Occupational fragmentation and sectoral employment adjustments

Malik Curuk* and Gonzague Vannoorenberghe†

September 2013

This paper shows that the geographic distribution of occupational employment within a country affects the sensitivity of an industry's employment to external shocks. We develop a theory-based measure of occupational similarity between industries and show that the geographic proximity to industries using similar occupations raises the ability of an industry to respond to aggregate shocks. Using data on the employment growth of region-industry pairs in the U.S., we confirm empirically that the employment of an industry responds more to national shocks in regions where other industries using similar occupations are located. Calibrating our model to the U.S. economy, we show that if workers could costlessly change occupation and migrate across regions, the extent of cross-sectoral labour reallocation following an external shock would be substantially stronger. We also find that employment in agricultural sectors and in apparel and textile manufacturing are among the least responsive to external shocks, implying that workers in these sectors are very "sector-specific".

Keywords: Occupations, Local labour markets, Factor specificity, Terms of trade

J.E.L. classification: F1, J2

*Tilburg University, Department of Economics, PO Box 90153 5000 LE Tilburg, Netherlands. Tel. +31 13 466 8053,e-mail:M.Curuk@uvt.nl.

†Tilburg University, Department of Economics, PO Box 90153 5000 LE Tilburg, Netherlands. Tel. +31 13 466 2511,e-mail:G.C.L.Vannoorenberghe@uvt.nl.

1 Introduction

A large literature has studied the role of technological progress and of international trade as drivers of productivity growth and of welfare gains. A central channel for these gains to materialise is the reallocation of factors towards sectors with large productivity growth or with a comparative advantage (McMillan and Rodrik (2011)). The presence of short-run frictions to the factor adjustment process can however substantially reduce the size of these gains (Lee and Wolpin (2006), Kambourov (2009)).

In this paper, we show that the ability of an industry to adjust its labour input in the short-run hinges on the availability of the relevant type of labour - occupations - in the regions where the industry is located. We emphasise the importance of two types of widely documented short-run rigidities in explaining labour market reactions to external shocks[1]: the geographical immobility of workers as well as their inability to change occupation in the short run. We show that these two dimensions are important determinants of the regional employment fluctuations of U.S. industries between 2003 and 2008.

We model the United States as a collection of small regional units which differ in their industry structure. For example, around 4% of employees in Detroit were working in the manufacturing of motor vehicle parts[2] in 2003 compared to a national average of 0.5%. We also assume that industries use occupations in different proportions. Electrical engineers for instance represent 4% of employees in the manufacturing of measuring instruments, but only 0.1% of the national labour force[3]. In our theoretical model, we derive an index of the "ease" with which an industry can adjust its employment in a particular region. This index, which we call the "employment responsiveness" of a particular region-industry pair, measures the relative size of the pool of labour with which the industry can exchange labour if it wants to adjust employment. The index captures two different effects. First,

[1]A large literature shows that regional mobility is imperfect in the short run (e.g. Blanchard, Katz, Hall, and Eichengreen (1992)) and has decreased over time to reach low levels in the 2000s (Partridge, Rickman, Rose, and Kamar (2012)). On the costs of changing occupations, see Kambourov and Manovskii (2009), Sullivan (2010) or Artuç and McLaren (2012).

[2]Metropolitan Statistical Area: Detroit-Warren-Livonia, industry: "Motor vehicle parts manufacturing" (NAICS 3363), source: County Business Patterns of the U.S. Census.

[3]Electrical engineers are occupation 17-2071 in the Standard Occupational Classification of the Bureau of Labor Statistics. The industry is NAICS 3345: "Navigational, measuring, electromedical and control instruments manufacturing". The figures are for 2003.

for a given regional industry composition, the share of an industry in the region's labour force should be negatively related to its capacity to respond to aggregate shocks. For example, if an industry employs a large fraction of a region's labour force, it will find it more difficult to expand as there are only relatively few workers it can attract from other industries. This directly results from the geographical immobility of labour. Second, for a given share of employment in the region's labour force, an industry should find it easier to expand if other industries in the region use a similar mix of occupations.

