The Effect of Immigration on Wages: Exploiting Exogenous Variation at the National Level

By Joan Llull[*§]
MOVE, Universitat Autònoma de Barcelona, and Barcelona GSE

This version: March 2015

I estimate the effect of immigration on wages of native male correcting for endogenous allocation of immigrants across education-experience cells. Exogenous variation is obtained from interactions of push factors, distance, and skill cell dummies: distance mitigates the effect of push factors more severely for some skill groups. I propose a two-stage approach (Sub-Sample 2SLS) that estimates the first stage regression with an augmented sample of destination countries, and the second stage with a restricted sub-sample of interest. Asymptotic properties are derived. Results show important OLS biases. For U.S. and Canada, Sub-Sample 2SLS elasticities average -1.2, very stable across alternative instruments.

Keywords: *Immigration, Wages, Sub-Sample Two-Stage Least Squares*
JEL Codes: *J61, J31, C26.*

I. Introduction

With the resurgence of large scale immigration into OECD countries since 1960s, economists have been trying to assess whether and by how much immigration affects wages of native workers. This immigration wave has attracted so much attention in part because of its magnitude, and in part because of its composition (Card, 2009). Despite the big effort, however, there is still no consensus on what are the consequences of such worker inflows for wages of native workers.

In order to estimate the effect of immigration on wages, the literature compares, in alternative ways, the evolution of wages in labor markets that are exposed to different immigration shocks. Early studies defined labor markets geographically,

[*] MOVE. Universitat Autònoma de Barcelona. Facultat d'Economia. Bellaterra Campus – Edifici B, 08193, Bellaterra, Cerdanyola del Vallès, Barcelona (Spain). URL: http://pareto.uab.cat/jllull. E-mail: joan.llull [at] movebarcelona [dot] eu.

[§] I wish to thank Manuel Arellano, Stéphane Bonhomme, George Borjas, Julio Cáceres-Delpiano, Giacomo De Giorgi, Susanna Esteban, Nezih Guner, Tim Hatton, Jenny Hunt, Stephan Litschig, Enrique Moral-Benito, Björn Öckert, seminar participants at Uppsala University, Universitat Autònoma de Barcelona, and Universitat de Barcelona, and participants at the SAEe (Vigo); Barcelona GSE Winter Workshop; UCL-Norface Conference on Migration: Global Development, New Frontiers; SOLE Annual Meetings (Boston); EEA-ESEM Annual Meetings (Gothenburg); and ENTER Jamboree (Stockholm) for helpful comments and discussions. Christopher Rauh provided excellent research assistance. Financial support from European Research Council (ERC) through Starting Grant n.263600, and from the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centers of Excellence in R&D (SEV-2011-0075), is gratefully acknowledged.

mainly as metropolitan areas. More recent papers, pioneered by Borjas (2003), define labor markets at the national level as skill (education-experience) cells. These two approaches have a common complication: immigrants are not randomly allocated across labor markets. Because labor migration is mainly an economic decision, markets experiencing positive wage shocks tend to attract more immigrants. As a result, a positive correlation between immigration and wages is spuriously generated, which may bias upward the estimates of wage effects of immigration. This concern was already raised in the context of geographical studies by Altonji and Card (1991), who used past settlements of immigrants as instruments for current inflows, a strategy that became very popular since then. On the other hand, the literature has mostly ignored this issue when the analysis is done at national cross-skill cell level. Borjas (2003) acknowledges that "the immigrant share may also be endogenous [...] [if] the labor market attracts foreign workers mainly in those skill cells where wages are relatively high, [...] [in which case] results [...] should be interpreted as lower bounds of the true impact of immigration" (p.1349).

In this paper, I propose a novel approach to identify the effect of immigration on native male wages correcting for the non-random allocation of immigrants across skill cells. The identification strategy uses exogenous variation obtained from the interaction of three sources. First, push factors, which provide time-series variation. Four push factors are separately considered: wars, political regimes, natural disasters, and economic variables. Second, distance, which mitigates the effect of push factors, adding destination country variation. For instance, a war in the Balkans pushes more migrants to neighboring EU countries than to countries that are further away (e.g., see Angrist and Kugler, 2003). And third, skill-cell dummies, to capture that the mitigating effect of distance after push factor is more severe for specific groups of workers. Empirically, this happens to be the case for less educated and middle-aged (-experienced) individuals. The resulting interactions provide exogenous variation in immigration across skill cells, destination countries, and over time, which allows identification of wage elasticities to immigration in very demanding models.

The usage of the variation in distance for identification requires cross-destination country data. Still, for different reasons, the researcher might be interested in a single destination country (e.g. United States), or in a limited set of neighboring countries (e.g. United States and Canada), which limits this variation. This is the case in the present paper. The motivation for this restricted focus includes comparability with the existing literature, and data availability (information on wages in harmonized census microdata is only available for the United States and

Canada). I propose an alternative to the standard 2SLS approach (which I refer to as Sub-Sample 2SLS) that allows me to circumvent this complication. In the estimation of the first stage equation, I use all available European countries, the United States, and Canada. Then, the second stage sample is restricted to the subset of countries of interest (i.e. either the United States and Canada, or United States alone). Under not very restrictive (and partially testable) assumptions, this estimator provides consistent estimates of the effect of immigration on wages.

Theoretical properties and inference for the Sub-Sample 2SLS estimator are discussed. The estimator builds on two existing approaches in the literature: Two-Sample 2SLS (Angrist and Krueger, 1992; Arellano and Meghir, 1992), and Split-Sample IV (Angrist and Krueger, 1995). These methods combine moment conditions obtained from two independent samples. Unlike them, the Sub-Sample 2SLS uses two samples that are, by construction, not independent as they partially overlap. The Sub-Sample 2SLS estimator can generally be implemented in the analysis of data sets in which instruments and endogenous regressors are available for the whole sample, but the dependent variable is only available for a random sub-sample. This situation is very common in cross-country data, and in data sets that include supplements, like the March Supplement of the Current Population Survey, or the long list of supplements of the Panel Study of Income Dynamics, as these supplements are only available for a sub-sample of observations.

Results show that existing cross-skill cell analyses in the literature are substantially biased. OLS wage elasticities to immigration are estimated to be between -0.3 and -0.4, consistent with the literature (e.g. Borjas, 2003; Aydemir and Borjas, 2007, 2011). Sub-Sample 2SLS estimates average around -1.2, more than three times OLS counterparts. Interestingly, this result is very stable to the use of alternative push factors, which is remarkable because four push factors that are very uncorrelated with each other are considered. Even if wars, political regimes, and natural disasters were selecting a specific group of migrants (emergency-type), economic variables would select a very different one (economic migrants), still producing the same result. The strong similarity across local average treatment effects identified with so different instruments suggests that the proposed instruments may be consistently estimating the average treatment effect.

Two additional controversies from the literature can be analyzed under the current framework. First, Borjas (2003, 2006) find that, if labor markets are defined geographically and in terms of skills, estimated wage elasticities to immigration are smaller the more disaggregated is the geographical classification. Borjas (2006) provides evidence suggesting that local labor market impacts of immigration are arbitraged out through internal migration decisions. However, Card (2001) finds that intercity mobility rates of natives and earlier immigrants are insensitive to immigrant inflows. As an alternative explanation, Aydemir and Borjas (2011) propose measurement error as a potential source for these differences: immigrant shares calculated from public use Census microdata are computed with considerable noise, which is increasing with geographical disaggregation. This measurement error creates attenuation biases that are consequently larger at lower geographical levels. As the proposed instruments are uncorrelated with this measurement error, Sub-Sample 2SLS estimates should not suffer from attenuation bias, which provides a test of the measurement error hypothesis against the alternative of spatial arbitrage. To implement it, I reproduce the baseline analysis at a more disaggregated geographical level (nine divisions in the United States, and five big regions in Canada). The OLS gap between national and regional level estimates is partially closed in Sub-Sample 2SLS results. This suggests that measurement error is a relevant source of discrepancies between national and regional level results. Yet, even though estimates are not precise enough to reject that the difference between national and regional level Sub-Sample 2SLS estimates is zero, point estimates differ, suggesting that some role might be left to spatial arbitrage.

Second, despite being widely used in the literature, the networks instrument proposed by Altonji and Card (1991) has also been criticized (see Borjas, 1999). Baseline estimates in this paper are a reasonable benchmark to evaluate the performance of the networks instrument in the skill-cell analysis at national and regional levels. Results from different versions of the networks instrument are compared to Sub-Sample 2SLS estimates. In general, the networks instrument performs very poorly at the national level (which is not its natural application), in the sense that it produces estimates that are very similar to OLS and very different from baseline Sub-Sample 2SLS results. At the regional level, some modified versions of the instrument partially correct the endogeneity bias, although, in general, point estimates are below any of the regional level Sub-Sample 2SLS results.

The literature provides a wide range of estimates of wage elasticities to immigration, which are surveyed in Friedberg and Hunt (1995), Borjas (1999), Card (2005), and Kerr and Kerr (2011). Some studies, like Grossman (1982), Card (1990, 2001, 2005), LaLonde and Topel (1991), or Friedberg (2001) find, in general, small effects of immigration on native wages. Borjas (2003) and Aydemir and Borjas (2007, 2011) estimate wage elasticities to be between -0.3 and -0.4. Altonji and Card (1991), Goldin (1994), and Borjas, Freeman and Katz (1992), using different approaches, find elasticities that average around -1.2, similar to the estimated elasticities in this paper. The effect of immigration on other outcomes is also analyzed in the literature using cross-labor market comparisons. Examples of these outcomes include employment (Angrist and Kugler, 2003), prices of goods and services (Cortés, 2008), aggregate productivity (Llull, 2011), and housing rents (Saiz, 2007). Saiz (2007) and Cortés (2008) use the networks instrument. Angrist and Kugler (2003) and Llull (2011) use push-distance interactions as instruments (the former uses dummies for different episodes of Balkans War interacted with distance to the Former Yugoslavia). This variation (cross-country-time) would not provide identification for the models estimated in this paper, as it would be completely absorbed by country-time fixed effects. The effect of immigration on any of these outcomes could be estimated using the strategy proposed in this paper.

The rest of the paper is organized as follows. Section II gives a detailed description of the identification strategy, and discusses the theoretical properties and inference for the Sub-Sample 2SLS estimator. Section III presents some description of the data, including data sources, variable definitions, and a short exploration of some facts. Section IV presents the central results from the paper. Section V revisits some controversies in the literature. Section VI concludes.

II. Exogenous Variation at the National/Cross-Skill Level

A. Wage effects of immigration

Identification of wage effects of immigration requires the comparison of wages in labor markets that experience different immigration shocks. As labor markets are not observed experiencing counterfactual sequences of shocks, the comparison is made across similar labor markets with different levels of immigration. These labor markets can be defined in terms of skills, geographic regions, and/or time.

The standard approach in the literature estimates the following regression:

$$\ln w_s = \vartheta p_s + \boldsymbol{x}_s' \boldsymbol{\phi} + \upsilon_s, \tag{1}$$

where $\ln w_s$ is the log wage of natives in labor market s; $p_s \equiv M_s/(M_s + N_s)$ is the fraction of immigrants in the workforce; $\boldsymbol{x}_s = (x_{1s}, ..., x_{Hs})'$ is a vector of control variables that may include period, region, and skill dummies, their interactions, and/or any other variable that generates differences in wage levels across labor markets; and υ_s is an i.i.d. error term (Aydemir and Borjas, 2011).[1] Similar specifications have been used to estimate the effect of immigration on other outcomes:

[1] The wage elasticity is then $\frac{\partial \ln w_s}{\partial \tilde{p}_s} = \frac{\theta}{(1+\tilde{p}_s)^2}$, where $\tilde{p}_s \equiv M_s/N_s$ (see Borjas, 2003). This wage elasticity assumes labor markets are closed. More specifically, Equation (1) does not allow immigration into a given labor market s to affect wages in a different labor market s'. In a

employment (Angrist and Kugler, 2003), prices of goods and services (Cortés, 2008), aggregate productivity (Llull, 2011), and housing rents (Saiz, 2007).

A common problem with this approach is that immigrants are not randomly allocated across labor markets. As immigrants are moving in search of better economic opportunities, they are more likely to penetrate labor markets that experience positive wage shocks. As a result, υ_s and p_s may be positively correlated, which biases OLS estimates of ϑ upward. The literature that uses a geographical definition of labor markets have addressed this concern using past settlements of immigrants to instrument current inflows. This so-called "networks instrument" was first introduced by Altonji and Card (1991) and has been widely used ever since.[2] Despite its widespread usage, though, the instrument have also generated some controversy: if regional wage shocks are persistent over time, the instrument would be correlated with current wage shocks through past shocks, which would break the exclusion restriction (Borjas, 1999). Yet, an alternative instrument have been hard to find, with the exception of natural experiments.[3]

Partially driven by this concern, recent papers, starting by Borjas (2003), have changed the definition of labor markets to skill cells.[4] A general practice when using this definition of labor markets is to disregard the potential endogeneity of immigrant inflows in specific skill groups. However, as acknowledged by Borjas (2003, p.1349), a similar endogeneity problem may apply in this framework, which again would bias OLS estimates upward. Several papers in the literature analyze self-selection of immigrants in terms of skills (e.g. Borjas, 1987; Chiquiar and Hanson, 2005; Fernández-Huertas Moraga, 2011), and even though they do not agree on the exact pattern of self-selection, a general conclusion is that migrants are not randomly distributed across skill cells. They agree in that the differential

nested CES environment, like the one used in Borjas (2003, sec. VII) or Ottaviano and Peri (2012), this implies that the estimated elasticity would be an estimate of the *own* wage elasticity, provided that fixed effects for all nesting levels except the last one are included in the regression (Ottaviano and Peri, 2012). Thus, in that case the estimated elasticities measure the effect of increasing immigration on a given cell keeping the stock of immigrants in other cells constant on the wages of natives employed in cell of interest. Depending on the elasticity of substitution across labor markets, cross-effects are typically expected to be either less negative or positive.

[2] Recent examples are Card (2001), Card and Lewis (2007), Saiz (2007), Cortés (2008), Peri and Sparber (2009), Cortés and Tessada (2011), and Dustmann, Frattini and Preston (2013).

[3] Card (1990), Hunt (1992), Friedberg (2001), Glitz (2012), Monràs (2014), and Dustmann, Schönberg and Stuhler (2014) use different geopolitical events as natural experiments.

[4] Aydemir and Borjas (2007, 2011), Borjas (2008), Borjas, Grogger and Hanson (2010), Bratsberg and Raaum (2012), Bratsberg, Raaum, Røed and Schøne (2014), Carrasco, Jimeno and Ortega (2008), and Steinhardt (2011), among others, estimate a similar regression to that in Borjas (2003). Dustmann et al. (2013) combine regional and skill variation. Other papers, like Borjas, Freeman and Katz (1997), Ottaviano and Peri (2012) and Manacorda, Manning and Wadsworth (2012), undertake a more structural approach, using a production function with different skill groups in the spirit of Borjas (2003, sec.VII).

returns to skills in origin and destination countries are important determinants of migration decisions, which self-selects immigrants into specific skill cells as a reaction of cell-specific wage shocks.

