Opening the Floodgates: Immigration and Structural Change[*]

Bernt Bratsberg[†]   Andreas Moxnes[‡]   Oddbjørn Raaum[§]

Karen Helene Ulltveit-Moe[¶]

May 2017

PRELIMINARY AND INCOMPLETE

Abstract

This paper investigates the impact of a large shock to labor supply on industry growth and structural change. The EU enlargement of 2004 and 2007 lead to an unprecedented migration wave to Norway. The country received the largest number of migrants relative to country size, compared to all other developed countries, over the ensuing decade. We develop a simple factor-proportions theory and sufficient statistic approach that can be used to identify the aggregate impact of a labor supply shock across occupations on industry growth. Using detailed data on industry performance, immigration by occupation and occupational characteristics, we introduce a new instrument that exploits the fact that language barriers in the Norwegian labor market are significant for foreign workers and that they vary across occupations and source countries. Our results point to migration leading to large adjustments in industry size, and in particular to sectors of the economy that are intensive in the use of immigrant occupations.

1   Introduction

What is the impact of a large immigration induced labor supply shock on the industry mix of the economy? In many countries, immigration is the major factor driving changes in labor supply. Studies of the immigration impact on receiving countries typically focus on the wage structure (Dustmann et al., 2016), although recent contributions also characterize employment adjustments. Still, there is relatively scant evidence on how industries expand or contract in response to immigration shocks.[1] This paper attempts to reduce this gap in the literature.

Our starting point is the 2004 and 2007 expansions of the EU, which lifted migration restrictions for roughly 100 million individuals from the EU accession countries.[2] Over the ensuing decade, Norway, as a member of the European Economic Area (EEA) and therefore part of the EU single market, was the country that received the largest inflow of migrants, relative to country size, compared to all other developed countries.[3] Norway became a popular migrant destination because real wages there are among the highest, and unemployment among the lowest, in Europe.[4] When the floodgates opened in 2004, the immigrant share of employment was 7 percent. Nine years later, by the end of our period of analysis, the immigrant share was 10 percentage points higher. In addition to the sheer magnitude of the immigration shock, the Norwegian case is particularly useful to study since the policy change was completely exogenous: as a member of EEA but not the EU, Norway is bound to accept immigrants from all EU countries (and adopt most EU legislation), but is not represented in neither the European Parliament nor the European Commission. The policy change was therefore rapid, comprehensive and externally imposed, providing a unique setting to study the impact of immigration on structural change because it sidesteps the endogenous nature of policy changes that typically presents challenges to empirical studies.

If immigrants were more or less identical to the native population, then there would be no reason to expect that the industry mix would change. This was far from the case however. As we document below in Section 5, immigrants were highly concentrated in certain types of occupations. Our hypothesis is therefore that a supply shock to an occupation lowers relative wages there, in line with much of the "partial elasticity" evidence in the wage

---

[1]We review the previous literature below.

[2]The EU accession countries are: Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, Slovenia (2004) and Bulgaria and Romania (2007).

[3]The EEA agreement extends the EU single market to the three EFTA countries Iceland, Norway and Liechtenstein. According to OECD data, migration to Norway was higher than migration to all other OECD countries except Luxembourg between 2003 and 2013 (measured as the change in the foreign-born population relative to total population (OECD, 2017).

[4]In addition, there were few transitional restrictions on immigration from the accession countries, in contrast to the practice among most EU countries (Dølvik and Eldring, 2008).

structure literature. This in turn will benefit industries which are intensive in the use of that occupation, causing them to grow faster than other industries.

We formalize this idea in the simplest possible economic framework. The labor market consists of O occupations and the cost of switching occupation is prohibitive in the short run. Within a narrowly defined occupation, natives and immigrants are perfect substitutes. There are I industries which use occupation-specific labor with different intensities. The supply side of the labor market is governed by a Roy-type model, similar to Lagakos and Waugh (2013). Workers can choose which industry to work in, and because of idiosyncratic worker-industry productivity shocks, individuals with the same occupation may choose to work in different industries. This structure gives rise to a simple general equilibrium relationship between employment growth of industry i and the weighted average change in labor supply of occupation o, where the weights are the initial factor intensities for that sector. Simply put, after sorting out all the general equilibrium effects in the model, a sufficient statistic for the change in industry size is the weighted average change in labor supply, across all occupations o. An important theoretical finding is that we can use the sufficient statistic to estimate the aggregate (and not just relative) effect of a labor supply shock on industry size. Hence, we make important headway relative to traditional reduced-form analyses, where one can only hope to identify relative (and not general equilibrium) effects. Our approach mirrors Donaldson and Hornbeck (2016) who use the structure of a trade model to identify aggregate effects of the U.S. railroad network on land values.

We test our hypothesis by using detailed Norwegian data on industries, occupations and the immigrant share within occupations. A major identification challenge is that occupation o workers may migrate because occupation o-intensive industries are booming. We therefore propose a new Bartik-style instrument and methodology to overcome the identification problem (Bartik, 1991). Our instrument is based on the premise that the cost of migrating to a destination country and working in occupation o depends on the language intensity of that occupation. For example, working as a journalist, a language intensive occupation, requires extensive local language training and practice. On the other hand, working as a carpenter, a relatively low language intensive occupation, requires only rudimentary Norwegian skills.[5] The migration cost, and language training time, is therefore lower in the carpenter than the journalism occupation. Moreover, the cost of learning the local language depends on one's mother tongue. While Norwegian is linguistically very different from the languages of the EU accession countries, it is relatively similar to the other Scandinavian languages, Swedish and Danish. We therefore expect immigrants from linguistically similar countries to sort into

---

[5]For the vast majority of occupations and industries, Norwegian is the main workplace language.

more language intensive occupations than immigrants from linguistically remote countries.[6]

Using standard data on language intensity (across occupations) and linguistic distance (across countries), we predict the change in the immigrant share for every occupation from 2004 until 2013. Across alternative specifications, the first stage is powerful: The interaction between language intensity and linguistic distance is a strong and robust predictor of the change in the immigrant share across industries. Those predicted changes will then enter the sufficient statistic derived from theory described above. Our 2SLS results are in line with what theory predicts - industries that are intensive in the use of occupations with high immigration grow faster than other industries, contributing to an adjustment in the industry mix and structural change. In our model, industry size adjusts because relative unit costs across industries change. Hence, a second testable hypothesis is that average wage costs decline in industries intensive in occupations with high immigration. Using the same methodology as described above, we find economically and statistically significant adjustments on industry wages as well, consistent with theory.

The exclusion restriction of the instrument is violated if the interaction between language intensity and linguistic distance has an impact on industry size other than the effect going through labor supply. We perform two sets of robustness checks. First, a concern is that the language intensity of occupations/industries is correlated with other occupation/industry characteristics that also determine industry growth. We control for unobserved industry trends by including detailed 2-digit industry and municipality fixed effects. We also add controls for 5-digit observable pre-sample characteristics, such as the skill intensity of the industry. Second, if language intensity is systematically related to industry growth, even in the absence of immigration, then we should obtain significant estimates when regressing the right-hand side on industry growth in the period before the floodgates opened in 2004. Reassuringly, we find no such relationship.

This paper makes several contributions. First, we develop a new methodology for estimating the aggregate impact of immigration on industry growth. This includes building a new model that delivers a testable reduced-form expression derived from general equilibrium theory. Second, we propose a new identification strategy based on exogenous characteristics of occupations and source countries, which turns out to be a powerful predictor of immigration flows across occupations. We believe that this methodology can be used in many different contexts, such as for other time periods and other countries.

There are only a handful of papers exploring the relationship between immigration and

---

[6]We do not follow other Bartik-style studies of using past location of immigrants. The reason is that the pre-enlargement immigrant workers from the accession countries where typically either seasonal workers in specific agricultural areas or highly educated dissidents from the communist area. None of these groups represented any network that would facilitate immigration after the migration barriers were removed.

industry adjustment. Early contributions are Hanson and Slaughter (2002) and Gandal et al. (2004), who develop a decomposition framework to study how changes in labor supply are absorbed in the economy. This approach has since been extended and improved in various directions (Dustmann and Glitz, 2015, González and Ortega, 2011, Lewis, 2003). This body of research is different than ours in several respects. First, our unit of analysis is occupations, instead of skills and/or geographic regions, giving considerable variation in the magnitude of the supply shock (given the large observed heterogeneity in immigration across occupations). Second, we develop a sufficient statistic approach derived from general equilibrium theory, while the previous literature has focused on decomposition frameworks. Third, as described above, our instrumental variable approach is new to the literature.

