Labour markets

SUMMARY

The continuing inflow of hundreds of thousands of refugees into many European countries has ignited much political controversy and raised questions that require a fuller understanding of the determinants and consequences of refugee supply shocks. This paper revisits four historical refugee shocks to document their labour market impact. Specifically, we examine: The influx of Marielitos into Miami in 1980; the influx of French repatriates and Algerian nationals into France at the end of the Algerian Independence War in 1962; the influx of Jewish émigrés into Israel after the collapse of the Soviet Union in the early 1990s; and the exodus of refugees from the former Yugoslavia during the long series of Balkan wars between 1991 and 2001. We use a common empirical approach, derived from factor demand theory, and publicly available data to measure the impact of these shocks. Despite the differences in the political forces that motivated the various flows, and in economic conditions across receiving countries, the evidence reveals a common thread that confirms key insights of the canonical model of a competitive labour market: Exogenous supply shocks adversely affect the labour market opportunities of competing natives in the receiving countries, and often have a favorable impact on complementary workers. In short, refugee flows can have large distributional consequences.

JEL codes: J2, J23, J30, J61, J68

—George J. Borjas and Joan Monras

Economic Policy July 2017 Printed in Great Britain

The labour market consequences of refugee supply shocks

George J. Borjas and Joan Monras*

Harvard University, NBER and IZA; Centro de Estudios Monetarios y Financieros (CEMFI) and IZA

1. INTRODUCTION

The recent inflow of hundreds of thousands of Syrian refugees into many European countries has inevitably rekindled interest in documenting the determinants and consequences of such "refugee supply shocks." Although the war in Syria started in 2011, and refugee camps formed in the area soon thereafter, the refugees initially moved mainly to Lebanon, Jordan, and Turkey. As the Syrian conflict continued and escalated, the refugees began to move to Europe through Greece, with alternate routes quickly emerging in Hungary, Austria, and the Balkans. It is difficult to enumerate precisely just how many refugees have already entered the continent, but many news reports claim that over one million asylum seekers arrived in Europe in calendar year 2015.

*We are grateful to Dante Amengual, Andrea Ichino, Joan Llull, Jan Stuhler, and four referees for valuable comments on a previous draft of this paper.

The Managing Editor in charge of this paper was Andrea Ichino.

Economic Policy July 2017 pp. 361–413 Printed in Great Britain

This inflow of refugees has already generated a great deal of political conflict in all the receiving countries, and exposed major fissures in the economic, social, and cultural fabric that hold together the European Union. Much of the controversy surrounds the long-term implications of the open-door policy implicit in German Prime Minister's Angela Merkel's unilateral assertion that "the fundamental right to asylum for the politically persecuted knows no upper limit" (Alexe, 2015).

The full consequences of the epochal events now reverberating throughout Europe will not be known for many years (or perhaps even decades). Nevertheless, the persistent influx of large numbers of refugees raises fundamental questions about their impact that encourage a "revisiting" of other refugee supply shocks in other countries and at other times to determine whether there are universal lessons to be learned from such shocks.

This paper provides such a revisiting. Despite the obvious differences in the factors that have motivated refugee shocks throughout history—including the size and timing of the flows, the skills of the refugees, and the countries and localities affected—there are similarities as well, and these similarities can help provide a unifying framework for how to think about the labour market consequences of current or future supply shocks.

Almost by definition, refugee supply shocks are exogenous along a number of important dimensions. The timing of the supply shock typically has little to do with economic conditions in the receiving countries. The size of the supply shock depends at least partly on the circumstances that created the exogenous political turmoil. And the skill composition of the refugees often hinges on the nature of the political conflict that motivated the exodus. In some cases, these political events lead to an outflow of high-skill workers, while in other cases they lead to an outflow of low-skill workers.

The paper reexamines the evidence surrounding some key historical refugee supply shocks. In particular, we document the labour market consequences of four distinct shocks, each of which has been analyzed separately in previous research:

1. The flow of Cuban refugees in the Mariel boatlift in 1980, a shock that affected mainly the city of Miami (Card, 1990; Peri and Yasenov, 2015; Borjas, 2016, 2017).
2. The flow of Jewish émigrés to Israel following the collapse of the Soviet Union in the early 1990s (Friedberg, 2001).
3. The flow of refugees into France, both French repatriates and Algerian nationals, that followed the conclusion of the Algerian War of Independence in 1962 (Hunt, 1991).
4. The flow of refugees into several European countries from the long Yugoslav Wars during the 1990s (Angrist and Kugler, 2003).

Table 1 summarizes some of the essential details that characterize these supply shocks. The Mariel shock involved a total of about 120,000 refugees; the exodus created by the Yugoslav Wars involved 250,000 persons; the shock of Soviet émigrés into Israel involved almost 500,000 refugees; and nearly 1.5 million refugees entered France after the end of the Algerian conflict. The different shocks also differed in their skill composition. The Mariel shock consisted mainly of very low-skill workers, with most of the refugees

Table 1. Overview of the four refugee supply shocks

Refugee supply shock:	Number of refugees (in 1000s) All	Number of refugees (in 1000s) Men aged 25–59 years	Localities/ occupations most affected	Predominant skills of the refugees	Increase in supply of most affected group
1. Mariel, 1980	120.6	47.9	Miami	High school dropouts	31.9% (male high school dropouts in Miami)
2. Soviet émigrés to Israel, 1990	476.5	101.6	Skilled workers in industry and construction	College graduates	267.9% (male college graduates in "skilled workers in industry and construction")
3. Algerian Independence War, 1962					
A. French Repatriates	1358.9	302.0	Provence–Alpes–Cote d'Azur	Balanced across groups	10.2% (men in Provence–Alpes–Cote d'Azur)
B. Algerian Nationals	162.1	77.1	Provence–Alpes–Cote d'Azur	Less than primary schooling	5.8% (men with less than primary schooling in Ile de France)
4. The Yugoslav Wars, 1991–2001	258.6	65.1	Some cities in Austria and Switzerland	Secondary schooling completed	4.6% (men in Vienna)

lacking a high school education; the Soviet émigrés entering Israel were disproportionately high-skill, with most of them having at least a college degree; and the refugee flow exiting Algeria consisted of both extremes, with many low-skill Algerian nationals and many at least moderately skilled French repatriates.

Although each of these shocks has been examined independently in prior research, our analysis differs in three crucial ways. The existing studies "pick and choose" a particular methodological approach, often based on the data available or on the idiosyncratic characteristics of a particular shock. An obvious problem with this piecemeal approach is that it is unclear if the empirical findings truly reveal universal insights about the impact of refugee supply shocks, or instead reflect that a particular researcher chose a particular methodological approach to study the impact of a particular episode. Put bluntly, are the findings documented in the literature sensitive to the choice of methodological approach used to examine the impact of a particular supply shock?

Our analysis instead derives a *single* empirical approach based on the short-run implications of factor demand theory. In principle, this methodological approach can be applied to measure the consequences of any refugee supply shock. The theoretical derivation indicates precisely the correlation between labour market outcomes and the number of refugees that should be estimated in any specific context. And it also delineates the conditions under which that correlation can be interpreted as measuring a causal short-run impact of the refugee-induced increase in the supply of labour.

Second, our analysis plays close attention to isolating the particular groups that are most likely to be affected by refugee supply shocks. As noted earlier, the supply shocks sometimes consist of high-skill workers, while in other cases they consist of low-skill workers. One important lesson from our examination of the evidence is that the adverse labour market impact of refugee supply shocks can only be properly estimated when the analysis closely matches the skills of the refugees with those of the native workers who are most likely competing in the same labour market.

Equally important, the emphasis on the skill distributions of natives and of refugees implies that we can also examine the short-run impact of the shocks on potentially complementary native groups. The low-skill *Marielitos* may have increased the wage of high-skill Miamians, and the high-skill émigrés may have increased the wage of low-skill Israelis. These potential complementarities are an important part of any complete assessment of the economic consequences of refugee supply shocks. Our analysis provides the first estimates of the cross-effects of immigration that are based entirely on observed data and are not contaminated by extraneous assumptions about the functional form of the aggregate production technology.

Finally, rather than rely on proprietary or confidential data, we use the publicly available censuses maintained at IPUMS. Although these data are sometimes less than ideal, they can be easily adapted to measure the labour market consequences of refugee supply shocks on both competing and complementary workers. In view of the very contentious policy debate over the economic impact of immigration, the use of publicly available data has one non-trivial implication: Our results are fully reproducible.

The empirical analysis reported below uses the theory-derived empirical specification to estimate the impact of the *Marielitos*, of the French repatriates and Algerian nationals moving to France, of the Soviet émigrés in Israel, and of the refugees from the Yugoslav wars into several European countries. Despite the differences in the historical events that triggered the refugee flows, in the skill composition of the refugees, and in the countries and localities affected by the shocks, the evidence reveals a common thread: Exogenous refugee supply shocks have an adverse effect on the labour market opportunities of competing natives in the destination countries. Depending on the episode and the data, we document that the shock sometimes reduces the wage of competing workers; sometimes it reduces their employment rates; and sometimes it reduces both. At the same time, exogenous supply shocks often have a beneficial impact on the employment opportunities of *complementary* native workers. In short, refugee supply shocks have sizable short-run distributional consequences in the labour markets of receiving countries.

2. FRAMEWORK

It is instructive to begin by considering how one would go about estimating the labour market impact of immigration if one had an ideal empirical setting and ideal data. In particular, suppose that the receiving country has a competitive labour market and that volatile political conditions abroad randomly generate a flow of refugees. It is crucial to emphasize that the refugee supply shock is random along *all* relevant dimensions, including the timing, the size and skill composition of the flow, and the eventual geographic sorting of the refugees in the receiving country.

The economy of the receiving country is composed of r isolated labour markets. These markets can be defined along a number of characteristics commonly shared by groups of workers. To fix ideas, and because this is the context most often seen in the existing literature, it is useful to think of the index r as indicating a regional labour market (although our approach can be applied to alternative classifications, such as an occupation). In this ideal setting, workers cannot move from one market r to another in response to either supply or demand shocks. The production technology in the firms populating each of these markets uses s different types of workers that are defined along another characteristic, such as their educational attainment. Pairs (r, s) of labour markets and factor types define each of the k different "cells" in which the national labour market can be subdivided and for which data are available.

We can derive a standard isoelastic labour demand function for each of these k cells by assuming that competitive firms maximize profits. Prior to the refugee supply shock ($t = 0$), there are L_{rs0} workers in region r of skill type s. The pre-shock CES aggregate production function for region r is given by:

$$Q_{r0} = \left[\sum_s \alpha_{s0} L_{rs0}^{\delta} \right]^{1/\delta}, \tag{1}$$

where $\delta = (\sigma - 1)/\sigma$; and σ is the elasticity of substitution across worker types. Note that the weights attached to the various skill groups (i.e., the α's) can vary over time, due perhaps to technological shifts that may favor one skill group over another.

Profit maximization implies that the wage paid to workers in cell (r, s) at $t = 0$ is:

$$\log w_{rs0} = \log p_{r0} + \log \alpha_{s0} + \eta \log Q_{r0} - \eta \log L_{rs0}, \tag{2}$$

where p_{r0} is the price level in region r prior to the supply shock, and η $(= 1/\sigma)$ is the wage elasticity.

It is useful to think of L_{rs0} as giving the number of pre-existing workers in cell (r, s) prior to the supply shock. We will often refer to this pre-existing workforce as "natives," but it should be obvious that L_{rs0} could potentially include both native- and foreign-born workers. In the short run, with the quantity of other factors of production held constant, economic theory predicts that an increase in the size of the workforce in a particular cell reduces the "own" wage.[1] Note also that w_{rs0}, the equilibrium wage prior to the

1 Differentiating Equation (2) with respect to L_{rs0} yields $\partial \log w_{rs0}/\partial \log L_{rs0} = -(1-\kappa_s)/\sigma$, where κ_s is the share of income accruing to skill group s.

refugee supply shock, incorporates the impact of all immigration-induced supply shocks that occurred before the entry of the new flow of refugees.

The labour markets in the receiving country are then "shocked" by the political upheaval abroad. This upheaval sends an influx of M_{rs} new refugees into each region–skill cell. We can write the post-shock marginal productivity condition as:

$$\log W_{rs1} = \log p_{r1} + \log \alpha_{s1} + \eta \log Q_{r1} - \eta \log (L_{rs1} + M_{rs}). \tag{3}$$

The wage change observed in cell (r, s) can then be written as:

$$\begin{aligned} \Delta \log w_{rs} &= \Delta \log p_r + \eta \Delta \log Q_r + \Delta \log \alpha_s - \eta \log \frac{L_{rs1} + M_{rs}}{L_{rs0}}, \\ &= \theta_r + \theta_s - \eta \log \frac{L_{rs1}(1 + m_{rs})}{L_{rs0}}, \\ &= \theta_r + \theta_s - \eta \log \frac{L_{rs1}}{L_{rs0}} - \eta\, m_{rs}, \end{aligned} \tag{4}$$

where $\theta_r = \Delta\log p_r + \Delta\log Q_r$, and is captured by a region-specific fixed effect; $\theta_s = \Delta\log \alpha_s$, and is captured by a skill-specific fixed effect; and $m_{rs} = M_{rs}/L_{rs1}$.[2] Note that m_{rs} gives the relative size of the supply shock: the percent increase in the number of workers due to the entry of refugees into cell (r, s).

In addition to the fixed effects θ_r and θ_s, Equation (4) has two regressors. The wage change obviously depends on the refugee supply shock. Although there is much confusion in how this shock should be measured (compare Borjas, 2003, and Card and Peri, 2017), the marginal productivity condition that underlies the theory-based empirical approach clearly indicates that the measure of the supply shock should give the percent by which immigrants increased the size of the workforce, with the base being *the number of native workers in the post-shock period*.[3]

Equation (4) also shows that the wage may have changed because the number of native workers in cell (r, s) might have risen or fallen between the two periods. Some of the change may be due to demographic factors that are unrelated to changes in economic conditions during the relevant period, such as mortality in the pre-existing workforce or secular trends in the skill mix of the native population. But some of the change may be endogenous, induced by the refugee supply shock itself. In other words, the entry of the M_{rs} refugees may generate a labour supply response in the native population.

Assume initially that the change in the supply of pre-existing workers is exogenous, due to long-term demographic factors. We have already assumed that the refugee

2 The derivation of Equation (4) uses the approximation log (1 + *mrs*) ≈ *mrs*, which is appropriate as long as the refugee supply shock is "small."

3 Card and Peri (2017) argue that it is preferable to use the pre-shock period workforce as base (see also Dustmann *et al.*, 2017). The bias induced by any particular specification is related to the endogenous native labour supply response. We discuss this response in greater detail below.

supply shock is, by definition, exogenous. The correct specification of a regression model that estimates the impact of the refugee supply shock would then relate the wage change in a particular labour market to the percent change in supply in the native population and to the percent change in supply due to the refugees (as well as region- and skill-fixed effects). The two "supply" regressors should have identical coefficients, and those coefficients, as indicated by Equation (4), should equal the wage elasticity η.

3. STATISTICAL DIFFICULTIES

It is obvious that the real-world data typically available to measure how immigration affects labour markets do not meet the ideal conditions of the refugee supply shock discussed above. Although the timing of the shock may be independent from economic conditions in the receiving country, the actual number of refugees as well as their distribution across the (r, s) cells will be affected by those conditions. After all, only those persons who have the most to gain by leaving will be the ones likely to end up as refugees. Moreover, those self-selected refugees will tend to settle in those regions of the receiving country that offer the most favorable economic opportunities.

Natives will also respond to the refugee supply shock. These responses imply that the region–skill cells cannot be thought of as isolated islands, and that supply shocks that affect one cell have spillover effects on other cells. In the short run, for example, native workers or firms might move from one regional labour market to another to take advantage of the changes in the wage structure. In the long run, the demographic variables that may be the "fundamentals" determining endowments of each factor of production are no longer exogenous, as natives might pursue particular types of human capital investments and avoid others.

In addition to these endogeneity issues, there is a measurement problem inherent in this type of analysis that might generate substantial bias: the skills that refugees acquired prior to the political upheaval might not be very valuable to employers in the receiving country. For instance, a college degree acquired abroad might not have the same "knowledge content" as a college degree acquired in the receiving country. Similarly, language difficulties might impose a barrier for migrants wishing to enter certain occupations. As a result, the observable skills of the refugees, as measured by years of educational attainment or professional certificates, provide erroneous information about which specific factors of production they are truly competing with or complementing. This measurement error in the size of the supply shock in cell (r, s) will, in general, bias the estimate of the wage elasticity.