The main prediction of our model is that an industry which faces a positive (negative) shock at the national level should expand (contract) its employment more in regions where the value of our responsiveness index is high, as the industry finds it easier to recruit (shed) labour[4]. To clarify our insight on the importance of occupations, consider two regional statistical areas in the U.S.: Grants Pass in Oregon and San Jose in California. In both regions, the manufacture of measuring instruments accounts for about 2.5% of total employment. Apart from this industry, Grants Pass is very reliant on the health care sector, tourism and on the metal industry. San Jose on the other hand, where the Silicon Valley is located, has a substantial employment in other industries which also employ many electrical engineers (e.g. the computer industry). In such a case, our model predicts that the manufacture of measuring instruments can respond to aggregate shocks more easily in San Jose than in Grants Pass as it has a access to a larger pool of electrical engineers. Our index combines this reasoning for each occupation that an industry uses to calculate an industry-region specific measure of the responsiveness of employment to national shocks.

The mechanism behind our theory implies that employment growth in a given region-industry pair does not only depend on national shocks in that particular industry, but also on the shocks to all other industries. We compute a measure of the impact of national shock to industry j on the growth of employment of industry $i \neq j$ in a particular region. In terms of our example, we expect that a boom in the national demand for computers will substantially reduce employment in the manufacturing of measuring instruments in

[4]For contracting industries, employment decreases more if employees can easily find a job in another industry in the same region and occupation, i.e. if our index is large. If they lack good outside options, employees might be more willing to take wage cuts than be fired, making employment less reactive to negative shocks. We elabourate on the difference between expanding and contracting industries in section 4.2.

San Jose, as electrical engineers flock to the neighboring booming computer industry[5]. In Grants Pass on the other hand, the virtual absence of the computer industry suggests that employment in the manufacturing of measuring instruments will be insensitive to the good fortune of the computer industry at the national level.

We assess the importance of the spatial and occupational frictions at two different levels[6].

First, we test our model empirically by exploiting the variation in employment growth between different U.S. regions within an industry. We combine data on occupations from the U.S. Bureau of Labor Statistics and on employment from the County Business Patterns of the Census Bureau and observe (i) the size of each 4-digit industry in each metropolitan and micropolitan state area (MSA) between 2003 and 2008 and (ii) the relative use of different occupations in different industries. We capture the nationwide shock to an industry by its national employment growth, and interact it with our measure of employment responsiveness in an MSA-industry pair. The interaction is a positive, significant and robust determinant of the short-run employment growth of an MSA-industry pair.[7] This is in line with our model: within an industry, employment responds more to national shocks in MSAs where our index of employment responsiveness is larger. We also show that the cross-industry effects are important determinants of short-run employment changes.

Second, we use our model to assess the degree to which U.S. workers are "specific" to an industry. An important strand of the international trade literature examines the distributional consequences of trade when factors of production differ in their ability to switch industry. In our model, the degree to which a worker is specific to an industry depends on the presence of other industries using his occupation in his region. The more other industries use his occupation, the easier it is for a worker to switch industry and the

[5]There might of course be other channels linking industries, such as input-output linkages which we discuss further and control for in section 4.2.

[6]Other types of frictions, such as search and matching, may also be important determinants of the short-run reaction of employment to shocks. These seem however insufficient to explain the slow short-run adjustment to trade liberalisation in a country like Brazil (Cosar (2013)). In practice, our geographic and occupational immobility may interact with search and matching in creating a region-occupation specific labour market, where the number of vacancies for an occupation depends on a region's industry composition.

[7]Our results are robust to excluding the employment of the state where the MSA is located while computing the national employment growth of the industry.

less specific he is[8] to his industry. Aggregating across all regions and occupations used by an industry, we determine the average specificity of workers in an industry at the national level. The calibration of our model to the U.S. economy shows that workers in agricultural sectors and in textile manufacturing are on average most specific to their industry, and are therefore likely to be hit most by a negative shock to the price of their industry's output[9].