A related and complementary literature analyzes to what extent investment and production techniques also respond to immigration, see e.g. the survey by Lewis (2013) as well as Lewis (2011). Our paper also relates to the extensive literature on how immigration affects the wage structure, see e.g. Card (2001), Borjas (2003), Dustmann et al. (2005) and Manacorda et al. (2012). Recent contributions also include Burstein et al. (2017) and Ottaviano et al. (2013), while Dustmann et al. (2016) offer a review of different approaches and provides a framework for discussing why parameter estimates differ and how they should be interpreted.
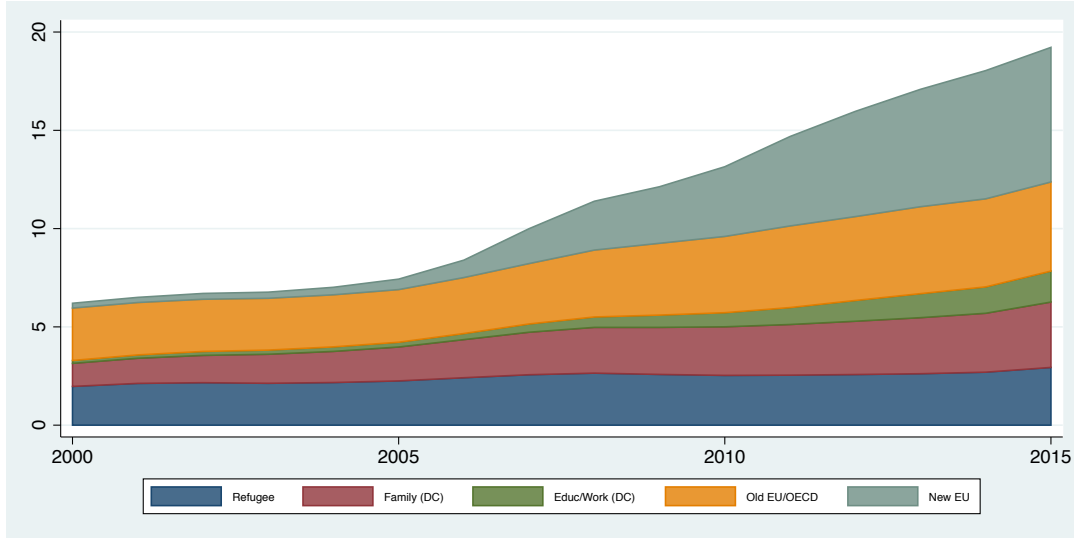
The rest of the paper is organized as follows. Section 2 presents basic facts on the migration shock Norway experienced after the Eastern Enlargement. Section 3 develops the theoretical framework that we use to guide the empirical analysis. Section 4 presents the empirical strategy and show we approach the identification challenge. Section 5 describes the data, Section 6 presents the empirical results, while in Section 7 we test for robustness and discuss the empirical evidence for underlying assumptions. Section 8 concludes.

2 Immigration and Industry Reallocation: Basic Facts

The enlargement of the EU in 2004 and 2007 led to substantial labor migration from east to west as migration restrictions were lifted for about 100 million individuals. Due to favorable macroeconomic conditions, and unlike many EU countries, lax transitional restrictions, Norway was the country with the highest immigration rate, relative to country size. Before the 2004 EU enlargement, accession country citizens had very limited access to the Norwegian labor market. Work permits were provided via domestic employers in need of specialist competence, or on a temporary 3-month seasonal basis, typically for agricultural work.

The Figure 1 shows how the number of immigrant employees relative to total employment developed over the period 2000 to 2015. Only private sector employment and employees be-

Figure 1: Immigrant Employment Shares. Private sector. 2000 to 2015.

tween the age of 20 and 61 are included. Immigrant employees are one of the following groups: (i) Refugees, (ii) Family reunion from developing countries (DC), (iii) Education/work from DC, (iv) Old EU/OECD countries, (v) new EU countries. The immigrant share rose from 7 to 17 percent over just ten years (2004-2013) and has continued to raise until today. About 60 percent of migrants came from EU accession countries. Not surprisingly, immigration had a large impact on aggregate population growth over this period. Almost 70 percent of population growth was due to net immigration.
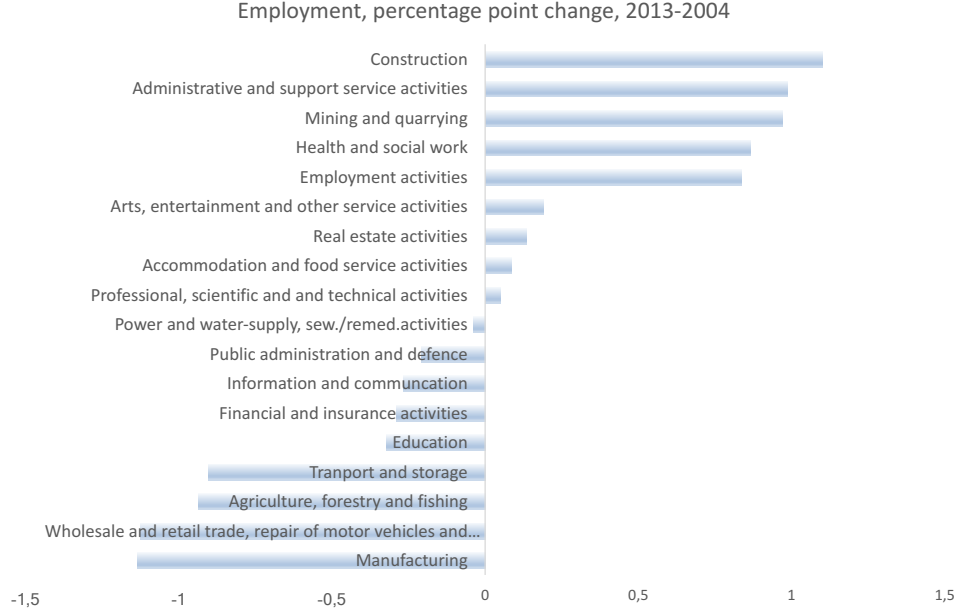
In the aftermath of the EU enlargement, Norway did not only experience a major migration shock. Between 2004 and 2013 the country also faced uneven employment growth across different industries (Figure 2) Among high-growth sectors were the construction and mining sectors. While it is quite clear that the growth in the resource extraction industry ("mining and quarrying") was driven by increased oil production fueled by a booming oil price, other forces must have been responsible for the growth in the construction sector. We aim to investigate the extent to which the observed structural change in the Norwegian economy was driven by the immigration shock.

3 Theoretical Framework

3.1 Model

We introduce a simple theoretical framework to guide the empirical part of the paper. The main objective of the model is to show how a labor supply shock to a given occupation affects employment and wages across different industries while accounting for all general equilibrium

Figure 2: Employment, Percentage Point Change, 2004-2013.

Note: The figure shows the percentage point change in employment by industry from 2004 to 2013.

effects. We consider a factor proportions model where production takes place in I industries that are indexed $i = 1, ..., I$ using labor from $o = 1, ..., O$ occupations. The labor supply side features a Roy-Frechet type model similar to Lagakos and Waugh (2013).[7] Departing from the existing literature, we show how labor supply shocks translate into general equilibrium adjustments in industry size, by using the "exact hat algebra" approach from Dekle et al. (2007).

Production and Labor Demand. Production in each industry requires the use of various occupations. Industries differ according to the intensity with which they use different occupations. The production function in industry i is given by

$$y_i = \varphi_i \prod_o E_{io}^{\omega_{io}},$$

where E_{io} is the number of efficiency units of labor in occupation o, φ_i is industry productivity and ω_{io} are non-negative weights that sum to one, $\sum_o \omega_{io} = 1$. Consumer preferences across sectors are Cobb-Douglas with expenditure shares β_i. Product and labor markets are

---

[7]Other recent contributions using a Roy framework to model the choice of industry or occupation are Burstein et al. (2017), Curuk and Vannoorenberghe (2014) and Galle et al. (2016).

perfectly competitive. Demand for efficiency units of occupation o labor in sector i is thus

$$E_{io}^D = \omega_{io} \frac{\beta_i Y}{w_{io}}, \tag{1}$$

where w_{io} is the wage per efficiency unit paid to a worker with occupation o in industry i, total sales of industry i is $p_i y_i = \beta_i Y$, and Y is aggregate income.

Labor supply. Workers in an occupation differ in terms of productivity. Each worker h with occupation o independently draws a number of efficiency units z_{hi} for each industry i from a Fréchet distribution

$$F_{io}(z) = e^{-A_{io} z^{-\kappa}},$$

with location parameter $A_{io} > 0$ and shape parameter $\kappa > 1$. A greater A_{io} implies that a high efficiency draw in industry i for occupation o workers is more likely. A_{io} also captures the notion that, on average, some occupations are more valuable in certain industries, e.g., that professors are more productive in the education sector than in the agricultural sector. The parameter κ reflects the heterogeneity of productivity draws across industries and captures the degree to which workers are industry-specific. For a small κ, a worker typically has very different draws of productivity across industries, and the loss in productivity incurred by changing industry is relatively large. For a large κ, on the other hand, the productivity draws across industries are relatively close to each other, and changing industry does not result in a large loss of productivity.[8] We assume that the costs of switching occupation are in the short run prohibitive. The assumption that workers are tied to an occupation is supported by evidence on the costs of changing occupation, see e.g. Sullivan (2010) and Kambourov and Manovskii (2009). It is moreover supported by our own evidence on inter-occupational mobility in Section 7.

A worker h with occupation o faces a nominal income in industry i that is a product of her productivity draw and the wage paid: $y_{hio} = z_{ih} w_{io}$. Workers choose an industry as to maximize their income and they offer their entire labor endowment to this industry. Since indirect utility is a monotonic function of the efficiency draw z_{ih}, indirect utility also has a Frechet distribution with shape parameter κ and a location parameter A_{io}. Following Eaton and Kortum (2002) and building on Lagakos and Waugh (2013), we exploit the properties of the productivity distribution and express the share of workers with occupation o choosing to work in industry i as

$$\pi_{io} \equiv \frac{L_{io}}{L_o} = \frac{A_{io} w_{io}^\kappa}{\Phi_o^\kappa}, \tag{2}$$

---

[8]The assumption that workers are mobile across industries but that moving incur a productivity loss is in line with the evidence that there are substantial costs of changing industry, see e.g. Lee and Wolpin (2006) and Brascoupe et al. (2010) due to the loss of industry specific human capital.

where $\Phi_o^\kappa \equiv \sum_j A_{jo} w_{jo}^\kappa$ is the "earnings potential" of occupation o across all industries and $L_o = \sum_j L_{jo}$ is the total mass of workers with occupation o. The supply of efficiency units of this occupation to industry i is moreover given by

$$E_{io}^S = \eta \frac{\Phi_o}{w_{io}} \pi_{io} L_o, \tag{3}$$

where $\eta \equiv \Gamma \left(1 - 1/\kappa \right)$ is a constant with Γ being the gamma function.