We use the empirical counterpart of Equation (4) to estimate the wage effects of the refugee supply shock and to discuss various identification problems. Our basic empirical regression specification is given by:

$$\Delta \log w_{rs} = \theta_r + \theta_s - \eta \, \log \frac{L_{rs1}}{L_{rs0}} - \eta \, m_{rs} + \epsilon_{rs}. \tag{5}$$

A key requirement for estimating the wage elasticity η is that the residual ϵ_{rs} be independent from both the size of the refugee supply shock and from the size of the native response. It is easy to imagine many real-world situations in which such a restriction will fail to hold.

3.1. Endogenous native labour supply

One statistical problem that affects estimates of the wage elasticity arises from the endogeneity of native labour supply. Remarkably, the existing literature has, at best, only superficially addressed the biases created by this type of native response.[4]

The endogeneity of the change in native labour supply in a particular region–skill cell, $\Delta \log L_{rs}$, can arise due to two distinct factors. First, the amount of labour that native persons already participating in the labour market will offer to employers likely depends on the wage. Put differently, the refugee supply shock affects native labour supply at the intensive margin. Second, the number of natives who choose to offer their services in a particular labour market will respond to changes in the market wage, creating a native response to the refugee supply shock at the extensive margin as well.

Regardless of which margin we are referring to, it is easy to see how endogenous native labour supply contaminates estimates of the wage elasticity by taking a first-order Taylor's expansion of the log change in the size of the native workforce. Equation (5) can then be rewritten as:

$$\Delta \log w_{rs} = \theta_r + \theta_s - \eta \ \frac{L_{rs1} - L_{rs0}}{L_{rs1}} - \eta \ m_{rs} \ + \ \epsilon_{rs}. \tag{5 ′}$$

We can then posit a standard model of the labour supply response of natives by writing:

$$\frac{L_{rs1} - L_{rs0}}{L_{rs1}} = \gamma \frac{M_{rs1}}{L_{rs1}} + \ u_{rs}, \tag{6}$$

where the parameter γ measures the native labour supply response. If the refugee supply shock lowers the market wage, the supply parameter γ is negative as long as the substitution effect dominates the income effect in the neoclassical labour supply framework. In other words, as the entry of refugees lowers the price of leisure, not only do fewer natives

4 There are some exceptions. For example, Borjas (2003, Table III) estimates the wage impact of immigration using a regression model that includes a variable giving the number of native workers in the skill group (which is then differenced by adding appropriate fixed effects to the model). However, the properties of the wage elasticities resulting from this particular specification have not been examined in the subsequent literature, despite the widespread adoption of the “skill-cell” approach. Similarly, Monras (2015*a*) includes the changes in the level of regional GDP and in native labour supplies of the various skill groups in his main regression specification.

work, but those who do remain in the workforce work fewer hours. We can substitute the labour supply response in Equation (6) to obtain the reduced form:

$$\Delta \log w_{rs} = \theta_r + \theta_s - \eta\,(1+\gamma)\,m_{rs} + \epsilon^*_{rs}. \tag{7}$$

Equation (7) shows that if we simply exclude the change in the native-born workforce from the estimated regression model (as almost all studies in the literature do), the regression coefficient that relates wage changes to the supply shock measures an amalgam of the wage elasticity η and the labour supply parameter γ. As long as $\gamma < 0$, the OLS estimate of the factor price elasticity is biased toward 0, suggesting that the refugee supply shock had a relatively weak impact on wages. The intuition is obvious: the wage impact of the supply shock is attenuated by the fact that natives supplied less work effort to the labour market, and as a result the *real* supply shock was not as large as implied by mechanically calculating the number of refugees. Equation (7) also illustrates the interesting case where the displacement effect is one-to-one (or $\gamma = -1$). The wage change in cell (r, s) is then uncorrelated with the refugee supply shock because the "complete" native response ensured that there was no supply shock to speak of.

It is worth noting that the magnitude of the supply parameter γ, which determines the size of the downward bias in estimates of the wage elasticity, depends on how the isolated labour markets (r, s) are defined. For example, Borjas *et al.* (1997) documented that the estimated wage elasticity is more negative the larger the geographic size of the labour market (e.g., states as opposed to cities). This result follows easily from Equation (7) because it is probably more costly to move across states than across cities (i.e., γ is more negative the smaller the geographic area). Similarly, in some contexts it may be sensible to define labour markets in terms of occupations, rather than regions. Because it may be more difficult for natives to switch occupations (implying γ is closer to zero), the resulting bias should be relatively small.

3.2. Endogenous migrant locations

A positive spurious correlation between ϵ_{rs} and m_{rs} may arise because migrants *choose* in which localities to settle in the receiving country. Suppose region 1 is thriving (i.e., wages are growing fast), while region 2 is not. Income-maximizing refugees are then more likely to end up in region 1, creating a positive correlation between the change in the wage observed in cell (r, s) and the refugee supply shock, and making it more difficult to detect any potential wage depression caused by the supply shock itself.

The search for an instrument that corrects for this specific type of endogeneity dominates the existing discussion of the statistical problems that arise when measuring the wage impact of immigration. Beginning with Altonji and Card (1991), the typical study uses what has become known as the "migration networks" instrument. Altonji and Card proposed that an instrument for m_{rs} could be the geographic sorting of an earlier wave of immigrants, arguing that the new immigrants would most likely end up in those regions where the earlier immigrants settled because family networks reduce the costs of migration. If labour market conditions in particular areas were uncorrelated over time, this means that new migrants enter particular regions for reasons that are unrelated to current labour market conditions. The migration networks instrument has been refined (Card, 2001) by using a more sophisticated lag based on national origin: the new immigrants from country *j* settle in those cities where earlier waves of type-*j* immigrants settled.

It is widely recognized that using a "lagged supply shock" as an instrument is invalid if economic conditions in local labour markets are serially correlated. The initial waves of type-*j* immigrants chose to settle in region *r* for a reason (including faster wage growth), and if this reason persists over time, the serial correlation violates the condition that the instrument should be independent of the error term in Equation (5).

Although the migration networks instrument is widely used, few studies examine the validity of the zero serial correlation assumption. Jaeger *et al.* (2016) provide a rare and important exception, showing that the non-zero correlation found in real-world local labour markets badly contaminates IV estimates of the wage elasticity. The Jaeger–Ruist–Stuhler solution to the problem, however, makes exacting data demands, requiring that we observe local labour market conditions for a long span of time prior to the supply shock. Such data are not available in the context of the refugee supply shocks we examine. Instead, our empirical analysis adopts the approach introduced by Monras (2015*a*). He argues that the combination of a networks instrument with a supply shock that occurred at time *t* for truly exogenous reasons (combined with adequate controls for the trend in local economic conditions) provides a "compromise" solution that can help identify the effect of migration even in the presence of serial correlation.

It is worth stressing that this particular endogeneity issue remains even if the cells were demarcated by occupation rather than region. Economic conditions will likely attract refugees who have skills in high-demand occupations, creating a spurious positive correlation between the residual in the wage growth regression and the size of the refugee supply shock, and biasing the estimate of η toward zero. We will use an analogous "employment networks" logic to construct an instrument in this context, arguing that the costs of entering an occupation for a new refugee are likely to be lower when that occupation has already been penetrated by their compatriots, who can provide valuable (and cheap) information about job opportunities. The empirical analysis reported below uses this alternative approach when analyzing the Israeli labour market, where the small geographic size of the country hampers the use of geographic variation.

3.3. Downgrading of immigrant skills

A particularly challenging problem arises when the pre-migration skills of immigrants are not a good predictor of the group of native workers with whom they will compete in the receiving country. Some of the training that the refugees acquired prior to the move may be specific to the country of origin, reducing the stock of human capital that is marketable in the post-migration period. As a result, the observation that a particular supply shock had many high-skill workers "on paper" does not necessarily imply that it is the high-skill natives who will be adversely affected by this shock. As demonstrated in Dustmann *et al.* (2013), the classification issues raised by this type of "skill-downgrading" can contaminate estimates of the wage impact of immigration.

It is easy to determine the nature of the bias by considering the generic regression model that allocates immigrants and natives to region–skill cells. Suppose the pre-existing size of the workforce remains constant after the supply shock and that there are two types of workers in each of r regional labour markets: high-skill (h) and low-skill (u). The data, therefore, consist of two observations in each of r locations. Equation (5) then implies that the wage change for each of the two types of workers is given by[5]:

$$\Delta \log w_{rh} = \theta - \eta \frac{M_{rh}}{L_{rh1}} + e_{rh}, \tag{8a}$$

$$\Delta \log w_{ru} = \theta - \eta \frac{M_{ru}}{L_{ru1}} + e_{ru}. \tag{8b}$$

If the pre-migration skills of group h survived the move to the receiving country, Equations (8a) and (8b) would correctly specify the regression model that estimates the wage elasticity η. Suppose, however, that a fraction π of the high-skill refugees "lose" their skills during the move.[6] The regression model that would correctly estimate the wage impact of immigration is then given by:

$$\Delta \log w_{rh} = \theta - \eta \frac{(1-\pi)M_{rh}}{L_{rh1}} + e_{rh} = \theta - \eta \frac{M_{rh}}{L_{rh1}} - \eta\pi \frac{-M_{rh}}{L_{rh1}} + e_{rh}, \tag{9a}$$

$$\Delta \log w_{ru} = \theta - \eta \frac{M_{ru} + \pi M_{rh}}{L_{ru1}} + e_{ru} = \theta - \eta \frac{M_{ru}}{L_{ru1}} - \eta\pi \frac{M_{rh}}{L_{ru1}} + e_{ru}. \tag{9b}$$

By comparing Equations (8a) and (8b) with Equations (9a) and (9b), it is easy to see that the downgrading of skills effectively adds a regressor to the generic regression model. This additional regressor takes on a value of $(-M_{rh}/L_{rh1})$ for the high-skill labour markets, and (M_{rh}/L_{ru1}) for the low-skill labour markets. The coefficient of this additional regressor would equal $-\eta\pi$.

Put differently, the bias introduced by unobserved skill downgrading can be easily reinterpreted as an omitted variable bias. A straightforward application of the omitted-variable bias formula (see the Appendix in Borjas and Monras, 2016) shows that the

5 To simplify the discussion, suppose that the wage growth has been deflated by the observed wage growth observed for each region and for each skill group, so that the regression need not include the vectors of fixed effects θr and θs.

6 More generally, we can think of L as giving the number of efficiency units of a particular group of pre-existing workers, and π would be the rate at which the efficiency units depreciate after the move.

OLS coefficient of the refugee supply shock variable resulting from estimating the misspecified model in Equations (8a) and (8b) is[7]:

$$\operatorname{plim} \hat{\eta} = \eta - \eta\pi \frac{\sigma_h^2}{\sigma_h^2 + \sigma_u^2}\left[1 - \rho_{hu}\frac{\sigma_u}{\sigma_h}\right], \tag{10}$$

where σ_s^2 is the variance in the measure of the supply shock for type s workers across markets; and ρ_{hu} is the correlation between the high-skill and the low-skill supply shocks. Equation (10) implies that if the refugee supply shock does not affect wages ($\eta = 0$), the misclassification of some high-skill immigrants into low-skill cells does not generate any bias. The OLS coefficient measuring the wage elasticity will still be zero.

If the true wage elasticity η is negative, however, skill downgrading biases the estimated wage elasticity, and the nature of the bias depends on how the high- and low-skill supply shocks are distributed across markets. One interesting special case arises when the supply shocks for high-skill and low-skill workers are equally spread out (so that $\sigma_h^2 = \sigma_u^2$). It is then easy to show that the wage elasticity is biased toward zero regardless of the value of ρ_{hu}. Another interesting special case occurs when the correlation ρ_{hu} equals zero, so that (roughly) the cities where high-skill refugees end up provide no information about where the low-skill refugees settle. It is obvious from Equation (10) that the estimated wage elasticity will again be biased toward zero.

3.4. Complementarities across skill groups

Up to this point, our discussion has focused on estimating the impact of a refugee supply shock in a particular region–skill cell on the wage of natives who belong to that same region–skill cell—in other words, the identification of the "own" wage effect of immigration. We have shown that, under certain conditions, the functional form assumption of an aggregate CES production function in a regional labour market produces a very simple regression model that identifies the own wage effect by relating the wage change observed in a particular cell to the refugee supply shock in that cell, even while ignoring the changes that might have occurred in the quantities of other factor inputs. This regression model has become the *de facto* generic regression in the literature (although it is not often linked to a factor demand theoretical framework).

The entry of refugees into a particular skill group obviously has ramifications for the wages of workers in other skill groups, and a full accounting of the impact of the supply shock would require documenting not just the own wage effect of

7 The derivation assumes that the native workforce is equally split between high- and low-skill workers.

immigration, but the "cross-effects" as well. Because the number of potential cross-effects explodes as the number of skill groups increases, the existing literature, including both the early work of Grossman (1982) and the framework introduced in Borjas (2003), reduces dimensionality by exploiting functional form assumptions about the production technology. For example, Borjas (2003) classifies workers into 32 skill groups (four education groups and eight experience groups). If capital is also a factor input, there are then a potential 1,089 wage effects that need to be estimated. The imposition of a nested CES framework on the data, where various skill groups are aggregated into efficiency units, leads to a remarkable reduction in the number of primitive parameters (i.e., the elasticities of substitution). In Borjas (2003), only *three* distinct elasticities of substitution are sufficient to derive all 1,089 potential own- and cross-wage effects.

This reduction in the parameter space, however, comes at great cost. The nested CES framework greatly limits the types of cross-group complementarities that are allowable. Moreover, the functional form assumptions introduce *numerical* constraints on the value of the wage effects. For example, a linear homogeneous Cobb–Douglas production function that has capital and labour efficiency units as inputs implies that the long-run wage effect of immigration averaged across all skill groups is identically equal to zero while the short-run effect equals the negative of the income share of capital. These numerical constraints then cascades over to all other wage effects estimated in such a framework, raising questions about whether the results accurately reflect the underlying data and greatly reducing their value for policy analysis.

To minimize the role of such extraneous assumptions on the estimated cross-effects of immigration, we only assert the existence of a generalized production function. Consider the production function, $F(L_{rh}, L_{ru})$, where L_{rh} gives the number of high-skill workers in region r, and L_{ru} gives the number of low-skill workers. The production function F has the typical properties (i.e., concave, twice differentiable, etc.). If we set the price level as the numeraire, we can then write a general characterization of what happens to wages in region r and skill group s ($s = h, u$) as:

$$\Delta \log w_{rs} = \alpha_{sh} \Delta \log L_{rh} + \alpha_{su} \Delta \log L_{ru}. \tag{11}$$

where α_{sj} gives the factor price elasticity defined by $\partial \log w_{rs} / \partial \log L_{rj}$.[8]

We exploit the fact that the refugees in many of the episodes examined in this paper were often concentrated in one particular skill group. Low-skill refugees made up a large

8 The marginal productivity condition (say, for high-skill workers) is $w_{rh} = F_h(L_{rh}, L_{ru})$. Totally differentiating the first-order condition yields: $\mathrm{d}w_{rh} = F_{hh}\mathrm{d}L_{rh} + F_{hu}\mathrm{d}L_{ru}$. Equation (11) then follows easily from this differential, where the factor price elasticity $\alpha_{ij} = \kappa_i c_{ij}$; κi is the share of income accruing to skill group i; and cij is the elasticity of complementarity ($cij = FijF/FiFj$) between groups i and j.

fraction of the *Marielitos* in Miami and of the Algerian nationals in France, while college graduates dominated the influx of Soviet émigrés in Israel. Consider an episode where all refugees belong to the low-skill group. We can then rewrite Equation (11) as:

$$\Delta \log w_{rs} = \alpha_{sh} \Delta \log L_{rh} + \alpha_{su} \Delta \log L_{ru} + \alpha_{su} m_{ru}, \tag{12}$$

where the Δ log L_{rs} variables are interpreted as the change in the number of *native* workers in cell (r, s); and $m_{ru} = M_{ru}/L_{ru1}$, the measure of the refugee supply shock. We can then estimate Equation (12) separately for each skill group.