The baseline version of Equation (1) implemented in this paper follows the base regression estimated by Borjas (2003) when combining geographical and skill-cell definitions of labor markets (Column 1, Table V, p.1353), which I later expand with additional combinations of fixed effects in the robustness section. Specifically:

$$\ln w_{ijkt} = \theta p_{ijkt} + \eta_i + \kappa_j + \iota_{kt} + \xi_{ik} + \zeta_{it} + \chi_{jt} + \varepsilon_{ijkt}. \tag{2}$$

Labor markets are defined by education $i = 1, ..., I$, experience $j = 1, ..., J$, country/region $k = 1, ..., K$, and time $t = 1, ..., T$. Different boundaries are used in the geographical component of the labor market definition: a single geographical market (United States), different countries (United States and Canada), and regions within countries (nine United States divisions and five big regions in Canada). Systematic differences across labor markets are captured by a set of market-specific effects: η_i, κ_j, ι_{kt}, ξ_{ik}, ζ_{it}, and χ_{jt}, which also capture unobserved persistence. Additional sets of fixed effects are included in some regressions. The remaining unobserved error term, ε_{ijkt}, is standard zero mean econometric error, potentially with $\mathbb{E}[p_{ijkt}\varepsilon_{ijkt}|\eta_i, \kappa_j, \iota_{kt}, \xi_{ik}, \zeta_{it}, \chi_{jt}] \neq 0$, as argued above.

B. Exogenous variation of immigration

Given the set of market-specific effects included in Equation (2), a valid instrument for p_{ijkt} needs to have variation across skill cells, destination countries/regions, and time. For instance, push factors will not identify θ by themselves, since they only provide variation over time, as neither will do distance, which only provides variation across geographical labor markets.

Analyzing the effect of immigration on employment, Angrist and Kugler (2003) interact dummies for three different episodes of the Balkans War in 1990s with distance between each destination country and the Former Yugoslavia, in order to get cross-country and time variation in the instrument. In a similar spirit, Llull (2011) uses interactions of wars/political regimes and distance to estimate how immigration affects aggregate productivity. In both cases, the relevance of the instrument comes from the fact that distance mitigates the effect of the push factor (e.g. a war in the Balkans is more likely to push migrants to European countries than to the United States). In the present context, these exogenous variables would still not provide enough variation to identify θ in Equation (2), as they are invariant across skill cells (and hence all their variation would be absorbed by the country/region-time effect, ι_{kt}).

Building on this idea, the relevant variation in the present paper comes from the observation that distance have a stronger mitigating effect on a push factor for individuals in some skill cells than in others. Section III provides suggestive evidence indicating that this is the case for less educated and middle-aged (middle-experienced) workers. For instance, less educated and middle-aged workers were overrepresented among migrants that moved from the Former Yugoslavia to European countries after the Balkans War, and underrepresented among those who migrated to the United States. In other words, European countries received more migrants from the Balkans than from any other destination in general, but especially so for less educated and middle-aged individuals.[5]

More formally, first stage coefficients are allowed to vary across skill cells. In particular, the first stage equation (at the bilateral level) is:

$$p_{ijqkt} = \alpha_{ij} r_{qt} \ln g_{qk} + \mu_i + \lambda_j + \varrho_{kt} + \psi_{ik} + \varsigma_{it} + \varphi_{jt} + \nu_{ijqkt}, \tag{3}$$

where p_{ijqkt} is the stock of immigrants with education i and experience j, from country q (for $q = 1, ..., Q$), living in country/region k in year t; $\ln g_{qk}$ is the log of the physical distance between origin country q and destination country/region k; r_{qt} is an exogenous push factor; α_{ij} is the coefficient associated to $r_{qt} \ln g_{qk}$ for education-experience cell ij; μ_i, λ_j, ϱ_{kt}, ψ_{ik}, ς_{it}, and φ_{jt} are fixed effects; and ν_{ijqkt} is a zero mean error term. Once this first stage regression is estimated, the 2SLS procedure implies obtaining the (excluded part of the) aggregate exogenous prediction of immigrant shares as:

$$\hat{p}_{ijkt} = \sum_q \hat{\alpha}_{ij} r_{qt} \ln g_{qk}. \tag{4}$$

In the empirical analysis below, I use four alternative push factors: wars, political regimes, natural disasters, and economic variables. The presence of wars, natural disasters, or bad economic conditions fosters migration. Regarding political regimes, well developed democracies are attractive locations to live in, and even though strong authoritarian countries might be unattractive, out-migration is often legally bounded; countries with weak political systems typically offer an environment of instability and uncertainty that encourages individuals to move. Wars, political regimes, and natural disasters, even though not very correlated

[5] Even though disentangling the underlying reasons that motivate this result is not a primary goal of this paper, a potential explanation could be that less educated individuals may be more likely to be financially constrained, and middle-aged may be more likely to carry dependant family members with them, which, in both cases, increase the cost of distance.

with each other, they all could be associated with emergency migration.[6] Economic variables, instead, are more connected to economic migration.

The exclusion restriction is such that:

$$\mathbb{E}\left[\left(\sum_q \alpha_{ij} r_{qt} \ln g_{qk}\right) \varepsilon_{ijkt} \middle| \eta_i, \kappa_j, \iota_{kt}, \xi_{ik}, \zeta_{it}, \chi_{jt}\right] = 0. \tag{5}$$

This implies that the differential projection of the push-distance interaction on the share of immigrants in different skill cells (but not necessarily the interaction itself) should be uncorrelated with the second stage error term ε_{ijkt}. This seems plausible for either of the four push factors.

C. An aggregated first stage

The natural way of estimating Equation (3) is by using bilateral migration data. However, computing immigrant shares for each country pair, skill cell, and point in time requires very large sample sizes. Even using census data like in this paper, sample sizes are in general too small to accurately compute immigrant shares for many country pairs.[7] The use of so noisy immigrant shares, although does not cause a bias in the estimation of θ (as long as the measurement error is uncorrelated with the instrument), reduces precision drastically.[8]

To address this issue, I estimate an aggregate version of Equation (3):

$$p_{ijkt} = \alpha_{ij} \left(\sum_q r_{qt} \ln g_{qk}\right) + \tilde{\mu}_j + \tilde{\lambda}_k + \tilde{\varrho}_t + \tilde{\psi}_{ik} + \tilde{\varsigma}_{it} + \tilde{\varphi}_{jt} + \tilde{\kappa}_{kt} + \nu^*_{ijkt}, \tag{6}$$

where the tildes indicate that the fixed effects from Equation (3) are multiplied by the total number of countries of origin, Q, and $\nu^*_{ijkt} = \sum_q \nu_{ijqkt}$. The two approaches are asymptotically equivalent, but in a finite sample, they provide different precision for the reasons described in the previous paragraph.

D. Sub-Sample Two-Stage Least Squares

Although, theoretically, parameter θ would be identified in the approach described above using data on a single destination country, identification based on

[6] The aggregate amount of immigrants in a country can be seen as a sum of binary individual decisions of whether to migrate or not. What we are wondering is whether the group of *compliers* selected by each of these instruments is representative of the population of interest.

[7] Aydemir and Borjas (2011) argue that a similar problem occurs in the computation of immigrant shares at the state or metropolitan area by skill group.

[8] Whether estimating the first stage regression at the bilateral level reduces or increases precision of the estimates is not clear *a priori*. Bilateral shares are noisily measured because of the aforementioned sample size concerns; however, the regression at the bilateral level exploits additional variation from the data, and the sample size used to estimate the first stage regression becomes larger. For the data used in this paper, the first effect seems to dominate.

multiple destinations exploits the variation provided by distance in the first stage, which increases efficiency. However, in this paper, comparability with existing literature and data availability (as wages are only available for United States and Canadian censuses) motivates focusing on the United States and Canada. For this purpose, I propose a two-stage approach that allows me to identify θ exploiting the variation in distance in the first stage but without need of using cross-country variation in the second stage. This approach, referred hereinafter as Sub-Sample 2SLS, estimates the first stage regression (6) using an expanded sample that includes the United States, Canada, and several European countries, and then estimates the structural Equation (2) with the restricted sample of destination countries (the United States and Canada, or the United States alone) using the predicted exogenous immigrant shares obtained from the first stage regression.

The approach builds on previous work in the literature that combines moments from different samples in estimation. Angrist and Krueger (1992) and Arellano and Meghir (1992) provide seminal work on the topic —the former introduce the Two-Sample IV estimator, and the latter propose a two-step method that combines moments from two different samples in a similar vein (Two-Sample 2SLS).[9] Angrist and Krueger (1995) introduce the Split-Sample IV estimator, which divides a sample into two independent sub-samples, and combines them in a Two-Sample IV to correct weak instruments bias. The identification strategy proposed here is comparable to Split-Sample IV in that it makes use of two different sub-samples of the same data set, but it differs in that these two sub-samples are, by construction, not independent of each other, as they partially overlap.

For notational simplicity, let $s = 1, ..., N$ be a general subindex for each unique combination $ijkt$, such that $N \equiv I \cdot J \cdot K \cdot T$. Then, let $y_s \equiv \ln w_{ijkt}$, $\boldsymbol{x}_s \equiv \left(p_{ijkt}, \boldsymbol{d}_{ijkt}^{fe\;\prime}\right)'$, where $\boldsymbol{d}_{ijkt}^{fe}$ is a vector of dummy variables to capture all fixed effects included in Equation (2), and $\boldsymbol{z}_s \equiv \left(\left(\sum_q r_{qt} \ln g_{qk}\right) \boldsymbol{d}_{ij}^{sc\prime}, \boldsymbol{d}_{ijkt}^{fe\;\prime}\right)'$, where $\boldsymbol{d}_{ij}^{sc}$ is a vector of skill cell dummies. Additionally, let $\boldsymbol{\beta} \equiv (\theta, \boldsymbol{\eta}', \boldsymbol{\kappa}', \boldsymbol{\iota}', \boldsymbol{\xi}', \boldsymbol{\zeta}', \boldsymbol{\chi}')'$, $\boldsymbol{\pi}_1 \equiv (\boldsymbol{\alpha}', \boldsymbol{\mu}', \boldsymbol{\lambda}', \boldsymbol{\varrho}', \boldsymbol{\psi}', \boldsymbol{\varsigma}', \boldsymbol{\varphi}', \boldsymbol{\nu}')'$, and Π be the projection matrix of $\boldsymbol{z}_s$ on $\boldsymbol{x}_s$, where $\boldsymbol{\pi}_1$ is the first column, an identity matrix of size $\dim\{\boldsymbol{d}_{ijkt}^{fe}\}$ is the bottom-right square block, and a matrix of zeros is the remaining block. And, finally, let $d_s \equiv \mathbb{1}\{k \in \{US, CAN\}\}$ (or eventually $d_s \equiv \mathbb{1}\{k = US\}$) be an indicator variable that takes a value of one if a given observation is included in the second

[9] Björklund and Jäntti (1997), Jappelli, Pischke and Souleles (1998), Currie and Yelowitz (2000), Dee and Evans (2003), Borjas (2004), and Almond, Doyle, Kowalski and Williams (2010) are examples of implementations. Inoue and Solon (2010) clarify a common confusion regarding their asymptotic distribution. Angrist and Pischke (2009) provide a textbook introduction.

stage sub-sample. Then, the Sub-Sample 2SLS estimator is given by:

$$\hat{\boldsymbol{\beta}}_{SuS2SLS} = \left(\sum_{s=1}^{N} d_s \hat{\boldsymbol{x}}_s \hat{\boldsymbol{x}}_s \right)^{-1} \sum_{s=1}^{N} d_s \hat{\boldsymbol{x}}_s y_s, \tag{7}$$

where:

$$\hat{\boldsymbol{x}}_s = \hat{\Pi}' \boldsymbol{z_s}, \text{ and } \hat{\Pi} = \left(\sum_{s=1}^{N} \boldsymbol{z}_s \boldsymbol{z}_s' \right)^{-1} \sum_{s=1}^{N} \boldsymbol{z}_s \boldsymbol{x}_s'. \tag{8}$$

In other words, the coefficients from the first stage equation, Π, are estimated with the full sample, and the resulting exogenous predictions of $\boldsymbol{x}_s$, $\hat{\boldsymbol{x}}_s$, are used to identify the structural parameters, $\boldsymbol{\beta}$, from the sub-sample selected by d_s.

E. Asymptotic properties and inference

Asymptotic results in the Two-Sample IV literature rely on the use of independent samples in estimation. These results are inapplicable here because, by construction, the two sub-samples are not independent from each other, as they partially overlap. Hence, the asymptotic properties of $\hat{\boldsymbol{\beta}}_{SuS2SLS}$ need to be explicitly discussed. The compact notation used in Equation (7) is convenient in the derivation of these asymptotic results following conventional arguments (standard 2SLS is, indeed, a special case of (7) in which $d_s = 1 \ \forall s$). This section highlights the main asymptotic results, and Appendix A provides detailed derivations.

In addition to the exclusion restriction in Equation (5), consistency requires that:

$$\mathbb{E}[d_s \boldsymbol{z}_s \varepsilon_s] = \mathbb{E}[d_s \boldsymbol{z}_s \nu_s] = 0. \tag{9}$$

This implies that the exclusion restriction is satisfied for the sub-sample selected by d_s, and that the relation between $\boldsymbol{x}_s$ and $\boldsymbol{z}_s$ is invariant across sub-samples.

If assumptions in Equation (9) hold, then:

$$\hat{\boldsymbol{\beta}}_{SuS2SLS} \underset{d}{\rightarrow} \mathcal{N}(\boldsymbol{\beta}, N^{-1} V_0) \tag{10}$$

with:

$$V_0 = \mathbb{E}[d_s \Pi' \boldsymbol{z}_s \boldsymbol{z}_s' \Pi]^{-1} \, \mathbb{E}[d_s \varepsilon_s^2 \Pi' \boldsymbol{z}_s \boldsymbol{z}_s' \Pi] \, \mathbb{E}[d_s \Pi' \boldsymbol{z}_s \boldsymbol{z}_s' \Pi]^{-1}, \tag{11}$$

and $\Pi \equiv \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1} \, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{x}_s]$.[10] V_0 is a version of the standard formula, computed for the sub-sample selected by d_s, where the regressor is $\hat{\boldsymbol{x}}_s$, provided that residuals are properly adjusted. By the analogy principle, a consistent estimator replaces expectations by sums, and ε_s by $\hat{\varepsilon}_s \equiv y_s - \boldsymbol{x}_s' \hat{\boldsymbol{\beta}}_{SuS2SLS}$.

[10] In all derivations, I follow the literature (e.g. Borjas, 2003) in assuming that p_{ijkt} is observed without error in the data. Sample sizes in different censuses are large enough for this assumption to be plausible. Additionally, sample size weights are used in the estimation.

The assumptions in Equation (9) are central in the derivation of the asymptotic results. The first condition, $\mathbb{E}[d_s \boldsymbol{z}_s \varepsilon_s] = 0$, is by construction not testable; it is not even so against the alternative that the exclusion restriction is only satisfied in the whole sample, because only $d_s y_s$ and not y_s is observed. Yet, the second condition, $\mathbb{E}[d_s \boldsymbol{z}_s \nu_s] = 0$, can be tested. Specifically, $\Delta = 0$ in the regression:

$$\boldsymbol{x}_s = \Gamma' \hat{\boldsymbol{x}}_s + \Delta' d_s \hat{\boldsymbol{x}}_s + \epsilon_s, \tag{12}$$

is a necessary and sufficient condition for $\mathbb{E}[d_s \boldsymbol{z}_s \nu_s] = 0$ (see Appendix A). Put differently, the relation between predicted and actual regressors needs to be stable across sub-samples. A significance test for Δ is implemented in the analysis below.

III. Data

A. Data construction and sample description

The empirical analysis below combines information from several data sets. Immigrant shares are computed from census microdata for different countries. These data are extracted from IPUMS-International (Minnesota Population Center, 2011), which includes harmonized variables across countries and years. Immigrant shares are calculated for Austria, Canada, France, Ireland, Switzerland, and the United States for years 1970, 1980, 1990, and 2000. An expanded (unbalanced) sample that includes additional countries (the Netherlands, Italy, Portugal, and Spain) and additional dates (1960) is used in some specifications. Immigrant shares are computed for men aged 18-64 who participate in the civilian labor force (women are also included in some specifications). Immigrants are defined differently across countries. Whenever birthplace and citizenship are available, a person is defined as an immigrant if she is foreign-born and either a noncitizen or a naturalized citizen. Otherwise, the available pieces of this rule are implemented. The definition used for each country is consistent across years. Skill cells are defined by education and experience. Education is divided in three harmonized groups: primary or less, secondary, and tertiary. Experience, defined as number of years since school completion, is divided into 5 eight-year categories: <8, 8-15, 16-23, 24-31, and 32+ years. This classification delivers 15 skill cells per year and country. Sample selection and variable definitions are described in more detail in Appendix B1.

