Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

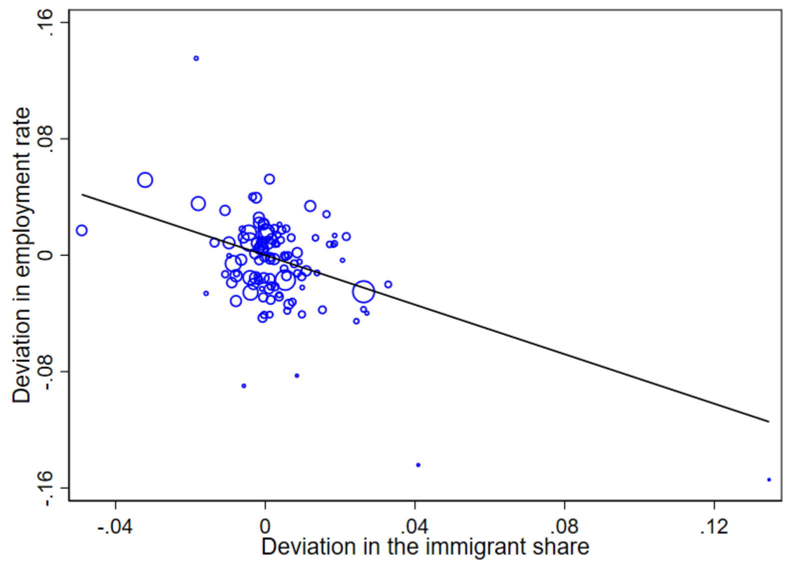
Online Appendix to "Gender, Selection into Employment, and the Wage Impact of Immigration"

George J. Borjas and Anthony Edo

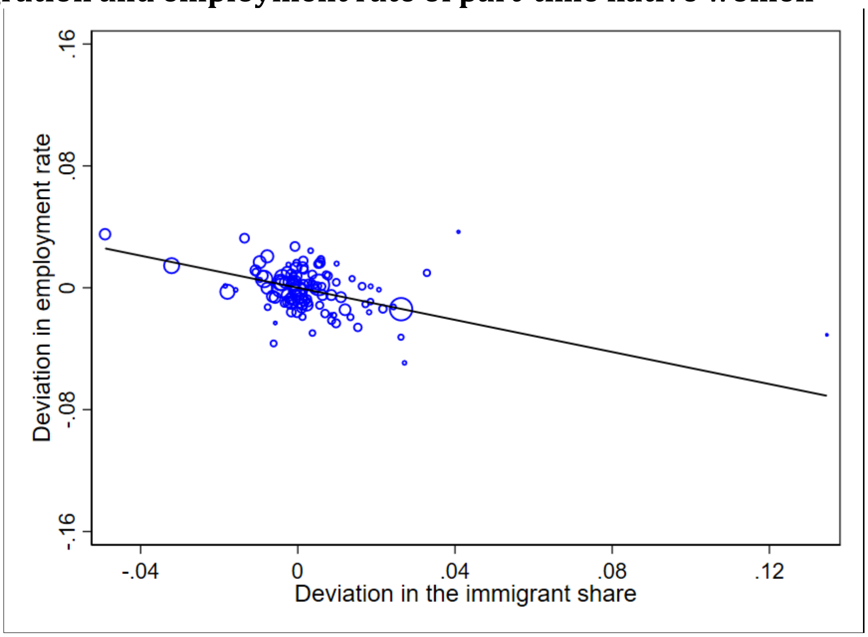
Appendix A: Impact of immigration on employment and internal mobility

Figure A. Immigration and employment of native women by full/part time status

1. **Immigration and employment rate of full-time native women**



2. **Immigration and employment rate of part-time native women**



Notes: The unit of observation in the scatter diagrams is a region-year cell over the 1982-2016 period. The employment rate of full-time (part-time) women is the ratio of full-time (part-time) employment to the population. The figures correlate the deviation in each employment rate to the deviation in the immigrant share after removing year-specific effects that are common to all regions in a given census year. The deviations in the 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. The size of the circles is proportional to the value of the weight.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Table A1: Impact of immigration on native employment rates

	Native women			Native men		
	1968-2016	1982-2016		1968-2016	1982-2016	
	OLS	OLS	IV	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
1. Regional level	-0.81***	-0.98***	-0.97***	0.10	-0.02	-0.06
	(0.14)	(0.10)	(0.10)	(0.09)	(0.10)	(0.12)
Wild cluster bootstrap p-value	*0.09*	*0.03*	*0.01*	*0.52*	*0.94*	*0.80*
Kleibergen-Paap F-test	-	-	28.53	-	-	26.17
Observations	154	110	110	154	110	110
2. Departmental level	-0.63***	-0.64***	-0.80***	-0.05	-0.12*	0.09
	(0.06)	(0.08)	(0.15)	(0.09)	(0.07)	(0.13)
Wild cluster bootstrap p-value	*0.00*	*0.05*	*0.00*	*0.92*	*0.41*	*0.54*
Kleibergen-Paap F-test	-	-	9.74	-	-	9.56
Observations	760	560	560	760	560	560

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The unit of observation is an area-year cell over the 1982-2016 period. Specifications 1 and 2 exploit variations across 22 regions and 94 departments, respectively. The dependent variable is the employment-to-population rate of native women in columns 1-3, and of native men in columns 4-6. Columns 3 and 6 instrument the share of immigrants with the shift-share instrument defined in equation (18) and computed using the 1968 French census. All regressions include area 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 A2: Impact of immigration on the mobility of natives across regions

	OLS	IV
	(1)	(2)
Immigrant inflow rate	1.33***	1.00
	(0.32)	(0.72)
Wild cluster bootstrap p-value	*0.02*	*0.14*
Kleibergen-Paap F-test		13.29

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. The table exploits 110 observations (22 regions and 5 years). The dependent and independent variables respectively give the change in the native and immigrant population (aged 18-64) in a given region between two points in time, both divided by the regional native population (aged 18-64) at time $t-1$. Column 2 instruments the immigration variable with the shift-share instrument computed using the 1962 French census. 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.

Discussion of Table A2:

Table A2 investigates the effect of immigration on the mobility of natives over the 1982-2016 period by estimating the following first-difference equation:

$$\frac{\Delta \mathrm{N}_{rt}}{\mathrm{N}_{rt-1}} = \rho_0 + \rho_1 \frac{\Delta \mathrm{M}_{rt}}{\mathrm{N}_{rt-1}} + \theta_r + \theta_t + \upsilon_{rt}\,. \tag{A1}$$

The dependent and independent variables respectively give the change in native and immigrant working-age population in region r between t and $t-1$, both divided by the regional native population in $t-1$. Region and time fixed effects are captured by θ_r and θ_t, respectively; υ_{rt} is the error term.

The coefficient ρ_1 gives the change in the size of the native population in response to one additional immigrant in a region when the stock of immigrants increases by one. A positive estimate thus implies that immigration tends to attract more natives in the areas targeted by immigrants. A negative estimate implies that immigration induces a crowd-out effect on native workers.