This approach exploits the natural experiment created by the refugee supply shock to measure not only the own wage effect, but also the cross-effects. Put differently, the cross-effects are identified by relying on the exogenous nature of refugee supply shocks and on the concentration of the refugees in a very small number of skill groups. The cross-effect is given by the coefficient that relates the labour market outcomes of skill groups untouched (at least directly) by the refugees to the measure of the supply shock in the skill group that was most directly affected by the political upheaval.

3.5. Other problems

In an ideal (from a researcher's point of view) supply shock, the migrants would be randomly selected from the population of the sending country, which is hardly ever the case. Fernandez-Huertas (2011) documents that Mexicans moving to the United States tend to be less skilled than the Mexicans who choose to remain behind. Although this type of selection may be due to a variety of factors, Borjas (1987) shows that differences in the returns to skills between the sending and receiving countries can systematically generate various patterns of selection. As a result, even in the context of refugee supply shocks, labour market conditions at the destination may influence migration incentives. The correlation between labour market conditions in the destination and the self-selection of the refugees would further attenuate the estimate of the wage elasticity η. Although it is recognized that the self-selection of immigrants contaminates the measured wage impact of immigration, there have not been any studies that attempt to quantify this bias.

The supply shock may also generate "general equilibrium" effects because the refugees might influence the average level of productivity in the aggregate economy. One such effect that has received some attention is the possibility that some immigrants, and particularly high-skill immigrants, bring new ideas and knowledge that expand the production frontier. Unfortunately, the typical attempt to estimate the impact of supply shocks on the average wage level in a receiving country has relied on extraneous functional form assumptions about the production technology. As we noted earlier, this approach builds in a numerical answer for the general equilibrium wage effects.

In sum, our discussion shows the importance of thinking carefully about both the underlying theory and the statistical issues created by real-world supply shocks when we measure the labour market impact of immigration. In one sense, the measurement of this wage impact is a trivial exercise. The canonical model of supply and demand, which is fundamental to our understanding of how real-world labour markets work, predicts that the refugees will obviously lower the wage of competing native workers in the short run. To conduct yet another study documenting that labour demand curves are downward sloping, therefore, would seem to be a rather pedestrian exercise.

However, measuring the wage impact of supply shocks introduces thorny measurement and statistical problems that have yet to be fully resolved. In fact, labour economists have devoted a disproportionate amount of time in the past three decades to document what is, in the end, a trivial empirical finding. The resulting confusion (and sometimes obfuscation) in the literature has not been a productive contribution to the immigration policy debate. The examination of refugee supply shocks—which are truly exogenous on at least some dimensions—can perhaps help clarify and increase our understanding of how immigration affects real-world labour markets.

4. MARIEL

On 20 April 1980, Fidel Castro declared that Cuban nationals wishing to emigrate could leave freely from the port of Mariel. Cuban-Americans living in the United States quickly organized a boatlift to bring their relatives. The first migrants arrived on 23 April, and over 100,000 had taken advantage of Castro's invitation by 3rd June. By the time the boatlift ended through an agreement between the US and Cuban governments in October 1980, about 125,000 Cubans had moved and Miami's workforce had grown by about 8%. The *Marielitos* were disproportionately low-skill, with most lacking a high school diploma. The Mariel supply shock increased the size of this low-skill workforce in Miami by nearly 20%.

We begin our empirical analysis of refugee supply shocks by reexamining the Mariel data using the factor demand framework introduced earlier. The Mariel context plays a prominent role in the literature. Card's (1990) landmark study of this particular supply shock was a pioneer in the now-common approach of exploiting natural experiments to measure parameters of great policy interest. Card looked at labour market conditions, including wages and unemployment, in Miami in the years before and after Mariel, and compared the change in those variables to what was happening in comparable cities that were presumably unaffected by the refugees. Surprisingly, the relative wage for the average worker in Miami remained unchanged, leading Card to conclude that even sizable supply shocks had little effect on the price of labour in the affected markets.

There has been a flurry of renewed interest in the Mariel supply shock in the past year, sparked by the Borjas (2017) reappraisal of the Mariel evidence.[9] Borjas argued that it is crucial to study the impact of the *Marielitos* by focusing specifically on the earnings of the workers most likely to be affected by the supply shock—namely, the low-skill workforce. It turns out that the comparison of low-skill wages in Miami and various control cities before and after Mariel overturns the perception that Mariel had a negligible effect, showing instead that the relative earnings of male high school dropouts in Miami fell and that the magnitude of the wage drop was substantial.

4.1. Summary statistics

Table 2 summarizes the available data on the size and skill composition of the Mariel supply shock. The Cuban refugees began to arrive only a few weeks after the 1980 decennial census was conducted, so that the first enumeration of the *Marielitos* in large-scale surveys was not done until 1990. Specifically, the 1990 census reports the number of Cuban-born persons who moved to the United States in 1980 or 1981.[10] We define this group to be the population of refugees resulting from the Mariel supply shock.

The 1990 census enumerated 120,605 such immigrants. That census also reports the geographic location of the refugees as of 1985, with 69.4 thousand of the *Marielitos* (or almost 60%) living in Miami five years after the shock. Note that although all existing studies of the impact of the Mariel supply shock focus on labour market outcomes in the Miami metropolitan area relative to some set of placebo cities, 40% of the *Marielitos* were living outside the Miami area within 5 years after the shock. The main alternative locations were New York City (which housed 13% of the refugees), Los Angeles (7%), and Tampa (3%).[11]

9 See Peri and Yasenov (2015) and Borjas (2016). The empirical debate hinges on whether women and non-Cuban Hispanics should be included in the sample when calculating the average wage in a local labour market. Borjas (2016) notes that the inclusion of women is problematic because the female labour force participation rate was rising rapidly in the 1980s, and it grew differentially in different metropolitan areas. The inclusion of women then inevitably changes the sample composition over time, contaminating wage trends. Similarly, nearly half of the additional observations that would be added by including non-Cuban Hispanics in the sample are of immigrants who arrived *after* Mariel, again changing the sample composition and contaminating wage trends in local labour markets. Finally, the original draft of the Peri and Yasenov (2015) study examined a sample of workers aged 16–61 years, but did not exclude persons enrolled in school. This led to the erroneous classification of high school students aged 16–18 years as "high school dropouts" because those students had not yet received their high school diploma. It is worth noting that the possibility that Card's evidence does not correctly convey what happened to the low-skill labour market in Miami was first noted in the Online Appendix of Monras (2014), which examined wage trends in the *pooled* sample of high school dropouts and high school graduates and documented a relative decline in the wage of Miami's low-skill workforce.

10 The data from later surveys, including the 2000 decennial census, indicate that the number of Cuban immigrants who arrived in 1981 is relatively small, so that this definition of the *Marielitos* should not create substantial measurement error; see the detailed discussion in Borjas (2017).

11 We identify the 1985 locations from the 1990 census data both because of the larger sample size of the decennial census *and* because the census specifically identifies persons born in Cuba (as opposed to

Table 2. Size and skill composition of the Mariel supply shock

	Census data, 1990		
	Marielitos	Natives	% increase in supply
All persons (in 1000s)	120.6	247,339.0	0.05
Aged 25–59 years	73.2	105,674.6	0.1
Men, 25–59 years	47.9	51,696.4	0.1
% of men aged 25–59 years with education:			
High school dropouts	62.2	20.1	0.3
High school graduates	17.3	27.7	0.1
Some college	13.8	26.3	0.0
College graduates	6.6	25.9	0.0
Sample size, men 25–59 years	2,211	2,577,549	
	Census data 1990, Miami counts		
	Marielitos	Natives	% increase in supply
All persons (in 1000s)	69.4	852.7	0.08
Aged 25–59 years	54.7	576.0	0.09
Male, aged 25–59 years	34.5	290.5	0.12
% of men aged 25–59 years with education:			
High school dropouts	62.4	23.2	31.9
High school graduates	15.9	21.7	8.7
Some college	14.5	27.5	6.3
College graduates	7.2	27.6	3.1
Sample size, men 25–59 years	833	6,692	
	March CPS data, pooled 1978–1984 surveys		
	Sample size	Δlog wage	Δunemployment rate
Outside Miami			
High school dropouts	8,718	−0.17	0.06
High school graduates	20,299	−0.15	0.04
Some college	12,431	−0.12	0.03
College graduates	19,898	−0.06	0.01
Miami			
High school dropouts	146	−0.41	0.04
High school graduates	218	0.02	0.03
Some college	107	−0.09	−0.06
College graduates	192	−0.02	−0.03

Notes: The top two panels report data from the 1990 census based on 1985 locations, with age levels referring to 1985. The Marielitos are Cuban immigrants who arrived in the United States in 1980 or 1981; the natives are persons who are neither non-citizens nor naturalized citizens. The bottom panel reports statistics calculated in the sample of non-Hispanic men aged 25–59 years, who live in one of the 38 metropolitan areas. The (Δlog wage) and (Δunemployment rate) variables give the average change between the pooled 1978–1980 CPS surveys and the pooled 1982–1985 surveys. The 1981 survey, which reports earnings for the 1980 calendar year, is not used in the calculations. The sample sizes give the number of observations in the pooled March CPS survey years between 1978 and 1984, excluding 1981, used to compute the change in the average wage before and after 1980.

having Cuban ancestry, which is the only information available in the CPS). To further increase the precision of our measure of the supply shock, we remove other immigrants arriving in the United States between 1980 and 1985 from the base population in each cell. Specifically, we measure $m_{rs} = C_{rs1}/(L_{rs1} - C_{rs1} - I_{rs1})$, where Crs gives the number of *Marielitos* in cell (r, s); Irs gives the number of "other" immigrants who arrived in 1980–1985; and $Lrs1$ gives the size of the cell in 1985. The regression results are almost identical if we do not exclude the other immigrants from the base.

The regression framework derived earlier allows for the *Marielitos* to have an impact on cities other than Miami. This fact marks one key distinction between our regression-based approach and the treated–untreated difference-in-differences methodology employed by both Card (1990) and Borjas (2017). We discuss the implications of this methodological distinction in greater detail below.

Table 2 also confirms the insight that motivated the Borjas (2017) reappraisal. The Mariel supply shock was composed of disproportionately low-skill workers. Over 60% of the refugees lacked a high school diploma, when compared with only 20% of the native-born workforce in Miami. In contrast, only 7% of the Marielitos had a college diploma, when compared with over 25% of native workers. As a result, even though the Marielitos increased Miami's population by only 8%, they increased the number of male workers without a high school diploma by 32%.

The bottom panel of the table shows that the rate of wage growth for high school dropouts was far lower in Miami than outside Miami. Interestingly, the table also shows that the rate of wage growth for high school graduates, a group whose size was only increased modestly by the *Marielitos*, is noticeably higher in Miami than outside Miami. Similarly, the unemployment rate of high-skill workers decreased in Miami, while increasing in the rest of the country. These patterns suggest that refugee supply shocks generate not only adverse own wage effects, but may also improve labour market conditions for complementary workers, a result that was overlooked in earlier studies.

The summary statistics reported in the bottom panel of the table show that the rate of wage growth for high school dropouts was far lower in Miami. The rate of wage growth is calculated by pooling several cross-sections of the March CPS data (thereby increasing sample size). Specifically, the pre-Mariel average wage is calculated from the pooled 1978–1980 surveys, while the post-Mariel average wage comes from the 1982 to 1985 surveys. The table also shows that the rate of wage growth for high school graduates, a group whose size was only increased modestly by the *Marielitos*, is noticeably higher in Miami than outside Miami. Similarly, the unemployment rate of high-skill workers decreased in Miami, but increased in the rest of the country. These patterns hint at the possibility that supply shocks generate not only adverse own wage effects, but also improve labour market conditions for complementary workers, a result that was overlooked in earlier studies.

4.2. Results

We initially use the regression models in Equations (5) and (7) to identify the own effects of the Mariel supply shock. The analysis uses the sample of 38 metropolitan areas, including Miami, which can be consistently matched over the 1977–1984 period.[12]

12 We use the aggregated three-digit version of the *metarea* variable in the IPUMS files (rather than the four-digit version that would generate a sample of 44 metropolitan areas) to avoid including in the analysis local labour markets that have very few observations. Note that the measure of the supply shock in the right-hand side of the regressions is drawn from the 1980 and 1990 census data.

Table 3. The impact of the Mariel supply shock on competing workers

	OLS		IV	
	(1)	(2)	(3)	(4)
A. First stage				
Lagged supply shock	1.260	1.262	—	—
	(0.053)	(0.053)		
Change in native population	—	−0.002	—	—
		(0.001)		
B. Change in log weekly wage				
Mariel supply shock	−1.313	−1.350	−1.264	−1.310
	(0.338)	(0.346)	(0.320)	(0.322)
Change in native population	—	0.039	—	0.039
		(0.045)		(0.038)
C. Change in unemployment rate				
Mariel supply shock	0.060	0.066	0.007	0.015
	(0.072)	(0.075)	(0.079)	(0.083)
Change in native population	—	−0.007	—	−0.006
		(0.019)		(0.016)
D. Change in employment rate				
Mariel supply shock	−0.001	−0.001	0.052	0.053
	(0.092)	(0.097)	(0.102)	(0.107)
Change in native population	—	−0.000	—	−0.001
		(0.025)		(0.021)

Notes: Robust standard errors are reported in parentheses. The unit of observation is a city–education cell, and the data consist of 38 metropolitan areas and 4 education groups. The "Mariel supply shock" variable gives the ratio of the number of *Marielitos* in the cell to the number of natives in the cell as of 1985. The "change in native population" variable gives the log difference in the number of native persons in the cell between 1980 and 1985. The first-stage regression in Panel A relates the relative inflow of *Marielitos* in the cell as of 1985 to the share of Cubans in the cell as of 1980. All regressions have 152 observations and include both education-fixed effects and metropolitan area-fixed effects.

We classify workers into four education groups: high school dropouts, high school graduates, some college, and college graduates. We examine the labour market outcomes of non-Hispanic men aged 25–59 years, a group that approximates the prime-age native-born workforce in Miami, the city most affected by the Marielitos, around 1980. The unit of observation in the regressions is a city-education cell, so that the identifying variation arises both from the fact that the Marielitos settled in a specific set of locations and were disproportionately represented in the least-skilled group.

Table 3 reports the coefficients that estimate the own effect of the Mariel supply shock. Throughout the analysis, the regressions are weighted by $(n_1 n_0)/(n_1 + n_0)$, where n_t gives the number of observations used to calculate the dependent variable in a particular city-education cell at time t.[13] As implied by the specification in Equations (5) and (7),

13 These are the optimal weights for aggregated first-differenced cells from micro-level data. The variance of the differenced average residual (assuming the variance of the person-specific error term is not serially correlated and has constant variance) is given by $(\sigma_\epsilon^2/n_0 + \sigma_\epsilon^2/n_1)$. The optimal weight takes into account the fact that measurement error of the mean value in the cell is less accurate if the number of observations in that cell is small.

all the regressions include education fixed effects and metropolitan area fixed effects.[14] We use three alternative dependent variables: the rate of wage growth in a city-education cell (where the wage variable measures weekly earnings); the change in the average unemployment rate (where the unemployment rate is defined as the ratio of the number of persons unemployed to the number of persons in the labour force); and the change in the average employment rate (defined as the ratio of the number of employed to the size of the corresponding population).

It is useful to begin by discussing the regression coefficients from the simplest OLS specifications reported in the first two columns of Panel B of the table, where the dependent variable is the change in the average log weekly wage. The theory-based specification derived earlier requires the inclusion of the variable $\Delta \log L_{rs}$, a regressor that gives the log change in the size of the native workforce in the cell. The table reports coefficients from two alternative models that address the endogeneity of this variable in different ways. First, we simply exclude the variable from the regression so that the estimated wage elasticity is a reduced-form coefficient that incorporates the native labour supply response (and is biased toward zero). Alternatively, we replace $\Delta \log L_{rs}$ with the corresponding change in the native-born *population* in that cell, so that the coefficient of this variable becomes a type of reduced-form coefficient.[15] Note that the coefficient of the variable measuring the size of the Mariel supply shock is about -1.3, and statistically significant, regardless of how we address the endogeneity of the native labour supply response.[16] Figure 1 visualizes the wage impact of the *Marielitos*. The negative own wage effect is driven mostly by the changing market conditions facing low-skill workers in the very small number of cities where most of the refugees settled.