Table 1 lists sample sizes for each census. The average sample size is around 500,000 observations (including natives and immigrants), and there is variation across countries and over time. This size is large enough to compute immigrant shares at the skill-cell level with precision, but it is too small to compute them

Table 1—Sample Sizes from Different Censuses

	1960	**1970**	**1980**	**1990**	**2000**
Austria	—	**180,780**	**192,059**	**208,620**	**214,646**
Canada	—	**52,939**	**136,637**	**225,622**	**216,167**
France	600,469	**629,309**	**670,378**	**583,945**	**714,001**
Greece	—	**207,688**	**231,274**	**238,808**	**263,091**
Ireland	—	**57,849**	**80,940**	**83,706**	**99,088**
Italy	—	—	—	—	751,678
Netherlands	—	36,356	—	—	54,640
Portugal	—	—	119,109	115,923	124,878
Spain	—	—	530,065	500,859	518,982
Switzerland	—	**86,699**	**88,760**	**104,870**	**99,010**
United States	437,305	**474,621**	**2,871,935**	**3,194,928**	**3,428,515**

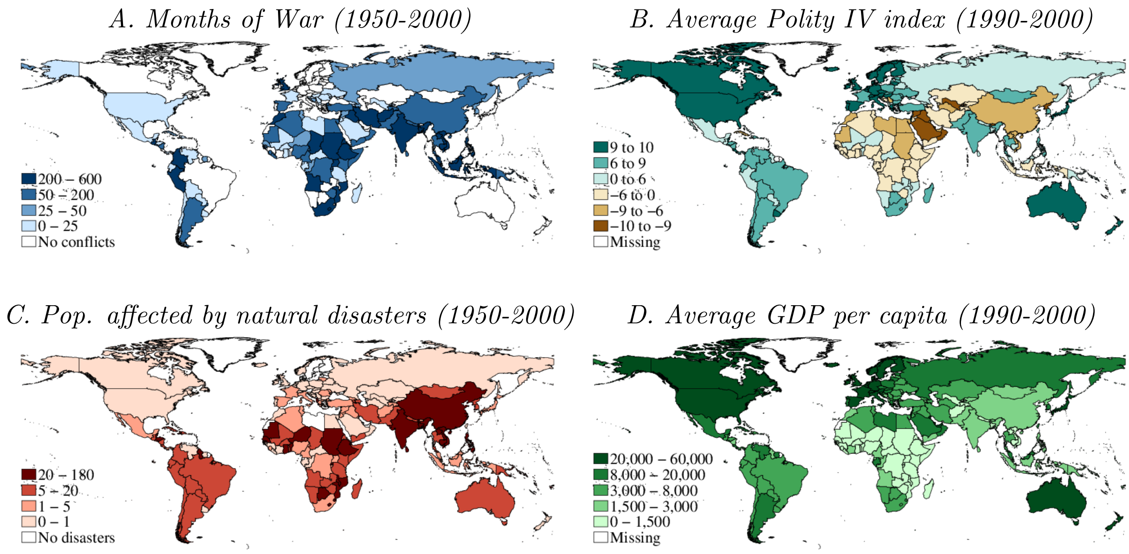
Note: The table reports the number of observations used in the computation of immigrant shares. Baseline balanced sample in bold. Samples are restricted to active male (working or unemployed) aged 18-64 with available information on country of origin and education.

for each country of origin. For example, if individuals were spread uniformly across skill cells, immigrant shares per year and destination country would be calculated, on average, with around 33,300 observations (500,000 individuals/15 cells), which would deliver very precise estimates. Even for the smallest samples, the shares would be computed with several thousands of observations. However, with an immigrant share of around 9% on average, if immigrants were uniformly distributed across countries of origin, even the average sample would only include (500,000 indiv.×9% immigrants)/(15 cells×188 countries)≈16 immigrants from each country. This situation justifies the use of an aggregated instead of a bilateral first stage regression, as discussed in Section II.C.

Native male earnings data, which are only available for the United States and Canada, are also drawn from IPUMS-International. To compute (monthly) average wages in each skill cell, the sample is further restricted to wage/salary employees who worked in the year prior to the survey, are not enrolled neither in school nor in the armed forces, and do not live in group quarters.

The instruments are built for 188 countries of origin, which are listed in Appendix B2, which also describes data sources and variable definitions. Distance is measured as the physical distance between the centroid of the most populated city of each country of a country-pair. Four push factors are considered: wars, political regimes, natural disasters, and economic variables. Wars, obtained from PRIO (Gleditsch, Wallensteen, Eriksson, Sollenberg and Strand, 2002), are measured as the number of months that a given country was involved in a civil war or a conflict in the preceding decade. Political regimes are represented by an indicator constructed from the Polity IV index (Marshall, Jaggers and Gurr, 2010),

Figure 1. Push Factors



Note: Top-left map: cumulative number of months in the 1950-2000 period that a country was involved in a civil war or conflict. Top-right map: average Polity IV index for the country during 1990-2000 (9 to 10 is Full Democracy, 6 to 9 is Democracy, 0 to 6 is Open Anocracy, -6 to 0 is Closed Anocracy, and -9 to -6 is Autocracy, and -10 to -9 is Strong Autocracy; see Marshall, Jaggers and Gurr (2010)). Bottom-left map: average fraction of the population (per 1,000 inhabitants) affected by natural disasters per year between 1950 and 2000. Bottom-right map: average GDP per capita for 1990-2000.

an index that ranges from -10 (strong autocracies) to 10 (full democracies). A value close to 0 indicates "anocracy", a regime-type where power is not vested in public institutions but spread amongst elite groups who are constantly competing with each other for power. As anocracies are typically the least resilient political system to short-term shocks (they create the promise but not yet the actuality of an inclusive and effective political system, and threaten members of the established elite), they generate uncertainty and are very vulnerable to disruption and armed violence; for this reason, they are more likely to foster migration. An indicator takes the value of 1 if the average of the index over the preceding decade is below -6 or above 6, and 0 otherwise is used. Natural disasters, calculated from EM-DAT database (EM-DAT, 2010), are measured as the fraction of the population affected (needed immediate assistance, displaced, or evacuated) by natural disasters (droughts, earthquakes, floods, and storms) over the preceding decade. And, economic conditions are measured as log average real GDP per capita in the preceding decade, obtained from Penn World Tables (Heston, Summers and Aten, 2012). Alternative push and distance variables are used as robustness checks.

Figure 1 plots the incidence of push factors across origin countries. Figure 1A shows the cumulative number of months of war in each country in years 1950-2000. Figure 1B presents average Polity IV indexes for 1990-2000. Figure 1C plots the

TABLE 2—REGIONAL DISTRIBUTION OF NET INFLOWS OF MIGRANTS ACROSS SELECTED COUNTRIES BY EDUCATIONAL LEVEL AND CONTINENT OF ORIGIN (1990-2000)

	Total	Primary	Secondary	Tertiary
i. Africa				
Australia/New Zealand	3.85	2.03	1.26	6.81
Europe	70.58	85.62	83.19	52.16
U.S./Canada	25.57	12.35	15.55	41.03
ii. Americas				
Australia/New Zealand	0.46	0.07	0.29	1.27
Europe	8.65	1.28	18.66	13.65
U.S./Canada	90.90	98.64	81.06	85.07
iii. Asia				
Australia/New Zealand	6.79	7.88	6.55	6.36
Europe	28.04	46.70	40.04	15.41
U.S./Canada	65.16	45.42	53.41	78.24
iv. Europe				
Australia/New Zealand	-4.90	—	-25.89	7.84
Europe	110.00	—	125.61	58.12
U.S./Canada	-5.10	—	0.28	34.04
v. Oceania				
Australia/New Zealand	51.56	113.57	39.01	45.92
Europe	25.73	-37.37	42.29	29.54
U.S./Canada	22.71	23.80	18.69	24.53

Note: The table shows the regional distribution of net inflows of migrants (differences in stocks) in selected destination countries by continent of origin. European destination countries include EU-15 (excluding Luxembourg and Ireland), Norway, and Switzerland. Primary educated migrants from Europe omitted due to negative aggregate inflow. *Data source*: Docquier and Marfouk (2006).

average fraction of the population affected by natural disasters per year between 1950 and 2000. And Figure 1D presents average real GDP per capita for 1990-2000. All plots show substantial variability across countries, and little overlap.

B. *Descriptive evidence for heterogeneous first stage coefficients*

The identification strategy described above exploits the presence of a differential mitigating effect of distance across skill cells. In the following lines, I briefly present some suggestive evidence that points towards this heterogeneity. I also propose some tentative examples on why this could happen. It is important to note that the orthogonality of the instruments does not hinge on these specific examples, as it is, in any case, unlikely that cell-specific wage shocks in a destination country are correlated with, say, wars or natural disasters in origin countries or the distance to them. Likewise, this suggestive evidence does not aim at establishing relevance for the instrument, which is more formally discussed in Section IV.

Table 2 presents the regional distribution of net inflows of immigrants across

Table 3—Differential Mitigation Effect of Distance on the Correlation between Push Factors and Migration at Different Educational Levels (1990-2000)

	Total	Primary	Secondary	Tertiary
Conflict dummy	-0.213	-0.771	-0.166	-0.017
	(0.114)	(0.460)	(0.175)	(0.156)
Political regimes	0.036	0.676	-0.302	-0.214
	(0.098)	(0.469)	(0.212)	(0.109)
Affected by natural disasters	0.245	-0.270	0.940	0.647
	(0.016)	(0.376)	(0.402)	(0.155)
GDP per capita growth	0.257	0.661	0.007	-0.072
	(0.126)	(0.251)	(0.225)	(0.117)

Note: The table reports estimated β_3 coefficients from the following regression fitted to different samples:

$$\Delta m_{qk} = \beta_0 + \beta_1 push_q + \beta_2 \ln dist_{qk} + \beta_3 push_q \times \ln dist_{qk} + u_{qk},$$

where q indicates origin country, k indicates destination country, $push_q$ is the corresponding push factor, $\ln dist_{qk}$ is the (log) distance between country q and country k, and m_{qk} is the change between 1990 and 2000 in the fraction of country k's workforce (of a given educational group) that is a migrant from country q. One push factor at a time is introduced in each panel. Different columns present estimates for different educational groups. Destination countries included in the sample are as in Table 2. Standard errors, in parenthesis, are clustered by origin country. *Data source*: Docquier and Marfouk (2006).

selected OECD countries by continent of origin and educational level. Given the aforementioned sample size limitations of the available census microdata, this information is obtained from Docquier and Marfouk (2006), who report immigrant stocks by educational level and country of origin across OECD countries in 1990 and 2000. The table presents the fraction of net migration flows (difference in stocks) absorbed by each group of destination countries. A first observation is that distance matters in determining where to migrate (e.g. migrants from Africa and Europe mostly move to European countries, migrants from the Americas move to the United States and Canada, and Oceanian migrants mostly go to Australia and New Zealand). More importantly, distance seems to play a more important role for primary educated compared to tertiary educated. For instance, Europe receives 86% of primary educated African migrants and only 52% of those with tertiary education, whereas the United States/Canada receive 12% and 41%. On the contrary, the United States and Canada receive 99% of all primary educated migrants from the Americas versus 85% of those with tertiary education, while European countries receive respectively 1% and 14%. An analogous pattern is observed for Oceania with Oceanian migrants.[11]

A question remains on whether the differential role of distance across educational levels operates on migrants that move in reaction to a push shock. Using the

[11] Table C1 in Appendix C provides some specific examples of migration from countries that suffered selected war or disaster episodes during 1990s, pointing in the same direction.

Table 4—The Relation Between Distance and Migration to the United States after Selected Push Factors by Skill Level

	Conflicts		Political regimes		Natural disasters		GDP p.c. growth	
i. By Education								
Primary	-0.618	(0.456)	-1.027	(0.705)	-0.437	(0.338)	-0.135	(0.070)
Secondary	-0.069	(0.048)	-0.098	(0.067)	-0.041	(0.034)	-0.023	(0.007)
Tertiary	0.002	(0.021)	-0.007	(0.019)	0.018	(0.015)	0.004	(0.008)
ii. By Experience								
0-7 years	-0.085	(0.065)	-0.173	(0.121)	-0.064	(0.060)	-0.027	(0.015)
8-15 years	-0.161	(0.127)	-0.239	(0.167)	-0.095	(0.083)	-0.035	(0.017)
16-23 years	-0.106	(0.075)	-0.165	(0.113)	-0.069	(0.056)	-0.025	(0.012)
24-31 years	-0.063	(0.043)	-0.115	(0.078)	-0.039	(0.040)	-0.017	(0.012)
31+ years	-0.062	(0.043)	-0.093	(0.063)	-0.032	(0.032)	-0.010	(0.020)

Note: The table reports estimated β_1 coefficients from the following regression fitted to different samples:

$$\Delta m_{qt} = \beta_0 + \beta_1 \ln dist_q + u_{qt},$$

where q indicates country of origin, t indicates Census year, $dist_q$ is the distance between country q and the U.S., and m_{qt} is the period t fraction of the workforce (with the given educational or experience level) that is from country q. Regressions are estimated with a sample of countries/periods in which there is: a war (first column), an anocracy regime type (second column), a natural disaster (third column), and negative average GDP per capita growth rate (fourth column). Each row is estimated for a given level of education or experience. Standard errors, in parenthesis, are clustered by origin country.

same data, this question is addressed in Table 3. The table presents the estimated interaction coefficients of a set of regressions of net migration for a pair of countries in 1990-2000 on a given push factor, (log) distance between origin and destination countries, and their interaction. These regressions are estimated separately for each educational level and push factor. Results are analogous to Table 2.

Data availability prevents the replication of the same exercise for experience levels. Instead, I focus on the United States as a destination country, for which I can compute immigrant shares by age level for a large fraction of origin countries. I take the sample of origin country-periods experiencing a *positive* push factor (i.e. a war, anocracy, a natural disaster, or negative GDP per capita growth), and, for each education or experience group, I regress the share of immigrants from that country on log distance. Results are presented in Table 4. In the upper panel, the same conclusions as in Table 3 are reached, except that, given the much smaller number of observations, precision is lower.[12] For experience, results suggest that the effect of the push shock is more mitigated by distance in the case of middle experienced (middle aged) individuals.

[12] Note that the signs of political regimes and GDP per capita switches because a *positive* push factor implies a smaller value of the instrument in both cases.

IV. Results at the National Level

We now turn into the estimation results. This section presents different estimates for parameter θ in Equation (2) obtained from applying the methodology described earlier to different second-stage sub-samples, using alternative instruments, and alternative combinations of fixed effects. Before that, first stage results are discussed, with emphasis on testing the validity of the above assumptions.