An important identification issue in estimating the impact of immigration on native internal mobility is that immigrants are not randomly allocated across regions. If immigrants choose the most economically prosperous regions, one would expect a spurious positive correlation between the population share of immigrants and the population share of natives. To address this concern, we rely on equations (16) and (17), and instrument $\Delta \mathrm{M}_{rt}/\mathrm{N}_{rt-1}$ by $\left(\widehat{M}_{rt} - \widehat{M}_{rt-1}\right)/\widehat{N}_{rt-1}$.

The IV estimated coefficient from Table A2 is not significant, indicating that immigration does not affect the mobility of natives across French regions. As a result, the estimated labor market effects induced by immigration are unlikely to be biased by native internal migration.

Appendix B: Imperfect substitution between men and women

We begin with the Cobb-Douglas aggregate production function in (6). For simplicity, assume that L_{kt} is composed of only native workers N_{kt} in the pre-immigration period:

$$N_{kt} = \left[H_{kt}^{\delta} + F_{kt}^{\delta}\right]^{\frac{1}{\delta}}, \qquad (B1)$$

where $\delta \leq 1$; and H_{kt} and F_{kt} represent the number of male and female native workers, respectively. The market wage of female native workers in the pre-immigration period ($t = 0$) is then given by:

$$\log w_{k0}^{F} = \varphi_k + [(1-\delta) - \eta] \log N_{k0} + (\delta - 1) \log F_{k0}\,. \qquad (B2)$$

Note that Equation (B2) collapses to Equation (7) in the text if male and female native workers are perfect substitutes (δ = 1). To fix ideas, we assume that the market-specific parameter φ_k is constant. An influx of M_k immigrants (of which M_k^F are female) enters the market, and we assume that their labor supply is perfectly inelastic. Native labor supply, however, might respond to the supply shock at the extensive margin, so that the total number of native workers in the market in the post-migration period ($t = 1$) is N_{k1}, changing total employment from N_{k0} to $(N_{k1} + M_k)$. Similarly, female employment changes from F_{k0} to $(F_{k1} + M_k^F)$.

The wage for native female workers in market k in the post-migration period (*after* natives have responded) can then be written as:

$$\log w_{k1}^{F} = \varphi_k + [(1-\delta) - \eta] \log(N_{k1} + M_k) + (\delta - 1) \log(F_{k1} + M_k^F)\,. \qquad (B3)$$

The change in the wage of female native workers is:

$$\Delta \log w_k^F = [(1-\delta) - \eta](\Delta \log N_k + m_k) + (\delta - 1)\,(\Delta \log F_k + m_k^F), \qquad (B4)$$

where the immigrant share $m_k = \log(1 + M_k/N_{k1})$ measures the total supply shock, and the share $m_k^F = \log(1 + M_k^F /\, F_{k1})$ gives the supply shock in the female labor market.

Note that we can express the change in total native employment as:

$$\Delta \log N_k \approx \frac{\Delta H_k}{H_k}\frac{H_k}{N_k} + \frac{\Delta F_k^F}{F_k}\frac{F_k}{N_k} = (1-\theta)\,\Delta \log H_k + \theta\,\Delta \log F_k\,, \qquad (B5)$$

where θ is the female fraction of the workforce. We find that male employment did not respond to the immigrant supply shock. To simplify the exposition, we assume that male native labor supply is perfectly inelastic and $\Delta \log H_k = 0$. We can then approximate $\Delta \log N_k \approx \theta\,\Delta \log F_k$.

To relate the change in the wage of female natives to the total supply shock (and obtain a “reduced-form” labor demand equation), we assume that $\Delta \log F_k = \gamma m_k$, where the parameter γ measures the crowd-out effect for women. We can then rewrite Equation $(B4)$ as:

$$\Delta \log w_k^F = [(1-\delta)-\eta]\,(1+\theta\gamma)\, m_k + (\delta - 1)\,(\gamma m_k + m_k^F). \qquad (B6)$$

In our data, there is a very strong positive correlation between the supply shock in the female labor market, m_k^F, and the total supply shock, m_k. In fact, the regression relating these two variables has an intercept equal to -0.008 (0.004), a slope of 0.92 (0.06), and an R-squared of 0.94. Without loss of generality, we approximate $m_k^F = \alpha^F m_k$, allowing us to express the wage change of female native workers solely as a function of the total supply shock:

$$\Delta \log w_k^F = \left[-\eta(1+\theta\gamma) + (1-\delta)[(1-\alpha^F) - \gamma(1-\theta)]\right] m_k. \qquad (B7)$$

Equation $(B7)$ shows that regressing the change in the wage of native women on the (total) immigrant supply shock does not identify the "reduced-form" wage elasticity when men and women are imperfect substitutes ($\delta \neq 1$). The measured wage effect is then contaminated by the labor supply response of native women, by the elasticity of substitution between men and women, by the share of the workforce that is female, and by the relationship between the total supply shock and the supply shock in the female labor market.

It is important to note that estimating the impact of an immigration-induced increase in labor supply on the *relative* wage of native women would identify another set of parameters. We continue to assume that the labor supply of men is inelastic at the extensive margin. The impact of the supply shock on the wage growth of men can then be obtained by using the analogous steps used to derive equation (B7), and the approximation $m_k^H = \alpha^H m_k$, where $m_k^H = \log(1 + M_k^H / H_{k1})$. It follows that the wage growth of men is given by:

$$\Delta \log w_k^H = \left[-\eta(1+\theta\gamma) + (1-\delta)[1-\alpha^H + \theta\gamma]\right] m_k. \qquad (B8)$$

We can use Equations (B7) and (B8) to derive the change in the gender wage gap:

$$\Delta \log w_k^F - \Delta \log w_k^H = (1-\delta)\,(\alpha^H - \alpha^F - \gamma)\, m_k. \qquad (B9)$$

Equation (B9) shows that the impact of the supply shock on the gender wage gap depends on the elasticity of substitution between the two groups and on the labor supply response of native women.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Appendix C: Test of Identifying Assumptions

This section discusses the validity of our IV strategy relying on Goldsmith-Pinkham, Sorkin, and Swift (2020). First, we compute the Rotemberg weights for each origin-specific share of immigrants used in the construction of our instrument. The results are displayed in the Panel A of Table C1. Portugal receives the highest weight (0.22), followed by naturalized immigrants (0.19), Algerian (0.16) and Moroccan (0.07) immigrants. These four nationality groups receive almost 60 percent of the overall weight. Given the relative importance of Portuguese and North-African immigration to France, it is not surprising to find that these groups play an important role in the construction of our shift-share instrument.

Second, we test whether the initial origin-specific shares used to build our instrument are correlated with initial period characteristics. More precisely, Panel B of Table C1 shows the correlation between the origin-specific shares and a set of regional characteristics in year 1968. We use the same regional variables as in Goldsmith-Pinkham, Sorkin, and Swift (2020, Table 5), plus the share of natives below 30.[1] Although 80 percent of the estimated coefficients are insignificant, it turns out that the regional distribution of immigrants tend to be positively correlated with regional population size and the share of high educated natives. Our estimates moreover indicates no correlation between the mean wages of natives and immigrants and the origin-specific shares, indicating the presence of plausible exogenous variation.

