Yes
Inverse Mills ratio	-	0.14***	-	0.14***	-	0.14***	-	0.14***
		(0.05)		(0.05)		(0.05)		(0.05)
Kleibergen-Paap F-test	-	-	-	-	48.73	48.73	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	89.22	89.22
SW multivariate F-test (log nat.)	-	-	-	-	-	-	303.77	303.77
	B. Impact on the wage of native men							
Immigrant share	-0.63***	-0.67***	-0.86***	-0.89***	-0.70***	-0.79***	-1.23***	-1.30***
	(0.21)	(0.23)	(0.25)	(0.27)	(0.20)	(0.19)	(0.15)	(0.15)
Wild cluster bootstrap p-value	0.00	0.00	0.00	0.00	0.05	0.04	0.00	0.00
Log of native labor force	-	-	-0.07***	-0.07***	-	-	-0.08***	-0.08***
			(0.01)	(0.01)			(0.01)	(0.01)
Wild cluster bootstrap p-value			0.00	0.00			0.00	0.00
Selectivity-corrected estimates	-	Yes	-	Yes	-	Yes	-	Yes
Inverse Mills ratio	-	0.05	-	0.05	-	0.05	-	0.05
		(0.04)		(0.04)		(0.04)		(0.04)
Kleibergen-Paap F-test	-	-	-	-	36.62	36.62	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	73.66	73.66
SW multivariate F-test (log nat.)	-	-	-	-	-	-	679.35	679.35

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered at the region-education-age level. The unit of observation is a region-education-age-year cell over the 1982-2016 period, and all regressions have 440 observations (22 regions, 2 education groups, 2 age groups and 5 years). The dependent variable is the wage of native women (Panel A) or men (Panel B). Columns 3, 4, 6 and 8 further adjust wages for sample selection. 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 time, and interacted region-education-age fixed effects, and are weighted by cell size. Wild bootstrap p-values in italics are computed using 1,000 bootstrap replications. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.