A crucial feature of the Mariel supply shock is that the refugees could only leave Cuba from the port of Mariel. Many of the Cuban-Americans who already lived in the United States bought or rented boats they would then take to Mariel to pick up their relatives (as well as other potential refugees) waiting at the port. Given the extreme clustering of the pre-Mariel Cuban refugees in the Miami metropolitan area, with about 50% of that population living in Miami in 1980, it is not surprising that about 60% of the Marielitos ended up there as well.

We address the potential endogeneity created by the geographic distribution of the *Marielitos* by using the geographic sorting of the pre-Mariel Cuban immigrants to predict where the new refugees would settle.[17] The first panel of Table 3 shows the relevant

14 The inclusion of these fixed effects do not saturate the regression because the nature of the Mariel supply shock led to very unbalanced supply shifts across both cities and education groups.

15 This reduced-form specification implies that the regression coefficient of the refugee supply shock will generally differ from that of the change in the population of the specific cell.

16 The wage elasticity estimated in the regression framework is very similar to the -1.5 elasticity produced by applying a difference-in-differences approach to the March CPS data; see Borjas (2017).

17 We use data from the 1980 census to calculate the distribution of pre-Mariel Cubans across the region–skill cells.

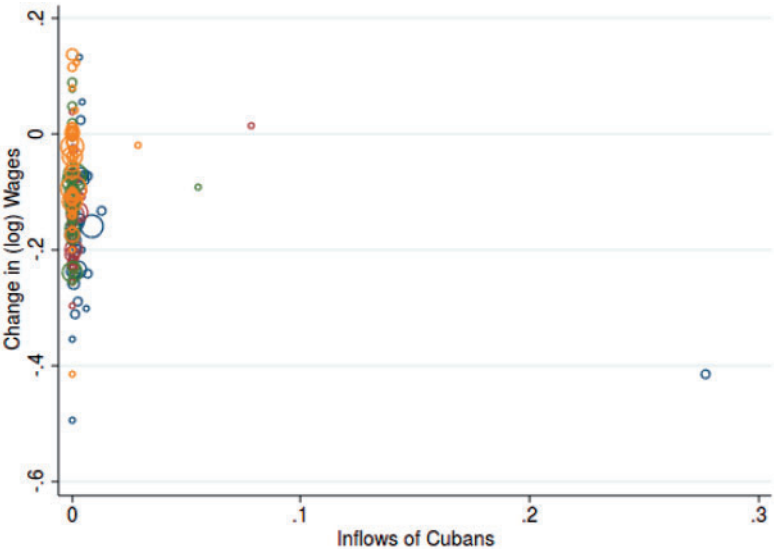


Figure 1. The impact of the Mariel supply shock on native wages

Notes: This figure plots the change in the wage of native men aged 25–59 years against the size of the migrant inflow in each city-education cell using the March CPS survey years of 1978–1980 as the pre-migration period and the March CPS survey years of 1982–1985 as the post-migration period. Each dot represents a city-education cell. The size of the dots represents the size of the cells. Blue, red, green, and yellow dots indicate "less than primary", "Primary completed", "Secondary completed", and "University completed", respectively. The figure exploits variation across 38 metropolitan areas and 4 education groups.

coefficient from the first-stage regression, summarizing the relation between the size of the Mariel supply shock in cell (r, s) and the share of Cubans in that cell prior to 1980. Not surprisingly, the coefficient is strongly positive.[18]

The IV estimates of the wage elasticity are again about -1.3, so that there is little indication that controlling for the endogeneity of the geographic sorting of the refugees plays any role in determining their labour market impact.[19] The similarity between the OLS and IV estimates in the Mariel context is not surprising. The physical characteristics of the boatlift ensure that the geographic sorting of the Marielitos after arrival had little to do with economic conditions *circa* 1980.

18 In other words, we use the lagged share of Cubans in the workforce as the instrument. The first-stage regression coefficient then essentially estimates by how much the pre-existing immigrant workforce in a given cell increased the supply of migrants to that cell as a result of the refugee shock. Some studies in the literature (e.g., Ottaviano and Peri, 2007) use the predicted location of the actual flows observed in the data. If the instrument is sufficiently strong in these alternative IV specifications, the first-stage coefficient, by construction, will hover around 1.0.

19 We also estimated the Mariel regressions using data from the Outgoing Rotation Group (ORG) CPS files rather than the March CPS, and obtained similar results, although, as in Borjas (2017), the point estimate of the own wage elasticity is smaller. The own wage elasticity estimated in the ORG when the regression model includes the change in the size of the native population is -0.51 (0.11) in the OLS regression, and -0.43 (0.16) in the IV regression. These elasticity estimates are again virtually identical to the -0.5 elasticity implied by applying a difference-in-differences methodology to the ORG data.

Table 4. Own and cross effects of the Mariel supply shock

	High school dropouts	High school graduate	Some college	College graduates
A. Change in log weekly wage	−0.906	0.589	0.131	0.245
	(0.352)	(0.311)	(0.445)	(0.368)
B. Change in unemployment rate	−0.083	−0.099	−0.344	−0.160
	(0.208)	(0.181)	(0.196)	(0.080)
C. Change in employment rate	0.083	0.125	0.316	0.050
	(0.282)	(0.228)	(0.291)	(0.138)

Notes: Standard errors are reported in parentheses. The unit of observation is a city, and there are 38 metropolitan areas in the analysis. The table reports the coefficient of the "Mariel supply shock for low-skill workers," which gives the ratio of the number of *Marielitos* who are high school dropouts to the number of natives who are high school dropouts in 1985 in the particular city. The regressions also contain regressors giving the change in the size of the native population of each of the four education groups. The regressions are estimated separately for each education group using IV and have 38 observations.

The bottom two panels of the table report analogous regressions using the change in the unemployment and employment rates as dependent variables. None of the coefficients are significantly different from zero. The own effects of the Mariel supply shock, therefore, seem to be restricted to changes in wages. This finding may be informative about how labour markets adjust to supply shocks during a period of very high inflation. The US inflation rate was 13.5% in 1980 and 10.3% in 1981.

Table 4 reports selected coefficients from the extended regression model that identifies both own- and cross-effects. We can carry out this analysis because most of the *Marielitos* were high-school dropouts, allowing us to examine their impact on the other skill groups. To do so, we estimate Equation (12) separately in each of the four education groups used in the analysis (i.e., high school dropouts, high school graduates, some college, and college graduates). The regression model also includes variables measuring the change in the native population in each of the groups. This regression specification enables us to detect potential complementarities across factor types without imposing any functional form restrictions on the production technology.

As before, the estimated own wage effect is negative and significant, with a wage elasticity of about −0.9. Similarly, the estimate of the own employment effect is not distinguishable from zero.[20] The table, however, shows that the cross-effects are numerically important. Although the supply shock of the predominantly low-skill Marielitos lowered the wage of high school dropouts, it raised the wage of workers with a high school education, and this effect is both numerically and statistically significant. The cross-wage elasticity is about +0.6. In addition, the unemployment rate of workers with more than a high school diploma also fell significantly.

In sum, our analysis of the Mariel supply shock yields an interesting result. As implied by the simplest model of a competitive market, supply shocks can have both negative

20 The instrument in this regression uses only variation in the location of the least-skilled Cubans.

and positive effects on the pre-existing workforce. Those workers who most resemble the refugees suffer the wage loss, while the wage gains accrue to those workers who complement the skills brought in by the refugees. The negative and positive effects of supply shocks, however, need not occur along the same dimensions. In the Mariel context, the own effects tend to show up as wage cuts, while cross-effects are observed in both wages and employment.

4.3. Alternative approaches to natural experiments

It is of interest to contrast the results summarized in Tables 3 and 4 with the evidence reported in existing studies of the Mariel supply shock. There is a key methodological difference between our theory-based regressions and the atheoretical approach exemplified in earlier studies. Both Card (1990) and Borjas (2017) pursued a difference-in-differences approach, comparing the changed conditions in the Miami labour market to the changed conditions in a set of control cities.

It is obviously very difficult to construct perfect control groups or "placebos" outside a laboratory setting. A good placebo needs to satisfy two distinct conditions. First, the treated and control groups must be comparable in important ways. Card (1990) compared Miami with a control group of four cities: Atlanta, Houston, Los Angeles, and Tampa. This particular placebo was partly selected by looking at employment dynamics in various cities both *before and after* the Mariel supply shock. The Borjas (2017) reappraisal showed that the construction of the control group plays a key role in any evaluation of the impact of the *Marielitos*. Using an alternative control group based on employment trends *prior* to the supply shock, as well as employing the Abadie, Diamond, and Hainmueller (2010) synthetic control method, consistently resulted in larger (i.e., more negative) estimates of the own wage effect of the supply shock.

A second condition that a good placebo must satisfy is that there should not be any spillovers between the treatment and control groups. This condition, although conceptually important, has been ignored in all existing Mariel studies. It is trivial to see how such spillovers arise in this context. Nearly 40% of the *Marielitos* chose to settle in cities outside Miami. In fact, two of the cities in the control group used in Card's (1990) study were cities that actually received many refugees: 7.4 thousand settled in Los Angeles and another 3.1 thousand settled in Tampa. The generic difference-in-differences approach, therefore, suffers from the fact that some of the cities in the control group were treated by the exogenous supply shock as well.

Unlike the traditional difference-in-differences calculations that compare Miami and a placebo, our regression-based analysis allows for the refugee supply shock to affect many different markets, along both the region and skill dimensions. We identify the impact by exploiting the different numbers of *Marielitos* settling in different cities and the different numbers of *Marielitos* in different education groups. In other words, the regression approach fully incorporates the fact that the refugee supply shock "treated" many markets and treated those markets differentially, and then uses that dispersion to identify how supply shocks alter labour market outcomes.

5. ÉMIGRÉS TO ISRAEL FROM THE FORMER SOVIET UNION

Prior to the late 1980s, it was extremely difficult for Soviet Jews to migrate to Israel (Buwalda, 1997). The pressures for such migration began soon after the Six-Day War in 1967, when Israel began to more forcefully state its demand that Soviet Jews be allowed to rejoin their families or build a new life in Israel. The Soviet reluctance to allow such migration became an important obstacle in attempts to improve relations between the Soviet Union and the West, and it was not until Michael Gorbachev's Glasnost initiative in 1986 that the Soviet Union allowed the emigration of its Jewish population.

In 1986 and 1987, a small number of visas were granted to Soviet Jews who wished to emigrate. Most of these émigrés, however, settled in the United States or Canada, and only a small fraction moved to Israel. The United States was a particularly appealing destination because it allowed Soviet emigrants to qualify for refugee status, making it relatively easy to obtain entry visas. By 1989, the United States had changed the rule that classified Soviet émigrés as refugees, making it almost impossible for Soviet Jews to move to that country unless an American relative could sponsor their entry. In contrast, Israel's Law of Return continued the open-door policy of welcoming all Jews. As Friedberg (2001) notes: "Between 1989 and 1995, 610,100 immigrants arrived from the [former Soviet Union], increasing the size of the Israeli population by 13.6%."

5.1. Summary statistics

We use data drawn from the 1983 and 1995 Israeli census microdata maintained by IPUMS. Each of these data files represents a 10% random sample of the Israeli population. The censuses report information on country of birth and the year of migration (if born abroad). Using the 1995 census, we define a Soviet émigré as someone born in the former Soviet Union who migrated to Israel between 1990 and 1995. For convenience, we refer to the pre-existing population of Israeli citizens as "natives" even though a large fraction (42%) was born outside Israel. As Table 5 shows, the Soviet émigrés made up almost 10% of the population in 1995.[21]

The table also summarizes key characteristics of the Soviet émigrés. In contrast to the *Marielitos*, the émigrés were highly skilled. Few of them (11%) lacked a secondary

21 The inflow of Soviet émigrés was so large that we measure the size of the shock as $m_k = \text{Soviet}_k/(L_{k1} + \text{Soviet}_k)$, where Soviet_k gives the number of émigrés in cell k. The point estimates of the wage impact do not change significantly if we exclude the number of émigrés from the denominator, but the IV coefficient is less precisely estimated. This is probably because the derivation of our estimating equation uses the approximation that the refugee supply shock is "small," an assumption that is false for many cells in the Israeli context.

Table 5. Size and skill composition of Soviet émigrés in Israel, 1995

	Émigrés	Natives	% increase in supply
All persons (in 1000s)	476.4	4,924.4	9.7
Aged 25–59 years	223.6	1,898.7	11.8
Men aged 25–59 years	101.6	934.4	10.9
% of men aged 25–59 years with education:			
Less than primary	4.5	11.7	4.2
Primary completed	6.2	20.7	3.3
Secondary completed	46.0	49.6	10.1
University completed	43.2	18.0	26.2
% of men aged 25–59 years working as:			
Academic professionals	14.7	11.6	15.3
Associate professionals and technicians	8.2	9.5	10.4
Managers	1.0	9.4	1.3
Clerical workers	3.6	10.1	4.3
Agents, sales workers, and service workers	5.7	13.9	5.0
Skilled agricultural workers	1.3	3.3	4.9
Skilled workers in industry and construction	51.1	35.0	17.6
Unskilled workers	14.3	7.2	24.1
% of men aged 25–59 years with university education working as:			
Academic professionals	29.9	50.5	17.1
Associate professionals and technicians	11.6	10.4	32.2
Managers	1.6	18.0	2.6
Clerical workers	4.8	8.4	16.4
Agents, sales workers, and service workers	5.3	6.6	23.3
Skilled agricultural workers	1.0	1.2	24.0
Skilled workers in industry and construction	35.6	3.8	267.9
Unskilled workers	10.2	1.2	249.3
Sample Size: Men aged 25–59 years	10,160	93,443	

	Δlog earnings	Sample size
Average in Israel, excluding skilled workers in industry and construction		
Less than primary	0.12	12,470
Primary completed	0.13	25,717
Secondary completed	0.07	58,279
University completed	0.17	24,770
Skilled workers in industry and construction		
Less than primary	0.00	2,271
Primary completed	0.01	3,123
Secondary completed	−0.12	2,748
University completed	−0.39	542

Notes: The sample of Soviet émigrés consists of persons born in the former Soviet Union who did not reside in Israel in 1990. The sample of Israeli natives consists of persons who were not born in the former Soviet Union. The bottom panel reports statistics calculated in the sample of Israeli native men aged 25–59 years. The (Δlog earnings) variable gives the average change in annual earnings between the 1983 and 1995 censuses for the particular group.

education, when compared with a third of the Israel population. In contrast, 43% had a university education, when compared with only 18% of the Israeli natives.

The table documents interesting differences between the occupational distributions of the émigré and native populations. Note that 14% of the émigrés, despite their very high educational attainment, ended up as "unskilled workers," even though only 7% of the native Israeli population worked in such jobs. Similarly, over 50% of the émigrés worked as “skilled workers in industry and construction,” again a far higher representation than the 35% of natives in that occupation.[22]

The bottom panel of the table presents summary statistics giving the wage growth of native Israelis observed in selected education/occupation categories. We examine the change in annual earnings (the wage measure that is available in both Israeli censuses). It is telling that the education–occupation cell that was most affected by the Soviet émigrés, university graduates who end up as “skilled workers in industry and construction,” experienced a remarkably large drop in earnings during the period.

We show below that this mismatch between the pre-existing skills of the émigrés (as measured by their educational attainment) and the type of job they actually ended up doing in Israel may have played an important role in generating Friedberg’s (2001) conclusion that the émigrés did not affect the Israeli wage structure. A reexamination of the data that allows for the very high-skill émigrés to influence the earnings of workers employed in occupations that typically employ low-skill workers overturns this result and demonstrates that the émigrés indeed adversely affected the wage of “truly competing” workers, and likely increased the wage of complementary workers.

5.2. Results

Israel is a small country; its land size is roughly the size of El Salvador or New Jersey. As a result, it makes little sense to define a labour market in terms of a region–skill classification. The short commuting distance from one city to another would generate sufficient spillovers across markets to make it difficult to measure the impact of a supply shock by exploiting dispersion at the regional level. Not surprisingly, Friedberg’s (2001) examination of the Soviet supply shock focused on the impact of the émigrés on wages across occupations, so that these “markets” are less likely to be affected by the spillovers resulting from native internal migration.

Although we adapt Friedberg’s choice of an occupation (rather than a local labour market) to define the relevant unit of analysis, our analysis differs in a crucial way. The educational attainment of the émigrés provides an additional measure of skills that is likely to affect productivity and wages—even if the émigrés must initially work in jobs that do not reflect their credentials. Therefore, we define a labour market as a particular occupation–education pairing.