Third, Table C2 test the presence of correlation between the pre-1982 growth in native labor market outcomes and the 1982-2016 change in the predicted immigrant shares. As dependent variables, columns 1-2 use the change in the mean regional wage of native women and men between 1976 and 1982, while columns 3-4 use the change in the employment rate of native women and men between 1975 and 1982.[2] The regressor of interest is the change in the predicted share of immigrants, as defined by $\widehat{m}_{rt} - \widehat{m}_{rt-1}$, between 1982 and 1990, as well as between 1982 and 2016. Panel A and B respectively use the historical 1962 and 1968 censuses to compute the predicted regional change in immigrant shares.

The estimated coefficients on the pre-1982 changes are not statistically significant. This result implies that there is no correlation between the pre-1982 labor market opportunities of natives and the subsequent predicted share of immigrants across regions. This supports the validity of our instrument as it indicates a lack of persistent regional trend affecting native wages and employment.

[1] All regional variables are taken from the 1968 French census data, except for the wage data, which are taken from a wage survey published in 1970 ("Enquête Formation et Qualification Professionnelles" or FQP). More precisely, the 1970 FQP reports the net annual wages of employees for the years 1969. We exclude Corsica from our exercise since the 1970 FQP data do not report any wage information for that region.

[2] The wage data in 1976 comes from the 1977 FQP, while the employment data comes from the 1975 census. Because the wage data only covers 80 native workers in Corsica, we exclude this region from our analysis.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Table C1: Rotemberg weights and correlation between origin country shares and regional characteristics

	Origin country shares											
	Portuguese immigrants	Naturalized immigrants	Algerian immigrants	Moroccan immigrants	Rest of the world	Other Europ. imm.	Other African immigrants	Spanish immigrants	Tunisian immigrants	Italian immigrants	Turkish immigrants	Natives
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Rotemberg weights	0.22	0.19	0.16	0.07	0.07	0.07	0.06	0.06	0.04	0.04	0.01	-
B. Correlation between origin country shares and regional characteristics in 1968												
Log population size	0.03	0.05***	0.05**	0.07*	0.03	0.07**	0.02	0.04***	0.03	0.05**	0.02	0.04***
	(0.03)	(0.01)	(0.02)	(0.04)	(0.03)	(0.03)	(0.04)	(0.01)	(0.03)	(0.02)	(0.02)	(0.01)
Share of high educated natives	3.07**	0.51	1.63*	1.31	3.56**	-0.07	3.27	2.12***	3.62**	0.40	3.16***	0.60
	(1.33)	(0.51)	(0.81)	(1.70)	(1.64)	(1.25)	(2.28)	(0.53)	(1.50)	(0.98)	(1.05)	(0.38)
Share of young natives	0.21	-0.00	0.39	-0.80	0.93	-0.79	0.83	0.01	1.06	0.40	0.60	-0.03
	(0.94)	(0.44)	(0.77)	(1.36)	(1.21)	(1.05)	(1.47)	(0.34)	(1.20)	(0.55)	(0.81)	(0.34)
Mean wage of natives	-0.01	0.21*	0.26	0.12	0.29	0.20	0.38	-0.00	0.30	0.28	0.15	0.02
	(0.29)	(0.12)	(0.17)	(0.29)	(0.31)	(0.21)	(0.46)	(0.12)	(0.29)	(0.22)	(0.21)	(0.08)
Mean wage of immigrants	0.12	-0.08	0.02	0.12	0.14	0.05	0.20	-0.06	0.12	-0.11	0.11	0.04
	(0.14)	(0.05)	(0.07)	(0.11)	(0.14)	(0.07)	(0.19)	(0.06)	(0.13)	(0.08)	(0.10)	(0.03)
Share of manufacturing emp.	0.10	-0.06	0.01	0.17	-0.23	0.24	-0.23	-0.11	-0.22	0.05	-0.01	0.02
	(0.28)	(0.12)	(0.11)	(0.30)	(0.28)	(0.22)	(0.44)	(0.14)	(0.23)	(0.15)	(0.18)	(0.07)

Notes: Panel A provides the Rotemberg weights for each origin share used to compute the shift-share instrument following Goldsmith-Pinkham, Sorkin and Swift (2020). Each column in Panel B reports the estimated coefficients of a regression of an origin-specific share in 1968 on regional characteristics in 1968. Below the point estimate, the standard errors in parentheses are clustered at the regional level.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Table C2: Relationship between past labor market outcomes and predicted immigrant shares across regions

	Dependent variable			
	Difference in mean log wage between 1975 and 1982		Difference in employment rate between 1975 and 1982	
	Native women	Native men	Native women	Native men
	(1)	(2)	(3)	(4)
A. Difference in predicted immigrant share using the 1968 census				
1. Between 1982 and 1990	2.15	2.22	0.22	0.29
	(1.63)	(1.82)	(0.51)	(0.46)
2. Between 1982 and 2016	0.43	-0.52	-0.26	0.08
	(0.72)	(1.14)	(0.25)	(0.24)
B. Difference in predicted immigrant share using the 1962 census				
3. Between 1982 and 1990	1.95	1.38	0.06	0.06
	(1.31)	(1.67)	(0.48)	(0.42)
4. Between 1982 and 2016	0.45	-0.42	-0.23	-0.08
	(0.49)	(0.69)	(0.15)	(0.15)

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. Due to wage data limitations, we exclude Corsica. The unit of observation is a region, and all regressions have 21 observations. Each column reports results of a single regression of the change in the wage in columns 1-2 and employment-to-population rate in columns 3-4 of natives between 1975 and 1982 on the change in the predicted immigrant population between 1982 and 1990, and 1982 and 2016. To predict the share of immigrants, we use the 1968 and 1962 French censuses in specifications 1-2 and 3-4, respectively. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Appendix D: Immigration and Full-time Employment

Table D: Determinants of working in a full-time job

	Full-time v. part-time indicator			
	Native women		Native men	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Immigrant share	-0.74*	-1.00***	0.02	0.27
	(0.41)	(0.36)	(0.47)	(0.65)
Marginal effect	-0.24	-0.32	0.00	0.03
Married	-0.22***	-0.22***	0.27***	0.27***
	(0.02)	(0.02)	(0.02)	(0.02)
Marginal effect	-0.07	-0.07	0.03	0.03
Presence of children below 6	-0.35***	-0.35***	0.11***	0.11***
	(0.03)	(0.03)	(0.04)	(0.04)
Marginal effect	-0.11	-0.11	0.01	0.01
Home ownership	-0.06***	-0.06***	-0.04*	-0.04*
	(0.01)	(0.01)	(0.02)	(0.02)
Marginal effect	-0.02	-0.02	-0.00	-0.00
Observations	104,417	104,417	115,182	115,182

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 native worker is a full-time employee and zero otherwise. Columns 2 and 4 instrument the share of immigrants with the shift-share instrument computed using 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.