3.2   Labor Market Equilibrium

In equilibrium, labor demand must equal supply for each occupation-industry pair, i.e. that (1) equals (3), which yields

$$A_{io} w_{io}^\kappa = \frac{\omega_{io} \beta_i Y}{L_o \eta} \left(\sum_j A_{jo} w_{jo}^\kappa \right)^{\frac{\kappa-1}{\kappa}}.$$

Summing $A_{io} w_{io}^\kappa$ across all industries and rearranging, we get an explicit expression for the equilibrium wage in industry i for occupation o,

$$w_{io} = \left(\frac{\omega_{io} \beta_i}{A_{io}} \right)^{\frac{1}{\kappa}} \left(\frac{Y}{\eta L_o} \right) \left(\sum_j \omega_{jo} \beta_j \right)^{\frac{\kappa-1}{\kappa}}. \tag{4}$$

Appendix A.1 provides detailed derivations. Hence, the wage received by occupation o in industry i is greater the higher the demand faced by this industry, the higher the total demand for occupation o across all industries in the economy, and the smaller the mass of workers with occupation o.

From the labor supply side of the model, we know that that share of o workers in industry i, π_{io}, depends on the wages in every industry. Inserting the expression for equilibrium wages from equation (4) in equation (2), we get that

$$\pi_{io} = \frac{\omega_{io} \beta_i}{\sum_i \omega_{io} \beta_i}, \tag{5}$$

so that, in equilibrium, π_{io} is only a function of the parameters of the model. Appendix A.2 provides detailed derivations.

3.3 A Labor Supply Shock

Let us now consider a shock to the labor supply of one or more occupations o, keeping all other exogenous variables constant. We are interested in the impact of the shock on labor allocation, industry size and factor returns. To simplify notation, we let $\hat{x} \equiv x'/x$ express the relative change in a variable, where x and x' denote the values in the initial and counterfactual equilibrium, respectively.

Using "exact hat algebra" from Dekle et al. (2007), the identity $L_i = \sum_o \pi_{io} L_o$, and the fact that π_{io} is constant in equilibrium, we get our first proposition:

Proposition 1. *Consider a change in the labor supply of occupations o, L_o, keeping all other parameters constant. In general equilibrium, the change in industry employment, L_i, is*

$$\hat{L}_i = \sum_o \lambda_{io} \hat{L}_o, \tag{6}$$

where $\lambda_{io} = L_{io}/L_i$ and using the notation $\hat{x} = x'/x$, where x and x' denote the value in the initial and counterfactual equilibrium, respectively.

Proof. See Appendix A.3. □

Hence, the intensity of occupation o in industry i will determine the extent to which a labor supply shock to this occupation translates into industry growth.

In the model, occupation wages adjust according to equation (4), so $\hat{w}_{oi} = \hat{w}_{oj} = \hat{Y}/\hat{L}_o$. At the industry level, the average wage is $W_i \equiv \sum_o w_{io} E_{io}/L_i = \beta_i Y/L_i$. Hence, in relative changes, we get $\hat{W}_i = \hat{Y}/\hat{L}_i$. This leads us to our second proposition:

Proposition 2. *Consider a change in the labor supply of occupations o, L_o, keeping all other parameters constant. In general equilibrium, the change in average industry wages, W_i, is*

$$\hat{W}_i = \frac{\hat{Y}}{\sum_o \lambda_{io} \hat{L}_o}. \tag{7}$$

Therefore, a positive labor supply shock to occupation o will ceteris paribus lead to a larger decline in industry wages in those sectors that use occupation o intensively.

A useful feature of Proposition 1 is that all general equilibrium effects are purged from the expression in equation (6). This means that we can use Proposition 1 to estimate the *aggregate* effect of a labor supply shock on industry size, accounting for all general equilibrium effects. Typically, reduced-form analyses can only identify *relative* magnitudes. Aggregate wage effects are, however, not identified. Proposition 2 shows clearly that in

general equilibrium, average wage growth also depends on total income growth $\hat{Y}$. Hence, our framework can only identify relative wage effects.

3.4 Aggregate Income

In remains to determine the change in occupation income $\hat{Y}_o$ and aggregate income $\hat{Y}$. Appendix A.4 shows that the change in earnings potential $\hat{\Phi}_o$ is determined by a fixed point that depends on the vector of supply shocks $\hat{L}_o$ and the initial income shares of each occupation, Y_o/Y:

$$\hat{\Phi}_o = \frac{\sum_p \frac{Y_p}{Y} \hat{L}_p \hat{\Phi}_p}{\hat{L}_o}.$$

Given the solution to this fixed point, the change in aggregate income is simply $\hat{Y} = \hat{\Phi}_o \hat{L}_o$ (Appendix A.4), and nominal occupation and industry wages are then pinned down by equations (4) and (7).

4 Empirical Strategy

4.1 Empirical Specification

We set out to analyze the impact of a significant labor migration shock on industry adjustments. Our point of departure is Proposition 1, which states that, in general equilibrium, industry employment growth is a simple weighted average of labor supply growth. As emphasized above, this expression has fully internalized all general equilibrium effects. This means that we can overcome a standard problem in reduced form analysis - that only relative effects can be identified. Using the insight from Proposition 1, we can in fact identify the aggregate effect of the labor supply shock on industry size.

A problem with taking Proposition 1 to the data, is that $\hat{L}_o$ is observed with measurement error because many occupational codes are missing in the raw data in the first few years of our panel. To alleviate this, we use the immigrant share $\mu_o \equiv M_o/L_o$, where M_o is the number of immigrants in occupation o, as our basic independent variable instead.[9] Assuming that (i) natives and migrants are perfect substitutes within an occupation, i.e. $L_o = N_o + M_o$, where N_o is the number of natives, and (ii) the supply shock is entirely driven by immigrants, i.e.,

---

[9]Consider that observed $L_o = \epsilon_o \tilde{L}_o$ and $M_o = \epsilon_o \tilde{M}_o$, where ϵ_o is measurement error and $\tilde{L}_o$ and $\tilde{M}_o$ are the true variables. Then $\mu_o = \tilde{M}_o/\tilde{L}_o$, so there is no measurement error in the immigrant share μ_o.

$\Delta L_o = \Delta M_o$, Proposition 1 can be approximated by:

$$\Delta \ln L_i = \sum_o \lambda_{io} \Delta \check{\mu}_o, \tag{8}$$

where $\Delta \check{\mu}_o = \Delta \mu_o / (1 - \mu_o)$.

The derivations are found in Appendix A.5. While much of the previous literature assumes imperfect substitutability between natives and migrants within similar skill groups (Ottaviano et al. 2013), our approach is to assume that they are perfect substitutes within narrowly defined occupations. Note that our framework would also imply imperfect substitutability across natives and immigrants at higher levels of aggregation, such as broad occupation or skill groups. Section 7 presents empirical evidence supporting the assumption that natives and immigrants are perfect substitutes within occupation. Section 7 also explores the case when the number of natives can adjust as well (so-called native flight).

Final Specification. In our data, we have variation across both industries i (5-digit NACE) and municipalities r. In order to allow for regional and/or industry trends in industry size (which are potentially correlated with $\sum_o \lambda_{io} \Delta \check{\mu}_o$), we include 2-digit industry-municipality pair fixed effects α_{jr}. Adding a term ϵ_{ir} for measurement error and a slope coefficient β, we get the estimating equation

$$\Delta \ln L_{ir} = \alpha_{jr} + \beta \sum_o \lambda_{io} \Delta \check{\mu}_o + \epsilon_{ir}. \tag{9}$$

4.2   Instrumental Variables

Estimating equation (8) is not trivial because high growth industries may also attract immigrants with occupations that are intensively used in that industry. We therefore need an instrument. Our instrument exploits the fact that: (i) occupations are more or less language intensive, and (ii) migrants differ in the linguistic distance between their mother tongue and Norwegian, which is the typical workplace language. The cost of entering an occupation therefore varies across occupations and immigrant groups because immigrants have to invest in language training to master the local language sufficiently well. Hence, the interaction between these two variables is a supply shifter that varies across occupations.

We proceed as follows. Let $\mathbb{L}_o$ be a measure of language requirements in occupation o and $\mathbb{D}_n$ the linguistic distance from source n to Norwegian. We define these variables precisely in the next section. If linguistic distance $\mathbb{D}_n$ is high, then language intensity $\mathbb{L}_o$ should matter more and, everything else being equal, lead to a more uneven supply of migrants across occupations. A straightforward way to capture this complementarity is to use a logit-type

Table 1: Language Intensity and Linguistic Distance.