We classify workers into four education groups: less than primary schooling, completed primary schooling, completed secondary schooling, and completed a university education. We also use the occupation classification available in the IPUMS files of the Israeli census, which are the eight broad occupation groups listed in Table 5. We restrict

22 The full name of the occupation is “skilled workers in industry, and construction, and other skilled workers.”

Table 6. The impact of Soviet émigrés on competing workers in Israel

	OLS		IV	
	(1)	(2)	(3)	(4)
A. First stage				
Lagged supply shock	2.686	2.659	—	—
	(0.514)	(0.552)		
Change in native population	—	0.007	—	—
		(0.037)		
B. Change in log annual earnings				
Émigré supply shock	−0.730	−0.740	−0.616	−0.611
	(0.266)	(0.298)	(0.316)	(0.334)
Change in native population	—	0.009	—	−0.004
		(0.083)		(0.071)

Notes: Robust standard errors are reported in parentheses. The unit of observation is an occupation–education cell, and the data consist of eight occupations and four education groups. The "émigré supply shock" variable gives the ratio of the number of Soviet émigrés in the cell to the total size of the cell as of 1995. The "change in native population" variable gives the log difference in the number of native persons in the cell between 1983 and 1995. The first-stage regression in Panel A relates the share of Soviet émigrés in the cell as of 1995 to the share of Soviet immigrants in the cell as of 1983. All regressions have 32 observations and include education-fixed effects.

the empirical analysis to male Israeli natives aged 25–59 years. Finally, the nature of the Israeli census data implies that we can only use the change in log annual earnings as the dependent variable. The occupation of employment is only available for persons who work so that we cannot examine the impact of the supply shock on either the employment or the unemployment rate.

Table 6 summarizes the main regression results using the simpler specification that focuses on identifying the own wage effects. Consider initially the OLS results in the first two columns of the bottom panel. The estimated coefficient is about −0.73 (with a standard error of 0.27). Note that the estimate of the wage elasticity of about −0.7 is unchanged when we add a regressor giving the change in the size of the native population in the particular cell.

The table also reports the wage effect from the IV specification. The instrument is the share of earlier Soviet migrants (who were observed in the 1983 census) employed in a particular occupation–education pairing. The key coefficient in the first stage of the IV is highly significant, so that the new émigrés found employment in roughly the same occupations that employed the compatriots that arrived prior to the collapse of the Soviet Union. The IV estimate of the wage elasticity is −0.62 (0.32), similar to the coefficient obtained in the OLS regression. In short, a regression analysis based on the notion that a labour market consists of an occupation–education cell unambiguously indicates that the Soviet émigrés adversely affected the earnings of comparable workers.

Figure 2 illustrates this insight, showing a large "cloud" of occupation–education cells unaffected by the Soviets. It also shows that the relatively few cells that "welcomed" the émigrés are the source of the negative wage effect. Those cells are composed of the select

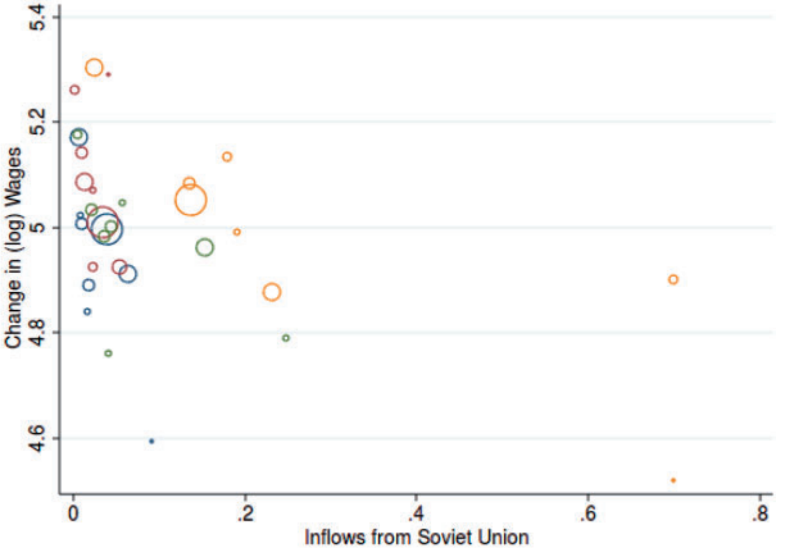


Figure 2. The impact of the Soviet émigrés on native wages

Notes: This figure plots the 1983–1995 change in the annual earnings of native Israeli men aged 25–59 years against the size of the migrant inflow in each cell. Each dot represents an education–occupation cell. Blue, red, green, and yellow dots indicate "less than primary", "Primary completed", "Secondary completed", and "University completed", respectively. The size of the dots represents the size of the cells, measured by the optimal weights used in the regression tables. The figure exploits variation across four educations groups and eight occupation categories.

occupations that attracted high-skill émigrés. Not surprisingly, those occupations were the ones that experienced the lowest wage growth between 1983 and 1995.[23]

As shown in Table 5, the Soviet supply shock was largest for workers who had a university degree, where the inflow of the refugees represented a 26% increase in supply. As with our analysis of the Mariel supply shock, we exploit this fact to identify the potential cross-effects. Specifically, we estimate the cross-effects model given by Equation (12) separately for each of the education groups, where the key regressor gives the supply shock experienced by university graduates in a particular occupation group.[24]

23 There was a currency change in Israel in 1986, one year after the 1985 census that we use to establish the baseline in the pre-shock period. The wage data in 1985 is denominated in Shekels while the post-shock period data are denominated in New Shekels. Therefore, the vertical axis in Figure 2 reflects both the change in the currency as well as real wage growth. Our analysis examines *relative* differences across education–occupation cells. The currency change is absorbed by the constant in the regressions.

24 The regression also includes a variable that controls for the change in the size of the native population in the "own" occupation-schooling group. Because of the small number of observations within each education group, we exclude the "cross" changes in native labour supply. The results reported below would be very similar if, instead, we aggregated the data to two education groups and included both the own- and cross-native supply responses.

Table 7. Own and cross effects of the Soviet émigrés in Israel

	Less than primary	Primary completed	Secondary completed	University completed
Change in log annual earnings	0.350 (0.184)	−0.070 (0.117)	−0.083 (0.121)	−0.739 (0.208)

Notes: Standard errors are reported in parentheses. The unit of observation is an occupation, and there are eight occupations in the analysis. The table reports the coefficient of the "émigré supply shock for high-skill workers," which gives the ratio of the number of Soviet émigrés who completed a university education relative to the number of natives who also completed a university education in 1995 in the particular occupation. The regressions also contain regressors giving the change in the size of the native population for the own education group. The regressions are estimated separately for each occupation group using IV and have eight observations.

Table 7 reports IV coefficients from the cross-effects specification. The results again indicate that the own-effects of the high-skill Soviet émigrés are negative, with a wage elasticity of about −0.7. The table also reveals, however, that there were some positive complementarities between the high-skill émigrés and the least skilled Israeli natives who had not completed their primary education. The earnings of the lowest education group increased after the refugee supply shock, with a cross-elasticity of +0.35 (0.18).

The Israeli evidence is comparable to that obtained in the Mariel context. The entry of the low-skill *Marielitos* increased the wage of natives who were more highly skilled, while the entry of the high-skill Soviet émigrés increased the wage of natives who were least skilled. These cross-effects document the distributional consequences that refugee supply shocks can have on the receiving country's labour market.

5.3. Skill downgrading

In important ways, the evidence summarized in Table 6 is both similar to and very different from the evidence reported in Friedberg (2001), the study that has most carefully examined the consequences of this specific supply shock. As we noted earlier, the Friedberg analysis uses an occupation as the unit of analysis and examines the trend in education-adjusted wages within an occupation. Friedberg also reports both OLS and IV estimates of the own wage effect attributable to the Soviet influx.

In fact, the (OLS) own wage effects that Friedberg estimated are very similar to those reported in Tables 6 and 7, showing a significant reduction in the wage of those occupations most affected by the Soviet émigrés. For example, Friedberg (1990, Table II) reports an own wage elasticity of −0.616 (0.206). Friedberg then argued that the occupational sorting of the Soviet émigrés was endogenous, as income-maximizing émigrés would obviously gravitate toward the highest paying occupations.

To control for this endogeneity, Friedberg used the migrant's occupation in the Soviet Union, *prior to migration*, to instrument for the migrant's eventual occupation in Israel, arguing that the pre-migration occupational choice was unaffected by the Israeli wage structure. The use of this particular instrument, which is available in a small survey of Soviet émigrés used by Friedberg but not in the IPUMS files, leads to an IV estimate of the wage elasticity that is positive and insignificant, leading her to conclude that "the influx of Russians to a given occupation in Israel does not appear to have adversely affected the wage growth of natives working in that occupation" (Friedberg, 2001, p. 1395). It is important to emphasize that the difference between the OLS and IV results in the Friedberg study is puzzling, and remains unexplained.[25] As long as émigrés are income-maximizers, the endogeneity created by the self-sorting of the émigrés into high-paying jobs should bias the OLS coefficient *toward zero*.

Table 6 shows that our IV estimate of the wage elasticity remains negative and significant, and is, in fact, about the same magnitude as the OLS coefficient. There is a crucial difference, however, between the two instruments: Friedberg's instrument is based on the occupation that the Soviet émigrés held in the Soviet Union prior to migration; our instrument is based on the actual occupations that earlier waves of émigrés with similar education pursued in Israel. The difference in the estimates of the wage elasticity implied by the two instruments is, of course, related to the possibility that the skills the émigrés acquired in the Soviet Union may not be completely transferable to the Israeli labour market. In fact, it is easy to document that the pre-existing skills of the émigrés are *not* a very good predictor of the type of job they actually end up doing.

Table 5 reports the occupation distributions of the émigrés and natives who have a university degree. Recall that half of the émigrés are in this particular education category. The table clearly shows a substantial downgrading in the type of job that a high-skill émigré held in Israel. Only 1% of native university graduates, for example, end up as "unskilled workers." Among émigrés, however, the probability of working in such jobs increases 10-fold. Only 4% of native university graduates are "skilled workers in industry and construction." Among the émigrés, however, the probability increases 9-fold, to 36%. In short, the data clearly indicate that pre-existing educational skills, although obviously correlated with the type of job that the émigrés will do in Israel, can generate very large errors in predicting the post-migration allocation of émigrés across occupations. The skill downgrading can help explain not only the different elasticities produced by the two instruments, but also the puzzling result in the Friedberg study where the use of instrumental variables leads to a more positive wage elasticity.

Consider the case where fluency in Hebrew forms a barrier into certain occupations. For instance, suppose there are two occupations in Israel, one where workers need to be fluent in Hebrew (e.g., a TV personality) and one where workers do not (e.g., working at a manufacturing assembly line). As a result of the difference in language requirements, the typical Soviet émigré, even though he might hold a university degree, will inevitably end up in occupations where Hebrew fluency is unimportant.

It is easy to document that type of sorting in the Israeli labour market. The 1983 Israeli census reports whether Hebrew was a first language for each enumerated person.

25 Cohen-Goldner and Paserman (2011) make a related point, arguing that Friedberg used a weak instrument that led to an understatement of the wage impact of the émigré supply shock.

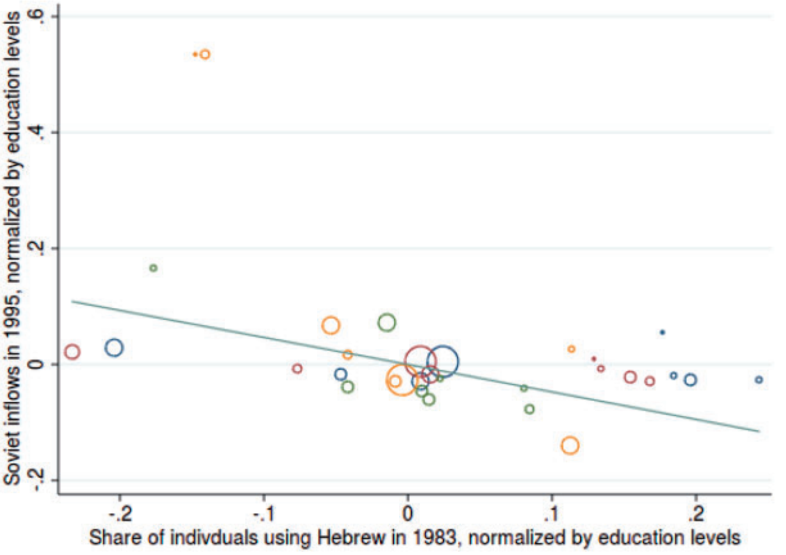


Figure 3. Soviet inflows and language use prior to the supply shock

Notes: The figure plots the inflow of Soviet émigrés against the share of workers who speak Hebrew as their first language in 1983 and who were not from the Soviet Union. We remove education-fixed effects from both variables. Each dot represents an education–occupation cell. Blue, red, green, and yellow dots indicate "less than primary", "Primary completed", "Secondary completed", and "University completed", respectively. The figure exploits variation across four educations groups and eight occupation categories.

We can construct a measure of the share of workers who have Hebrew as a first language (excluding migrants from the Soviet Union) in each occupation–education cell. Figure 3 reveals a strong negative relationship between the size of the Soviet supply shock in each cell and the fraction of people who speak perfect Hebrew. The Soviet supply shock was far smaller in those occupations that require Hebrew fluency.

Suppose that the wage elasticity η is indeed negative. The OLS wage elasticity reported in Table 6 is then essentially measuring the wage change in the occupations where Hebrew fluency is irrelevant (i.e., the occupations actually affected by the supply shock) relative to the wage change in occupations where Hebrew fluency is required (i.e., the occupations less affected by the supply shock). Our IV estimate is picking up exactly the same effect. In contrast, Friedberg's (2001) instrument would predict a sizable shock in Hebrew-intensive occupations—occupations that, in fact, were not much affected by the refugee inflow. Put differently, Friedberg's instrument gives more weight to those occupations that received the fewest immigrants. This misallocation tends to downward bias the estimate of the wage elasticity.

To document the importance of skill downgrading when measuring wage impacts in the Israeli context, we carried out two distinct exercises. First, for each education level we can distribute the émigrés according to the native distribution across occupations in 1983. This prevents the across-occupation skill downgrading and plays a role similar to the instrument used in Friedberg's study. Alternatively, we can keep the distribution of émigrés across occupations as observed in the data, allowing émigrés to enter

Table 8. Sensitivity tests to skill downgrading

Regressor:	(1)	(2)	(3)	(4)
Measure of supply shock:				
Predicted inflow using native occupational distribution within an education group	0.257 (0.683)	−0.228 (0.937)	—	—
Predicted inflow using native educational distribution within an occupation group	—	—	−0.718 (0.251)	−0.729 (0.320)
Change in native population	—	−0.087 (0.104)	—	0.006 (0.100)

Notes: Robust standard errors are reported in parentheses. The unit of observation is an occupation–education cell, and the data consist of eight occupations and four education groups. The dependent variable is the change in log annual earnings between 1983 and 1995 for native Israelis in each cell. The regressor giving the "predicted inflow using native occupational distribution within an education group" gives the émigré supply shock calculated after assigning Soviet émigrés in each education group to occupations according to the occupational distribution of natives within each education group. The regressor giving the "predicted inflow using native educational distribution within an occupation group" gives the émigré supply shock calculated after assigning Soviet émigrés in each occupation to an educational category based on the education distribution of natives in each occupation category. All regressions have 32 observations and include education-fixed effects.

occupations normally performed by lower educated workers, but within each of those occupations we can then assume that the émigré inflow was distributed into the different education groups according to the native distribution of education. This allows for skill downgrading along both the occupation and education dimensions.

Table 8 reports the results from both exercises. When using the distribution of natives across occupations to assign the Soviet émigrés, the wage elasticity is 0.257 (0.683), which resembles the Friedberg (2001, Table II) IV estimate of 0.549 (1.28). If we instead keep the occupation distribution of migrants as observed, but assign them to different education levels using the native distribution within each occupation, the estimate is −0.72 (0.25). This own wage elasticity is virtually identical to the estimates reported in Tables 6 and 7. In sum, the evidence indicates that Soviet émigrés "landed" in occupations that were quite different from the occupations they held in the former Soviet Union, leading them to compete with natives in the occupations they actually entered rather than in the occupations they held before migration.