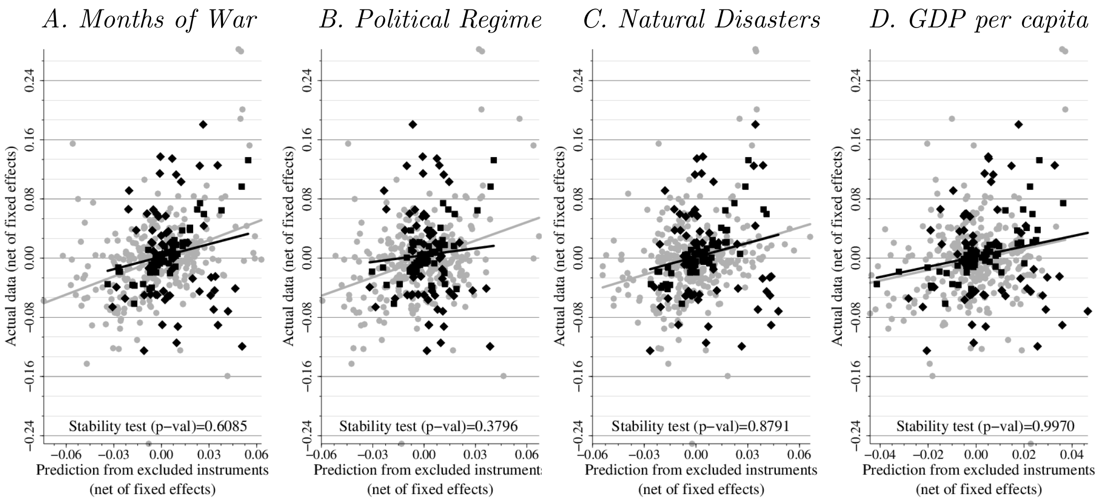
A. First stage results

Because the second stage results presented in the paper correspond to over a hundred different first stage regressions, this section discusses the main general results, emphasizing the baseline specifications.[13] Coefficients for the four alternative excluded instruments in the baseline specifications are displayed in Table D1 in Appendix D. Point estimates are consistent with the evidence described in Section III. The coefficients should be interpreted relative to a base category, that is 0-8 years of potential experience, primary educated. For push factors that are positively associated with migration probabilities (months of war and natural disasters), a negative coefficient for a given cell means that distance mitigates the effect of the push factor by less than the baseline category. For push variables constructed such that they are negatively associated with migration, the reverse is true. For wars and natural disasters, the mitigating effect of distance is particularly severe for primary educated with 9-16 and 17-24 years of potential experience, and the least severe for tertiary educated with 0-8 and +32 years of potential experience. A similar pattern emerges for natural disasters. For the political regime indicator and for log GDP per capita, the mitigating effect of distance is again clearly marked for primary educated compared to other education levels, but the patterns across potential experience levels are flatter.

The Sub-Sample 2SLS approach requires the assumptions in Equation (9) to be satisfied, in addition to the standard conditions. Equation (12) proposes a simple test for the condition $\mathbb{E}[d_s \boldsymbol{z}_s \nu_s] = 0$: the relation between predicted and actual immigrant shares (net of fixed effects) should be stable across sub-samples. This test is implemented in Figure 2 for the balanced sample. In the figure, scatter diagrams plot residuals from regressions of actual and predicted immigrant shares on education, experience, country-period, education-period, experience-period, and education-country dummies. Black points indicate observations for

[13] Detailed first stage results for any regression estimated in the paper are available from the author upon request.

Figure 2. Stability of First Stage Predictions Across Sub-Samples



Note: Black: United States (squares) and Canada (diamonds). Gray: other countries included in the balanced panel. Scatter diagrams relate the share of immigrants in each education-experience-period-country cell with the corresponding prediction using the indicated set of instruments. Both actual and predicted shares are net of education, experience, country-period, education-period, experience-period, and education-country fixed effects. Lines represent a fitted regression for each sub-sample. P-values of the stability test described in the text are presented at the bottom of each figure.

the United States (squares) and Canada (diamonds), and gray points indicate observations for Austria, France, Greece, Ireland, and Switzerland, which are the countries included in the balanced sample. The relation between actual and predicted immigrant shares is very stable across sub-samples. Plotted regression lines for each sub-sample have very similar slopes, and the position of the different points throughout the plots overlap substantially. More formally, the p-values of the test, presented at the bottom of each figure, clearly cannot reject the null hypothesis of stability across sub-samples in any of the cases. Similarly, stability cannot be rejected for any of the baseline first stage regressions, as shown in Table D1. This suggests that the stability condition of the first stage regression is satisfied, and, hence, that the approach is valid in this context.

For the baseline first stage regressions, F-tests of joint significance of the coefficients of the excluded regressors fluctuate around an average of 5.3, with some differences across alternative instruments (they average 7.7, 5.0, 5.3, and 3.1 for wars, political regimes, disasters, and GDP per capita in the three different specifications considered as baseline, presented in Table D1). As a reference, the relevant Stock and Yogo (2005) critical values for the weak instruments test are 4.67, 6.45, and 11.52 for maximum relative biases of 0.3, 0.2, and 0.1 respectively. Because the weak instruments bias of IV is towards OLS, these F-statistics imply that the Sub-Sample 2SLS estimates presented below could still be a lower bound of the negative immigration, as a maximum relative bias of 0.2–0.3 could be committed.

This maximal bias would imply that, for instance, if OLS elasticities were -0.4 and Sub-Sample 2SLS counterparts were -1.2, the true elasticities would be between -1.2 and -1.54 (i.e. with a relative bias of 0.3, the Sub-Sample 2SLS bias would be equal to $(1.2 - 0.4) \times (1 - 1/0.3) = -0.34$). Therefore, the estimates presented below, are, if anything, conservative. Nonetheless, notice that, as shown in Table 8, the estimation results below are stable to the use of a host of different instruments, in some cases with excluded F statistics ranging up to above 16.

B. Benchmark estimation: United States and Canada

Estimation results for parameter θ from a sample that includes both the United States and Canada as destination countries are presented in Table 5. Each parameter estimate (and standard error) in the table is obtained from a different regression. Different rows include different specifications for Equation (2), and different columns are estimated with different instruments, as indicated. All regressions are weighted by the sample size used to calculate average wages in each skill cell, except those in second and third rows, unweighted and weighted using sample sizes used to compute immigrant shares respectively. Regressions in the fourth row are estimated with the expanded unbalanced panel. In the fifth row, annual instead monthly wages are used as a dependent variable. In the last row, both male and female are used to compute immigrant shares, unlike in other rows, where these are computed counting only males.

The first column presents OLS results. Point estimates are very similar to previous estimates in the literature. The baseline coefficient is -0.556, with a standard error of 0.130. Borjas (2003) finds a point estimate of -0.572 for weekly earnings in the United States, and Aydemir and Borjas (2007, 2011) find -0.507 in Canada. This estimate implies a wage elasticity evaluated at the mean value of the immigrant supply increase in the United States of -0.38 (-0.35 if it is evaluated at the average supply increase in Canada).[14] With this elasticity, a 10 percent immigrant-induced increase in the number of workers in a particular skill group would reduce the wage of that group by 3.5-3.8%. Point estimates are very similar across different specifications. In general, the implied elasticities range between -0.28 and -0.45.[15]

[14] This elasticity is computed as in footnote 1. By year 2000, immigration had increased male labor force in the United States by 16.8 percent, and, as a result, the wage elasticity is obtained multiplying the coefficient by approximately 0.7 (Borjas, 2003). For Canada, this increase was of 25.8 percent, which implies multiplying the coefficient by 0.63 (Aydemir and Borjas, 2007).

[15] As noted in Ottaviano and Peri (2012), this elasticity is an estimate of the "own" wage elasticity. In other words, it describes how the wages of natives in a given cell would be affected by the increase of immigration in that cell.

Table 5—The Effect of Immigration on Native Male Wages: U.S. and Canada

	OLS	Sub-Sample 2SLS			
		Months of war	Political regime	Natural disasters	GDP per capita
Baseline	-0.556	-1.655	-1.856	-1.691	-1.774
	(0.130)	(0.648)	(0.706)	(0.723)	(0.796)
Unweighted regression	-0.400	-1.187	-1.190	-1.272	-1.154
	(0.157)	(0.809)	(0.754)	(0.847)	(0.727)
Weighs are sample sizes for shares	-0.563	-1.670	-1.856	-1.709	-1.782
	(0.133)	(0.978)	(0.980)	(0.971)	(0.924)
Unbalanced panel	-0.558	-2.067	-2.285	-2.015	-2.216
	(0.132)	(1.109)	(1.201)	(1.176)	(1.433)
Log annual wages	-0.639	-1.653	-1.929	-1.784	-1.818
	(0.235)	(0.819)	(0.918)	(0.891)	(0.989)
Includes female in LF counts	-0.621	-1.666	-1.850	-1.694	-1.773
	(0.134)	(0.567)	(0.623)	(0.635)	(0.707)

Note: The table reports the coefficient of the immigrant share from regressions where the dependent variable is the average log wage for native males aged 18-64 in each education-experience-period-country cell (monthly wage, except otherwise indicated). Each row is a different specification; each column uses a different set of instruments. All regressions include 120 observations in the second stage, except those estimated with the unbalanced panel (135). All regressions are weighted by the sample size used to compute wages in each cell, except otherwise indicated. All regressions include education, experience, country-period, education-period, experience-period, and education-country fixed effects. Standard errors, in parenthesis, are computed as derived in Appendix A.

Sub-Sample 2SLS estimation results are presented in the remaining four columns. Each column uses the instruments generated by the push factor indicated at the top row. Baseline point estimates range between -1.655 (s.e. 0.648), using conflicts as push variation, and -1.856 (s.e. 0.706), using political regimes; the estimated coefficient from natural disasters is -1.691 (s.e. 0.723), and the one for GDP per capita is -1.774 (s.e. 0.796), exactly in the center of the range. These estimates imply elasticities ranging from -1.15 to -1.30, between 3 and 3.5 times larger than OLS estimates. A similar pattern is sustained across different specifications. In general, Sub-Sample 2SLS estimates are around 3 times larger than OLS counterparts, and implied wage elasticities average around -1.2.[16]

One of the key features of the results presented in Table 5 is that point estimates are very stable across specifications that use different instruments. This is surprising because the instruments used across columns are very different. Table 6 shows cross-origin country/time correlation between the four push factors used

[16] Altonji and Card (1991) indeed find that "a 1 percentage point increase in the fraction of immigrants in an SMSA reduces less-skilled native wages by roughly 1.2 percent" (p.226). That paper is among the very few geographical level studies in the literature that find a substantial effect of immigration on wages. Other studies that obtain a similar result include Borjas et al. (1992), using a time series approach, as shown by calculations in Friedberg and Hunt (1995), and Goldin (1994), using data for 1890-1921.

Table 6—Cross-Origin Country Correlation Across Push Factors

	Months of war	Political regime	Natural disasters	GDP per capita
Months of War	1.000			
Political regimes	-0.184	1.000		
Natural Disasters	0.111	-0.091	1.000	
GDP per capita	-0.263	0.289	-0.239	1.000

Note: The table reports correlation coefficients across origin countries and time between the four push factors used to generate the different sets of instruments.

to generate the instruments. The correlation between factors is very low, even between wars and political regimes. This result reinforces the validity of the four variables as instruments, because they are all highly correlated with migration despite being uncorrelated between themselves.

An important implication of the stability of the Sub-Sample 2SLS coefficients across columns in Table 5 is in the interpretation of the results. Following Imbens and Angrist (1994), each of these estimates can be interpreted as a local average treatment effect (LATE). Even if one considers that wars, political regimes, and natural disasters might select a specific group of *compliers* (emergency-type migrants), economic variables select a very different group of them (economic migrants), still producing the same result. The similarity across estimates obtained from so different instruments suggests that the resulting coefficients may be consistent estimates the average treatment effect (ATE).

C. Results for the United States

Even though the finding of similar OLS estimates to those in Borjas (2003) for the United States and in Aydemir and Borjas (2007, 2011) for Canada suggests that effect of immigration on wages is similar in the two countries, the estimation θ restricting the second stage sample to a single country (namely, the United States) is interesting for two reasons. First, to illustrate that the proposed instruments are suitable to identify the effect of immigration on different outcomes of a single destination country. And second, to evaluate the performance of the Sub-Sample 2SLS estimator under the selection of different sub-samples.

Table 7 presents estimation results for a second stage sub-sample that only includes the United States. Point estimates are slightly larger both for OLS and Sub-Sample 2SLS, and precision drops a bit as a consequence of the reduction in the number of observations. OLS coefficients fluctuate around an average of -0.76 (with implied elasticities of about -0.53), and Sub-Sample 2SLS coefficients

Table 7—The Effect of Immigration on Native Male Wages: United States Only

	OLS	Sub-Sample 2SLS			
		Months of war	Political regime	Natural disasters	GDP per capita
Baseline	-0.695	-1.805	-2.028	-1.883	-1.934
	(0.223)	(0.787)	(0.861)	(0.901)	(0.955)
Unweighted regression	-0.801	-1.706	-1.727	-1.954	-1.648
	(0.273)	(0.704)	(0.776)	(0.910)	(0.721)
Weighs are sample sizes for shares	-0.718	-1.824	-2.029	-1.904	-1.946
	(0.224)	(0.815)	(0.941)	(0.964)	(0.896)
Unbalanced panel	-0.698	-2.236	-2.465	-2.236	-2.388
	(0.222)	(1.344)	(1.446)	(1.469)	(1.699)
Log annual wages	-0.911	-1.815	-2.124	-1.988	-1.998
	(0.435)	(0.934)	(1.048)	(1.042)	(1.135)
Includes female in LF counts	-0.763	-1.813	-2.015	-1.872	-1.928
	(0.213)	(0.684)	(0.754)	(0.781)	(0.843)

Note: The table reports the coefficient of the immigrant share from regressions where the dependent variable is the average log wage for native males aged 18-64 in each education-experience-period (monthly wage, except otherwise indicated). Each row is a different specification; each column uses a different set of instruments. All regressions include 60 observations in the second stage, except those estimated with the unbalanced panel (75). All regressions are weighted by the sample size used to compute wages in each cell, except otherwise indicated. All regressions include education, experience, period, education-period, experience-period, and education-country fixed effects. Standard errors, in parenthesis, are computed as derived in Appendix A.

average -1.97 (implying an elasticity of -1.38). Across specifications, Sub-Sample 2SLS are, on average, 2.6 times larger than OLS counterparts. The similarity of these results with those obtained for the United States and Canada suggest that the effect of immigration is relatively homogeneous across these two countries, as discussed in Aydemir and Borjas (2007). It is also suggestive evidence in favor of the validity of the Sub-Sample 2SLS approach, because similar results are obtained running the second stage in two different (even though correlated) sub-samples.