	Norway	Sweden	Poland
Carpenter	.5	.64	.84
Pre-school teacher	.5	.36	.16
SUM	1	1	1

functional form. The predicted share of immigrants from country n in occupation o is

$$\zeta_{no} = \frac{e^{-\mathbb{L}_o \mathbb{D}_n}}{\sum_p e^{-\mathbb{L}_p \mathbb{D}_n}}. \tag{10}$$

The intuition is straightforward: suppose that we only have two types of occupations, carpenters and pre-school teachers, and two source countries of immigration, Sweden and Poland. The cost of learning the local language for a Polish worker is high, due to a very different mother tongue. The labor supply of Polish immigrants will thus be skewed towards occupations that only needs rudimentary skills in the local language, say e.g. carpenters. For a Swedish worker, the cost of learning the local language is rather low, and we would therefore expect the labor supply of Swedish immigrants to be relatively more evenly distributed across occupations. Table 1 provides a numerical example illustrating how the complementarity affects the predicted shares across occupations and source countries, using actual data on $\mathbb{L}_o$ and $\mathbb{D}_n$. Of course, in the limit when $\mathbb{D}_n$ is zero, then $\mathbb{L}_o$ should not matter at all, because there are no costs of language training. Hence, natives will be evenly assigned across occupations, meaning that our instrument has no power to predict native labor supply.

The next step is to predict the immigrant share across occupations. We predict the number of immigrants and natives in occupation o as

$$\tilde{M}_{ot} = \sum_n \zeta_{no} M_{nt} \tag{11}$$

$$\tilde{N}_{ot} = \zeta_{NO,o} N_t, \tag{12}$$

where M_{nt} and N_{nt} is data on observed total stocks of immigrants and natives (across all occupations), respectively. Note, as discussed above, that $\zeta_{NO,o}$ is identical across all occupations because language barriers are zero for natives. The predicted immigrant share in occupation o is therefore $\tilde{\mu}_{ot} = \tilde{M}_{ot} / \left(\tilde{M}_{ot} + \tilde{N}_{ot} \right)$. The final step is to use $\Delta \tilde{\mu}_o$ instead of $\Delta \check{\mu}_o$ in equation (8). The instrument for $\sum_o \lambda_{io} \Delta \check{\mu}_o$ is therefore

$$\sum_o \lambda_{io} \Delta \tilde{\mu}_o. \tag{13}$$

Our main specification uses the time period 2004 to 2013, so the difference operator Δ refers to the change over this period. As a robustness check we also construct a simpler instrument that only exploits the variation in language intensity across occupations. In the Appendix Section D we provide details on this alternative instrument and empirical results based on this.

Identification. The proposed methodology closely resembles a Bartik-style instrument (Bartik, 1991). In this literature, predicted immigrant flows to a region r are calculated using weights based on historical regional settlement of immigrants from source n and overall immigration by source country n. The idea behind equations (11) and (12) is similar, except that (i) our unit of observation is occupation-source country instead of region-source country, and that (ii) we use exogenous characteristics of occupations and source countries, instead of historical settlement patterns, to calculate the weights ζ_{no}.

The instrument predicts higher growth in immigrant shares ($\Delta\tilde{\mu}_o$) among occupations that are less language intensive. The gradient between $\Delta\tilde{\mu}_o$ and language intensity will be determined by the origin mix of the immigration shock. Since a dominant share of immigrants came from EU accession countries (see Section 2), which are linguistically distant, the instrument predicts a relatively steep gradient. However, it is worth pointing out that the validity of this instrument is likely to extend beyond this specific context and may also work well in other countries or time periods, where the mix of immigrants is different compared to this particular episode.

The exclusion restriction of our instrument is that the weighted average language intensity of occupations in an industry is not systematically related to industry growth other than through the impact of immigration. Since our final specification includes 2-digit industry and municipality fixed effects, the identifying variation comes from within 2-digit (across 5-digit) differences in average language intensity. A potential concern is that, even within 2-digit industries, differences in language requirements are systematically related to e.g. skill intensity, and it may happen that industry growth is correlated with skill intensity. We deal with this issue by including a vector of 5-digit industry characteristics, based on worker and industry data from before the immigrant shock (2003). Specifically, we include pre-sample skill intensity, measured as the share of workers with a completed high school education or higher, average wages, value added, employment, export intensity, measured as total exports relative to total revenue, and the wage share, measured as wage costs relative to total costs (all in logs). Section 7 also presents evidence supporting the exclusion restriction.

5 Data and Variables

Our empirical analysis of the migration shock is based on four main data sets. The first data set is balance sheet data from Statistics Norway for all private non-financial joint-stock companies for the period 1999 to 2013. The balance sheet data is based on data from annual reports that according to Norwegian law must be handed in to a public Register of Company Accounts. The data set contains key account figures related to a firm's income statement and balance sheet including employment, wages, sales and value added. We use the balance sheet data to construct a panel of industry-municipality variables (NACE 5-digit industries). There are in total 441 municipalities and 671 NACE 5-digit industries in our dataset. Guided by the empirical predictions derived from the theory, see Propositions 1 and 2, we focus on two main industry outcomes, employment and average wages. We measure the change in employment and wages between 2004 and 2013. In our discussion of results and mechanisms we also consider other measures of industrial activity.

The second data set is employer-employee data, which includes information on wages and occupations by person-firm-year as well as immigration status (country of birth). Measuring wages and employment, we only include full-time employees. Information on hours worked is given in brackets which are too wide to calculate full-time equivalents based on all workers. The employer-employee data is used to construct the factor-intensity matrix λ_{io} for 214 NACE 3-digit industries and 325 STYRK 4-digit occupational codes, using 2004 values. Table 12 in the Appendix provides a snapshot of the factor intensity matrix for a few different occupations and industries. The dataset is also used to construct the 2004-2013 change in immigrant shares, $\Delta\mu_o$, for each occupation o. Figure 3 illustrates the relationship between employment and change in immigrant shares across 3-digit industries (left panel) and occupations (right panel) in our sample. There is no obvious association between size of the industry/occupations and the migration shock. A few industries and occupations stand out. The immigrant share in construction increased substantially, such as NACE 45.2 and 45.4 where the immigrant share increased by more than 20 percentage points. There was also a significant increase among temp agencies (NACE 74.5, "Labour recruitment and provision of personnel"). The most impacted occupations were helpers and cleaners (STYRK 9132), unskilled workers in construction and maintenance (STYRK 9310) and various carpenter occupations (STYRK 7125 AND 7421).

Third, the O*Net Resource Center offers information of occupational characteristics.[10] Occupations are ranked with respect to a set of requirements. The value 1 means that a given type of skill is not important for the type of work carried out within this occupation,

---

[10]http://www.onetcenter.org/content.html

Figure 3: Changes in Immigration Shares and Industry Size: By Industry and Occupation.
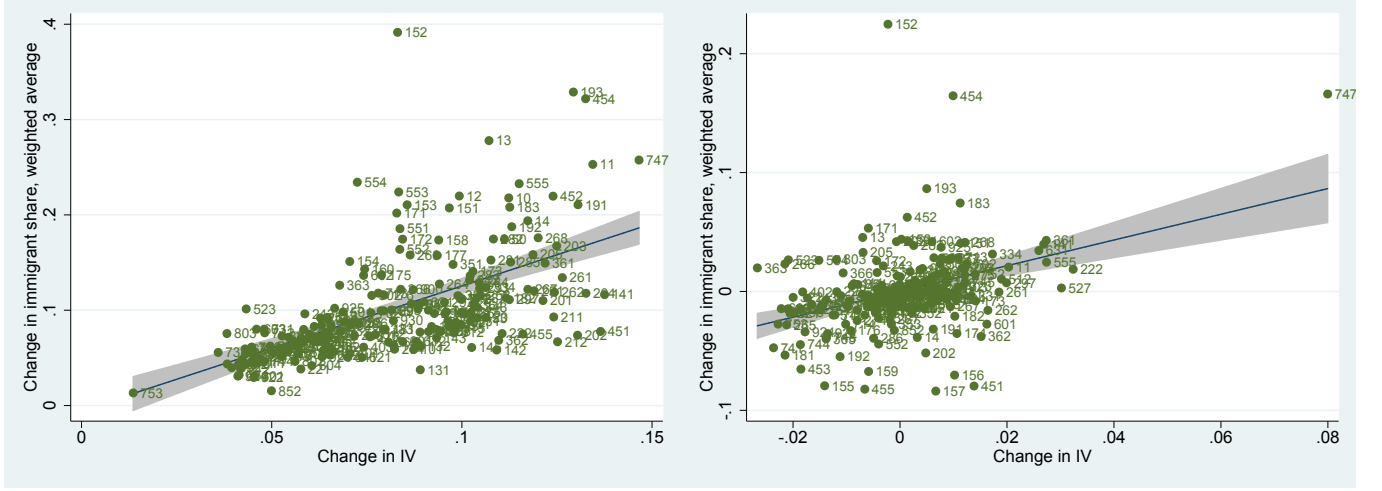

Note: The figure shows the percentage point change in the share of immigrant relative to total employees on the x-axis, and total 2013 employment on a log scale on the y-axis (in 1000s persons). The unit of observation is 3-digit NACE sector (left figure) and 4-digit STYRK occupation code (right figure). Industries/occupations with 2013 employment < 1000 persons are omitted from the figures.

while the value 5 means that it is extremely important. Based on the O*Net data we construct a measure of occupation specific language requirements, which we refer to as language intensity. This measure is based on an average of oral and written comprehension and expression requirements. We use the crosswalk provided by Hoen (2016) in order to match the O*Net data with the occupational codes used in the Norwegian data.