Finally, our paper contributes to the regional science literature in two distinct ways. First, we relate to the literature mapping national shocks to regional labour market outcomes. Our strategy is related to Blanchard, Katz, Hall, and Eichengreen (1992) or Bound and Holzer (2000), in that we map national employment growth to its regional counterparts. In contrast to them however, we highlight the importance and the effectiveness of the occupational dimension in this mapping. A fast-growing literature also uses variation across local labour markets to identify the effects of trade shocks on labour market outcomes (Chiquiar (2008), Autor, Dorn, and Hanson (2012), Kovak (2013), McLaren and Hakobyan (2012), Topalova (2010) among others) but remains silent about the role of occupations. One notable exception, Ebenstein, Harrison, McMillan, and Phillips (2011), shows that wages in occupations more exposed to international trade are lower, which highlights the importance of occupational immobility. Our contribution is to develop a theory-based measure of short-run employment frictions, which also incorporates the multilateral relationships among industries, using spatial variation in industry specialisation and variation in the occupational mix across industries. Second, we shed a new light on the relevance of the Marshallian argument for labour market pooling. Recent studies (e.g. Ellison, Glaeser, and Kerr (2010)) have shown that industries using similar occupations tend to collocate in space. We take a different perspective and ask whether one of the main arguments behind labour market pooling, which is that an industry's employment can better adapt to shocks if it is located close to the pool of skills it needs (Overman and

[8]Following the seminal paper of Mussa (1974), we think of specificity not only as a technological concept capturing whether a factor is needed in the production function of different industries, but also as an economic concept reflecting the relative size of industries in which a factor can be used.

[9]In the appendix, we provide an additional perspective on the aggregate consequences of our model for the U.S.. We consider a large shock to the U.S. terms of trade - the large increase in the price of tradables over the period 2005-2008 - and compare the growth rate of GDP predicted by our model to the growth rate predicted by a model with no frictions. Assuming no change in the price of non-tradables, we find that the geographic and occupational frictions cost 1.15 percentage points of GDP growth, thereby halving the growth rate which would obtain with no frictions.

Puga (2010)), indeed occurs in practice. Our affirmative answer confirms that particular rationale for labour market pooling.

Apart from the aforementioned papers, we also relate to the the large literature studying the impact of external shocks on labour reallocation between industries. In developing countries, the sectoral allocation of labour does not seem to respond much to trade shocks (Wacziarg and Wallack (2004), Godlberg and Pavcnik (2007), Kambourov (2009), Topalova (2010)), suggesting that an important fraction of the gains from trade are not reaped by developing countries. In the U.S., an early literature suggests that employment in an industry is moderately reactive to changes in import penetration (Grossman (1986), Freeman and Katz (1991), Revenga (1992) and Gaston and Trefler (1997)). We contribute to this literature by incorporating occupational immobility as a source of short-run employment frictions to aggregate shocks and provide a richer set of predictions on the heterogeneous reaction of regions and industries[10].

The paper is structured as follows. Section 2 develops the model and derives the computation of our measure of responsiveness. Section 3 describes the empirical strategy and section 4 presents the results. Section 5 computes the measures of worker specificity implied by our model using U.S. data. Section 6 concludes.

2 The model

2.1 The setup

The economy consists of a mass one of workers and I goods, each produced by a different industry. Labor is the only factor of production in the economy. We consider a national economy divided in N regions. Each region is a local labour market in the sense that workers cannot migrate between regions. Goods markets are however integrated and the price of a good is identical in all regions. We think of the regions as small open economies, which take the price of each good as given.