D. Robustness

Results in Table 5 and Table 7 already show that estimated elasticities are very similar regardless of which of the four push factors are used to construct the instrument. Table 8 further explores this similarity combining different definitions for both push factors and distance. Across columns, the distance measure is changed. The first two columns use physical distance, one as in the baseline (first column), and the other weighting origin countries by log physical area (second column). The two right columns use two alternative measures of linguistic distance: the probability that two randomly selected individuals in a country speak the same language, and a linguistic distance index constructed by Mélitz and Toubal (2014). Different rows use different push factors. In particular, three measures of wars,

Table 8—Robustness to Alternative Choices of Instruments

	Distance measure:			
Push factor measure:	Physical distance (baseline)	Phys.dist. (weighted by area)	Prob. rand. speaking same lang.	Linguistic distance index
i. Wars:				
Months of war (baseline)	-1.655	-1.663	-1.463	-1.557
	(0.648)	(0.648)	(0.635)	(0.691)
War dummy	-1.673	-1.683	-1.725	— [b]
	(0.692)	(0.691)	(0.900)	
Casualties (per 1,000 inhabitants)	-1.740	-1.778	-1.497	— [b]
	(0.621)	(0.629)	(0.780)	
ii. Political regimes:				
Political regime (baseline)	-1.856	-1.853	-1.188	-0.847
	(0.706)	(0.696)	(0.349)	(0.222)
Absolute value of Polity IV Index	-1.792	-1.784	-1.146	-0.878
	(0.731)	(0.720)	(0.318)	(0.227)
Democracy level	-1.684	-1.642	-1.076	-1.178
	(0.765)	(0.750)	(0.322)	(0.379)
Autocracy level	-1.951	-1.949	— [b]	— [b]
	(0.688)	(0.686)		
iii. Natural disasters:				
Natural disasters (baseline)	-1.691	-1.671	-1.434	-1.380
	(0.723)	(0.706)	(0.524)	(0.472)
Disaster dummy	-0.795	-0.901	-1.438	— [b]
	(0.233)	(0.272)	(0.760)	
Disaster damage per capita	-1.968	— [a]	-0.824	— [b]
	(1.390)		(0.246)	
Killings (per 1,000 inhabitants)	— [a]	— [a]	-1.303	— [a]
			(0.444)	
Drought only	-1.638	-1.633	-1.164	-1.198
	(0.704)	(0.694)	(0.582)	(0.442)
Earthquakes only	-2.262	-2.254	— [a,b]	-1.746
	(0.893)	(0.873)		(0.615)
Flood only	-1.645	-1.633	-1.555	— [b]
	(0.723)	(0.709)	(0.705)	
Storms only	-1.811	-1.801	— [b]	— [b]
	(0.810)	(0.851)		
iv. Economic variables:				
GDP per capita (baseline)	-1.774	-1.774	-1.375	-1.418
	(0.796)	(0.796)	(0.458)	(0.503)
Population density	-1.692	-1.666	-1.145	-1.133
	(0.702)	(0.677)	(0.318)	(0.320)
Real exchange rate	-1.715	-1.701	-1.445	— [a]
	(0.729)	(0.729)	(0.534)	
Employment rate	-1.694	-1.715	-1.238	-1.398
	(0.722)	(0.694)	(0.372)	(0.445)
GDP per capita growth	-2.061	-2.135	-1.421	— [a]
	(0.810)	(0.848)	(0.420)	

[a] Null hypothesis of joint insignificance of coefficients from excluded regressors is not rejected at 5%.
[b] Stability across sub-samples is rejected at 5%.

Note: The table reports the coefficient of the immigrant share from regressions that follow the baseline specification in Table 5, each point estimate being obtained using a different set of instruments (rows and columns vary push and distance variables as indicated). Standard errors, in parenthesis, are computed as derived in Appendix A.

four of political regimes, eight of natural disasters, and five different economic variables are alternatively used (see Appendix B2 for a detailed description).

Results presented in the table are very stable regardless of the definition of push factors and distance used to construct the instruments. When physical distance is used, with only one exception, all point estimates range between -1.638 and -2.267, and implied elasticities still average around -1.2. Very similar estimates are obtained when origin countries are additionally weighted by log physical area. With linguistic distance, results are also in line, and implied elasticities, even though slightly smaller (they average around -0.9) are still 2.5 times larger than in OLS. By push factors, the most remarkable case is that of economic variables: even with very different economic variables (GDP per capita, population density, real exchange rate, employment rate, and economic growth) all point estimates range between -1.692 and -2.061 when the baseline distance measure is used (between -1.133 and -2.135 across different distance measures), and implied wage elasticities still average around -1.2 (-1.1 across all distance measures). In sum, the noticeable stability of estimates across different choices of instruments that are very uncorrelated with each other, and that have different levels of relevance, with F statistics ranging up to above 16, reinforces the credibility of the estimates.

All regressions presented above include country–period, education–period, experience–period, and education–country fixed effects, mimicking the baseline estimation in Borjas (2003) when he combines geographical (state in his case) and skill definitions of labor markets, as in this paper (Column 1, Table V, p.1353). This minimal specification responds to the reduced number of observations used in estimation (120 observations in the baseline sample, 135 in the unbalanced case, when the United States and Canada are considered). Despite this, however, the model is technically identified with additional interactions of fixed effects.

Table 9 presents estimates in which all two-way and three-way combinations of fixed effects are progressively introduced. Even though precision of the estimates dramatically decreases with the inclusion of additional fixed effects (the most demanding model estimates 88 parameters with 120 observations), results in Table 9 are consistent with the results presented above. Both OLS and Sub-Sample 2SLS elasticities are slightly reduced, but Sub-Sample 2SLS estimates are still around 2.8 times OLS counterparts. OLS parameter estimates average around -0.42, implying an elasticity of about -0.29, and Sub-Sample 2SLS estimates average around -1.2, with an implied elasticity of about -0.84.

Table 9—Robustness to Different Combinations of Fixed Effects

	OLS	Sub-Sample 2SLS: Months of war	Political regime	Natural disasters	GDP per capita
1a. Baseline: balanced	-0.556	-1.655	-1.856	-1.691	-1.774
	(0.130)	(0.648)	(0.706)	(0.723)	(0.796)
1b. Baseline: unbalanced	-0.558	-2.067	-2.285	-2.015	-2.216
	(0.132)	(1.109)	(1.201)	(1.176)	(1.433)
2a. (1a)+experience-period	-0.611	-1.074	-1.131	-1.122	-1.064
	(0.163)	(0.282)	(0.262)	(0.321)	(0.297)
2b. (1b)+experience-period	-0.614	-1.327	-1.347	-1.347	-1.233
	(0.164)	(0.434)	(0.383)	(0.511)	(0.417)
3a. (2a)+education-experience	-0.294	-0.411	-0.638	-0.968	-0.888
	(0.185)	(0.663)	(1.120)	(0.733)	(1.055)
3b. (3b)+education-experience	-0.325	-1.201	-0.281	-1.592	-1.322
	(0.183)	(0.919)	(1.478)	(1.133)	(1.014)
4a. (3a)+education-country-period	-0.346	— [a]	— [a]	— [a]	-1.301
	(0.245)				(1.132)
4b. (3b)+education-country-period	-0.383	-1.343	— [a]	-1.745	-1.462
	(0.235)	(0.939)		(1.035)	(0.889)
5a. (4a)+experience-country-period	-0.391	-0.577	— [a]	— [a,b]	-1.575
	(0.282)	(0.446)			(1.045)
5b. (4b)+experience-country-period	-0.427	— [b]	— [a]	— [b]	-1.537
	(0.265)				(0.787)
6a. (5a)+education-experience-period	-0.416	— [b]	— [b]	— [b]	— [b]
	(0.423)				
6b. (5b)+education-experience-period	-0.416	— [b]	— [b]	— [b]	— [b]
	(0.449)				

[a] Null hypothesis of joint insignificance of coefficients from excluded regressors is not rejected at 5%.
[b] Stability across sub-samples is rejected at 5%.

Note: The table reports the coefficient of the immigrant share from regressions where the dependent variable is the average log wage for native males aged 18-64 in each education-experience-period-country cell (monthly wage, except otherwise indicated). Each row introduces a different set of fixed effects; each column uses a different set of instruments. Regressions estimated with the balanced sample include 120 observations in the second stage, and those estimated with the unbalanced panel include 135. All regressions are weighted by the sample size used to compute wages in each cell. The baseline regression education, experience, country-period, education-period, experience-period, and education-country fixed effects. Standard errors, in parenthesis, are computed as derived in Appendix A.

V. Revisiting the Literature

A. *Spatial Correlations vs Factor Proportions: Measurement Error?*

Results so far consistently suggest a negative wage elasticity to immigration of around -1.2, once endogeneity is corrected for in cross-skill cell comparisons at the national level. This is above three times the OLS estimate. The literature have shown that OLS elasticities are much larger if they are estimated at the national level than across more disaggregated geographical units. Borjas (2003) finds smaller wage elasticities at the state level than at the national level, and Borjas (2006) and Cortés (2008) estimate smaller elasticities at the metropolitan area level than at the state level.[17] Two explanations have been given in the literature for this discrepancy: spatial arbitrage, due to interregional flows of labor, that tend to equalize opportunities for workers of given skills across regions (Borjas, 2006);[18] and measurement error, as there is substantial sampling error in the construction of immigrant shares, which is negatively related to the size of labor markets, creating a larger attenuation bias when smaller labor markets are considered (Aydemir and Borjas, 2011).[19] The instruments used in this paper allow some assessment on which of the two explanations prevails, because attenuation bias will be corrected for by the instrument (given that the instrument is uncorrelated with measurement error), whereas spatial arbitrage will not.

Table 10 estimates a similar regression as in Table 5, except that the United States is divided in nine divisions, and Canada in five big regions, which are, in general, at least as sizeable as many European countries.[20] OLS estimates confirm the results in the literature. In the baseline case, the point estimate for θ is -0.324 (s.e. 0.069), with an implied elasticity of around -0.2, almost a half of the elasticity obtained at the national level, and in line with the results in Borjas (2006). This result is robust across specifications.

Sub-Sample 2SLS results prove again stable. In the baseline regression, point estimates average around -1.2 (more precisely estimated than at the national level), which implies an elasticity of around -0.8. This elasticity is now four times the elasticity implied by OLS (instead of three as at the national level), but somewhat smaller than that estimated at the national level (it is scaled by a factor of two thirds, as opposed to the one half in OLS). As a result, a reasonable conclusion seems to be that the discrepancy is mainly driven by measurement error, but some potential role might still be open for spatial arbitrage, although estimates are not precise enough to reject that national and regional level elasticities coincide.

[17] Estimates (std.err.) in Borjas (2006) are -0.532 (0.189), -0.352 (0.061), -0.266 (0.037), and -0.057 (0.024) respectively at the national, division, state, and metropolitan area levels.

[18] Arbitrage can also happen across skills, as noted by Llull (2014), but *a priori* there is no reason to think of differences in cross-skill adjustments at different geographical levels.

[19] Dustmann et al. (2013) suggest that, in the United Kingdom, immigrants downgrade upon entry. Extrapolating their result to the U.S. and Canada, downgrading of immigrants could be an additional source of measurement error. While this measurement error would lead to an underestimation of low skilled immigration and overestimation of more skilled immigration, it is less clear whether variation at the relevant level (e.g. overtime or cross-age variation within a given educational group) would be systematic. In any event, the instruments would correct the resulting bias, provided they are uncorrelated with the measurement error, which is plausible. While a country experiencing a war in a given period may have lower quality of schooling, and subsequent immigrants could downgrade more, it is not clear that they would do so more intensively in different countries depending on distance (the relevant variation).

[20] The nature of the instrument impedes further geographical disaggregation, as distance will hardly play a role in the decision to migrate to, say, New York City versus Philadelphia.

Table 10—The Effect of Immigration on Wages at the Regional Level

		Sub-Sample 2SLS			
	OLS	Months of war	Political regime	Natural disasters	GDP per capita
Baseline	-0.324	-1.201	-1.251	-1.193	-1.647
	(0.069)	(0.394)	(0.392)	(0.368)	(0.529)
Unweighted regression	-0.338	-0.842	-0.816	-0.835	-0.868
	(0.073)	(0.340)	(0.338)	(0.306)	(0.344)
Weighs are sample sizes for shares	-0.309	-1.104	-1.159	-1.110	-1.562
	(0.069)	(0.369)	(0.375)	(0.329)	(0.436)
Unbalanced panel	-0.323	-1.402	-1.279	-1.408	-1.917
	(0.067)	(0.481)	(0.438)	(0.438)	(0.718)
Log annual wages	-0.292	-0.568	-0.775	-0.766	-1.086
	(0.098)	(0.453)	(0.477)	(0.439)	(0.568)
Includes female in LF counts	-0.352	-1.303	-1.303	-1.292	-1.687
	(0.073)	(0.395)	(0.389)	(0.366)	(0.509)

Note: The table reports the coefficient of the immigrant share from regressions where the dependent variable is the average log wage for native males aged 18-64 in each education-experience-period-region cell (monthly wage, except otherwise indicated). The United States is divided in nine regions (divisions) and Canada in five (details in the text). Each row is a different specification; each column uses a different set of instruments. Different specifications are as in Table 5 except for replacing countries by regions. All regressions include education, experience, region-period, education-period, experience-period, and education-region fixed effects. Standard errors, in parenthesis, are computed as derived in Appendix A.

B. *Alternative Variation Used in the Literature: the Networks Instrument*

The literature that uses a geographical definition of labor markets have addressed endogeneity concerns by using past settlements of immigrants from each country of origin in each city of destination as instruments for current inflows into the city. This so-called "networks instrument" was fist introduced by Altonji and Card (1991). In that paper, they use the fraction of immigrants in a city in 1970 to predict the change in the fraction of immigrants over the following decade. Card (2001) constructs a modified version of the instrument that multiplies the total number of immigrants from a source country q that entered in the United States between 1985 and 1990, by the fraction of immigrants from the same origin country that entered in earlier cohorts and are observed living in city k in the base year 1985, and by the fraction of all 1985-1990 immigrants from a source country k that work in a given occupation. This measure provides time, city, and occupational variation. Other variations have been used in the literature.

Even though the suitability of network instruments for the analysis at the national level might be limited, it is useful to compare results in this paper with those obtained using different definitions of the network instrument, both at the national and at the regional level. Table 11 provides this comparison for the United States. The first two panels of the table provide OLS and Sub-Sample

Table 11—Comparison With Networks Instrument for the United States

	National level		Regional level	
Base year $t_0 =$	1960	1970	1960	1970
OLS	-0.695 (0.223)		-0.323 (0.083)	
Sub-Sample 2SLS:				
Months of war	-1.805 (0.787)		-1.299 (0.439)	
Political regime	-2.028 (0.861)		-1.344 (0.431)	
Natural disasters	-1.883 (0.901)		-1.310 (0.411)	
GDP per capita	-1.934 (0.955)		-1.791 (0.603)	
Network instruments:				
$M_{ijkt-10}$	-0.513 (0.205)		-0.485 (0.172)	
$\sum_q \frac{M_{ijqt}}{M_{qt}} \frac{M_{kqt_0}}{M_{qt_0}} M_{qt}$	-0.668[a] (0.161)	-0.667[a] (0.166)	-0.582 (0.126)	-0.752 (0.140)
$\sum_q \frac{M_{ijqt_0}}{M_{qt_0}} \frac{M_{kqt_0}}{M_{qt_0}} M_{qt}$	—[b]	-0.555[c] (0.396)	-0.986 (0.579)	-1.173 (0.349)
$\sum_q \frac{M_{ijqkt_0}}{M_{qt_0}} M_{qt}$	—[b]	-0.555[c] (0.396)	-0.955 (0.529)	-1.223 (0.338)

[a] This number should coincide with OLS, as, at the national level, the central ratio is equal to 1, and, hence, the instrument and the instrumented variables coincide. They do not exactly coincide because immigrants for which the exact country of birth was unknown or missing are not included in the instrument, but they are included in the instrumented variable.
[b] The t statistic for the first stage coefficient of this equation is very close to zero, and, hence, the second stage coefficient is not well identified using this instrument.
[c] These estimates coincide by construction.

Note: The table reports the coefficient of the immigrant share from regressions that follow the baseline specification. The sample of destination countries/regions is restricted to the United States. Different network instruments are self-explained, and defined in the text. Standard errors in parenthesis.

2SLS results for the United States. Results at the national level reproduce the first row of Table 7, whereas results at the regional level, which had not been presented above, compare to the first row of Table 10, with similar results.[21]

The third panel in Table 11 presents 2SLS estimation results obtained using different definitions of the networks instrument. The relevance of the instrument is provided by the higher likelihood to migrate into cells in which a larger network is available (namely, a larger stock of country fellows). Given that a cell is defined by education-experience and, potentially, region, the instrument appeals to the stronger link of new immigrants with country fellows of a similar skill level. Except

[21] Data availability in Canadian Census PUMS impedes doing this analysis for Canada.

where otherwise indicated, first stage results are sufficiently strong to justify this argument. The exclusion restriction would be provided by past settlements being uncorrelated with wage innovations in a given cell beyond its connection with current immigration (and with the set of dummies included in the regression).