6. THE ALGERIAN WAR OF INDEPENDENCE

The Algerian War ended with the signing of the Evian Accords on 19 March 1962. France "insisted that the settler citizens stay and become a part of the Algerian nation" (Choi, 2016, p. 2, 4), but the settler citizens and other Algerians had other ideas, and Algerian independence quickly sparked a flow of refugees to France. In the summer of 1962 alone, "750,000 French citizens including 100,000 naturalized Jews and several thousand pro-French Muslim Algerians fled the nationalist takeover." Over time, the number of refugees increased, as the pre-independence population of Algeria included "900,000 white colonials of mixed European descent known otherwise as *pieds noirs*."

As this very brief summary of the events that followed the Evian Accords suggests, the independence of Algeria sparked *two* distinct types of refugee flows into France. The first consisted of the French repatriates, the French nationals who lived in Algeria and returned to France after 1962. The second consisted of Algerian nationals. In fact, the number of Algerian nationals moving to France increased sharply in 1964 "with the arrival of over 75,000 *harkis*, the Muslim Algerian soldiers who had fought on the French side during the War of Independence" (Choi, 2016, p. 6).

Hunt (1992) examined the labour market impact of the first of these refugee flows, consisting of repatriates, on the French labour market. Her analysis suggests that the repatriates had only a small (but adverse) impact on the unemployment rate or wage of French native workers. Hunt's analysis, however, ignored that the supply shock of repatriates may have been particularly large in some skill groups, and much smaller for other groups. In particular, her empirical exercise consisted of essentially correlating the change in some measure of labour market outcomes in a particular city on a measure of the total supply shock of repatriates affecting that city. It is crucial to carefully match the skill level of French natives with the skill level of the refugees to correctly measure the labour market impact of the supply shock. The Hunt study also overlooked the fact that the end of the Algerian war ignited a sizable and concurrent flow of Algerian nationals. It is unlikely that the two refugee flows are uncorrelated. If nothing else, the timing of both flows was motivated by the same political upheaval.

6.1. Summary statistics

We use the 1962 and 1968 French census microdata maintained at IPUMS to determine the size and skill composition of the two refugee shocks originating in Algeria.[26] Each census enables us to count and document the characteristics of persons who were not living in France at the time of the earlier census, which occurs 6 years prior to the enumeration.

Table 9 reports the counts of persons in three key demographic groups: the number of French repatriates (or French nationals who were not living in France at the time of the last census); the number of Algerian refugees (or Algerian nationals who were not living in France at the time of the last census); and the number of French natives (or French nationals who were living in France at the time of the last census).

The 1968 French census enumerated 1.4 million persons of French nationality who were not living in France in 1962. Although the census data maintained at IPUMS do not indicate where these persons resided in 1962, the historical context suggests that a

26 It is important to note that the 1962 census enumeration was carried out before the bulk of refugees arrived in France after the Evian accords were signed in late March 1962. In other words, it is unlikely that the labour supply decisions of French natives enumerated in the 1962 census were affected by the influx of refugees who would soon enter the country.

Table 9. Size and skill composition of the Algerian supply shock, 1968

	French repatriates	Algerian nationals	French natives	% increase in supply: French repatriates	% increase in supply: Algerian nationals
All persons (in 1000s)	1,358.9	162.1	45,732.6	3.0	0.4
Aged 25–59 years	595.0	87.8	18,610.0	3.2	0.5
Male, aged 25–59 years	302.8	77.1	9,079.9	3.3	0.8
% of men aged 25–59 years with education:					
Less than primary	26.1	96.3	37.2	2.3	2.2
Primary completed	36.8	2.5	36.4	3.4	0.1
Secondary completed	25.8	1.0	20.2	4.3	0.0
University completed	11.2	0.2	6.1	6.1	0.0
% of men aged 25–59 years living in:					
Ile de France	21.3	35.2	18.9	3.8	1.6
Lorraine	3.1	7.3	4.4	2.3	1.4
Rhone Alpes	9.7	16.4	8.8	3.7	1.6
Provence–Alpes–Cote d'Azur	19.2	15.6	6.3	10.2	2.1
% of low-skill men aged 25–59 years living in:					
Ile de France	17.3	34.6	13.1	3.1	5.8
Lorraine	2.8	7.5	4.3	1.5	3.9
Rhone Alpes	10.6	16.8	8.2	3.0	4.5
Provence–Alpes–Cote d'Azur	22.9	15.8	6.2	8.6	5.6
Sample size of men, aged 25–59 years	15,139	3,857	453,993		

	Δunemployment rate	Δemployment rate	Sample size
Average in France:			
Less than primary	0.01	0.01	462,579
Primary completed	0.01	0.01	330,784
Secondary completed	0.00	0.00	159,991
University completed	0.00	0.00	49,703
Average in Provence–Alpes–Cote d'Azur:			
Less than primary	0.02	−0.03	32,188
Primary completed	0.01	0.00	20,872
Secondary completed	0.01	0.01	9,886
University completed	0.01	0.01	3,466

Notes: The sample of French repatriates consists of French citizens who were not living in France in 1962; the sample of Algerian nationals consists of Algerians who were not living in France in 1962; and the sample of French natives consists of French citizens who were living in France in 1962. The bottom panel reports statistics calculated in the sample of French native men aged 25–59 years. The (Δunemployment rate) and (Δemployment rate) variables give the average change between the 1962 and 1968 censuses for the particular group.

sizable fraction of this group originated in Algeria.[27] The supply shock of repatriates increased the size of the native French population by about 3%.

27 We experimented with alternative definitions of the repatriate population, such as including French nationals who had moved to France prior to 1962. The analysis reported below opts for the most conservative definition in the sense that it leads to relatively weak labour market impacts of the repatriates. It is important to note that, concurrent with the repatriate and Algerian supply shocks, there was also a lot of churn in the number of foreign-born persons in France due to the entry and exit of over a million guest workers, mainly from Spain, Portugal, and Italy.

In addition, 162,000 Algerian nationals migrated to France between 1962 and 1968, so that this supply shock increased the size of the population by about 0.4%. Note, however, that almost half of the Algerian nationals were men in their prime work years, compared to only 20% of French natives (and 23% of the repatriates).

Our empirical analysis of the impact of the two supply shocks focuses on the group of native French men aged 25–59 years, a group that had 9.1 million persons in 1968. The repatriates increased the size of this population by 3.3%, while the Algerian refugees increased its size by almost 1%.

Table 9 shows the difference in the skill composition of the various groups. Practically all (96%) of the Algerian refugees had less than a primary education, when compared with only 37% of the French natives and 26% of the repatriates. The extreme concentration of the Algerian refugees in the least skilled category implies that this specific supply shock increased the number of low-skill workers in the aggregate French labour market by 2.2%. In contrast, the skill composition of the French repatriates was much more balanced, and comparable with that of French natives, with a slight skew toward a more skilled composition. Among men aged 25–59 years, for example, 26% of French natives and 37% of French repatriates had at least a secondary education.

The two refugee flows also differed in their geographic settlement in France. The geographic sorting of the French repatriates very much resembled that of French natives, except that the region Provence–Alpes–Cote d'Azur received a somewhat larger share. However, a much larger number of the Algerian refugees settled in the Paris and southern regions. For example, 16% of Algerians settled in Rhone-Alpes, a region that hosted only 9% of French natives, and an additional 35% settled in Ile de France, but only 19% of French natives lived in the Paris metropolitan area.

We exploit variation across region–education cells to estimate the labour market impact of the two supply shocks. The available census data enable us to define 88 such cells (22 regions and 4 education groups). The extreme bunching of the Algerian refugees, both in terms of their educational attainment and their geographic distribution, into a small number of cells creates a great deal of dispersion in the size of the supply shock across labour markets. This variation helps to more precisely measure the impact of the Algerian nationals. In contrast, the similarity in the skills and (to some extent) geographic sorting of the French nationals and the French repatriates suggests that there may not be sufficient variation to precisely identify the impact of this supply shock.[28]

The bottom panel of Table 9 hints at the nature of the evidence. The Provence–Alpes–Cote d'Azur region, where a large number of the very low-skilled Algerian nationals eventually settled, witnessed an increase in the unemployment rate of French natives with less than a primary education of 2 percentage points (double the national

28 This fact explains why we do not include location-fixed effects in the regressions reported in Table 10. Including these fixed effects does not change the results for the Algerian inflows, but makes the coefficients for the French repatriates supply shock very unstable.

Table 10. The impact of French repatriates and Algerian nationals on French natives

	OLS		IV	
	(1)	(2)	(3)	(4)
A. First stage: share of refugees				
Lagged repatriate supply shock	1.283	0.067	—	—
	(0.115)	(0.053)		
Lagged Algerian supply shock	−0.063	0.558	—	—
	(0.076)	(0.027)		
B. Change in unemployment rate				
Repatriate supply shock	0.063	0.067	0.089	0.096
	(0.040)	(0.041)	(0.038)	(0.039)
Algerian supply shock	0.270	0.265	0.247	0.240
	(0.067)	(0.069)	(0.067)	(0.069)
Change in native population	—	−0.006	—	−0.009
		(0.011)		(0.011)
C. Change in employment rate				
Repatriate supply shock	−0.075	−0.057	−0.100	−0.083
	(0.066)	(0.069)	(0.077)	(0.081)
Algerian supply shock	−0.647	−0.666	−0.636	−0.651
	(0.206)	(0.211)	(0.222)	(0.226)
Change in native population	—	−0.022	—	−0.019
		(0.029)		(0.027)

Notes: Robust standard errors are reported in parentheses. The unit of observation is a region–education cell, and the data consist of 22 regions and 4 education groups. The "repatriate supply shock" and "Algerian supply shock" variables give the ratio of the number of French repatriates or the number of Algerian nationals in the cell to the number of French natives in the cell as of 1968. The "change in native population" variable gives the log difference in the number of native persons in the cell between 1962 and 1968. The first-stage regression in Panel A relates these shares to the respective shares as of 1962. All regressions have 88 observations and include education-fixed effects.

average), and a decrease in the employment rate of 3 percentage points (in contrast to an *increase* in the employment rate of 1 percentage point in the national labour market).

6.2. Results

Table 10 reports the regression coefficients obtained from alternative specifications of the generic regression model that identifies the own effect of supply shocks in Equations (5) and (7). Because there are two distinct, though concurrent, supply shocks, the regression specification is expanded to include the measure of the supply shock for each of the two types of refugees.

The French census data do not report a worker's earnings. We use two alternative variables to measure the impact of the refugees: the unemployment rate (defined as the fraction of the labour force participants in a particular cell who are unemployed); and the employment rate (defined as the fraction of the population in the cell that is employed). The dependent variables used in the regressions give the change in each of these employment indicators for each region–education cell between 1962 and 1968.

The OLS coefficients are reported in the first two columns of the bottom two panels of the table. The Algerian refugees had a sizable and significant positive effect on the unemployment rate of French natives, as well as a negative and significant effect on their employment rate. In other words, the supply shock of Algerian refugees drove competing French natives out of the labour market, and made the job-finding process more difficult for those natives who stayed in the market. Because the Algerian refugee flow was disproportionately low-skill and clustered in a small number of locations, the regressions are essentially indicating that very low-educated native workers in a small number of French cities were indeed adversely affected by the Algerian supply shock.

It is easy to see the positive impact of the Algerian supply shock on the unemployment rate of comparable French natives in the raw data. The bottom panel of Figure 4 shows the scatter diagram illustrating the relation between the change in the unemployment rate in a particular region–education cell and the size of the corresponding Algerian supply shock. It is obvious that the unemployment rate increased most for French workers in those region–education cells most affected by the entry of the Algerian refugees.

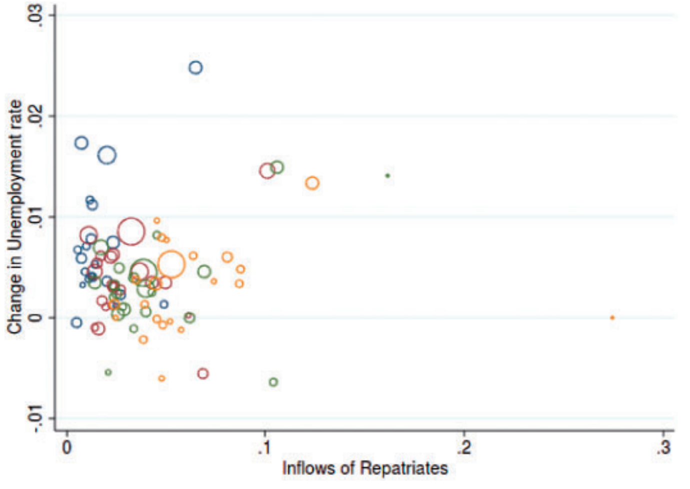
Moreover, the effect of this supply shock on the unemployment rate is numerically large. A 5% increase in the size of the cell due to an influx of Algerian refugees, which is roughly the size of the shock in the most affected region–education cell (natives living in Provence–Alpes–Cote d'Azur who did not complete their primary education) increased the unemployment rate of this group by 1.3 percentage points. The French unemployment rate for prime-age, low-skill men in the mid-1960s was only 2%, so that the supply shock had a substantial impact on French unemployment.

In contrast, the OLS estimate of the impact of the repatriates on the employment and unemployment rates of French natives is typically insignificant, having a significant effect on unemployment only after we account for the endogeneity of the geographic distribution of the repatriates.[29] The fact that the supply shock of French repatriates was "balanced" across education cells makes it difficult to estimate the resulting cross-effects using the regression framework derived earlier. Our identification of cross-effects exploits that refugee supply shocks are often very unbalanced in their skill composition, so that we need only look at how the labour market outcomes of different skill groups relate to the supply shock experienced by the one skill group that was most affected. We can carry out this exercise to estimate the cross-effects resulting from the supply shock of Algerian nationals, but it is not possible to use the methodology to estimate the corresponding cross-effects from the supply shock of French repatriates.

As Table 9 shows, the supply shock of Algerian nationals was *extremely* unbalanced. Almost all the Algerians had not completed their primary education, suggesting that we

29 The flow of repatriates, unlike the flow of Algerian nationals, began prior to the end of the war in 1962. French nationals are recorded to have returned to France as early as 1954. The French census data available at IPUMS does not enable us to determine the origin of these repatriates. However, their skill and location distribution is similar to the larger flow of repatriates that followed the conclusion of the Algerian independence war.

A The French repatriates



B The Algerian nationals

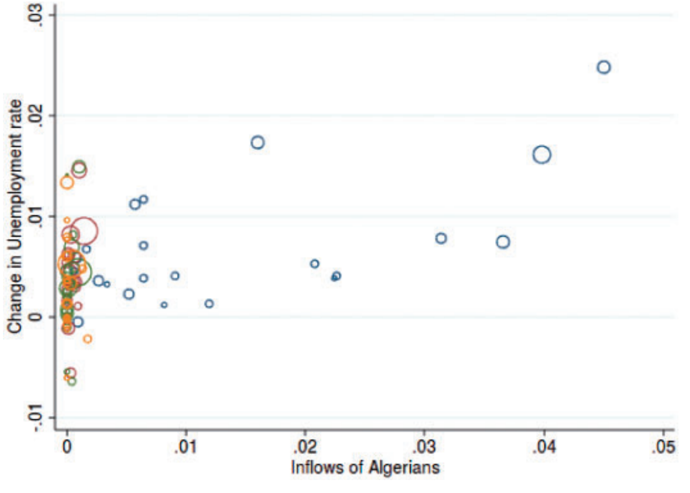


Figure 4. The impact of the supply shocks after the Algerian War on the native unemployment rate

Notes: The figure plots the 1962–1968 change in the unemployment rate of French native men aged 25–59 years against the size of the migrant inflow in each cell. Each dot represents a region–education cell. The top figure shows the impact of the French repatriates and the bottom figure shows the impact of the Algerian nationals. Blue, red, green, and yellow dots indicate "less than primary", "Primary completed", "Secondary completed", and "University completed", respectively. The size of the dots represents the size of the cells. The figure exploits variation across 4 education groups and 22 locations.

can use Equation (12) to determine how the low-skill Algerians affected the employment opportunities of more skilled French natives.