Each industry consists of a large number of firms, which produce a homogeneous good

[10] Another strand of the literature examines the reallocation of employment between firms within an industry after a trade liberalisation, see Revenga (1997), Trefler (2004) or Menezes-Filho and Muendler (2011).

and behave in a perfectly competitive manner. In each industry, production requires the use of a set of occupations (e.g. cook, accountant or chemical engineer), combined in proportions which are specific to the industry. We therefore allow industries to differ in the intensity with which they use different occupations. The production function of industry i is given by:

$$y_i = \left[\sum_o \alpha_{io}^{\frac{1}{\epsilon}} \Lambda_o^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon-1}} \tag{1}$$

where $\epsilon > 0$, o stands for occupations and Λ_o is the number of units of effective labour in occupation o.

Each worker in the economy inelastically supplies one unit of labour of one of the occupations. We assume that workers cannot choose the occupation in which they are active, and have no possibility to change occupation[11]. Since workers are immobile between occupations and regions, the mass of workers in occupation o and region r is exogenous and given by L_{or}. Workers in an occupation differ in terms of productivity. Each worker independently draws a productivity parameter z for each industry from a Fréchet distribution:

$$F(z) = e^{-z^{-\nu}}. \tag{2}$$

Worker h in region r faces a vector $\{z_{hi}\}_{i \in I}$, summarizing the number of effective labour units he can provide in each industry[12]. The parameter $\nu > 0$ affects the heterogeneity of productivity draws between industries and captures the degree to which workers are industry-specific. For small ν, a worker typically has very different draws of productivity in different industries, and the percentage loss in productivity incurred by changing industry is large. For a large ν, on the other hand, the productivity draws of a worker in different industries are relatively close to each other. In this case, changing industry does not typically result in a large change of productivity.

The assumption that workers are tied to an occupation and that they can move between industries by incurring a productivity loss offers a stylised representation of the evidence that (i) there are substantial costs of changing industry (Lee and Wolpin (2006), Artuç, Chaudhuri, and McLaren (2010)) due to the loss of industry specific human capital, and

[11]The Theory Appendix 7.3.3 relaxes this assumption.

[12]The use of Fréchet distribution in modeling the heterogeneity in productivity has been popularised by Eaton and Kortum (2002) in the trade context. For the details on its use in models of industry choice, see Hsieh, Hurst, Jones, and Klenow (2013) and Vannoorenberghe and Janeba (2013).

that (ii) the costs of changing occupation are at least as large (Sullivan (2010), Artuç and McLaren (2012), Kambourov and Manovskii (2009)).

2.2 Equilibrium

Firms in industry i located in region r take the price of good i (p_i) as given and maximise their profits, given by:

$$\max_{\{\Lambda_{ior}\}_{o\in O}} p_i y_{ir} - \sum_o w_{ior}\Lambda_{ior} \tag{3}$$

where w_{ior} is the wage paid per unit of effective labour to occupation o in the industry-region pair ir. The first order condition of the maximisation problem can be rearranged to show that:

$$\Lambda_{ior} = \alpha_{io} w_{ior}^{-\epsilon} p_i^{\epsilon} y_{ir} \tag{4}$$

where Λ_{ior} denotes the demand for effective units of labour in industry i, occupation o and region r. Plugging (4) in (1) shows that in equilibrium, if industry i produces in region r, the price p_i must be equal to the marginal costs of production:

$$p_i = \left[\sum_o \alpha_{io} w_{ior}^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}. \tag{5}$$

A worker h in occupation o observes his vector of productivity in all industries $\{z_{hi}\}_{i\in I}$, as well as the wage per effective unit of labour paid in each of the industries for his occupation $\{w_{ior}\}_{i\in I}$. Based on this information, he decides to work in the industry which gives him the highest income $z_{hi}w_{ior}$. As shown in the appendix, the number of workers choosing industry i in region r and the effective labour this corresponds to are:

$$L_{ior} = \frac{w_{ior}^{\nu}}{\sum_{j\in I} w_{jor}^{\nu}} L_{or} \tag{6}$$

$$\Lambda_{ior} = \Delta w_{ior}^{\nu-1} \left(\sum_{j\in I} w_{jor}^{\nu}\right)^{\frac{1-\nu}{\nu}} L_{or} \tag{7}$$