The first row of the panel specifies the instrument following the original implementation in Altonji and Card (1991): lagged stock of immigrants in a given cell is used to instrument current inflows. Point estimates are very similar to OLS (the magnitude is even slightly smaller at the national level), and the difference is not statistically significant. Therefore, the instrument seems to fail at correcting the endogeneity bias. Interestingly, the estimated effect at the regional level is now very similar to that at the national level. This result is consistent with the measurement error explanation for the smaller effects at the regional level: if measurement error is independent and identically distributed, the instrument is only correlated with the true immigrant share, but not with the measurement error, and, therefore, it corrects the attenuation bias.

The second definition reassembles the version of the instrument proposed in Card (2001): the total inflow of immigrants is distributed across regions according to the settlement of immigrants from each country of origin in the base year (1960 or 1970 as indicated), and to skill cells according to the current fraction of immigrants in a given skill cell. At the national level, estimates using this instrument are, by construction, numerically equal to OLS (with the clarification indicated in table note a). Like in the original version, the instrument provides similar results at the regional and at the national level.

Finally, the last two rows introduce to refinements of the instrument proposed by Card (2001). In the first case, the share of immigrants from a given origin country by skill cells is set to the base year. In the second, the product of the two ratios is replaced by the overall base-year share of immigrants from a given origin in a given skill-region cell. At the national level, the two refinements are equivalent, and provide again results that are very similar to OLS. At the regional level, point estimates are now more negative, around -1.1 on average, still smaller than (but closer to and not statistically different from) Sub-Sample 2SLS results.

All this suggests that the networks instrument generates estimates that are shifted towards OLS when compared to the benchmark results from this paper, especially for earlier versions of the networks instrument. Additionally, the fact that the difference is larger when the instrument is not entirely based on the distribution in the baseline year (first two versions in Table 11) highlights the relevance of endogeneity in immigrant inflows across different skill cells.

VI. Conclusions

This paper proposes a novel strategy to identify the effect of immigration on wages at the national cross-skill cell level. While they have not been addressed in the literature, endogeneity biases may arise in the OLS estimation of the effect of immigration on wages as immigrants are not randomly allocated across skill cells. To correct for them, the strategy uses the heterogeneous role played by distance in mitigating a push factor across different types of workers as a source of exogenous variation. Distance seems to mitigate the effect of a push factor more strongly for less educated and middle-aged individuals. Consequently, push-distance interactions with heterogeneous first stage coefficients for different skill cells are used as instruments in the estimation. Four push factors are alternatively considered: wars, political regimes, natural disasters, and economic variables.

In order to exploit the variation in distance in the identification of first stage coefficients, a cross-destination country analysis is required. Because of data availability (as wages are only available in United States and Canadian censuses), and also for comparability with existing results in the literature, the interest of this paper is in the United States and Canada. In order to exploit the cross-country variation in distance, the proposed strategy identifies the first stage coefficients with an expanded sample of destination countries, and restricts the second stage sample to the subset of countries of interest. This approach leads to the Sub-Sample 2SLS estimator that have not been proposed so far in the literature. This estimator, for which I discuss theoretical properties and inference, is useful in contexts in which endogenous regressors and instruments are available for a given sample, but the outcome of interest is only available for a random sub-sample (like in cross-country data, or in data supplements for commonly used data sets like the Current Population Survey or the Panel Study of Income Dynamics).

Sub-Sample 2SLS estimated wage elasticities to immigration average -1.2, which is above three times OLS counterparts. This result is very stable across alternative push factors and definitions of distance, suggesting that the resulting wage elasticity may be an estimate of the average treatment effect. The main conclusion is that, even when the national level cross-skill cell approach is used, endogenous allocation of immigrants across labor markets creates a substantial bias in the estimation of wage effects of immigration.

REFERENCES

Almond, Douglas, Joseph J. Doyle Jr., Amenda E. Kowalski, and Heidi

Williams, "Estimating Marginal Returns to Medical Care: Evidence from at-Risk Newborns," *Quarterly Journal of Economics*, May 2010, *125* (2), 591–634.

Altonji, Joseph G. and David E. Card, "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives," in John M. Abowd and Richard B. Freeman, eds., *Immigration, Trade and the Labor Market*, Chicago: The University of Chicago Press, 1991, chapter 7, pp. 201–234.

Angrist, Joshua D. and Adriana Kugler, "Protective or Counter-Productive? Labour Market Institutions and the Effect of Immigration on EU Natives," *Economic Journal*, June 2003, *113* (448), F302–F331.

_ **and Alan B. Krueger**, "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples," *Journal of the American Statistical Association*, June 1992, *87* (418), 328–336.

_ **and** _ , "Split-Sample Instrumental Variables Estimates of the Return to Schooling," *Journal of Business and Economic Statistics*, April 1995, *13* (2), 225–235.

_ **and Jörn-Steffen Pischke**, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press, 2009.

Arellano, Manuel and Costas Meghir, "Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets," *Review of Economic Studies*, July 1992, *59* (3), 537–559.

Aydemir, Abdurrahman and George J. Borjas, "Cross-Country Variation on the Impact of International Migration: Canada, Mexico, and the United States," *Journal of the European Economic Association*, June 2007, *5* (4), 663–708.

_ **and** _ , "Attenuation Bias in Measuring the Wage Impact of Immigration," *Journal of Labor Economics*, January 2011, *29* (1), 69–112.

Björklund, Anders and Markus Jäntti, "Intergenerational Income Mobility in Sweeden Compared to the United States," *American Economic Review*, December 1997, *87* (5), 1009–1018.

Borjas, George J., "Self-Selection and the Earnings of Immigrants," *American Economic Review*, September 1987, *77* (4), 531–553.

_ , "The Economic Analysis of Immigration," in Orley C. Ashenfelter and David E. Card, eds., *Handbook of Labor Economics*, Vol. 3A, Amsterdam: North-Holland Publishing Company, 1999, chapter 28, pp. 1697–1760.

_ , "The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of

Immigration on the Labor Market," *Quarterly Journal of Economics*, November 2003, *118* (4), 1335–1374.

_ , "Food Insecurity and Public Assistance," *Journal of Public Economics*, July 2004, *88* (7-8), 1421–1443.

_ , "Native Internal Migration and the Labor Market Impact of Immigration," *Journal of Human Resources*, Spring 2006, *41* (2), 221–258.

_ , "Labor Outflows and Labor Inflows in Puerto Rico," *Journal of Human Capital*, Spring 2008, *2* (1), 32–68.

_ , Jeffrey T. Grogger, and Gordon H. Hanson, "Immigration and the Economic Status of African-American Men," *Economica*, April 2010, *77* (306), 255–282.

_ , Richard B. Freeman, and Lawrence F. Katz, "On the Labor Market Impacts of Immigration and Trade," in George J. Borjas and Richard B. Freeman, eds., *Immigration and the Work Force: Economic Consequences for the United States and Source Areas*, Chicago: The University of Chicago Press, 1992, chapter 7, pp. 213–244.

_ , _ , and _ , "How Much Do Immigration and Trade Affect Labor Market Outcomes?," *Brookings Papers on Economic Activity*, Spring 1997, *1997* (1), 1–67.

Bratsberg, Bernt and Oddbjørn Raaum, "Immigration and Wages: Evidence from Construction," *Economic Journal*, December 2012, *122* (565), 1177–1205.

_ , _ , Marianne Røed, and Pål Schøne, "Immigration Wage Effects by Origin," *Scandinavian Journal of Economics*, April 2014, *116* (2), 356–393.

Card, David E., "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, January 1990, *43* (2), 245–257.

_ , "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, January 2001, *19* (1), 22–64.

_ , "Is the New Immigration Really So Bad?," *Economic Journal*, November 2005, *115* (507), F300–F323.

_ , "Immigration and Inequality," *American Economic Review: Papers and Proceedings*, May 2009, *99* (2), 1–21.

_ and Ethan G. Lewis, "The Diffusion of Mexican Immigrants during the 1990s: Explanations and Impacts," in George J. Borjas, ed., *Mexican Immigration to the United States*, Chicago: The University of Chicago Press, 2007, chapter 6, pp. 193–227.

Carrasco, Raquel, Juan F. Jimeno, and A. Carolina Ortega, "The Effect of Immigration on the Labor Market Performance of Native-Born Workers: Some Evidence from Spain," *Journal of Population Economics*, July 2008, *21* (3), 627–648.

Chiquiar, Daniel and Gordon H. Hanson, "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States," *Journal of Political Economy*, April 2005, *113* (2), 239–281.

Cortés, Patricia, "The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data," *Journal of Political Economy*, June 2008, *116* (3), 381–422.

_ **and José Tessada**, "Low-Skilled Immigration and the Labor Supply of Highly Skilled Women," *American Economic Journal: Applied Economics*, July 2011, *3* (3), 88–123.

Currie, Janet and Aaron Yelowitz, "Are Public Housing Projects Good for Kids?," *Journal of Public Economics*, January 2000, *75* (1), 99–124.

Dee, Thomas S. and William N. Evans, "Teen Drinking and Educational Attainment: Evidence from Two-Sample Instrumental Variables Estimates," *Journal of Labor Economics*, January 2003, *21* (1), 178–209.

Docquier, Frédéric and Abdeslam Marfouk, "International Migration by Educational Attainment, 1990-2000," in Çağlar Özden and Maurice W. Schiff, eds., *International Migration, Remittances and the Brain Drain*, New York: Palgrave Macmillan, 2006, chapter 5, pp. 151–200.

Dustmann, Christian, Tommaso Frattini, and Ian Preston, "The Effect of Immigration along the Distribution of Wages," *Review of Economic Studies*, January 2013, *80* (1), 145–173.

_ **, Uta Schönberg, and Jan Stuhler**, "The Impact of Immigration on Local Labor Markets: Evidence from the Opening of the Czech-German Border," mimeo, University College London, 2014.

EM-DAT, "The OFDA/CRED International Disaster Database," CRED, Université Catholique de Louvain, 2010.

Fernández-Huertas Moraga, Jesús, "New Evidence on Emigrant Selection," *Review of Economics and Statistics*, February 2011, *93* (1), 72–96.

Friedberg, Rachel M., "The Impact of Mass Migration on the Israeli Labor Market," *Quarterly Journal of Economics*, November 2001, *116* (4), 1373–1408.

_ **and Jennifer Hunt**, "The Impact of Immigrants on Host Country Wages, Employment and Growth," *Journal of Economic Perspectives*, Spring 1995, *9* (2), 23–44.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Hårvard Strand, "Armed Conflict 1946-2001: A New Dataset," *Journal of Peace Research*, September 2002, *39* (5), 615–637.

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

Goldin, Claudia, "The Political Economy of Immigration Restriction in the United States, 1890 to 1921," in Claudia Goldin and Gary D. Libecap, eds., *The Reglated Economy: A Historical Approach to Political Economy*, Chicago: The University of Chicago Press, 1994, chapter 7, pp. 223–257.

Grossman, Jean B., "The Substitutability of Natives and Immigrants in Production," *Review of Economics and Statistics*, November 1982, *64* (4), 596–603.

Heston, Alan, Robert Summers, and Bettina Aten, "Penn World Table Version 7.1," Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, 2012.

Hunt, Jennifer, "The Impact of the 1962 Repatriates from Algeria on the French Labor Market," *Industrial and Labor Relations Review*, April 1992, *45* (3), 556–572.

Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, March 1994, *62* (2), 467–475.

Inoue, Atsushi and Gary Solon, "Two-Sample Instrumental Variables Estimators," *Review of Economics and Statistics*, August 2010, *92* (3), 557–561.

Jappelli, Tullio, Jörn-Steffen Pischke, and Nicholas S. Souleles, "Testing for Liquidity Constraints in Euler Equations with Complementary Data Sources," *Review of Economics and Statistics*, May 1998, *80* (2), 251–262.

Kerr, Sari Pekkala and William R. Kerr, "Economic Impacts of Immigration: A Survey," *Finnish Economic Papers*, Spring 2011, *24* (1), 1–32.

LaLonde, Robert J. and Robert H. Topel, "Labor Market Adjustments to Increased Immigration," in John M. Abowd and Richard B. Freeman, eds., *Immigration, Trade and the Labor Market*, Chicago: The University of Chicago Press, 1991, chapter 6, pp. 167–200.

Lemieux, Thomas, "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *American Economic Review*, June 2006, *96* (3), 461–498.

Llull, Joan, "Reconciling Spatial Correlations and Factor Proportions: A Cross-Country Analysis of the Economic Consequences of Immigration," mimeo, CEMFI, June 2011.

_ , "Immigration, Wages, and Education: A Labor Market Equilibrium Structural Model," Barcelona GSE Working papers n. 711, January 2014.

Manacorda, Marco, Alan Manning, and Jonathan Wadsworth, "The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain," *Journal of the European Economic Association*, February 2012, *10* (1), 120–151.

Marshall, Monty G., Keith Jaggers, and Ted Robert Gurr, "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2010," Version 2010. College Park, MD: Center for International Development and Conflict Management, University of Maryland., 2010.

Mélitz, Jacques and Farid Toubal, "Native language, spoken language, translation and trade," *Journal of International Economics*, July 2014, *93* (2), 351–363.

Minnesota Population Center, "Integrated Public Use Microdata Series, International: Version 6.1 [Machine-readable database]," Minneapolis: University of Minnesota, 2011.

Monràs, Joan, "Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis," minemo, Columbia University, January 2014.

Ottaviano, Gianmarco I. P. and Giovanni Peri, "Rethinking the Effect of Immigration on Wages," *Journal of the European Economic Association*, February 2012, *10* (1), 152–197.

Peri, Giovanni and Chad Sparber, "Task Specialization, Immigration, and Wages," *American Economic Journal: Applied Economics*, July 2009, *1* (3), 135–169.

Saiz, Albert, "Immigration and Housing Rents in American Cities," *Journal of Urban Economics*, March 2007, *61* (2), 345–371.

Steinhardt, Max F., "The Wage Impact of Immigration in Germany - New Evidence for Skill Groups and Occupations," *The B.E. Journal of Economic Analysis and Policy*, June 2011, *11* (1), 1–33.

Stock, James H. and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression," in Donald W. K. Andrews and James H. Stock, eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge: Cambridge University Press, 2005, chapter 5, pp. 80–

108.

The Conference Board, "The Conference Board Total Economy Database," http://www.conference-board.org/data/economydatabase, January, 2014.

Vincenty, Thaddeus, "Direct and Inverse Solutions of Geodesics on the Ellipsoid with Application of Nested Equations," *Survey Review*, April 1975, *23* (176), 88–93.

Appendix A: Asymptotic Properties and Inference. Derivations

Consider the following model:

$$y_s = \boldsymbol{x}_s'\boldsymbol{\beta} + \varepsilon_s, \quad s = 1, ..., N \tag{A1}$$

with:

$$\boldsymbol{x}_s = \Pi \boldsymbol{z}_s + \nu_s, \quad \text{and} \quad \mathbb{E}[\boldsymbol{z}_s \varepsilon_s] = 0. \tag{A2}$$

Let $d_s \equiv \mathbb{1}\{s \in \text{2nd stage sub-sample}\}$. The dependent variable y_s is only observed for the second stage sub-sample, i.e. $d_s y_s$ is observed instead of y_s.