Fourth, we use linguistic proximity data from Adsera and Pytlikova (2015) to measure linguistic distance. They develop an index of language proximity depending on how many levels of the linguistic family tree that the two languages share for 223 countries. The linguistic proximity index equals 0.1 if two languages are only related at the most aggregated level of the linguistic, for example Indo-European languages; it equals 0.25 if two languages belong to the same first and second-linguistic tree level, for example Germanic languages; it equals 0.45 if two languages share up to the third linguistic tree level, for example Germanic North languages; and 0.7 if both languages share the first four levels, for example Scandinavian East (Danish, Norwegian and Swedish). Our measure of linguistic distance, is calculated as $Dist_n = 1 -$ proximity is the linguistic distance from n to NO.

16

Figure 4: 1st Stage Regression.

Note: The figures show scatter plots between the 2004-2013 change in the instrument on the x-axis and the percentage point change in the weighted immigrant share ($\sum_o \tilde{\lambda}_{io} \Delta \mu_o$) on the y-axis. The unit of observation is a 3-digit NACE sector. The left figure shows the raw scatterplot and the right figure shows the scatterplot after demeaning both variables by 2-digit industry averages. The line represents the linear regression line and the gray area the 95 percent confidence interval.

6 Results

We estimate equation (9) instrumenting $\sum_o \lambda_{io} \Delta \check{\mu}_o$ with equation (13), as described above. Figure 4 illustrates the first stage regression, i.e. the relationship between $\sum_o \lambda_{io} \Delta \check{\mu}_o$ (vertical axis) and the instrument (horizontal axis). The left figure shows the raw data, whereas in the right figure both variables are demeaned by 2-digit industry averages, similar to including 2-digit industry fixed effects. Hence, even within 2-digit sectors, the instrument is highly correlated with $\sum_o \lambda_{io} \Delta \check{\mu}_o$. As discussed in Section 4.2, the across-industry variation in the instrument mostly comes from across-industry differences in average language intensity (within 2-digit, across 3- or 5-digit sectors). Appendix D therefore reports the 1st stage and 2SLS results when using the weighted average language intensity, $\sum_o \lambda_{io} \mathbb{L}_o$, as the instrument instead.

The impact on industry employment growth is reported in Table 2. Column (1) presents the 2SLS results in absence of any controls, while in column (2) controls for trends in industry and regional output by including 2-digit industry and municipality pair fixed effects. A potential concern is that language intensity and industry growth are otherwise related. We therefore include pre-sample characteristics of the industry and its workforce. In column (3) we add pre-sample industry controls in order to account for differences across industries in

17

Table 2: Immigration and Employment Growth. 2SLS Estimates.

Dependent variable: $\Delta \ln L_i$	(1)	(2)	(3)	(4)
Δ Immigrant share $(\sum_o \lambda_{io}\Delta\tilde{\mu}_o)$	$.60^a$	1.19^a	1.75^a	2.02^a
	$(.14)$	$(.36)$	$(.35)$	$(.35)$
Pre-sample industry controls	No	No	Yes	Yes
Pre-sample worker controls	No	No	No	Yes
Industry (2-digit)-municipality FE	No	Yes	Yes	Yes
		1st Stage Estimates		
$\sum_o \lambda_{io}\Delta\tilde{\mu}_o$	1.62^a	1.81^a	1.81^a	1.81^a
	$(.02)$	$(.04)$	$(.04)$	$(.04)$
Number of observations	16,763	16,763	16,763	16,763

Note: Robust standard errors clustered by industry-municipality in parentheses. Changes refer to the time period 2004 to 2013. The unit of observation is a 5-digit industry (NACE)-municipality pair. The independent variable and the instrument is constructed at the 3-digit industry level. Industry controls are: Log value added, log employment, log average wages, the share of exports in total sales and the share of wages in total costs (2003 values). The workers control is the share of workers with a completed high school education or higher (2003 values, averaged across firms in a 5-digit industry). ^{a} p< 0.01, ^{b} p< 0.05, ^{c} p< 0.1.

terms of e.g. openness (tradability) and technology,[11] while in column (4) we add pre-sample workers control to account for differences across industries in the skill composition of the labor force.[12] As illustrated in the scatterplot in Figure 4, the 1st stage is precisely estimated across all specifications. The 2SLS estimates suggest that, consistent with theory, the migration shock led to industry growth among those sectors that intensively use occupations that experienced a large labor supply shock. The empirical results are robust to the inclusion of industry and region trends as well as industry and worker controls.

Our theory suggests that industry size adjusts because average wages across industries are changing, see Proposition 2. We therefore estimate the impact on industry average wages and report 2SLS results in Table 3. Average wages in an industry are defined as the total wage bill of the industry relative to the number of employees. As above, we show results with and without fixed effects and controls. We find that the immigration shock led to reduced

---

[11]The industry controls are log value added, log employment, log average wages, the share of exports in total sales and the share of wages in total costs. All variables are calculated based on 2003 values.

[12]The workers control is the share of workers with a completed high school education or higher based on 2003 values and averaged across firms in a 5-digit industry.

Table 3: Immigration and Industry Wage Growth. 2SLS Estimates.

Dependent variable: $\Delta \ln W_i$	(1)	(2)	(3)	(4)
Δ Immigrant share ($\sum_o \lambda_{io} \Delta \tilde{\mu}_o$)	-.56[a]	-.47[b]	-.94[a]	-.74[a]
	(.08)	(.19)	(.19)	(.20)
Pre-sample industry controls	No	No	Yes	Yes
Pre-sample worker controls	No	No	No	Yes
Industry (2-digit)-municipality FE	No	Yes	Yes	Yes
Number of observations	16,763	16,763	16,763	16,763

Note: Robust standard errors clustered by industry-municipality in parentheses. Changes refer to the time period 2004 to 2013. The unit of observation is a 5-digit industry (NACE)-municipality pair. The independent variable and the instrument is constructed at the 3-digit industry level. Industry controls are: Log value added, log employment, log average wages, the share of exports in total sales and the share of wages in total costs (2003 values). The workers control is the share of workers with a completed high school education or higher (2003 values, averaged across firms in a 5-digit industry). [a] $p< 0.01$, [b] $p< 0.05$, [c] $p< 0.1$.

wage growth in the industries most intensive in the use of occupations that experienced a large labor supply shock. The result is robust to the inclusion of industry and worker controls.

Economic Magnitudes. What are the economic magnitudes of the migration shock? As discussed in Section 3.3, our framework does not only identify the relative effect of the labor supply shock on employment growth, which is standard in reduced form analyses, but it also identifies the aggregate effect. The reason is that the expression in Proposition 1 is purged of all general equilibrium effects. Splitting industries into percentiles according to their change in the weighted immigrant share, we get that $\sum_o \lambda_{io} \Delta \mu_o$ is five and 17 percentage points for the 10*th* and 90*th* percentile industry, respectively. Based on our estimates, this suggest that the migration shock led roughly 30 and 10 percent growth in industry employment in the most affected and least affected industries.[13]

According to Proposition 2, the wage impact is not purged of general equilibrium effect, so here our framework can only identify relative effects. Splitting industries into percentiles as above, we find that the most affected (90*th* percentile) industries faced 7 percent lower average wages compared to the least affected (10*th* percentile).[14] Over the sample period, the mean average wage, i.e. the average of $\Delta \ln W_i$, increased nominally by 40 percent. Hence,

---

[13]Calculated as .17 × 2.0 and .5 × 2.0 log points.
[14]Calculated as $.12x\,(-0.58) \approx .07$ log points

Table 4: Immigration and Industry Employment. Falsification Test.

Dependent variable: $\Delta \ln L_i$ (1999-2003)	(1)	(2)
Δ Immigrant share ($\sum_o \lambda_{io} \Delta \check{\mu}_o$) (2004-2013)	-.09	-.33
	(.12)	(.34)
Pre-sample industry controls	No	No
Pre-sample worker controls	No	No
Industry (2-digit)-municipality FE	No	Yes
	1st Stage Estimates	
$\sum_o \lambda_{io} \Delta \tilde{\mu}_o$ (2004-2013)	1.62^a	1.81^a
	(.02)	(.04)
Number of observations	14,315	14,315

Note: Robust standard errors clustered by industry-municipality in parentheses. Changes refer to the time period 2004 to 2013 for the instrument and 1999 to 2003 for the dependent variable. The unit of observation is a 5-digit industry (NACE)-municipality pair. The independent variable and the instrument is constructed at the 3-digit industry level. [a] $p < 0.01$, [b] $p < 0.05$, [c] $p < 0.1$.

although relative wages declined, even the most affected industries experienced nominal (and real) wage growth. We conclude that the migration shock lead to economically large adjustment of industry size. Sectors intensive in the use of occupations especially affected by immigration grew significantly faster. Our results suggests that these adjustments were triggered by changes in relative wage costs across industries.

7 Robustness and Discussion of Assumptions

7.1 Falsification Test

A potential concern is that industries with exposure to the migration shock are industries with in general higher employment growth than other industries. To address this concern, we perform a placebo test and regress 1999-2003 employment and wage growth on 2004-2013 immigration. Results are reported in Tables 4 and 5 for employment growth and wage growth respectively. The coefficients of interest are not significant, suggesting that there are no differential industry-specific pre-trends.[15]

---

[15]Note that we cannot include pre-sample controls here because we do not have data on industry outcomes before 1999.