where $\Delta \equiv \Gamma\left(1 - \frac{1}{\nu}\right)$ and $\Gamma()$ is the gamma function. We assume that $\nu > 1$ for the rest of the analysis. Equations (6) and (7) are respectively the labour supply and the supply of effective labour in occupation o in a particular industry-region pair. Both are increasing in the wage paid by that industry and are decreasing in the average wage paid by the other industries using occupation o in the region. The extent to which the labour supply in a particular occupation reacts to the wage differential between industries depends on ν, which indexes the degree of mobility of workers between industries. The larger the ν, the less important the worker-specific productivity differences between industries and the more workers react to wage differentials between industries. Equations (6) and (7) further show that the supply of labour in any region-industry pair is positive for any $w_{ior} > 0$. This property guarantees that each industry produces a positive amount in each region.

Equating the demand and supply of effective labour in each ior tuple (given by (4) and (7) respectively), we show in the Theory Appendix that:

$$w_{ior} = (\alpha_{io}\zeta_{ir})^{\frac{1}{\nu+\epsilon-1}} (\Delta L_{or})^{-\frac{1}{\epsilon}} \left(\sum_{j \in I} (\alpha_{jo}\zeta_{jr})^{\Omega} \right)^{\frac{\nu-1}{\epsilon\nu}}, \tag{8}$$

where $\zeta_{ir} \equiv p_i^{\epsilon} y_{ir}$ and $\Omega \equiv \nu/(\nu + \epsilon - 1)$. For expositional convenience, and although it is only correct if $\epsilon = 1$, we will refer to ζ_{ir} as the value of production of the industry in region r. The wage per effective labour unit in the ior tuple is (i) increasing in the demand for the occupation in ir, determined by the parameter α_{io} and by the value of production of industry i in region r (ζ_{ir}), (ii) decreasing in the supply of occupation o in region r (L_{or}), and (iii) increasing in the "outside option" of workers, which depends on the demand for their occupation in other industries. When ν is large, workers can easily move between industries and the wage in i becomes very sensitive to the outside option and insensitive to the demand conditions and technology parameters in the own industry.

Plugging the equilibrium condition (8) for w_{ior} in (5) and in (6), we obtain respectively:

$$\zeta_{ir}^{1-\Omega} = \Delta^{\frac{\epsilon-1}{\epsilon}} p_i^{\epsilon-1} \left[\sum_o \alpha_{io}^{\Omega} L_{or}^{\frac{\epsilon-1}{\epsilon}} \left(\sum_{j \in I} (\alpha_{jo}\zeta_{jr})^{\Omega} \right)^{\frac{1-\nu}{\nu}\frac{\epsilon-1}{\epsilon}} \right], \tag{9}$$

$$L_{ir} = \sum_o L_{ior} = \sum_o \frac{(\alpha_{io}\zeta_{ir})^{\Omega}}{\sum_{j \in I} (\alpha_{jo}\zeta_{jr})^{\Omega}} L_{or}. \tag{10}$$

Equations (9) and (10) are the two key relationships of interest in the model. They allow us to pin down the employment in a particular industry-region pair as a function of the exogenous parameters of the model. The first of these two conditions, in equation (9), establishes how the value of production in each industry-region pair depends on the value of production of other sectors in the region, on the exogenous price vector $\{p_i\}_{i \in I}$ and on exogenous region and industry characteristics (α_{io} and L_{or}). The second, equation (10), shows how the vector of $\{\zeta_{ir}\}_{i \in I}$ in a region maps to the number of employees in each industry-occupation pair in that region.

2.3 Comparative statics

We now perform a comparative statics exercise on the two relationships (9) and (10) to determine how exogenous changes to the prices of particular goods affects the employment in each industry-region pair, a relationship which is at the core of our empirical analysis. The present section shows the main results of the comparative statics exercise, the details of which can be found in the Theory Appendix 7.3.2.