The probability limit of $\hat{\boldsymbol{\beta}}_{SuS2SLS}$ can be derived following standard arguments:

$$\begin{aligned}
\operatorname*{plim}_{N\to\infty} \hat{\boldsymbol{\beta}}_{SuS2SLS} &= \operatorname*{plim}_{N\to\infty} \left(\textstyle\sum_s d_s \hat{\boldsymbol{x}}_s \hat{\boldsymbol{x}}_s'\right)^{-1} \textstyle\sum_s d_s \hat{\boldsymbol{x}}_s y_s = \\
&= \operatorname*{plim}_{N\to\infty} \left[\textstyle\sum_s \boldsymbol{x}_s \boldsymbol{z}_s' \left(\sum_s \boldsymbol{z}_s \boldsymbol{z}_s'\right)^{-1} \sum_s d_s \boldsymbol{z}_s \boldsymbol{z}_s' \left(\sum_s \boldsymbol{z}_s \boldsymbol{z}_s'\right)^{-1} \sum_s \boldsymbol{z}_s \boldsymbol{x}_s'\right]^{-1} \times \\
&\qquad \times \textstyle\sum_s \boldsymbol{x}_s \boldsymbol{z}_s' \left(\sum_s \boldsymbol{z}_s \boldsymbol{z}_s'\right)^{-1} \sum_s d_s \boldsymbol{z}_s y_s = \\
&= \left(\mathbb{E}[\boldsymbol{x}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1}\, \mathbb{E}[d_s \boldsymbol{z}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1}\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{x}_s']\right)^{-1} \times \\
&\qquad \times \mathbb{E}[\boldsymbol{x}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1}\, \mathbb{E}[d_s \boldsymbol{z}_s y_s] = \\
&= \left(\mathbb{E}[\boldsymbol{x}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1}\, \mathbb{E}[d_s \boldsymbol{z}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1}\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{x}_s']\right)^{-1} \mathbb{E}[\boldsymbol{x}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1}\, \mathbb{E}[d_s \boldsymbol{z}_s x_s']\boldsymbol{\beta} + \\
&\quad + \left(\mathbb{E}[\boldsymbol{x}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1}\, \mathbb{E}[d_s \boldsymbol{z}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1}\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{x}_s']\right)^{-1} \mathbb{E}[\boldsymbol{x}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1}\, \mathbb{E}[d_s \boldsymbol{z}_s \varepsilon_s].
\end{aligned} \tag{A3}$$

Consistency requires that:

$$\mathbb{E}[d_s \boldsymbol{z}_s \varepsilon_s] = 0, \tag{A4}$$

and that:

$$\begin{aligned}
\left(\mathbb{E}[\boldsymbol{x}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1}\, \mathbb{E}[d_s \boldsymbol{z}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1}\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{x}_s']\right)^{-1} \mathbb{E}[\boldsymbol{x}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']^{-1}\, \mathbb{E}[d_s \boldsymbol{z}_s x_s'] &= I_M \\
\Leftrightarrow \mathbb{E}[d_s \boldsymbol{z}_s \boldsymbol{z}_s']\, \mathbb{E}[d_s \boldsymbol{z}_s x_s'] &= \mathbb{E}[\boldsymbol{z}_s \boldsymbol{z}_s']\, \mathbb{E}[\boldsymbol{z}_s x_s'] \\
\Leftrightarrow \mathbb{E}[d_s \boldsymbol{z}_s \nu_s] &= 0,
\end{aligned} \tag{A5}$$

where I_M is a size $M \equiv \dim\{\boldsymbol{x}\}$ identity matrix.

The asymptotic distribution of $\hat{\boldsymbol{\beta}}_{SuS2SLS}$ can also be derived in the standard way. The Central Limit Theorem is applicable, and it implies that:

$$\sqrt{N}\left(\hat{\boldsymbol{\beta}}_{SuS2SLS} - \boldsymbol{\beta}\right) \xrightarrow[d]{} \mathcal{N}(0, V_0), \tag{A6}$$

where:

$$
\begin{aligned}
V_0 = \operatorname*{plim}_{N\to\infty} & \left[\textstyle\sum_s \boldsymbol{x}_s\boldsymbol{z}_s' \left(\sum_s \boldsymbol{z}_s\boldsymbol{z}_s'\right)^{-1} \sum_s d_s\boldsymbol{z}_s\boldsymbol{z}_s' \left(\sum_s \boldsymbol{z}_s\boldsymbol{z}_s'\right)^{-1} \sum_s \boldsymbol{z}_s\boldsymbol{x}_s'\right]^{-1} \times \\
& \times \textstyle\sum_s \boldsymbol{x}_s\boldsymbol{z}_s' \left(\sum_s \boldsymbol{z}_s\boldsymbol{z}_s'\right)^{-1} \sum_s d_s\boldsymbol{z}_s\varepsilon_s\varepsilon_s'\boldsymbol{z}_s' \left(\sum_s \boldsymbol{z}_s\boldsymbol{z}_s'\right)^{-1} \sum_s \boldsymbol{z}_s\boldsymbol{x}_s' \times \\
& \times \left[\textstyle\sum_s \boldsymbol{x}_s\boldsymbol{z}_s' \left(\sum_s \boldsymbol{z}_s\boldsymbol{z}_s'\right)^{-1} \sum_s d_s\boldsymbol{z}_s\boldsymbol{z}_s' \left(\sum_s \boldsymbol{z}_s\boldsymbol{z}_s'\right)^{-1} \sum_s \boldsymbol{z}_s\boldsymbol{x}_s'\right]^{-1} =
\end{aligned}
$$

$$
\begin{aligned}
= & \left(\mathbb{E}[\boldsymbol{x}_s\boldsymbol{z}_s']\,\mathbb{E}[\boldsymbol{z}_s\boldsymbol{z}_s']^{-1}\,\mathbb{E}[d_s\boldsymbol{z}_s\boldsymbol{z}_s']\,\mathbb{E}[\boldsymbol{z}_s\boldsymbol{z}_s']^{-1}\,\mathbb{E}[\boldsymbol{z}_s\boldsymbol{x}_s']\right)^{-1} \times \\
& \times \mathbb{E}[\boldsymbol{x}_s\boldsymbol{z}_s']\,\mathbb{E}[\boldsymbol{z}_s\boldsymbol{z}_s']^{-1}\,\mathbb{E}[d_s\varepsilon_s^2\boldsymbol{z}_s\boldsymbol{z}_s']\,\mathbb{E}[\boldsymbol{z}_s\boldsymbol{z}_s']^{-1}\,\mathbb{E}[\boldsymbol{z}_s\boldsymbol{x}_s'] \times \\
& \times \left(E[\boldsymbol{x}_s\boldsymbol{z}_s']\,\mathbb{E}[\boldsymbol{z}_s\boldsymbol{z}_s']^{-1}\,\mathbb{E}[d_s\boldsymbol{z}_s\boldsymbol{z}_s']\,\mathbb{E}[\boldsymbol{z}_s\boldsymbol{z}_s']^{-1}\,\mathbb{E}[\boldsymbol{z}_s\boldsymbol{x}_s']\right)^{-1} =
\end{aligned}
$$

$$
= \mathbb{E}[d_s\Pi'\boldsymbol{z}_s\boldsymbol{z}_s'\Pi]^{-1}\,\mathbb{E}[d_s\varepsilon_s^2\Pi'\boldsymbol{z}_s\boldsymbol{z}_s'\Pi]\,\mathbb{E}[d_s\Pi'\boldsymbol{z}_s\boldsymbol{z}_s'\Pi]^{-1}. \tag{A7}
$$

Interestingly, if we assume $\mathbb{E}[d_s|\boldsymbol{x}_s, \boldsymbol{z}_s, \varepsilon_s] = \tau$, which is a sufficient though not necessary condition for assumptions (A4) and (A5) to hold, the variance-covariance matrix in (A7) is the standard 2SLS variance-covariance matrix, scaled by a factor τ^{-1} which, indeed, is the (inverse of) the fraction of N that is used in the second stage sample. Put differently, under that stricter assumption, $\text{AsVar}\left(\hat{\boldsymbol{\beta}}_{SuS2SLS}\right) = N^{-1}V_0 = (N\tau)^{-1}V_{0,2SLS}$, where $V_{0,2SLS}$ is the standard 2SLS variance-covariance matrix, which is scaled by the inverse of the *correct* sample size.

The following consistent estimator of (A7) is used:

$$
\widehat{\text{AsVar}}\left(\hat{\boldsymbol{\beta}}_{SuS2SLS}\right) = N^{-1}\left(\textstyle\sum_s d_s\hat{\boldsymbol{x}}_s\hat{\boldsymbol{x}}_s'\right)^{-1} \sum_s d_s\hat{\varepsilon}_s^2\hat{\boldsymbol{x}}_s\hat{\boldsymbol{x}}_s' \left(\sum_s d_s\hat{\boldsymbol{x}}_s\hat{\boldsymbol{x}}_s'\right)^{-1}, \tag{A8}
$$

where $\hat{\boldsymbol{x}}_s \equiv \hat{\Pi}'\boldsymbol{z}_s$, and $\hat{\varepsilon}_s \equiv y_s - \boldsymbol{x}_s'\hat{\boldsymbol{\beta}}_{SuS2SLS}$.

Assumption (A4) is, by construction, not testable in this model, not even against the alternative that only $\mathbb{E}[\boldsymbol{z}_s\varepsilon_s] = 0$ is satisfied, because only d_sy_s and not y_s is observed. However, assumption (A5) can be tested. Consider the following regression:

$$
\boldsymbol{x}_s = \Gamma'\hat{\boldsymbol{x}}_s + \Delta' d_s\hat{\boldsymbol{x}}_s + \epsilon_s, \tag{A9}
$$

With some tedious algebra, it follows that Δ is given by:

$$
\Delta = \left[\mathbb{E}[d_s\hat{\boldsymbol{x}}_s\hat{\boldsymbol{x}}_s']\,(I_M - \Xi)\right]^{-1}\mathbb{E}[(d_sI_M - \Xi')\,\hat{\boldsymbol{x}}_s\boldsymbol{x}_s'], \tag{A10}
$$

where $\Xi \equiv \mathbb{E}[\hat{\boldsymbol{x}}_s\hat{\boldsymbol{x}}_s']^{-1}\,\mathbb{E}[d_s\hat{\boldsymbol{x}}_s\hat{\boldsymbol{x}}_s']$ is the regression coefficient of $\hat{\boldsymbol{x}}_s$ on $d_s\hat{\boldsymbol{x}}_s$. Testing that assumption (A5) is equivalent to test whether $\Delta = 0$ is satisfied. In other words, $\Delta = 0$ is a necessary and sufficient condition for $\mathbb{E}[d_s\boldsymbol{z}_s\nu_s] = 0$:

$$
\Delta = 0 \quad \Leftrightarrow \quad \mathbb{E}[d_s\hat{\boldsymbol{x}}_s\boldsymbol{x}_s'] = \mathbb{E}[d_s\hat{\boldsymbol{x}}_s\hat{\boldsymbol{x}}_s'] \quad \Leftrightarrow \quad \mathbb{E}[d_s\boldsymbol{z}_s\nu_s] = 0. \tag{A11}
$$

Note that Δ is identified, as $\mathbb{E}[d_s\hat{\boldsymbol{x}}_s\hat{\boldsymbol{x}}_s'] \neq \mathbb{E}[\hat{\boldsymbol{x}}_s\hat{\boldsymbol{x}}_s']$ (and, hence, $\Xi \neq I_M$) in general.

Appendix B: Variable Definitions and Sources

B1. Immigrant shares and wages

Immigrant shares and average wages by skill cell are computed using individual data from Public Use Microdata Samples from censuses of different countries. These data are extracted from IPUMS-International (Minnesota Population Center, 2011). The following lines describe definitions and data construction for the different variables, as well as sample selection protocols.

Countries and periods Immigrant shares are computed for a balanced panel that covers censuses around 1970, 1980, 1990, and 2000, including Austria, Canada, Greece, and Ireland (1971, 1981, 1991, 2001), Switzerland and the United States (1970, 1980, 1990, 2000), and France (1968, 1982, 1990, 1999), and for an unbalanced panel that additionally includes the Netherlands (1971, 2001), Portugal and Spain (1981, 1991, 2001), and Italy (2001), as well as year 1960 (France 1962 and United States 1960). Wage data are only available for United States and Canada. In the regional analysis, European countries are considered as single regions, the United States is divided in nine divisions (New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific divisions), and Canada is divided in five regions (Atlantic, Quebec, Ontario, The Prairies, and British Columbia).

Sample selection Immigrant shares are computed for men (some specifications also include women) aged 18-64 who participate in the civilian labor force (except for Netherlands 1971 because labor force status is not available), with available information on region when applicable. Individual weights are adjusted to account for individuals with no available information to compute immigrant status, education, and/or experience, such that they sum to the correct aggregates. The sample for wages additionally restricts to wage/salary employees who worked in the year prior to the survey, were not enrolled neither in school nor in the armed forces, and did not live in group quarters.

Immigration status Immigrants are defined with different criteria across countries. Citizenship and place of birth are used for that purpose. For a majority of countries in which citizenship and place of birth are available (Canada, France, Italy, Netherlands, Portugal, Spain, Switzerland, and the United States), both variables are combined. In such a case, an individual is considered to be native if she is citizen by birth, native-born citizen not specified, and native-born with unknown citizenship; she is considered as an immigrant if she is a naturalized citizen, not a citizen, a foreign-born citizen not specified, and foreign-born with unknown citizenship; if both citizenship and place of birth are unknown, including stateless, she is left to missing. For the two countries in which only citizenship is consistently available across years (Austria and Greece), citizens are natives, and non-citizens are immigrants. When only place of birth is consistently available (Ireland) native-born are natives, and foreign-born are immigrants.

Education Educational attainment is measured by the international recode provided by IPUMS-International for all countries in the sample, with the exception of the Netherlands, for which it is not available. The harmonized variable includes three categories (primary or less, secondary, and university) that are recoded by IPUMS based on the information provided in each country census. In the case of the Netherlands I coded education as follows: less than primary education includes no education, primary school and less (pre-primary and primary), and lower level; secondary education includes upper lower level, and intermediate level (upper secondary and post-secondary non university); primary and tertiary education are already defined in the original variable.

Experience Experience is calculated as age minus imputed age of entry in the labor market. The imputed age of entry in the labor market is 16 for primary educated, 18 for secondary educated, and 23 for tertiary educated. Five 8-year experience categories are created: <8, 8-15, 16-23, 24-31, 32+ years.

Wages The definition of native male wages includes wage and salary income. As noted above, the sample of wage earners is further restricted, compared to the general sample used to compute immigrant shares, and, additionally, they are only available for the United States and Canada. Top coded wages (for the United States, 25,000US$ in 1960, 50,000US$ in 1970, and 75,000US$ in 1980; for Canada 75,000C$ in 1970, 100,000C$ in 1980, and 200,000C$ in 1990 and 2000) are multiplied by 1.4 (e.g. Lemieux, 2006). Monthly wages are computed combining this information on annual earnings with (harmonized) number of months worked during the preceding year. Log-wages are averaged for each skill cell.

Immigrant shares Immigrant shares in each education-experience cell are defined as the total number of immigrants in the cell divided by the total number of individuals in the cell. Aggregates are computed using sampling weights.

B2. Instruments

Different instruments are obtained from different sources, as detailed below. In all cases, the sample of origin countries considered includes 188 countries (of which the corresponding destination country is dropped).[22] Detailed sources and definitions are provided for each variable.

Distance Physical distance is own calculated using the Vincenty method (Vincenty, 1975), which assumes that the figure of the Earth is an oblate spheroid. This measure is known to deliver a more accurate measurement than other measures, like the great-circle distance, which assumes that the Earth is spherical. Distance is computed between the centroid of the most populated city of each country in a country-pair. In the regional analysis, each region in the destination country is treated as a destination country itself.

In the robustness section, two different measures of linguistic distance are used alternatively. Both variables are obtained from Mélitz and Toubal (2014). The first one measures the probability that two randomly drawn individuals from a given pair of countries speak the same language. The second variable is an index constructed by these authors that combines information on common spoken, native and official language probabilities, and two linguistic proximity measures based respectively on *Ethnologue* language family tree and the Automated Similarity Judgement Program, which is a databank created by international ethnolinguists and ethnostatisticians covering the lexical aspects of more than 2400 of the world's nearly 7000 languages (Mélitz and Toubal, 2014, p.3).