Appendix E: Robustness of Results in Table 4

Table E1 reports the reduced-form estimates for the same set of regressions implemented in Table 4 – i.e. our main explanatory variables are the instruments used in Table 4. Table E1 confirms that controlling for the changing composition of the sample of female workers leads to more significant and negative estimated wage effects. In contrast, the estimated wage effects in the male sample are always significantly negative, and their magnitude does not depend on the selection correction.

A potential concern with the IV framework may be that the different sectoral composition across French regions in 1968 (which likely produced regional differences in immigrant inflows before 1968) can predict the regional industry structure in subsequent years. This correlation might then account for part of the regional differences in the evolution of native wages over the 1982-2016 period (Jaeger, Ruist and Stuhler, 2018). To address this potential identification issue, Table E2 shows that our estimated results are robust to including three additional regressors in the spirit of Bartik (1991): the predicted shares of native employment in manufacturing, construction, and service industries.[3] The inclusion of these controls should reduce concerns that differences in industrial composition across French regions could be problematic for the exclusion restriction. The table shows that the inclusion of the Bartik variables makes the estimated IV wage elasticities estimated for women even more negative.

We also carried out two additional exercises to show that our IV strategy is unlikely to suffer from serial correlation in economic outcomes across regions. First, we demonstrate that our IV results are robust to exploiting the spatial distribution of immigrants and natives in baseline year 1962 (rather than 1968) to compute the shift-share instrument in equation (18) and/or restrict the analysis to the 1990-2016 period. This strategy reduces the correlation between the baseline shares used to build the instrument and current economic shocks, and is more likely to satisfy the exogeneity assumption (Dustmann, Fabbri and Preston, 2005). Table E3 shows that our results are nearly identical when we lengthen the span of the interval between the baseline cross-section used to construct the shift-share instruments and the sample period.

Second, even in the absence of serial correlation in regional economic outcomes, shift-share instruments can be invalid if previous immigrant inflows (which are correlated with contemporaneous immigrant shares) are correlated with contemporaneous wage changes because of their effects on labor market dynamics. Jaeger, Ruist, and Stuhler (2018) suggest minimizing this correlation by exploiting periods with substantial changes in the national origin mix of immigrants. Edo, Giesing, Poutvaara and Öztunc (2019) and Ortega and Verdugo (2022) 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.

---

[3] For each sector j = {Agriculture, Manufacturing, Construction, Service}, the predicted employment share in region r is obtained by taking the ratio $\hat{E}_{jrt}/\sum_j \hat{E}_{jrt}$. The variable $\hat{E}_{jrt} = E_{jr}(1968) \cdot \Delta^{t-1968} E_j$ gives the predicted number of native workers in a region-sector cell at time t. It is obtained by multiplying the initial number of native workers in a region-sector cell with the change in national employment in sector j between 1968 and t.

In addition, we can include past immigrant inflows as an additional regressor in the labor demand equation and run IV regressions that instrument both for current and lagged immigration (Jaeger, Ruist and Stuhler, 2018). Table E4 applies this strategy by simultaneously including the variables m_{rt} and $log(M_{rt-1})$ in the reduced-form version of equation ($14b$). As instruments for the two immigration variables, we use $\widehat{m}_{rt}$ and the lagged (log) predicted number of immigrants based on equation (16). The results again indicate that correcting for selection bias leads to a more negative wage elasticity in the female sample.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Table E1: Immigration and wages, reduced-form estimates

	Native women				Native men			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted immigrant share in population	-0.14	-0.74**	-	-	-1.39***	-1.38***	-	-
	(0.22)	(0.26)			(0.37)	(0.37)		
Wild cluster bootstrap p-value	*0.48*	*0.16*			*0.18*	*0.18*		
Log predicted immigrant population	-	-	-0.05***	-0.09***	-	-	-0.08***	-0.08***
			(0.02)	(0.02)			(0.03)	(0.02)
Wild cluster bootstrap p-value			*0.02*	*0.00*			*0.01*	*0.01*
Log predicted female native labor force	-	-	-0.04**	0.01	-	-	0.11**	0.11**
			(0.02)	(0.02)			(0.04)	(0.04)
Wild cluster bootstrap p-value			*0.07*	*0.65*			*0.15*	*0.16*
Selectivity-corrected estimates	-	Yes	-	Yes	-	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, 4, 6 and 8 further adjust wages for sample selection. Our main explanatory variables are the instruments used in Table 4, such that the estimated coefficients represent reduced-form estimates instead of two-stage least square estimates. 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 E2: Immigration and wages controlling for local industry shares

	Sample period: 1982-2016				Sample period: 1990-2016			
	OLS	IV estimates			OLS	IV estimates		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Impact on the wage of native women								
Immigrant share	-0.26**	-0.29*	-0.62***	-1.32*	-0.29**	-0.24	-0.56***	-1.56**
	(0.12)	(0.15)	(0.17)	(0.70)	(0.13)	(0.19)	(0.18)	(0.73)
Wild cluster bootstrap p-value	*0.02*	*0.13*	*0.01*	*0.08*	*0.01*	*0.22*	*0.01*	*0.13*
Log of native labor force	-	-	-	-0.31	-	-	-	-0.36
				(0.27)				(0.28)
Wild cluster bootstrap p-value				*0.24*				*0.21*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.18***	0.18***	-	-	0.19***	0.19***
			(0.02)	(0.02)			(0.02)	(0.02)
Kleibergen-Paap F-test	-	33.67	33.67	-	-	18.22	18.22	-
SW multivariate F-test (imm. share)	-	-	-	2.92	-	-	-	3.27
SW multivariate F-test (log nat.)	-	-	-	2.53	-	-	-	2.56
B. Impact on the wage of native men								
Immigrant share	-0.64***	-0.88***	-0.88***	-0.70***	-0.75***	-0.97***	-0.97***	-0.74**
	(0.17)	(0.19)	(0.19)	(0.27)	(0.18)	(0.22)	(0.22)	(0.34)
Wild cluster bootstrap p-value	*0.05*	*0.01*	*0.01*	*0.05*	*0.04*	*0.00*	*0.00*	*0.11*
Log of native labor force	-	-	-	0.02	-	-	-	0.08
				(0.10)				(0.13)
Wild cluster bootstrap p-value				*0.85*				*0.48*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	-0.01	-0.01	-	-	-0.00	-0.00
			(0.07)	(0.07)			(0.07)	(0.07)
Kleibergen-Paap F-test	-	31.64	31.64	-	-	18.93	18.93	-
SW multivariate F-test (imm. share)	-	-	-	2.99	-	-	-	3.54
SW multivariate F-test (log nat.)	-	-	-	2.51	-	-	-	2.61

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 original 1982-2016 cross-sections and have 110 observations (22 regions and 5 years); the regressions in columns 5-8 use the 1990-2016 cross-sections and have 88 observations (22 regions and 4 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.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Table E3: Immigration and wages exploiting past local shares in 1962