Table 11 reports selected coefficients from the cross-effects regressions. Not surprisingly, the regressions still show an adverse own effect—employment rates are lower and unemployment rates are higher for French natives who do not have a primary

Table 11. Own and cross effects of the supply shock of Algerian nationals

	Less than primary	Primary	Secondary	University
A. Change in unemployment rate	0.319	0.059	0.014	−0.039
	(0.097)	(0.067)	(0.085)	(0.093)
B. Change in employment rate	−0.662	−0.295	−0.349	−0.229
	(0.239)	(0.131)	(0.139)	(0.169)

Notes: Standard errors are reported in parentheses. The unit of observation is a French region, and there are 22 regions in the analysis. The table reports the coefficient of the "Algerian supply shock for low-skill workers" variable, which gives the coefficient of the ratio of the number of Algerian nationals who have less than primary education to the number of natives without a primary education in 1968 in the particular city. The regressions also contain regressors giving the change in the size of the native population for each of the four education groups. The regressions are estimated separately for each education group using IV and have 22 observations.

education. However, we cannot detect any evidence of beneficial cross-effects in this episode. We have been unable to determine the reason for the absence of beneficial complementarities in the Algerian context. One obvious conjecture is that the supply shock of Algerian nationals was quite unique in terms of just how low-skill the refugees were relative to the baseline population.

7. THE BALKAN REFUGEES

For many of the people living in Europe during the 1990s, the names of Srebrenica, Sarajevo, Pristina, and Podgorica are indelibly associated with incidents from the last set of wars fought on European soil. After the collapse of communism, the former republic of Yugoslavia split into five new countries in 1991 and 1992: Slovenia, Croatia, Bosnia-Herzegovina, Serbia, and Macedonia. This breakup, however, was not without conflict. Various episodes of civil and military unrest hit the former communist country between 1990 and 2000. There were many casualties, and many more people lost their homes and sought refuge by moving elsewhere, either internally within the territory of the former Yugoslavia or to other countries in Europe.

It is difficult to estimate precisely how many Balkan refugees moved to European countries. The different timing and location of the various wars generated distinct waves of refugees. For example, the first wars started in northern Yugoslavia, when Slovenia and Croatia in 1991, and then Bosnia in 1992, declared independence. The war in Croatia and Bosnia lasted until 1995 when the Federal Republic of Yugoslavia recognized Croatia and Bosnia-Herzegovina as independent countries. In 1996, ethnic Albanians in Kosovo formed the Kosovo Liberation Army to fight for the creation of an ethnically separate Greater Albania. The War in Kosovo in 1998 and 1999 involved the southern region of former Yugoslavia, and affected large numbers of families. The calculation of the number of refugees created by this seemingly endless series of distinct conflicts is further complicated by the fact that many of the refugees eventually returned to parts of the former Yugoslavia once the wars ended.

The refugees from the former Yugoslavia tended to move to particular countries in Europe. The refugees then settled in particular regions within those countries. We rely on the variation across regions within different European countries to identify the labour market impact of this refugee supply shock.

7.1. Summary statistics

We use census data for seven European destination countries: Austria, Greece, Ireland, Portugal, Romania, Spain, and Switzerland. These countries were chosen based on the following criteria. First, we only use European countries with publicly available census data in the IPUMS archive. Second, we only use countries where we can construct a "before" and "after" snapshot of the relevant labour markets. The timing of the Yugoslav Wars suggests that the countries must have conducted a census around 1990 and another census around 2000.[30] Third, we need to enumerate and determine the skill distribution of the Balkan refugees, as well as measure their impact on the labour market opportunities of comparable natives. In other words, both the "pre" and "post" censuses for each country must report information on country of origin and labour market outcomes, and must report educational attainment in a manner that is comparable across countries.

Although other countries received large numbers of Balkan refugees at the time, the census data publicly available for those other countries do not satisfy our criteria. The French census, for example, does not report the country of origin of foreign-born persons. Similarly, the coding of educational attainment for the UK censuses differs significantly from that used by other countries. Finally, the relevant data are not available for either Germany or Sweden.

Table 12 shows that almost 260,000 persons born in the former Yugoslavia moved to the seven European countries in our sample during the 1990s.[31] This represents a very modest increase of only 0.3% in the aggregate population of those countries. However, as with many refugee supply shocks, the refugees clustered in a relatively small number of places. Almost all of the Balkan refugees settled in two of the countries included in our analysis, Austria and Switzerland, with Austria receiving 76% of the refugees and Switzerland receiving 17%. Within those two countries, the refugees were further clustered in specific regions, providing sufficient variation for the identification of their labour market impact. For example, Vienna received 34% of the refugees in Austria, but only 19% of the native population resided in that city.

We examine the labour market outcomes of native men aged 25–59 years in the various receiving countries. Table 12 reports that the Balkan refugees were disproportionately of intermediate skills: 44% of prime-age men in the receiving countries had completed a secondary education, when compared with 67% of the Balkan refugees.

30 We use the 1991 and 2001 censuses in our analysis of Austria, Greece, Portugal, and Spain; the 1991 and 2002 censuses for Ireland; the 1992 and 2002 censuses for Romania; and the 1990 and 2000 censuses for Switzerland.

31 We calculate this number by comparing the stock of migrants from Yugoslavia *circa* 1990 to the stock of migrants from the former Yugoslavia *circa* 2000.

Table 12. Size and skill composition of refugees from the former Yugoslavia, 2000

	Refugees	European natives	% increase in supply
All persons (in 1000s)	258.6	94,052.4	0.3
Aged 25–59 years	170.2	49,435.0	0.3
Male, aged 25–59 years	65.1	24,611.1	0.3
Education distribution, men aged 25–59 years			
Primary completed or less	25.1	44.4	0.1
Secondary completed	67.0	44.1	0.4
University completed	7.9	11.5	0.2
% of men aged 25–59 years living in:			
Austria	75.8	7.7	2.6
Greece	2.9	10.0	0.1
Ireland	0.9	3.6	0.1
Portugal	0.2	9.9	0.0
Romania	0.9	20.5	0.0
Spain	2.0	41.5	0.0
Switzerland	17.4	6.8	0.7
% of men aged 25–59 years in Austria living in:			
Burgenland, AUT	2.3	3.6	1.7
Niederosterreich, AUT	12.4	19.7	1.6
Wien, AUT	33.5	18.9	4.6
Karnten, AUT	6.8	6.9	2.6
Steiermark, AUT	11.8	14.9	2.1
Oberosterreich, AUT	16.1	16.9	2.5
Salzburg, AUT	7.4	6.3	3.1
Tirol	6.3	8.5	1.9
Vorarlberg, AUT	3.3	4.4	2.0
Sample size: men aged 25–59 years	5,871	1,744,826	
Average in Austria	Δunemployment rate	Δemployment rate	Sample size
Primary completed or less	0.03	−0.01	76,838
Secondary completed	0.01	−0.01	268,539
University completed	0.01	−0.00	30,759
Average in Vienna:			
Primary completed or less	0.02	−0.01	17,787
Secondary completed	0.03	−0.01	48,006
University completed	0.00	−0.00	10,100

Notes: The sample of refugees from the Balkan Wars consists of persons born in the former Yugoslavia, but who migrated to one of the seven European countries between 1990 and 2000. The sample of European natives consists of persons not born in the former Yugoslavia. The bottom panel reports statistics calculated in the sample of European native men aged 25–59 years. The (Δunemployment rate) and (Δemployment rate) variables give the average change between the 1990 and 2000 censuses for the particular group.

Our analysis of the Balkan episode defines a labour market as a particular region–education cell, where the region index now identifies a particular area within a particular country of destination. The publicly available census data allow us to identify 65 such geographic units (across seven different countries) in three different education groups, so that our analysis exploits variation across 195 cells.

Our analysis complements Angrist and Kugler's (2003) examination of this specific refugee supply shock. There are, however, several key differences. Angrist and Kugler used annual data from the Labor Force Survey maintained by the Eurostat, which

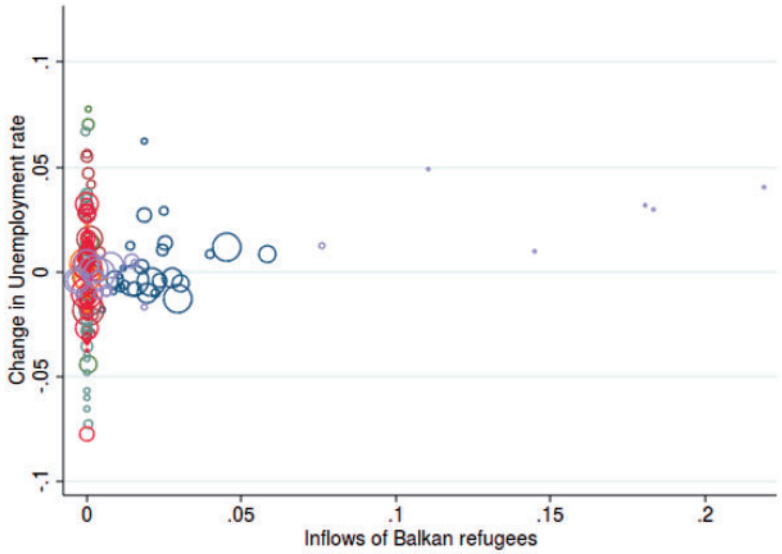


Figure 5. The impact of the Balkan refugees on the native unemployment rate

Notes: The figure plots the change in the unemployment rate of native men in seven European countries against the size of the migrant inflow in each cell between the census year closest to 1990 and the census year closest to 2000 for each of the countries used. Each dot represents a country of destination region within the country–education cell. Country-fixed effects are removed in the graph. Different colors represent the different countries used: Austria, Greece, Ireland, Portugal, Romania, Spain, and Switzerland. The size of the dots represents the size of the cells. The figure exploits variation across 3 education groups and 65 locations.

report labour market outcomes and migrant flows in all European countries. These data do not provide information on the educational attainment of the refugees or on the specific local labour markets (in a receiving country) that were affected by the shocks. Exploiting annual variation on labour market outcomes and on the size of supply shocks enables Angrist and Kugler to more precisely measure the labour market impact around the years of the Bosnian and Kosovo Wars. But the lack of information on educational attainment and the within-country regions most affected by the refugees implies that they must rely on aggregate differences across countries to identify the labour market impact.

Figure 5 shows that it is mainly some regions in Austria and Switzerland that witnessed significant supply shocks of refugees from the former Yugoslavia. In fact, there are some regions in Austria where the influx of Balkan refugees increased the size of the workforce by about 5% in some education groups, while in (some very small cells in) Switzerland the size of the supply shock sometimes neared 20%.

7.2. Results

Table 13 reports coefficients obtained from the "own effects" regression models in Equations (5) and (7). The pooled censuses from the seven European countries do not contain any information on a worker's earnings, so that our dependent variables are: the

Table 13. The impact of the Balkan supply shock on competing workers

	OLS		IV	
	(1)	(2)	(3)	(4)
A. First stage: share of refugees				
Lagged Balkan supply shock	0.152	0.144	—	—
	(0.036)	(0.036)		
Change in native population	—	0.004	—	—
		(0.005)		
B. Change in unemployment rate				
Balkan supply shock	0.209	0.209	0.456	0.487
	(0.078)	(0.103)	(0.311)	(0.376)
Change in native population	—	−0.000	—	−0.003
		(0.016)		(0.017)
C. Change in employment rate				
Balkan supply shock	−0.001	−0.000	−0.084	−0.091
	(0.020)	(0.022)	(0.109)	(0.116)
Change in native population	—	−0.000	—	0.001
		(0.002)		(0.002)

Notes: Robust standard errors are reported in parentheses. The table exploits variation across 3 education groups and 65 regions located in seven different countries. The unit of observation is a country–region–city cell. The "Balkan supply shock" variable gives the ratio of the number of Balkan refugees in the cell to the number of natives in the cell as of 2000. The "change in native population" variable gives the log difference in the number of native persons in the cell between 1990 and 2000. The first-stage regression in Panel A relates the share of Balkan refugees in the cell as of 2000 to the share of Yugoslavian migrants in the cell as of 1990. All regressions have 195 observations and include both education-fixed effects and country of destination-fixed effects.

change in the unemployment rate (defined as the fraction of the labour force participants in a region–education cell who are unemployed); and the change in the native employment rate (defined as the fraction of the population in the cell that is employed). Note that all regression specifications reported in the table include country-of-destination fixed effects, so that the impact of the supply shock is being identified from the variation across region–education cells *within* a particular country.[32]

The OLS coefficients of the "own" labour market impact are reported in the first two columns of the bottom two panels of the table.[33] The Balkan refugees had a positive and significant effect on the native unemployment rate and a negative (but insignificant) effect on the native employment rate. The point estimate suggests that a 5% refugee supply shock increases the unemployment rate of competing natives by about 1 percentage point. Because this particular supply shock was largest in Austria and Switzerland, where the unemployment rates in 2000 were 5.5% and 2.0%, respectively, the Balkan refugees had a sizable effect on native labour market opportunities in those two countries. Figure 5 illustrates the raw data that generate this positive correlation between the supply shock and the change in the unemployment rate.

32 It is important to control for country-fixed effects because different countries were trying to converge in macroeconomic conditions prior to entering the common currency union.

33 The regression results are similar if we estimate the regressions using only regional–education variation in the two countries (Austria and Switzerland) that received most of the Balkan refugees.

Table 14. Cross-effects of the Balkan supply shock

	Primary education or less	Secondary	University
A. Change in unemployment rate	−1.412	1.056	−0.008
	(0.951)	(0.754)	(0.411)
B. Change in employment rate	0.330	−0.103	−0.038
	(0.169)	(0.058)	(0.051)

Notes: Standard errors are reported in parentheses. The unit of observation is a country–region cell, and there are 65 regions located in seven different countries. The table reports the coefficient of the "Balkan supply shock for middle-skill workers," which gives the ratio of the number of Balkan refugees who had completed their secondary education to the number of natives who also completed a secondary education in 2000 in the particular country–region cell. The regressions also contain regressors giving the change in the size of the native population for each of the three education groups. The regressions are estimated separately for each education group using IV and have 65 observations.

Of course, the Balkan refugees may be endogenously choosing which particular labour market to move to (in terms of choosing both a particular country of destination and a particular region within that country), obviously preferring to settle in locations that offer the best employment opportunities. To address the endogeneity concern, we again use the migration network instrument. The use of IV to control for the geographic sorting of the Balkan refugees does not fundamentally alter our results, but the IV coefficients of the labour market impact are imprecisely estimated. It is worth noting that our estimates of the "own" employment effects seem weaker than those estimated by Angrist and Kugler (2003) using cross-country variation. Angrist and Kugler (2003, p. F318, F322) report that 100 more migrants lead to 35 or 83 fewer native jobs (depending on whether the impact is estimated using OLS or IV, respectively).[34]

As we noted earlier, the Balkan refugees were disproportionately located in the middle of the education distribution. We use the cross-effects regression model derived in Equation (12) to separately examine how this particular supply shock affected the employment outcomes of native persons at the two extremes of the skill distribution.

Table 14 summarizes the evidence on cross-effects. The expanded regression model from Equation (12) still yields the finding of adverse own effects. In other words, there was an increase in the unemployment rate of native workers with a secondary education *and* a decrease in their employment rate. Equally important, the disproportionately large number of intermediate-skill refugees lowered the unemployment rate and increased the employment rate of natives who had at most a primary education. Although the beneficial cross-effects on low-skill natives are often not statistically significant, the point estimates consistently suggest that their employment outcomes improved because of the entry of so many refugees in the next higher rung of the skill distribution.

34 The Angrist–Kugler point estimates assume a baseline immigration level of 5%. In contrast, using a setting similar to ours, Glitz (2012) examines variation across German regions that were differentially affected by the inflow of Soviet migrants after the collapse of the Soviet Union, and obtains employment effects that are quantitatively similar to those reported in Table 13.

8. SUMMARY

The recent entry of hundreds of thousands of refugees into many European countries has already generated a great deal of political controversy and raised many questions that require a fuller understanding of the determinants and consequences of refugee supply shocks. This paper revisited some of the historical refugee flows to document the labour market consequences of refugee-induced increases in labour supply.

Specifically, our analysis reexamines the evidence surrounding four episodes: (1) The influx of the *Marielitos* into Miami in 1980; (2) the influx of Jewish émigrés into Israel after the collapse of the Soviet Union in the early 1990s; (3) the influx of French repatriates and Algerian nationals into France at the end of the Algerian War of Independence in 1962; and (4) the influx of refugees from the former Yugoslavia into some European countries during the long series of Balkan wars between 1991 and 2001.