Table 5: Immigration and Industry Wage growth. Falsification Test.

Dependent variable: $\Delta \ln W_i$ (1999-2003)	(1)	(2)
Δ Immigrant share ($\sum_o \lambda_{io}\Delta\check{\mu}_o$) (2004-2013)	-.14^c	-.03
	(.07)	(.20)
Pre-sample industry controls	No	No
Pre-sample worker controls	No	No
Industry (2-digit)-municipality FE	No	Yes
Number of observations	14,315	14,315

Note: Robust standard errors clustered by industry-municipality in parentheses. Changes refer to the time period 2004 to 2013 for the instrument and 1999 to 2003 for the dependent variable. The unit of observation is a 5-digit industry (NACE)-municipality pair. The independent variable and the instrument is constructed at the 3-digit industry level. ^{a} p< 0.01, ^{b} p< 0.05, ^{c} p< 0.1.

7.2 Substitutability between Immigrants and Natives

The theoretical framework assumed that immigrants and natives are perfect substitutes within detailed occupation codes, while the previous literature typically assumes imperfect substitutability between natives and migrants within similar skill groups (Manacorda et al. (2012) and Ottaviano et al. 2013). To test our assumption, we explore to what extent wages are different between natives and immigrants within occupations. We use individual-level data on wages for natives and immigrants, and, using all full-time employment spells in 2014, we regress log wages on a dummy which takes the value one if the individual is an immigrant. Table 6 reports estimates of the immigrant-native log wage differentials. Columns (1)-(3) report results for males and (4)-(6) for females. Without any controls, the male wage gap is .29 log points (column 1). Controlling for age, experience and tenure, the wage gap drops considerably to .19 log points (column 2).[16] Adding 4-digit occupational fixed effects, the immigrant wage gap is zero (column 3). We also find a zero wage gap for females within occupations (column 6). Hence, we conclude that there is empirical support for the assumption of perfect substitutability between natives and immigrants within occupations.

7.3 Occupational Switching

Our model ruled out the possibility of endogenous occupational switching. If native workers respond to the migration shock by leaving occupations with less language requirements and

---

[16]In previous studies based on a CES tree structure, the degree of substitutability is typically defined within skill groups defined by education and experience.

Table 6: Immigrant-Native Wage Gap

Dependent variable: Log wage	Men			Women		
	(1)	(2)	(3)	(4)	5)	(6)
Immigrant dummy	-0.286[a]	-0.188[a]	0.017	-0.216[a]	-0.125[a]	0.014
	(0.027)	(0.025)	(0.014)	(0.039)	(0.029)	(0.014)
Age, experience and tenure controls	No	Yes	Yes	No	Yes	Yes
4-digit occupation fixed effects	No	No	Yes	No	No	Yes
No obs	791,163	791,163	791,163	356,715	356,715	356,715

Note: Robust standard errors clustered by 4-digit occupation. Data set restricted to full time employees in year 2014. [a] p< 0.01, [b] p< 0.05, [c] p< 0.1.

upgrade towards occupations with higher language requirements, this would trigger a labor supply adjustment not accounted for in our model. This section therefore investigates the impact of mobility across occupations.

Our first approach is to derive an estimating equation, similar to equation (8), under the assumption that native workers flee occupations with high immigration, so-called native flight. Consider the case that an increase ΔM_o of immigrant labor in occupation o leads to a decline $(1 - \psi) \Delta M_o$ of native labor in occupation o, $\Delta N_o = - (1 - \psi) \Delta M_o$. Appendix A.5 shows that the estimating equation then becomes

$$\Delta \ln L_i = \sum_o \frac{\psi}{1 - \psi \mu_o} \lambda_{io} \Delta \mu_o.$$

If the immigrant share is close to zero in the pre-period, i.e. $\mu_o \approx 0$, then this simplifies further to $\Delta \ln L_i = \psi \sum_o \lambda_{io} \Delta \mu_o$. High native flight (small ψ) therefore leads to a more muted industry employment response. When there if no native flight, i.e. $\psi = 1$, then we are back to the baseline case in equation (8). Therefore, high native flight will lead us to underestimate the true impact of a labor supply shock.

Our second approach is to quantify the extent of occupational switching in the data. We allocate each occupation into deciles according to language intensity. We then calculate a transition matrix showing the share of workers that switch between occupations belonging to different language intensity deciles between year $t - 1$ and t. We find that from one year to the next, about 90 percent of workers stay in their original language requirement decile, while another $3 - 4$ precent is found in the neighbor cell(s), see Tables 7 and 8. This evidence suggests there are large occupational switching costs, limiting the extent of native

flight, especially across occupations with very different language requirements.

8 Conclusions

TBW

Table 7: Occupational Switching by Decile of Language Importance. Natives.

t-1/t	1	2	3	4	5	6	7	8	9	10	Sum
1	90.4	2.4	1.7	0.9	1.6	0.6	0.4	1.0	0.6	0.5	100.0
2	2.2	91.3	1.6	0.9	1.0	1.0	0.3	0.8	0.6	0.4	100.0
3	1.1	1.2	91.4	1.1	1.2	1.2	0.7	0.8	0.9	0.4	100.0
4	0.7	0.7	1.5	90.1	1.3	0.4	0.8	2.0	1.2	1.3	100.0
5	0.9	0.6	1.1	0.8	89.5	1.4	1.2	1.3	1.9	1.2	100.0
6	0.2	0.4	0.9	0.2	0.9	89.9	2.4	2.2	2.0	1.0	100.0
7	0.2	0.2	0.4	0.4	0.8	1.8	90.4	1.7	2.8	1.4	100.0
8	0.3	0.2	0.3	0.9	0.8	1.6	1.1	88.8	3.3	2.7	100.0
9	0.2	0.2	0.3	0.6	1.2	1.2	2.0	2.9	88.5	3.1	100.0
10	0.1	0.1	0.2	0.5	0.7	0.6	1.0	2.5	3.1	91.2	100.0

Note: The deciles refer to the language intensity of the occupation. The population is all full-time workers over the period 2010 to 2015.

Table 8: Occupational Switching by Decile of Language Importance. Immigrants.

t-1/t	1	2	3	4	5	6	7	8	9	10	Sum
1	90.7	4.4	1.8	0.8	0.9	0.2	0.4	0.3	0.3	0.2	100.0
2	4.5	90.6	1.5	1.2	0.9	0.4	0.2	0.3	0.3	0.2	100.0
3	2.0	1.6	92.4	0.9	0.9	0.5	0.8	0.4	0.4	0.2	100.0
4	1.0	1.9	1.8	89.8	1.8	0.3	0.5	1.5	0.6	0.9	100.0
5	1.3	1.6	1.3	1.7	88.4	0.9	1.3	1.1	1.6	1.1	100.0
6	0.4	0.4	0.5	0.2	0.7	89.6	2.8	2.9	2.0	0.6	100.0
7	0.4	0.2	0.5	0.4	0.9	2.2	90.8	1.5	2.2	1.0	100.0
8	0.4	0.2	0.4	1.2	0.9	3.0	1.4	87.6	2.8	2.1	100.0
9	0.4	0.3	0.5	0.6	1.6	2.5	2.6	2.9	85.7	2.7	100.0
10	0.2	0.2	0.3	0.7	1.1	0.7	1.4	2.6	2.8	90.2	100.0

Note: The deciles refer to the language intensity of the occupation. The population is all full-time workers over the period 2010 to 2015.

References

Adsera, A. and M. Pytlikova (2015). The role of language in shaping international migration. *The Economic Journal 125*, F49–F81.

Bartik, T. J. (1991). Who benefits from state and local economic development policies? *Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.*.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the local labor market. *The Ouarterly Journal of Economics 118*(4), 1335–1374.

Brascoupe, C., S. Chaudhuri, and J. McLaren (2010). Trade shocks and labor adjustment: A structural empirical approach. *American Economic Review 100*(3), 1008–1045.

Burstein, A., G. Hanson, L. Tian, and J. Vogel (2017, apr). Tradability and the labor market impact of immigration: Theory and evidence from the u.s. *Mimeo*.

Card, D. (2001). Immigrant inflows, native outflows and the local labor market iimpact of higher immigration. *Journal of Labor Economics 19*(1), 22–64.

Curuk, M. and G. Vannoorenberghe (2014). Inter-sectoral labor reallocation in the short-run.

Dekle, R., J. Eaton, and S. Kortum (2007). Unbalanced trade. *American Economic Review: Papers and Proceedings 97*(2), 351–355.

Dølvik, J. E. and L. Eldring (2008). *Mobility of labour from new EU states to the Nordic Region: – Development trends and consequences*. Nordic Council of Ministers.

Donaldson, D. and R. Hornbeck (2016). Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics*.

Dustmann, C., F. Fabbri, and I. Preston (2005). The impact of immigration on the british labour market. *The Economic Journal 80*(1), 45–73.

Dustmann, C. and A. Glitz (2015). How do industries and firms respond to changes in local labor supply? *Journal of Labor Economics 33*(3), 711–750.

Dustmann, C., U. Schoenberg, and J. Stuhler (2016). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives 30*(4), 31–56.

Eaton, J. and S. Kortum (2002). Technology, geography and trade. *Econometrica 70*(5), 1741–1779.

Galle, S., A. Rodrigues-Clare, and M. Yi (2016). Slicing the pie: Quantifying the aggregate and distributional effects of trade. *mimeo*.