	Sample period: 1982-2016				Sample period: 1990-2016			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. Impact on the wage of native women							
Immigrant share	-0.23	-0.54***	-0.82**	-1.06***	-0.26	-0.61***	-0.94**	-1.23***
	(0.17)	(0.15)	(0.37)	(0.35)	(0.16)	(0.15)	(0.41)	(0.36)
Wild cluster bootstrap p-value	*0.20*	*0.04*	*0.09*	*0.00*	*0.19*	*0.03*	*0.01*	*0.02*
Log of native labor force	-	-	-0.23**	-0.20*	-	-	-0.23*	-0.19
			(0.11)	(0.10)			(0.13)	(0.12)
Wild cluster bootstrap p-value			*0.12*	*0.12*			*0.14*	*0.14*
Selectivity-corrected estimates	-	Yes	-	Yes	-	Yes	-	Yes
Inverse Mills ratio	-	0.18***	-	0.18***	-	0.19***	-	0.19***
		(0.02)		(0.02)		(0.02)		(0.02)
Kleibergen-Paap F-test	8.61	8.61	-	-	8.00	8.00	-	-
SW multivariate F-test (imm. share)	-	-	8.94	8.94	-	-	7.71	7.71
SW multivariate F-test (log nat.)	-	-	9.04	9.04	-	-	7.15	7.15
	B. Impact on the wage of native men							
Immigrant share	-0.87***	-0.86***	-0.76***	-0.77***	-0.95***	-0.95***	-0.87***	-0.87***
	(0.18)	(0.18)	(0.20)	(0.20)	(0.21)	(0.21)	(0.26)	(0.26)
Wild cluster bootstrap p-value	*0.02*	*0.02*	*0.01*	*0.01*	*0.04*	*0.04*	*0.02*	*0.02*
Log of native labor force	-	-	0.02	0.02	-	-	0.05	0.05
			(0.07)	(0.07)			(0.09)	(0.09)
Wild cluster bootstrap p-value			*0.72*	*0.75*			*0.49*	*0.49*
Selectivity-corrected estimates	-	Yes	-	Yes	-	Yes	-	Yes
Inverse Mills ratio	-	-0.01	-	-0.01	-	-0.00	-	-0.00
		(0.07)		(0.07)		(0.07)		(0.07)
Kleibergen-Paap F-test	6.05	6.05	-	-	5.31	5.31	-	-
SW multivariate F-test (imm. share)	-	-	9.80	9.80	-	-	8.52	8.52
SW multivariate F-test (log nat.)	-	-	9.95	9.95	-	-	7.96	7.96

Notes: Standard errors reported in parentheses are heteroscedasticity robust and clustered by region. The unit of observation is a region-year cell. The IV regressions in columns 1-4 use the original 1982-2016 cross-sections and have 110 observations (22 regions and 5 years); the IV regressions in columns 5-8 use the 1990-2016 cross-sections and have 88 observations (22 regions and 4 years). The dependent variable is the age- and education-adjusted wage of native women (Panel A) or men (Panel B). Columns 2-4 and 6-8 further adjust wages for sample selection. In columns 1-2 and 5-6, we instrument the share of immigrants with the shift-share instrument computed using the 1962 French census; columns 3-4 and 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 E4: Immigration and wages controlling for past immigrant inflows

	Sample period: 1982-2016				Sample period: 1990-2016			
	OLS estimates		IV estimates		OLS estimates		IV estimates	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Impact on the wage of native women								
Immigrant share	0.01	-0.36***	-0.02	-0.36***	-0.09	-0.51***	-0.09	-0.47***
	(0.07)	(0.08)	(0.08)	(0.09)	(0.07)	(0.08)	(0.09)	(0.10)
Wild cluster bootstrap p-value	*0.91*	*0.12*	*0.77*	*0.04*	*0.42*	*0.03*	*0.32*	*0.03*
Log of immigrant labor force in t-1	-0.03***	-0.02**	-0.04***	-0.04***	-0.03**	-0.02**	-0.04***	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Wild cluster bootstrap p-value	*0.29*	*0.29*	*0.00*	*0.02*	*0.30*	*0.29*	*0.00*	*0.03*
Selectivity-corrected estimates	-	Yes	-	Yes	-	Yes	-	Yes
Inverse Mills ratio	-	0.18***	-	0.18***	-	0.19***	-	0.19***
		(0.02)		(0.02)		(0.02)		(0.02)
SW multivariate F-test (imm. share)	-	-	69.82	69.82	-	-	63.84	63.84
SW multivariate F-test (log imm. in t-1)	-	-	64.41	64.41	-	-	56.23	56.23
B. Impact on the wage of native men								
Immigrant share	-0.74***	-0.73***	-0.82***	-0.82***	-0.85***	-0.85***	-0.90***	-0.90***
	(0.10)	(0.10)	(0.08)	(0.08)	(0.12)	(0.12)	(0.11)	(0.11)
Wild cluster bootstrap p-value	*0.17*	*0.17*	*0.00*	*0.00*	*0.24*	*0.24*	*0.00*	*0.00*
Log of immigrant labor force in t-1	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Wild cluster bootstrap p-value	*0.28*	*0.30*	*0.58*	*0.59*	*0.38*	*0.38*	*0.76*	*0.76*
Selectivity-corrected estimates	-	Yes	-	Yes	-	Yes	-	Yes
Inverse Mills ratio	-	-0.01	-	-0.01	-	-0.00	-	-0.00
		(0.07)		(0.07)		(0.07)		(0.07)
SW multivariate F-test (imm. share)	-	-	50.53	50.53	-	-	41.39	41.39
SW multivariate F-test (log imm. in t-1)	-	-	85.58	85.58	-	-	81.73	81.73

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 original 1982-2016 cross-sections and have 110 observations (22 regions and 5 years); the regressions in columns 5-8 use the 1990-2016 cross-sections and have 88 observations (22 regions and 4 years). The dependent variable is the age- and education-adjusted wage of native women (Panel A) or men (Panel B). Columns 2, 4, 6 and 8 further adjust wages for sample selection. Columns 3-4 and 7-8 instrument the share of immigrants and the log immigrant labor force with the shift-share instrument and the lagged predicted (log) size of immigrants computed using the 1968 French census. 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.

Appendix F: Additional Tests using Alternative Samples and Specifications

Tables F1 to F5 provide additional sets of robustness tests to assess the sensitivity of the baseline results reported in Table 4. These tables all have the same structure and reproduce (separately by gender) the regressions reported in columns 1, 5, 6 and 8 of Table 4 using alternative specifications, samples, variable definitions, and dependent variables.

Table F1 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 -0.12 (0.12) in the simplest IV model reported in column 2, increases to -0.49 (0.11) in column 3 when the regression adjusts for sample selection, and more than doubles to -1.06 (0.28) 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 -0.7 and -0.9.

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 F1 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 F2 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.[4] 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.[5] All the coefficients are similar to the baseline results in Table 4. The estimated effect of immigration on the female

[4] 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.