Although the labour market consequences of each of these shocks have been separately examined in prior studies, our study differs from the prior literature in three key ways. First, we use a common empirical approach, based on the implications of factor demand theory, to document the labour market impact of each of the supply shocks. Despite the obvious differences in the historical, economic, and political forces that motivated the various refugee flows, the use of the same empirical framework to study each of the episodes reveals a common thread in the evidence: Exogenous supply shocks adversely affect the labour market opportunities of competing natives in the destination countries.

This result implies that sometimes a refugee supply shock will harm low-skill workers in some regions of the receiving country (as was the case with the *Marielitos* in Miami or the Algerian nationals in France). In other cases, however, it is the high-skill workforce in the receiving country that bears the brunt of the impact (as was the case with high-skill Israelis competing with large numbers of high-skill Soviet émigrés).

Second, the very different skill distributions of natives and refugees in some of these episodes suggest that these natural experiments can be further exploited to identify the impact of the supply shocks on potentially complementary native groups. For example, the low-skill *Marielitos* may have increased the wage of high-skill Miamians, while the high-skill Soviet émigrés may have benefited low-skill Israelis. These complementarities should be an important part of any assessment of how refugee supply shocks alter the employment opportunities of native workers. Our empirical analysis documents that, in many cases, these beneficial effects do indeed exist and are numerically important.

Finally, rather than rely on proprietary or confidential data, we use the publicly available censuses maintained at IPUMS. Our use of easily accessible data implies that our results are fully reproducible. The reproducibility of the evidence is essential because the recent refugee supply shocks in Europe have already sparked extremely contentious policy debates in many receiving countries.

Our study of the four historical episodes of refugee supply shocks teaches an important lesson. Although the episodes differ in countless ways, a universal theme connects the evidence. The humanitarian principles that encourage receiving countries to accept as many migrants as possible have important distributional consequences, as predicted by the canonical model of supply and demand in the labour market.

Discussion

Andreas Madestam

Stockholm University

Background

George Borjas and Joan Monras (2017) provide a timely contribution to a research theme that strikes two cords currently attracting a great deal of attention in the policy debate: first, and foremost, by analyzing the impact of migration; and second, by adding to the discussion on the determinants of growing political polarization. Some argue that migration and inequality in the labour market has been major contributors to the trend toward increased political polarization in the United States (McCarty *et al.*, 2006; Autor *et al.*, 2016). The connection between labour market inequality and migration has also spurred a heated debate about the effects of migration flows on the receiving country at large, on the differential impact of migration on natives versus migrants and skilled versus unskilled, and whether there are complementarities across the skill and occupation gradient (see e.g., Borjas, 2003; Card, 2009; Ottaviano and Peri, 2012).

What this paper does

Against this background, Borjas and Monras investigate the effect of large inflows of refugees on receiving countries' labour markets by examining changes in wages and employment status. They first derive key predictions using a factor demand model showing that an expansion of a certain type of labour leads to a decrease in the wage of native labour of the same type. The predictions are tested using publically available IPUMS data, a generic empirical framework, and an instrumental variable strategy to assess the impact of four historical events—a Meta study—of the impact of migration on wages. The authors revisit the Mariel boatlift to Miami in 1980, the migration of Soviet Jews to Israel in 1989–1995, the influx to France in 1962–1968 following the Algerian War of Independence, and the Balkan wars and the subsequent mass flight to the rest of Europe in 1990–2000.

Specifically, the paper tests the prediction that wages and employment decrease in the skill and region cell where a larger proportion of migrants enter. Exploiting the unexpected shock of entry, together with the historical location or occupation of migrants coming from the same region as the refugees as an instrument, it finds that (i) wages and employment (unemployment) decreases (increases) following the refugee shock; and that (ii) cross-elasticities are important, showing positive effects on the skill and region cells that are indirectly affected by the refugee flows.

Comments

The paper's main contribution and selling point is that it refrains from, in its own words, a "pick and choose" empirical approach and instead relies on a generic framework using publically available data to offer accessible input to the ongoing discussion. I agree with the authors that they present a transparent overview, with the important caveat that the summary represents one side of debate. Much has been written as to why researchers arrive at such different conclusions on the impact of immigration. Dustmann *et al.* (2016) present a commendable review highlighting two important sources of the contradictory findings: the empirical specification used and the assumptions made about the labour supply elasticities across different skill groups and the downgrading of immigrants. I broadly agree with their conclusions, and rather than repeating Dustman *et al.*'s arguments I will take Borjas and Monras' empirical specification as given and focus more on issues related to empirical identification, sample selection, and inference.

To overcome the standard concerns in the migration literature of non-random migration, reverse causality, and omitted-variable bias, Borjas and Monras instrument for actual migration. They do so using the pre-shock location of historical migrants coming from the same country as the refugees currently under study. While much time is spent deriving the theory-based specification, neither the identifying assumption nor the exclusion restrictions are thoroughly spelled out in the paper. This is unfortunate as it becomes more difficult to assess the plausibility of the empirical exercise. Following this, the paper would have benefitted from an exogeneity check, showing that the instrument is "as good as randomly" assigned. Specifically, running the instrument (i.e., the pre-shock location) on a set of control variables, determined before the refugee shock, that we think are important drivers of post-shock wages and employment status. This test could include pre-shock wages, employment, education, income, population, etc. If the instrument is correlated with these outcomes, we run the risk of incorrectly attributing post-shock changes to the arrival of the refugees, as opposed to important secular trends. A complimentary test that examines the plausibility of the instrument would be to perform placebo regressions using other time periods (earlier censuses) when we expect no changes to occur. Again, if these placebos turn out to be significant, it signals that other things may be driving the labour market outcomes of interest.

The crux of the theoretical argument relies on heterogeneous treatment effects across the skill and occupation gradient. That is, we expect stronger effects in the skill/occupation and region cell where a larger proportion of migrants enter. For this to be empirically credible, however, we have to assume strict exogeneity. That is, that the partitioned categories (education, occupation, and labour market participation) do not change themselves as a result of the shock. If natives, for example, decide to acquire education because of the arrival of refugees, the endogenous response makes it difficult to interpret the estimated coefficients. The ideal solution would be to use pre-shock determinants (e.g., parental education). However, given the nature of the exercise (using publically available aggregate data), a second-best solution would be to present regressions with and without controls (no fixed effects at all and then including possibly endogenous responses). In addition, it would be instructive to put the potentially endogenous regressors on the left-hand side as outcomes rather than controls to check whether they are affected by the refugee shock (similar in spirit to Pei *et al.*, 2017).

The paper sets out to measure economy-wide effects but excludes 50% of the population, namely women. There is no theoretical argument provided as to why women should be dropped from the analysis. The empirical motivation put forward is that men have less volatile wage and employment trends and so the paper avoids possible compositional biases by excluding women from the analysis. However, I find this argument at odds with the one underlying the use of an instrument in the first place: that it solves the problem of existing pre-trends. More importantly, the exclusion might mask important interaction effects (between gender and employment) that could introduce a selection bias in the derived estimates. In addition, the overall welfare implications of the refugee shocks are less clear. A more transparent analysis would include women and then condition on gender to see how it affects the outcomes of interest. Related to the sample restriction, it is not clear whether the instrument is based on the previous location of the same group that it focuses on in the analysis (men ages 25–59 years) or the entire historical migrant community. In the latter, it (again) speaks in favor of including all people of working age in the study.

The data are collapsed at the relevant region–occupation/skill cell in all of the analysis. While this can be desirable from an inference standpoint it becomes problematic when there are few underlying cells. For example, the cross-elasticity results for Israel are estimated based on eight cells. For such a small number of cells there are no well-developed asymptotics. If the underlying population was very large, it would resemble stratified sampling (i.e., the error term should be normally distributed) in which case the exercise makes sense (see e.g., Donald and Lang, 2007). However, the problem is aggravated if the underlying population is small which may be the case for some of the cells exploited in the analysis, suggesting a bit of caution when interpreting the confidence intervals in these instances. In the same vein, the data are weighted using the underlying post-shock population in a given cell. Following Solon *et al.* (2015) it would be instructive to present the key results with and without weights, especially if there is a concern that the population changes endogenously with the refugee shocks.

Let me close the discussion with some comments regarding motivation and interpretations. As the paper correctly states, a strong selling point is the generic framework and the use of public data across four historical episodes that allow for an accessible analysis as well as future replications. Considering this, the paper would have benefited from a brief passage motivating the selection of the historical events included in the study. How representative are they of refugee supply shocks specifically and migration flows more generally? Similarly, what are the possible shortcomings of a "one-size fits all" approach? For example, how important are local labour market institutions, the degree of labour market competition, and ensuing state responses in explaining the outcomes?

I have two final comments regarding interpretation. While the paper provides a theoretical motivation for the main outcomes it is left to the reader to understand the intuition behind the cross-elasticities. More specifically, I would like to know what type of complementarities we should have in mind when different skill or education groups affect one another. For example, is the intuition different for upstream (low-skilled migrants affecting high-skilled natives) versus downstream complementarities (high-skilled migrants affecting low-skilled natives)? Finally, in the version of the paper discussed at the conference, there was some ambiguity as to whether the derived estimates were to be viewed as transitionary or long-run equilibrium outcomes. In the published version, the authors make clear that they derive the short-run impact of refugee labour supply shocks. In light of this, and the fact that all four episodes occurred between 15 to over 50 years ago, it would be highly policy relevant (and empirically feasible) to investigate the medium-to-long-term labour market effects. Does the adverse impact on competing natives persist? This should be the next important question for the authors to pursue.

Panel discussion

Following one of the points raised by Andreas Madestam during his discussion, Timothy Hatton suggested that the effect of refugee supply shocks on wages may dissipate over time due to long-run equilibrating mechanisms such as internal migration. Joan Monras agreed with this argument but clarified that this particular paper is focused on short-run effects. He also suggested that it may be interesting to examine other types of outcomes when analyzing longer-time horizons but this will be left for future research.

Fabian Waldinger noted that the implications of the results in the paper may be broader than what the title suggests, that is, the findings may not be specific to refugee supply shocks. Andrea Ichino and Refet Gürkaynak highlighted that the results are exactly what one would expect, while Kevin O'Rourke asked whether the authors knew if refugees are in the labour market after obtaining asylum. Richard Portes gave the example of the UK where migrants tend to fill gaps in the labour market and that this may have allowed the country to sustain a higher pressure of demand in recent years. Giulio Zanella asked how possible changes in the composition of the cells are factored into the authors' computations, that is, whether a change in wages can simply reflect a change in the composition of the cells.

George de Menil stressed this is a very important paper in terms of policy implications and emphasized the distributional effects of this particular form of increase in labour supply. Following the latter comment, Andrea Ichino asked in which sense the distributional effects are good or bad, given that these should ultimately depend on the welfare function.

Replying to comments from the audience and discussants, Joan Monras first clarified that they focused on males in order to select samples that would allow for measuring the unit price of labour in the best possible way. Nevertheless, he stated that they can potentially extend the analysis to include females as well. Regarding the general equilibrium effects of migration, Joan Monras argued that this is one of the greatest challenges in the migration literature and that more research is needed in this area.

REFERENCES

Abadie, A., A. Diamond and J. Hainmueller (2010). 'Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program', *Journal of the American Statistical Association*, 105, 493–505.

Alexe, D. (2015). 'Angela Merkel: no legal limit to refugee numbers', *World Affairs Journal*, September 11.

Altonji, J. and D. Card (1991). 'The effects of immigration on the labor market outcomes of less-skilled natives', in J. Abowd and R. Freeman (eds), *Immigration, Trade, and the Labor Market*, University of Chicago Press, Chicago, pp. 201–34.

Angrist, J. and A. Kugler (2003). 'Protective or counter-productive? European labor market institutions and the effect of immigrants on EU natives', *Economic Journal*, 113, F302–31.

Autor, D., D. Dorn, G. Hanson and K. Majlesi (2016). 'Importing political polarization? The electoral consequences of rising trade exposure', Working Paper.

Borjas, G. (1987). 'Self-selection and the earnings of immigrants', *American Economic Review*, 77, 531–53.

——. (2003). 'The labor demand curve is downward sloping: reexamining the impact of immigration on the labor market', *Quarterly Journal of Economics*, 118, 1335–74.

——. (2014). *Immigration Economics*, Harvard University Press, Cambridge, MA.

——. (2016). The wage impact of the Marielitos: additional evidence', No. w21850, National Bureau of Economic Research.

——. (2017). 'The wage impact of the Marielitos: a reappraisal', *Industrial and Labor Relations Review*, forthcoming.

Borjas, G., R. Freeman and L. Katz (1997). 'How much do immigration and trade affect labor market outcomes?', *Brookings Papers on Economic Activity*, 1, 1–67.

Borjas, G. and J. Monras (2016). The labor market consequences of refugee supply shocks, No. w22656, National Bureau of Economic Research.

Buwalda, P. (1997). *They Did Not Dwell Alone: Jewish Emigration from the Soviet Union 1967–1990*, Woodrow Wilson Center Press, Washington, DC.

Card, D. (1990). 'The impact of the Mariel boatlift on the Miami labor market', *Industrial and Labor Relations Review*, 43, 245–57.

——. (2001). 'Immigrant inflows, native outflows and the local labor market impacts of higher immigration', *Journal of Labor Economics*, 19, 22–64.

——. (2009). 'Immigration and inequality', *American Economic Review*, 99, 1–21.

Card, D. and G. Peri (2017). '*Immigration Economics* by George J, Borjas: a review essay', *Journal of Economic Literature*.

Choi, S.E. (2016). *Decolonization and the French of Algeria: Bringing the Settler Colony Home*, Palgrave Macmillan, New York.

Cohen-Goldner, S. and M. Paserman (2011). 'The dynamic impact of immigration on natives' labor market outcomes: evidence from Israel', *European Economic Review*, 55, 1027–104.

Donald, S.G. and K. Lang (2007). 'Inference with difference-in-differences and other panel data', *Review of Economics and Statistics*, 89, 221–33.
Dustmann, C., T. Frattini and I. Preston (2013) 'The effect of immigration along the distribution of wages', *Review of Economic Studies*, 80, 145–73.
Dustmann, C., U. Schönberg and J. Stuhler (2016). 'The impact of immigration: why do studies reach such different results?', *Journal of Economic Perspectives*, 30, 31–56.
Dustmann, C., U. Schonberg and J. Stuhler (2017). 'Labor supply shocks and the dynamics of local wages and employment', *Quarterly Journal of Economics*, forthcoming.
Fernandez-Huertas, J. (2011) 'New evidence on emigrant selection', *Review of Economics and Statistics*, 93, 72–96.
Friedberg, R. (2001). 'The impact of mass migration on the Israeli labor market', *Quarterly Journal of Economics*, 116, 1373–408.
Glitz, A. (2012). 'The labor market impact of immigration: a quasi-experiment exploiting immigrant location rules in Germany', *Journal of Labor Economics*, 2012, 30, 175–213.
Grossman, J. (1982). 'The substitutability of natives and immigrants in production', *Review of Economics and Statistics*, 54, 596–603.
Hunt, J. (1992). 'The impact of the 1962 repatriates from Algeria on the French labor market', *Industrial and Labor Relations Review*, 45, 556–72.
Jaeger, D., J. Ruist and J. Stuhler (2016). Are lagged immigrant shares valid as instruments?, working paper.
McCarty, N., H. Rosenthal and K.T. Poole (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. MIT Press.
Monras, J. (2014). Online appendix to 'Immigration and wage dynamics: evidence from the Mexican Peso crisis, Sciences Po.
——. (2015a). Immigration and wage dynamics: evidence from the Mexican Peso crisis, No. 8924, IZA Discussion Paper.
——. (2015b). Economic shocks and internal migration, No. 8840, IZA Discussion Paper.
Ottaviano, G. and G. Peri (2007). The effects of immigration on U.S. wages and rents: a general equilibrium approach, No. 13/07, Centre for Research and Analysis of Migration.
——. (2011). 'Rethinking *the effects of immigration on wages', Journal of the European Economic Association*, 10, 152–97.
——. (2012). 'Rethinking the effect of immigration on wages', *Journal of the European Economic Association*, 10, 152–97.
Peri, G. and V. Yasenov (2015). The labor market effects of a refugee wave: applying the synthetic control method to the Mariel boatlift, No. w21801, National Bureau of Economic Research.
Pei, Z., S. Pischke and H. Schwandt (2017). 'Poorly measured confounders are more useful on the left than on the right', NBER Working Paper 23232.
Solon, G., S. Haider and J.M. Wooldridge (2015). 'What are we weighting for?', *Journal of Human Resources*, 50, 301–16.