Gandal, N., G. H. Hanson, and M. J. Slaughter (2004). Technology, trade and adjustment to immigration in israel. *European Economic Review 48*(2), 403–428.

González, L. and F. Ortega (2011). How do very open economies absorb large immigration flows? recent evidence from spanish regions. *Labour Economics 18*(1), 57–70.

Hanson, G. and M. J. Slaughter (2002). Labor market adjustment in open economies: Evidence from the us states. *Journal of International Economics 57*(1), 3–29.

Hoen, M. B. (2016). Occupational crosswalk, dat and language requirements. Frisch Centre Working Paper 1/2016, Frisch.

Kambourov, G. and I. Manovskii (2009). Occupational specificity of human capital. *International Economic Review 50*(1), 63–115.

Lagakos, D. and M. E. Waugh (2013). Selection, agriculture, and cross-country productivity differences. *The American Economic Review 103.2*(948-980).

Lee, D. and K. I. Wolpin (2006). Intersectoral mobility and the growth of the service sector. *Econometrica 74*(1), 1–46.

Lewis, E. G. (2003). Local open economies within the u.s.: How do industries respond to immigration. Technical report, Federal Reserve Bank of Philadelphia Working Paper No. 04-1.

Lewis, E. G. (2011). Immigration, skill mix and capital skill complementarity. *The Quarterly Journal of Economics 126*, 1029–1069.

Lewis, E. G. (2013). Immigration and production technology. *Annu. Rev. Econ 5.1*(165-191).

Manacorda, M., A. Manning, and J. Wadsworth (2012). The impact of immigration on the structure of wages: Theory and evidence from britain. *Journal of the European Economic Association 10*(1), 120–151.

OECD (2017). Foreign-born population (indicator). doi: 10.1787/5a368e1b-en (accessed on 05 may 2017).

Ottaviano, G. I. P., G. Peri, and G. C. Wright (2013). Immigration, offshoring and american jobs. *American Economic Review 103*(5), 1960–1976.

Sullivan, P. (2010). Empirical evidence on occupation and industry specific human capital. *Labour Economics 17*(3), 567–580.

Appendix

A   Derivations

A.1   Equilibrium Wages

In equilibrium, labor demand must equal supply for each occupation-industry pair, i.e. that (1) equals (3), which yields

$$A_{io}w_{io}^{\kappa} = \frac{\omega_{io}\beta_i Y}{L_o \eta}\left(\sum_j A_{jo}w_{jo}^{\kappa}\right)^{\frac{\kappa-1}{\kappa}}. \tag{14}$$

Summing across all industries, we get

$$\sum_i A_{io}w_{io}^{\kappa} = \frac{Y}{\eta L_o}\left(\sum_j A_{jo}w_{jo}^{\kappa}\right)^{(\kappa-1)/\kappa}\sum_i \omega_{io}\beta_i$$

which can be rewritten to

$$\sum_i A_{io}w_{io}^{\kappa} = \left(\frac{Y}{\eta L_o}\sum_i \omega_{io}\beta_i\right)^{\kappa}. \tag{15}$$

Inserting (15) back into (14) yields

$$A_{io}w_{io}^{\kappa} = \omega_{io}\beta_i\left(\frac{Y}{\eta L_o}\right)^{\kappa}\left(\sum_i \omega_{io}\beta_i\right)^{\kappa-1}$$

$$w_{io} = \left(\frac{\omega_{io}\beta_i}{A_{io}}\right)^{\frac{1}{\kappa}}\left(\frac{Y}{\eta L_o}\right)\left(\sum_j \omega_{jo}\beta_j\right)^{\frac{\kappa-1}{\kappa}},$$

which is identical to equation (4) in the main text.

A.2 Sorting of Occupations to Industries

Inserting equations (4) and (15) into (2), we get

$$
\begin{aligned}
\pi_{io} &= \frac{A_{io} w_{io}^{\kappa}}{\sum_j A_{jo} w_{jo}^{\kappa}} \\
&= \frac{\omega_{io} \beta_i \left(\frac{Y}{\eta L_o}\right)^{\kappa} \left(\sum_i \omega_{io} \beta_i\right)^{\kappa-1}}{\left(\frac{Y}{\eta L_o} \sum_i \omega_{io} \beta_i\right)^{\kappa}} \\
&= \frac{\omega_{io} \beta_i}{\sum_i \omega_{io} \beta_i},
\end{aligned}
$$

which is identical to equation (5) in the main text.

A.3 Proposition 1

Consider a shock to the labor supply of one or more occupations o, keeping all other exogenous variables constant. We let $\hat{x} \equiv x'/x$ express the relative change in a variable, where x and x' denote the values in the initial and counterfactual equilibrium, respectively. In relative changes, identity $L_i = \sum_o \pi_{io} L_o$ becomes

$$
\begin{aligned}
\hat{L}_i &= \frac{\sum_o \pi'_{io} L'_o}{\sum_o \pi_{io} L_o} \\
&= \sum_o \frac{\pi_{io} L_o}{\sum_p \pi_{ip} L_p} \hat{\pi}_{io} \hat{L}_o \\
&= \sum_o \frac{L_{io}}{L_i} \hat{\pi}_{io} \hat{L}_o.
\end{aligned}
$$

Using the equilibrium expression for π_{io} in equation (5), we know that $\hat{\pi}_{io} = 1$. Hence, we get

$$
\hat{L}_i = \sum_o \lambda_{io} \hat{L}_o,
$$

where $\lambda_{io} = L_{io}/L_i$. This expression is identical to equation (6) in the main text.

A.4 Aggregate Income

Using equation (3), total income of occupation o can be written as

$$Y_o \equiv \sum_i w_{io} E_{io}$$

$$= \sum_i \eta \Phi_o \pi_{io} L_o$$

$$= \eta L_o \Phi_o.$$

Aggregate income is therefore $Y \equiv \sum_p Y_p = \eta \sum_p L_p \Phi_p$. In changes, we get

$$\hat{Y} = \sum_p \frac{Y_p}{Y} \hat{L}_p \hat{\Phi}_p.$$

Recall that $\Phi_o^\kappa \equiv \sum_i A_{io} w_{io}^\kappa$. Hence, from equation (15), $\Phi_o = \frac{Y}{\eta L_o} \sum_i \omega_{io} \beta_i$, or in changes $\hat{\Phi}_o = \hat{Y}/\hat{L}_o$. We therefore get the fixed point

$$\hat{\Phi}_o = \frac{\sum_p \frac{Y_p}{Y} \hat{L}_p \hat{\Phi}_p}{\hat{L}_o}. \tag{16}$$

We solve equation (16) numerically for each occupation o. Given the solution to $\hat{\Phi}_o$, we can back out occupation income $\hat{Y}_o$ and aggregate income $\hat{Y}$. Calculating the equilibrium in changes only requires data on initial income shares, Y_p/Y, in addition to the supply shock $\hat{L}_o$.

A.5 The Estimation Equation

No Native flight

We first discuss the case when $\Delta L_o = \Delta M_o$ (no native flight). Define the immigrant share $\mu_o \equiv M_o/L_o$. For marginal changes, the following must hold:

$$\Delta \mu_o = \frac{L_o \Delta M_o - M_o \Delta L_o}{L_o^2}$$

$$= \frac{\Delta M_o}{L_o} - \frac{M_o}{L_o} \frac{\Delta L_o}{L_o}$$

$$= \frac{\Delta M_o}{L_o} (1 - \mu_o).$$

The change in occupation employment L_o is then

$$\hat{L}_o = 1 + \frac{\Delta L_o}{L_o}$$
$$= 1 + \frac{\Delta M_o}{L_o}$$
$$= 1 + \frac{\Delta \mu_o}{1 - \mu_o}. \tag{17}$$

Using the log approximation $\ln(1 + x) \approx x$ or $e^x \approx 1 + x$ repeatedly, we can express Proposition 1 as

$$\hat{L}_i = \sum_o \lambda_{io} \hat{L}_o$$
$$= \sum_o \lambda_{io} e^{\Delta \ln L_o}$$
$$\approx \sum_o \lambda_{io} (1 + \Delta \ln L_o)$$
$$= 1 + \sum_o \lambda_{io} \Delta \ln L_o$$

or

$$\Delta \ln L_i = \ln\left(1 + \sum_o \lambda_{io} \Delta \ln L_o\right)$$
$$\approx \sum_o \lambda_{io} \Delta \ln L_o$$
$$= \sum_o \frac{\lambda_{io}}{1 - \mu_o} \Delta \mu_o,$$

where we used equation (17) and $\Delta \ln L_o = \ln\left[1 + \Delta \mu_o / (1 - \mu_o)\right] \approx \Delta \mu_o / (1 - \mu_o)$.

Native flight

Consider the case that an increase ΔM_o of immigrant labor in occupation o leads to a decline $(1 - \psi) \Delta M_o$ of native labor in occupation o, $\Delta N_o = -(1 - \psi) \Delta M_o$, so-called native flight. Then, $\Delta L_o = \psi \Delta M_o$ and

$$\Delta\mu_o = \frac{\Delta M_o}{L_o} - \psi\frac{M_o}{L_o}\frac{\Delta M_o}{L_o}$$
$$= \frac{\Delta M_o}{L_o}\left(1 - \psi\mu_o\right).$$

Hence, we have

$$\hat{L}_o = 1 + \frac{\Delta M_o}{L_o} + \frac{\Delta N_o}{L_o}$$
$$= 1 + \frac{\Delta M_o}{L_o} - (1-\psi)\frac{\Delta M_o}{L_o}$$
$$= 1 + \psi\frac{\Delta M_o}{L_o}$$
$$= 1 + \frac{\psi}{1-\psi\mu_o}\Delta\mu_o.$$

Using the same approximation as above, we get

$$\Delta\ln L_i \approx \sum_o \frac{\psi}{1-\psi\mu_o}\lambda_{io}\Delta\mu_o.$$

If the immigrant share is close to zero in the pre-period ($\mu_o = 0$), this can be simplified further to $\Delta\ln L_i \approx \psi\sum_o \lambda_{io}\Delta\mu_o$. Hence, a high degree of native flight (low ψ) will tend to dampen the industry size response of the immigrant supply shock.