[5] 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.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

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 F3 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 (14a) and probit equation (15) 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 0.09 to -0.35 in the unweighted regression model and from -0.35 to -0.56 in the full interaction model.[6] In contrast, within each of the two alternative specifications, the estimated wage impact is relatively stable for native men (the wage elasticity is about -0.55 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 workers. Following Mulligan and Rubinstein (2008), columns 1-4 of Table F4 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 -0.7 and -0.8, 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 F4 extend the analysis by calculating the hourly wage rate for each worker in the sample.[7] 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 -0.37 (0.09) to -0.78 (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.[8]

Table F5 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,

[6] 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.

[7] 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 variable likely contains substantial measurement errors, which may affect the estimated wage elasticities (Barrett and Hamermesh, 2019; Laroque and Salanié, 2002).

[8] 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.

we use the geographically smaller definition of a department (of which there are 94).[9] This sampling framework significantly increases the number of cells and introduces much more variation in immigration and wages into the analysis.[10]

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 (18)) *and* the predicted number of immigrants in a given region as defined in equation (16). This extension of the shift-share approach helps capturing 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 (19)) 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 -0.84 (0.20) for women and -0.93 (0.23) for men.

There is an important conceptual difference between the baseline region-level results in Table 4 and the department-level results in Table F5. As we noted earlier, there is little evidence linking immigration and native internal migration at the *regional* level. There is, however, some evidence linking immigration and internal migration at the geographically smaller *department* level (Edo, Giesing, Poutvaara and Öztunc, 2019). This migration response to supply shocks likely creates an additional source of selection bias that was ignored in Table F5, which only corrected for the selection bias produced by the labor force participation decision. The estimation of the wage impact of immigration when natives endogenously respond along several margins simultaneously is likely to be far more complex than the narrower type of selection examined in this paper.

---

[9] 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.

[10] 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 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.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Table F1: 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	-0.17**	-0.12	-0.49***	-1.06***	-0.09	-0.08	-0.43***	-0.92***
	(0.07)	(0.12)	(0.11)	(0.28)	(0.07)	(0.10)	(0.10)	(0.27)
Wild cluster bootstrap p-value	*0.09*	*0.46*	*0.08*	*0.00*	*0.18*	*0.49*	*0.06*	*0.00*
Log of native labor force	-	-	-	-0.19**	-	-	-	-0.20**
				(0.09)				(0.08)
Wild cluster bootstrap p-value				*0.08*				*0.09*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.19***	0.19***	-	-	0.18***	0.18***
			(0.02)	(0.02)			(0.02)	(0.02)
Kleibergen-Paap F-test	-	27.12	27.12	-	-	32.07	32.07	-
SW multivariate F-test (imm. share)	-	-	-	11.32	-	-	-	12.90
SW multivariate F-test (log nat.)	-	-	-	11.92	-	-	-	15.15
B. Impact on the wage of native men								
Immigrant share	-0.82***	-0.90***	-0.90***	-0.81***	-0.69***	-0.81***	-0.80***	-0.70***
	(0.14)	(0.11)	(0.11)	(0.22)	(0.12)	(0.09)	(0.08)	(0.17)
Wild cluster bootstrap p-value	*0.28*	*0.01*	*0.01*	*0.01*	*0.21*	*0.00*	*0.00*	*0.01*
Log of native labor force	-	-	-	0.05	-	-	-	0.02
				(0.08)				(0.06)
Wild cluster bootstrap p-value				*0.52*				*0.77*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	-0.00	-0.00	-	-	-0.01	-0.01
			(0.07)	(0.07)			(0.07)	(0.07)
Kleibergen-Paap F-test	-	17.13	17.13	-	-	21.00	21.00	-
SW multivariate F-test (imm. share)	-	-	-	12.41	-	-	-	14.09
SW multivariate F-test (log nat.)	-	-	-	13.13	-	-	-	16.53