Civil wars and conflicts The main conflict variable used in the fraction of months that a given country of origin was involved in a civil war or conflict during the preceding decade. This information is constructed based on starting

[22] These countries include: Afghanistan; Albania; Algeria; Angola; Antigua & Barbuda; Argentina; Armenia; Australia; Austria; Azerbaijan; Bahamas; Bahrain; Bangladesh; Barbados; Belarus; Belgium; Belize; Benin; Bermuda; Bhutan; Bolivia; Bosnia & Herzegovina; Botswana; Brazil; Brunei; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Canada; Cape Verde; Central African Rep.; Chad; Chile; China; Colombia; Comoros; Congo, Dem. Rep.; Congo; Costa Rica; Cote d'Ivoire; Croatia; Cuba; Cyprus; Czech Rep.; Denmark; Djibouti; Dominica; Dominican Republic; Ecuador; Egypt; El Salvador; Equatorial Guinea; Eritrea; Estonia; Ethiopia; Fiji; Finland; France; Gabon; Gambia; Georgia; Germany; Ghana; Greece; Grenada; Guatemala; Guinea; Guinea-Bissau; Guyana; Haiti; Honduras; Hong Kong; Hungary; Iceland; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Jamaica; Japan; Jordan; Kazakhstan; Kenya; Kiribati; Korea, North; Korea, South; Kuwait; Kyrgyzstan; Laos; Latvia; Lebanon; Lesotho; Liberia; Libya; Lithuania; Luxembourg; Macao; Macedonia; Madagascar; Malawi; Malaysia; Maldives; Mali; Malta; Mauritania; Mauritius; Mexico; Micronesia; Moldova; Mongolia; Morocco; Mozambique; Namibia; Nepal; Netherlands; Netherlands Antilles; New Zealand; Nicaragua; Niger; Nigeria; Norway; Oman; Pakistan; Palau; Panama; Papua New Guinea; Paraguay; Peru; Philippines; Poland; Portugal; Puerto Rico; Qatar; Romania; Russia; Rwanda; Samoa; Sao Tome & Principe; Saudi Arabia; Senegal; Serbia & Montenegro; Seychelles; Sierra Leone; Singapore; Slovak Republic; Slovenia; Solomon Islands; Somalia; South Africa; Spain; Sri Lanka; St. Kitts & Nevis; St. Lucia; St.Vincent & the Grenadines; Sudan; Suriname; Swaziland; Sweden; Switzerland; Syria; Taiwan; Tajikistan; Tanzania; Thailand; Togo; Tonga; Trinidad & Tobago; Tunisia; Turkey; Turkmenistan; Uganda; Ukraine; United Arab Emirates; UK; USA; Uruguay; Uzbekistan; Vanuatu; Venezuela; Vietnam; Yemen; Zambia; Zimbabwe.

and ending dates of conflicts obtained from PRIO (Gleditsch et al., 2002). An armed conflict is defined as "a contested incompatibility that concerns government and/or territory where the use of armed force between two parties (of which at least one is the government of a state) results in at least 25 battle-related deaths".

Alternative measures of conflict intensity used in the robustness section include a dummy for the presence of a conflict, and the estimated number of fatalities generated by a civil war or conflict in the origin country during the preceding decade per 1,000 inhabitants. Information on the number of casualties is also obtained from PRIO, and population figures come from Penn World Tables (Heston, Summers and Aten, 2012). Battle deaths are defined as deaths resulting from violence inflicted through the use of armed force by a party to an armed conflict during a contested combat (i.e. it excludes deaths outside the context of a reciprocal threat of lethal force, like execution of prisoners of war). The PRIO Battle Deaths Dataset provides a lower and an upper bound estimate, and in general, a "best estimate" of annual battle fatalities. I use the best estimate when available, and the average between the lower and the upper bound whenever the best estimate is not available.

Political regime Political regimes are identified based on the polity IV (pIV) index (Marshall, Jaggers and Gurr, 2010). The pIV index is constructed as the difference between two indices, one of democracy and the other of autocracy. Both indices combine information of competitiveness of executive recruitment and openness of executive recruitment, constraints on chief executive, and competitiveness of political participation. Competitiveness of executive recruitment adds two points to the democracy index (and hence to the pIV index) if executive recruitment happens through an election and one point if it is transitional, whereas it adds two points to the autocracy index (and hence subtracts them to the pIV index) if it happens by appointment. A similar gradient of contribution to either of the indices is given for all other items in the previous list (detailed scoring is presented in the pIV User's Manual). The resulting index ranges from -10 (strongly autocratic regime) to 10 (fully democratic regime). Intermediate levels of the index (e.g. larger than -6 and smaller than 6) indicate anocracies (see `http://www.systemicpeace.org/polity/polity4.htm`). Anocracies are regime-types where power is not vested in public institutions but spread amongst elite groups who are constantly competing with each other for power. As discussed in the main text, anocracies are more likely to foster migration.

The baseline political regime variable is an indicator that equals one if pIV index over the preceding decade averages between -6 and 6. Additional variables are constructed for robustness: the absolute value of the index, and the positive (democracy, $\max\{pIV, 0\}$), and the negative (autocracy, $\max\{-pIV, 0\}$) splines.

Natural disasters Information for natural disasters is obtained from the EM-DAT database (EM-DAT, 2010). The database provides information on occurrence and impact of natural disasters. For each disaster that is entered into the database, the following information is provided: dates, disaster type, country, region, the number of people reported killed, injured, homeless and affected, as well as estimates of infrastructure and economic damages. An event is defined as a disaster if at least one of the following occurs: ten people reported being killed by the event, 100 or more people affected (needed immediate assistance, displaced, or evacuated), declaration of a state of emergency, or there was a call for international assistance. I consider four types of natural disasters: earthquakes, floods, storms, and droughts. The baseline variable used as a push factor is the (cumulative) fraction of a country's population that was affected (needed immediate assistance, displaced, or evacuated) by the four considered disaster types over the preceding decade. This measure is computed by adding the number of individuals affected by each natural disaster that occurred over the preceding decade divided by the average population over the decade.

Several alternative measures are used: a disaster dummy; disaster damage per capita (in PPP US$), number of killings per 1,000 inhabitants, and fraction of the population affected by each type of natural disaster individually. To construct these variables, EM-DAT (2010) disaster information is combined with population data and PPP correction factors obtained from Penn World Tables.

Economic variables Log average GDP per capita at PPP US$ over the preceding decade is obtained from Penn World Tables, version 7.1. As data for the People's Democratic Republic of Korea and the Netherlands Antilles are not available in version 7.1, information for these countries is taken from version 6.2.

Several alternative economic variables are used in the robustness: population density, real exchange rate, employment rate, and GDP per capita growth. Employment data is obtained from the Total Economy Database (The Conference Board, 2014); all other variables are from Penn World Tables.

Appendix C: Regional Distribution of Net Migration Flows After Selected Push Factors

Table C1—Regional Distribution of Net Migration Flows to OECD Countries by Educational Level After Selected Conflicts and Natural Disasters (1990-2000)

	Total	Primary	Secondary	Tertiary	Weighted dist. (km)
	A. Conflicts				
i. Balkans War					
Australia/New Zealand	3.01	3.87	-6.71	8.38	15,846
Europe	77.99	90.16	64.68	58.99	1,027
U.S./Canada	19.00	5.97	42.03	32.63	7,185
ii. African Conflicts					
Australia/New Zealand	2.07	2.40	0.78	2.69	14,091
Europe	66.21	74.86	75.63	55.97	2,381
U.S./Canada	31.72	22.73	23.59	41.33	9,732
iii. Middle East Conflicts					
Australia/New Zealand	11.20	16.75	8.60	8.84	14,073
Europe	43.70	43.07	56.41	38.40	4,102
U.S./Canada	45.10	40.17	35.00	52.75	9,491
	B. Natural Disasters				
i. Cyclone Gorky (Bangladesh, 1991)					
Australia/New Zealand	4.40	1.18	0.21	9.72	9,258
Europe	53.93	77.04	44.97	30.77	7,838
U.S./Canada	41.68	21.78	54.82	59.51	12,660
ii. Manjil-Rudbar Earthquake (Iran, 1990)					
Australia/New Zealand	3.56	5.25	1.95	3.65	13,055
Europe	25.70	41.87	53.16	11.46	3,934
U.S./Canada	70.74	52.88	44.89	84.88	9,869
iii. İzmit Earthquake (Turkey, 1999)					
Australia/New Zealand	0.96	0.65	1.08	2.89	14,976
Europe	95.99	100.41	93.47	68.93	2,036
U.S./Canada	3.06	-1.07	5.45	28.18	8,093
iv. Vargas Tragedy (Venezuelan flood, 1999)					
Australia/New Zealand	0.56	0.38	0.11	0.85	15,149
Europe	42.90	42.92	71.86	26.90	7,223
U.S./Canada	56.54	56.70	28.04	72.25	3,459

Note: The table presents the regional distribution of net inflows of immigrants in different groups of destination countries from a selection of origin countries affected by conflicts (panel A) and natural disasters (panel B). European destination countries include EU-15 (excluding Luxembourg and Ireland), Norway, and Switzerland. Balkans War affected the countries that constituted the former Yugoslavia. African civil conflicts include those that happened in Algeria, Angola, Burundi, former Zaïre, Ethiopia/Eritrea, Liberia, Mali/Niger (Touareg rebellion), Rwanda, Sierra Leone, and Sudan during 1990s. Middle East conflicts comprise war episodes in Afghanistan, Iraq, Lebanon, Tajikistan, and Yemen. Distance is weighted by stock of immigrants from each country pair in 1990. *Data sources*: Docquier and Marfouk (2006) (migrant stocks), Gleditsch, Wallensteen, Eriksson, Sollenberg and Strand (2002) (conflicts), EM-DAT (2010) (disasters).

Appendix D: First Stage Coefficients of Baseline Regressions

Table D1—First Stage Coefficients: Baseline Estimation for Each Push Factor

	Months of War			Political Regime			Natural Disasters			Log GDP *per capita*		
	Balanced	Incl. female	Unbal.	Balanced	Incl. female	Unbal.	Balanced	Incl. female	Unbal.	Balanced	Incl. female	Unbal.
Primary Education												
[0-8] years	—	—	—	—	—	—	—	—	—	—	—	—
[9-16] years	0.336	0.411	0.360	0.048	0.063	0.053	0.605	0.735	0.679	0.005	0.006	0.005
	(0.184)	(0.171)	(0.156)	(0.031)	(0.029)	(0.030)	(0.325)	(0.295)	(0.286)	(0.003)	(0.003)	(0.003)
[17-24] years	0.264	0.413	0.267	0.031	0.063	0.034	0.523	0.782	0.570	0.003	0.006	0.003
	(0.161)	(0.162)	(0.142)	(0.028)	(0.028)	(0.028)	(0.288)	(0.290)	(0.263)	(0.003)	(0.003)	(0.003)
[25-31] years	0.176	0.387	0.186	0.017	0.060	0.019	0.356	0.739	0.421	0.002	0.006	0.002
	(0.138)	(0.150)	(0.118)	(0.024)	(0.025)	(0.023)	(0.267)	(0.285)	(0.229)	(0.002)	(0.003)	(0.002)
32+ years	-0.005	0.224	0.084	-0.002	0.037	0.020	-0.027	0.378	0.066	-0.000	0.004	0.002
	(0.195)	(0.178)	(0.158)	(0.033)	(0.030)	(0.030)	(0.398)	(0.363)	(0.326)	(0.003)	(0.003)	(0.003)
Secondary education												
[0-8] years	-2.274	-2.227	-2.467	0.702	0.696	0.753	-2.262	-2.250	-2.185	0.073	0.077	0.076
	(0.494)	(0.437)	(0.439)	(0.189)	(0.170)	(0.184)	(0.507)	(0.442)	(0.486)	(0.024)	(0.021)	(0.021)
[9-16] years	-1.983	-1.863	-2.141	0.743	0.752	0.801	-1.693	-1.553	-1.534	0.077	0.082	0.080
	(0.488)	(0.458)	(0.428)	(0.196)	(0.173)	(0.192)	(0.534)	(0.485)	(0.502)	(0.025)	(0.021)	(0.022)
[17-24] years	-2.052	-1.860	-2.230	0.727	0.752	0.783	-1.774	-1.506	-1.640	0.075	0.082	0.079
	(0.528)	(0.485)	(0.459)	(0.191)	(0.170)	(0.186)	(0.558)	(0.502)	(0.514)	(0.025)	(0.021)	(0.021)
[25-31] years	-2.130	-1.876	-2.300	0.714	0.751	0.770	-1.933	-1.540	-1.780	0.074	0.082	0.077
	(0.547)	(0.496)	(0.472)	(0.187)	(0.168)	(0.183)	(0.587)	(0.529)	(0.535)	(0.024)	(0.021)	(0.021)
32+ years	-2.285	-2.012	-2.382	0.700	0.733	0.774	-2.293	-1.878	-2.115	0.073	0.080	0.078
	(0.566)	(0.524)	(0.493)	(0.189)	(0.166)	(0.182)	(0.583)	(0.521)	(0.503)	(0.025)	(0.021)	(0.021)
Tertiary education												
[0-8] years	-2.397	-2.334	-2.462	0.754	0.750	0.766	-2.630	-2.500	-2.490	0.096	0.094	0.091
	(0.565)	(0.506)	(0.496)	(0.206)	(0.184)	(0.199)	(0.580)	(0.511)	(0.521)	(0.026)	(0.022)	(0.022)
[9-16] years	-2.100	-1.963	-2.131	0.796	0.806	0.815	-2.056	-1.798	-1.835	0.100	0.099	0.096
	(0.553)	(0.518)	(0.480)	(0.213)	(0.186)	(0.206)	(0.594)	(0.545)	(0.530)	(0.026)	(0.023)	(0.023)
[17-24] years	-2.169	-1.960	-2.220	0.780	0.807	0.797	-2.137	-1.750	-1.941	0.099	0.099	0.094
	(0.593)	(0.543)	(0.510)	(0.208)	(0.184)	(0.201)	(0.623)	(0.562)	(0.546)	(0.026)	(0.023)	(0.023)
[25-31] years	-2.250	-1.979	-2.293	0.767	0.805	0.783	-2.299	-1.787	-2.084	0.098	0.099	0.093
	(0.612)	(0.556)	(0.525)	(0.204)	(0.181)	(0.198)	(0.652)	(0.591)	(0.570)	(0.026)	(0.023)	(0.023)
32+ years	-2.403	-2.113	-2.373	0.753	0.787	0.788	-2.657	-2.123	-2.417	0.096	0.098	0.093
	(0.626)	(0.580)	(0.542)	(0.206)	(0.179)	(0.198)	(0.633)	(0.576)	(0.539)	(0.026)	(0.023)	(0.023)
Stability test (p-value)	0.609	0.837	0.755	0.380	0.773	0.873	0.879	0.711	0.722	0.997	0.551	0.637
Excluded F test (critical values below)	6.068[a]	7.215[a]	9.888[a]	4.134[a]	4.990[a]	5.863[a]	4.365[a]	5.228[a]	6.318[a]	2.473[a]	3.078[a]	3.847[a]

[a] Stock and Yogo (2005) critical values for Excluded-F are 4.67, 6.45, and 11.52 for maximum relative biases of 0.3 0.2, and 0.1.

Note: The table reports the first stage coefficients of the excluded instruments for the baseline first stage regressions (balanced and unbalanced samples, and including female in labor counts). A separate regression is estimated for each instrument. The regression relates the share of immigrants in each education-experience-period-country cell with the interaction of the corresponding push factor (listed on the top) with distance and education-experience cell dummies, as described in the text. The regression includes education, experience, country-period, education-period, experience-period, and education-country fixed effects.