B   Variable Construction

We construct a set of variables to measure industry responses to migration.

Sales To measure the sales of an industry (or firm) we use operating income.

Employment Is calculated based on account statistics.

Value added To measure value added in an industry we use operating income and subtract costs of intermediates. The account statistics split cost of intermediates into two groups: Costs of raw materials and consumables used and Other operating expenses.

Average wage To measure average wage we use data on employment and the wage costs calculated based on the Accounts statistics data.

Wage share To calculate wage share we use wage costs divided valued added (see above for construction).

Other costs share To calculate the share of other costs (typically the sourcing of services)

Table 9: Immigration and Employment Growth. OLS Estimates.

	(1)	(2)	(3)	(4)
Dependent variable: $\Delta \ln L_i$:				
Δ Immigrant share ($\sum_o \lambda_{io} \Delta \tilde{\mu}_o$)	$.27^a$	-.09	-.01	.04
	(.10)	(.17)	(.17)	(.17)
Dependent variable: $\Delta \ln W_i$:				
Δ Immigrant share ($\sum_o \lambda_{io} \Delta \tilde{\mu}_o$)	$-.20^a$	$-.18^b$	$-.35^a$	$-.29^a$
	(.06)	(.09)	(.08)	(.08)
Pre-sample industry controls	No	No	Yes	Yes
Pre-sample worker controls	No	No	No	Yes
Industry (2-digit)-municipality FE	No	Yes	Yes	Yes
Number of observations	16,763	16,763	16,763	16,763

Note: Robust standard errors clustered by industry-municipality in parentheses. Changes refer to the time period 2004 to 2013. The unit of observation is a 5-digit industry (NACE)-municipality pair. The independent variable is constructed at the 3-digit industry level. Industry controls are: Log value added, log employment, log average wages, the share of exports in total sales and the share of wages in total costs (2003 values). The workers control is the share of workers with a completed high school education or higher (2003 values, averaged across firms in a 5-digit industry). ^{a} p< 0.01, ^{b} p< 0.05, ^{c} p< 0.1.

in total costs we divide other operating expenses by the sum of wage costs, Costs of raw materials and consumables used and Other operating expenses.
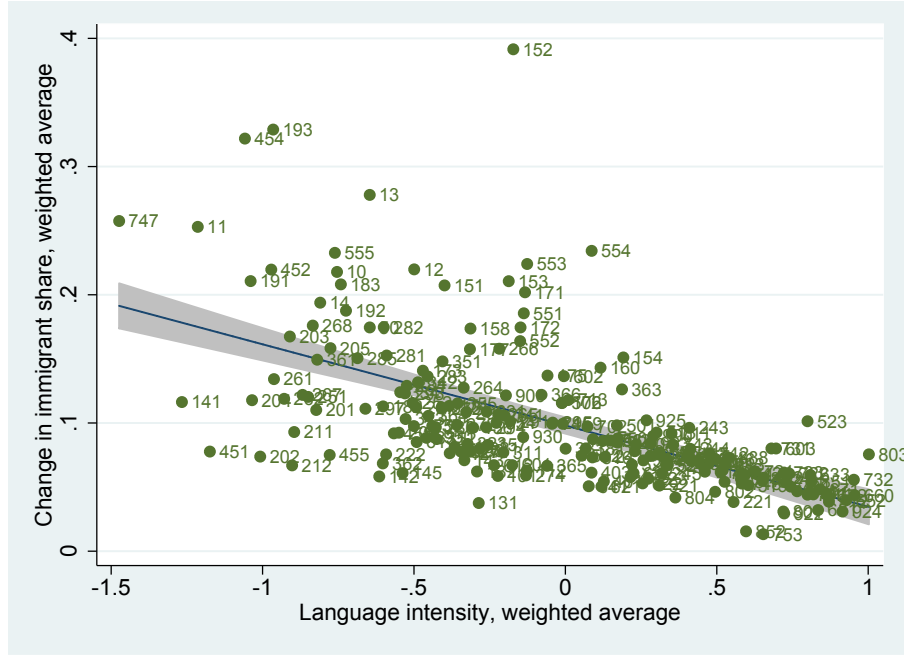
EBITDA To measure profits/returns to capital we use earnings before interests, taxes, depreciation and amortization, which is a standard measure in accounting.

Capital intensity To measure capital intensity we use data on Tangible fixed assets and divide this by employment.

C OLS Results

This section shows results when estimating (8) using OLS instead of 2SLS. Table 9 reports results on the impact of the migration shock on employment and average wages.

Figure 5: $1st$ Stage: Alternative Instrument

Note: The figures show scatterplots between the instrument $\sum_o \lambda_{io} \mathbb{L}_o$ on the x-axis and the percentage point change in the weighted immigrant share $(\sum_o \lambda_{io} \Delta \mu_o)$ on the y-axis. The unit of observation is a 3-digit NACE sector. The line represent the linear regression line and the gray area the 95 percent confidence interval.

D   Alternative Instrument

This section presents an alternative instrument that is only based on variation in language intensity across occupations. The alternative instrument for $\sum_o \lambda_{io} \Delta \check{\mu}_o$ is the weighted average language intensity, $\sum_o \lambda_{io} \mathbb{L}_o$, where $\mathbb{L}_o$ is the language intensity of occupation o. We re-estimate (8) with 2SLS using the alternative instrument. Figure 4 plots the first stage regression, i.e. the relationship between $\sum_o \lambda_{io} \Delta \tilde{\mu}_o$ and $\sum_o \lambda_{io} \mathbb{L}_o$. Tables 10 and 11 report 2SLS as well as first stage results based on the alternative instrument. The results are close to those reported based using our main instrument.

Table 10: Immigration and Employment Growth. 2SLS Estimates.

Dependent variable: $\Delta \ln L_i$	(1)	(2)	(3)	(4)
Δ Immigrant share $(\sum_o \lambda_{io} \Delta \check{\mu}_o)$	$.64^a$	1.24^a	1.88^a	2.14^a
	$(.14)$	$(.37)$	$(.36)$	$(.36)$
Pre-sample industry controls	No	No	Yes	Yes
Pre-sample worker controls	No	No	No	Yes
Industry (2-digit)-municipality FE	No	Yes	Yes	Yes
		1st Stage Estimates		
$\sum_o \lambda_{io} \mathbb{L}_o$	$-.08^a$	$-.08^a$	$-.08^a$	$-.08^a$
	$(.00)$	$(.00)$	$(.00)$	$(.00)$
Number of observations	16,763	16,763	16,763	16,763

Note: Robust standard errors clustered by industry-municipality in parentheses. Changes refer to the time period 2004 to 2013. The unit of observation is a 5-digit industry (NACE)-municipality pair. The independent variable and the instrument is constructed at the 3-digit industry level. Industry controls are: Log value added, log employment, log average wages, the share of exports in total sales and the share of wages in total costs (2003 values). The workers control is the share of workers with a completed high school education or higher (2003 values, averaged across firms in a 5-digit industry). ^{a} p< 0.01, ^{b} p< 0.05, ^{c} p< 0.1.

Table 11: Immigration and Industry Wage Growth. 2SLS Estimates.

Dependent variable: $\Delta \ln W_i$	(1)	(2)	(3)	(4)
Δ Immigrant share $(\sum_o \lambda_{io} \Delta \check{\mu}_o)$	$-.54^a$	$-.43^b$	$-.84^a$	$-.64^a$
	$(.08)$	$(.20)$	$(.19)$	$(.20)$
Pre-sample industry controls	No	No	Yes	Yes
Pre-sample worker controls	No	No	No	Yes
Industry (2-digit)-municipality FE	No	Yes	Yes	Yes
Number of observations	16,763	16,763	16,763	16,763

Note: Robust standard errors clustered by industry-municipality in parentheses. Changes refer to the time period 2004 to 2013. The unit of observation is a 5-digit industry (NACE)-municipality pair. The independent variable and the instrument is constructed at the 3-digit industry level. Industry controls are: Log value added, log employment, log average wages, the share of exports in total sales and the share of wages in total costs (2003 values). The workers control is the share of workers with a completed high school education or higher (2003 values, averaged across firms in a 5-digit industry). ^{a} p< 0.01, ^{b} p< 0.05, ^{c} p< 0.1.

Table 12: Factor Intensity Matrix

	Professor or similar STYRK 2310			Sum
NACE803 "Higher education"	.52	.	.	1
NACE 751 "Administration of the state (..)	.06	.	.	1
NACE 732 "Research (..) on natural sciences and engineering"	.05	.	.	1

	Carpenter or similar STYRK 7421			Sum
NACE203 "Manufacture of builders' carpentry and joinery"	.21	.	.	1
NACE205 "Manufacture of other products of wood (..)	.19	.	.	1
NACE454 "Building completion"	.15	.	.	1