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.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Table F2: 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	-0.12	-0.14	-0.49***	-0.87***	-0.05	-0.07	-0.34***	-0.94**
	(0.08)	(0.11)	(0.10)	(0.23)	(0.06)	(0.09)	(0.09)	(0.39)
Wild cluster bootstrap p-value	*0.13*	*0.27*	*0.03*	*0.00*	*0.49*	*0.48*	*0.07*	*0.01*
Log of native labor force	-	-	-	-0.17***	-	-	-	-0.30**
				(0.07)				(0.15)
Wild cluster bootstrap p-value				*0.05*				*0.15*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.18***	0.18***	-	-	0.18***	0.18***
			(0.02)	(0.02)			(0.02)	(0.02)
Kleibergen-Paap F-test	-	23.77	23.77	-	-	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	-0.64***	-0.83***	-0.82***	-0.67***	-0.71**	-1.04***	-1.04***	-0.72***
	(0.16)	(0.10)	(0.10)	(0.16)	(0.28)	(0.16)	(0.16)	(0.18)
Wild cluster bootstrap p-value	*0.34*	*0.00*	*0.00*	*0.01*	*0.38*	*0.01*	*0.01*	*0.01*
Log of native labor force	-	-	-	0.04	-	-	-	0.08
				(0.06)				(0.06)
Wild cluster bootstrap p-value				*0.48*				*0.22*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	-0.01	-0.01	-	-	-0.01	-0.01
			(0.07)	(0.07)			(0.07)	(0.07)
Kleibergen-Paap F-test	-	17.26	17.26	-	-	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 F3: 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	0.48	0.09	-0.35	-0.43	-0.32***	-0.35***	-0.56***	-0.72**
	(0.32)	(0.35)	(0.32)	(0.56)	(0.09)	(0.11)	(0.10)	(0.29)
Wild cluster bootstrap p-value	*0.73*	*0.85*	*0.26*	*0.52*	*0.08*	*0.09*	*0.02*	*0.02*
Log of native labor force	-	-	-	-0.03	-	-	-	-0.05
				(0.18)				(0.11)
Wild cluster bootstrap p-value				*0.92*				*0.65*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.18***	0.18***	-	-	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	-0.27*	-0.56***	-0.56***	-0.54***	-1.09***	-1.23***	-1.21***	-0.87***
	(0.15)	(0.15)	(0.15)	(0.21)	(0.18)	(0.12)	(0.12)	(0.22)
Wild cluster bootstrap p-value	*0.23*	*0.03*	*0.03*	*0.01*	*0.30*	*0.00*	*0.00*	*0.02*
Log of native labor force	-	-	-	-0.03	-	-	-	0.11
				(0.10)				(0.08)
Wild cluster bootstrap p-value				*0.74*				*0.24*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	-0.01	-0.01	-	-	-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.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Table F4: 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	-0.09	-0.08	-0.27***	-0.72***	-0.37***	-0.37***	-0.78***	-1.18***
	(0.07)	(0.10)	(0.10)	(0.26)	(0.04)	(0.09)	(0.09)	(0.24)
Wild cluster bootstrap p-value	*0.19*	*0.47*	*0.13*	*0.03*	*0.00*	*0.05*	*0.01*	*0.00*
Log of native labor force	-	-	-	-0.19**	-	-	-	-0.18**
				(0.08)				(0.08)
Wild cluster bootstrap p-value				*0.10*				*0.03*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.12***	0.12***	-	-	0.21***	0.21***
			(0.01)	(0.01)			(0.02)	(0.02)
Kleibergen-Paap F-test	-	32.06	32.06	-	-	25.66	25.66	-
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	-0.70***	-0.81***	-0.82***	-0.70***	-0.71***	-0.76***	-0.76***	-0.70***
	(0.12)	(0.09)	(0.09)	(0.17)	(0.10)	(0.10)	(0.10)	(0.17)
Wild cluster bootstrap p-value	*0.21*	*0.00*	*0.00*	*0.01*	*0.10*	*0.01*	*0.01*	*0.01*
Log of native labor force	-	-	-	0.03	-	-	-	0.02
				(0.06)				(0.06)
Wild cluster bootstrap p-value				*0.67*				*0.80*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.01	0.01	-	-	-0.01	-0.01
			(0.08)	(0.08)			(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.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Table F5: 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	-0.16	-0.45***	-0.17	-0.42***	-0.27	-0.65***	-0.58***	-0.84***
	(0.10)	(0.07)	(0.11)	(0.09)	(0.20)	(0.18)	(0.19)	(0.20)
Wild cluster bootstrap p-value	*0.35*	*0.00*	*0.26*	*0.00*	*0.25*	*0.01*	*0.01*	*0.00*
Log of native labor force	-	-	-0.01	0.03	-	-	-0.21***	-0.15***
			(0.04)	(0.04)			(0.06)	(0.06)
Wild cluster bootstrap p-value			*0.83*	*0.45*			*0.01*	*0.02*
Selectivity-corrected estimates	-	Yes	-	Yes	-	Yes	-	Yes
Inverse Mills ratio	-	0.22***	-	0.22***	-	0.22***	-	0.22***
		(0.04)		(0.04)		(0.04)		(0.04)
Kleibergen-Paap F-test	-	-	-	-	7.40	7.40	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	14.76	14.76
SW multivariate F-test (log nat.)	-	-	-	-	-	-	11.86	11.86
	B. Impact on the wage of native men							
Immigrant share	-0.58***	-0.58***	-0.49***	-0.50***	-0.70***	-0.70***	-0.93***	-0.93***
	(0.10)	(0.10)	(0.13)	(0.13)	(0.17)	(0.17)	(0.23)	(0.23)
Wild cluster bootstrap p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
Log of native labor force	-	-	0.07	0.07	-	-	-0.05	-0.05
			(0.05)	(0.05)			(0.07)	(0.07)
Wild cluster bootstrap p-value			*0.17*	*0.16*			*0.51*	*0.52*
Selectivity-corrected estimates	-	Yes	-	Yes	-	Yes	-	Yes
Inverse Mills ratio	-	0.01	-	0.01	-	0.01	-	0.01
		(0.04)		(0.04)		(0.04)		(0.04)
Kleibergen-Paap F-test	-	-	-	-	7.91	7.91	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	17.03	17.03
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 departments 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.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Appendix G: 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 estimation strategy (Borjas, 2003), where the wages of specific skill groups are linked directly to the influx of immigrants into the particular skill group.

Table G1 reports the coefficients resulting from an extension of the baseline analysis where we divide the sample into two education groups, workers who have a baccalaureate degree and those who have not.

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 G1 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 -0.85 (0.18) in column 2 to -1.02 (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 -0.7. 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:

$$\log w_{irst} = \alpha \, P_{it} + \theta_{rst} + \varphi \, \lambda_{it} + \mu_{it}, \qquad (F1)$$

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:

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

$$\hat{\theta}_{rst} = \theta_r + \theta_s + (\theta_r \times \theta_s) + \theta_t + \alpha_M\, m_{rst} + \alpha_N \log N_{rst} + \nu_{srt}, \qquad (F2)$$

where the dependent variable is the mean (adjusted) wage of natives in cell (r, s, t), and is estimated from equation $(F1)$. Note that equation $(F2)$ 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 G2 reproduces the structure of our baseline Table 4 by showing the OLS and IV regression coefficients from equation $(F2)$. 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.[11] 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 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 G2 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.

[11] In columns 5-6, our instrument for m_{rst} is $\hat{m}_{rst} = \log(1 + \hat{M}_{rst}/\hat{N}_{rst})$. We predict $\hat{M}_{rst}$ and $\hat{N}_{rst}$ 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 $\log \hat{M}_{rst}$.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

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 empirical framework that pays careful attention to the self-selection of the native workforce and to the labor supply response induced by the supply shock.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Table G1: 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	-0.70***	-0.85***	-1.02***	-1.22***	0.16	0.20	-0.14	-0.66
	(0.12)	(0.18)	(0.20)	(0.33)	(0.11)	(0.17)	(0.16)	(0.42)
Wild cluster bootstrap p-value	*0.04*	*0.04*	*0.02*	*0.00*	*0.31*	*0.33*	*0.53*	*0.23*
Log of native labor force	-	-	-	-0.16	-	-	-	-0.16
				(0.11)				(0.16)
Wild cluster bootstrap p-value				*0.11*				*0.38*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	0.08***	0.08***	-	-	0.21***	0.21***
			(0.02)	(0.02)			(0.03)	(0.03)
Kleibergen-Paap F-test	-	21.03	21.03	-	-	45.35	45.35	-
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	-1.25***	-1.47***	-1.36***	-1.10***	-0.55***	-0.61***	-0.68***	-0.43*
	(0.26)	(0.17)	(0.14)	(0.23)	(0.14)	(0.13)	(0.13)	(0.22)
Wild cluster bootstrap p-value	*0.39*	*0.00*	*0.00*	*0.01*	*0.06*	*0.04*	*0.03*	*0.02*
Log of native labor force	-	-	-	-0.02	-	-	-	0.17
				(0.12)				(0.11)
Wild cluster bootstrap p-value				*0.88*				*0.17*
Selectivity-corrected estimates	-	-	Yes	Yes	-	-	Yes	Yes
Inverse Mills ratio	-	-	-0.21***	-0.21***	-	-	0.15	0.15
			(0.05)	(0.05)			(0.13)	(0.13)
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.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Table G2: 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**	-0.36**	-0.85***
	(0.08)	(0.20)	(0.10)	(0.21)	(0.10)	(0.15)	(0.18)	(0.26)
Wild cluster bootstrap p-value	*0.01*	*0.01*	*0.01*	*0.00*	*0.68*	*0.12*	*0.05*	*0.01*
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	-	-	-	-	48.72	48.72	-	-
SW multivariate F-test (imm. share)	-	-	-	-	-	-	84.40	84.40
SW multivariate F-test (log nat.)	-	-	-	-	-	-	556.19	556.19
	B. Impact on the wage of native men							
Immigrant share	-0.63***	-0.67***	-0.86***	-0.89***	-0.70***	-0.79***	-1.20***	-1.30***
	(0.21)	(0.23)	(0.25)	(0.27)	(0.19)	(0.19)	(0.16)	(0.17)
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)	-	-	-	-	-	-	95.01	95.01
SW multivariate F-test (log nat.)	-	-	-	-	-	-	900.26	900.26

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.

Supplemental Material (not copyedited or formatted) for: George J. Borjas, Anthony Edo. 2026. "Gender, Selection into Employment, and the Wage Impact of Immigration."
Journal of Labor Economics 44(2). DOI: https://doi.org/10.1086/732774.

Appendix H: Exclusion Restrictions in the Baseline Selection Model

The exclusion restriction requires that the selection "instruments" (i.e., those variables included in the first-stage probit, but excluded from the second-stage wage regression) are independent of the error term in equation ($14a$). Although the exclusion restriction is not testable, Huber and Mellace (2014) derive a statistical procedure to jointly test the validity of the exclusion restriction and the assumption that the error in the selection equation is additively separable. A rejection of the test therefore implies the violation of at least one assumption.

Suppose that the instrument Z is a binary variable equal to 1 in the subsample that is "treated" by the instrument and 0 otherwise, and that the treatment increases labor force participation.[12] The exclusion restriction is satisfied if the mean wage of workers not treated by the instrument lies in the interval:

$$E\left[\log w \mid I, Z=1, \log w \leq y_q\right] \leq E[\log w \mid I, Z=0] \leq E\left[\log w \mid I, Z=1, \log w \geq y_{1-q}\right], \quad (G1)$$

where I represents the event that the person is employed; $q = Pr(I \mid Z=0)/Pr(I \mid Z=1)$; and y_q is the q^{th} quantile in the wage distribution of treated workers. The test estimates the potential mean wage of workers who will always choose to work (irrespective of the value of Z), and determines if its value lies between the lower and upper bounds of the wage distribution of workers treated by the instrument. The parameter q can also be interpreted as the share of "always selected" workers (i.e., those who work regardless of the value of the instrument) in the population of always selected workers and "compliers" (i.e., those persons whose labor force participation depends on the value of the instrument).

Equation ($G1$) yields two inequality constraints for the null hypothesis that the exclusion restriction holds:

$$H_0: \begin{pmatrix} E\left[\log w \mid I, Z=1, \log w \leq y_q\right] - E[\log w \mid I, Z=0] \\ E[\log w \mid I, Z=0] - E\left[\log w \mid I, Z=1, \log w \geq y_{1-q}\right] \end{pmatrix} \equiv \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (G2)$$

Table H reports the test statistic given by the maximum of the standardized mean constraints—i.e., the maximum of (θ_1, θ_2) divided by the standard deviation of the observed outcome. We implement the test for the selection variables used in our baseline model (marital status, presence of young children, and home ownership), three alternative years (1982, 1999, and 2016), and using the entire sample and various subsamples of native men and women. The test statistics are almost always negative or close to zero and have large p-values, indicating that we cannot reject the validity of both the exclusion restriction and additive separability in our selection model.[13]

[12] Vytlacil (2002) shows that additive separability is equivalent to assuming that the potential employment outcome of an individual is monotonically related to the value of the instrument.

[13] Hubert and Mellace (2014) and Maasoumi and Wang (2019) also fail to reject the identifying assumptions when using the presence of young children as a selection variable in related studies of female wages in Malaysia, Portugal, and the United States.

Table H: Huber-Mellace test statistics

	Native women						Native men	
	Married		Presence of children below 6		Home ownership		Home ownership	
	Test statistic	P-value	Test statistic	P-value	Test statistic	P-value	Test statistic	P-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Sample year: 1982								
All individuals	0.01	1.00	-0.13	1.00	0.07	1.00	-0.04	1.00
All individuals living outside the Paris region	0.10	1.00	-0.11	0.91	0.38	0.99	0.37	0.19
All individuals living in the Paris region	0.05	1.00	-0.11	1.00	0.06	1.00	-0.05	1.00
High educated individuals	0.06	1.00	-0.21	1.00	0.18	1.00	0.28	1.00
High educated individuals below age 40	0.14	1.00	0.00	1.00	-0.10	0.55	-0.06	1.00
High educated individuals above age 40	-0.36	1.00	-2.87	1.00	-0.05	1.00	0.20	1.00
Low educated individuals	0.06	1.00	-0.14	1.00	0.08	1.00	-0.08	1.00
Low educated individuals below age 40	0.10	1.00	-0.02	1.00	-0.09	0.54	-0.32	1.00
Low educated individuals above age 40	-0.03	1.00	0.34	0.40	0.03	1.00	-0.13	1.00
B. Sample year: 1999								
All individuals	0.06	1.00	-0.02	1.00	0.04	1.00	0.08	1.00
All individuals living outside the Paris region	0.17	1.00	-0.10	1.00	0.29	0.00	0.35	0.00
All individuals living in the Paris region	0.10	1.00	-0.04	1.00	0.05	1.00	0.08	1.00
High educated individuals	0.23	0.00	-0.10	1.00	0.10	1.00	0.11	1.00
High educated individuals below age 40	0.21	0.00	0.14	1.00	-0.30	1.00	-0.28	1.00
High educated individuals above age 40	-0.03	1.00	-3.17	0.98	0.04	1.00	0.05	1.00
Low educated individuals	0.03	1.00	-0.12	1.00	0.05	1.00	0.04	1.00
Low educated individuals below age 40	0.03	1.00	0.07	1.00	-3.55	1.00	-0.25	1.00
Low educated individuals above age 40	-0.08	1.00	-0.02	1.00	0.01	1.00	0.01	1.00
C. Sample year: 2016								
All individuals	-5.52	1.00	0.13	0.51	-0.17	1.00	-0.11	1.00
All individuals living outside the Paris region	-3.43	1.00	-4.77	1.00	-0.01	1.00	0.17	1.00
All individuals living in the Paris region	-4.85	1.00	0.11	0.51	-0.15	1.00	-0.12	1.00
High educated individuals	-5.30	1.00	-4.93	1.00	-0.11	1.00	-0.03	1.00
High educated individuals below age 40	-5.16	1.00	-4.80	1.00	-0.34	1.00	-0.49	1.00
High educated individuals above age 40	-0.09	1.00	-4.81	0.98	-0.13	1.00	-0.02	1.00
Low educated individuals	-4.70	1.00	0.03	1.00	-0.19	1.00	-0.21	1.00
Low educated individuals below age 40	-3.89	1.00	0.05	1.00	-0.15	0.53	-0.49	1.00
Low educated individuals above age 40	-0.19	1.00	-0.09	0.95	-0.21	1.00	-0.20	1.00

Notes: The table reports the test statistics and the corresponding *p*-value using the method from Huber and Mellace (2014). The distribution of the test statistic is estimated using 1,000 bootstrap replications. We implement the test for the entire samples of men and women, and various subsamples. High educated individuals have a baccalaureate degree, while low educated ones do not.

References

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.

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

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