Totally differentiating (10) gives:

$$\hat{L}_{ir} = \Omega \left[\hat{\zeta}_{ir} \underbrace{\left(\sum_o \frac{L_{ior}}{L_{ir}} \frac{L_{-ior}}{L_{or}} \right)}_{S^e_{ir}} - \sum_{m \neq i} \hat{\zeta}_{mr} \underbrace{\left(\sum_o \frac{L_{ior}}{L_{ir}} \frac{L_{mor}}{L_{or}} \right)}_{S^e_{imr}} \right] \tag{11}$$

where we use ˆ to denote percentage changes in variables and where L_{-ior} refers to the number of workers in occupation o and region r who are employed in all industries other than i. The above equation shows the effect of a change in the value of production of all industries in r ($\hat{\zeta}_{ir}$) on the employment of a particular industry i in r. Since we assume perfect competition, an increase in ζ_{ir} must be reflected in a combination of higher employment and/or higher wages in industry i. Equation (11) shows the extent to which employment reacts to such a change. The marginal effect of $\hat{\zeta}_{ir}$ on $\hat{L}_{ir}$ positively depends on Ω and S^e_{ir}, which both affect the "ease" with which industry i can recruit the workers it needs to expand. First, a higher Ω reflects that ν is large, meaning that workers within an occupation are very mobile between industries. This ensures that a small increase in w_{ior} induces many workers of that occupation to join the industry. Second, S^e_{ir} ($0 \leq S^e_{ir} \leq 1$) is an index which depends both on (i) the share of the industry's employment in the region's total employment - a relatively small industry should find it easier to attract additional workers - and (ii) on the similarity between the occupations used by the industry and the occupations used by other industries in the region[13]. The index S^e_{ir} therefore captures the

[13] This decomposition can be seen formally by rewriting: $S^e_{ir} = \left(\sum_o \frac{L_{ior}}{L_{ir}} \frac{L_{-ior}}{L_{-ir}} \frac{L_r}{L_{or}} \right) \frac{L_{-ir}}{L_r}$. The bracket is the correlation between the share of occupation o in industry i and the share of occupation o in all other industries, weighted by the inverse of the share of occupation o in the total regional employment. The weighting reflects the fact that occupations in short supply in the region are particularly constraining

similarity in the use of occupations between industry i and other industries in r, where the weights of each occupation are determined by their share of the industry's *employment*. The same intuition applies for an expansion in the value of another industry, denoted j. To expand, it will draw labour away from industry i, the more so the stronger the similarity of the occupation between i and j, and the more mobile are workers between industries.

Note that the same reasoning applies to the case where the value of industry i's production contracts (e.g. $\hat{\zeta}_{ir} < 0$). In this case, employment of industry i (L_{ior}) should decrease relatively more if workers can easily move to other industries, i.e. if ν is large, if the occupational intensity of other industries is similar to i's or if other industries are relatively large compared to i. If labour cannot easily be reallocated to other industries on the other hand, the model predicts that wages should take the bulk part of the adjustment. In reality, workers may also become unemployed if their industry contracts. Unemployment is a way to shed labour independently of the mobility of workers between industries, and the index S^e_{ir} may thus be of lesser importance for contracting than for expanding industries. However, we expect that workers should be more willing to accept wage cuts if they have less outside options, thereby making employment less sensitive to reductions in demand. In this light, we expect that the index S^e_{ir} should affect the extent to which employment responds to the growth rate of the industry even for contracting industries.

We now turn to the determination of the vector $\{\hat{\zeta}_{ir}\}_{i\in I}$ in region r as a function of exogenous changes to the price vector. Totally differentiating (9) gives:

$$\hat{\zeta}_{ir}\left[1+\frac{\nu-1}{\epsilon}\underbrace{\left(\sum_{o}\omega_{ior}\frac{L_{ior}}{L_{or}}\right)}_{S^c_{iir}}\right]=(\nu+\epsilon-1)\hat{p}_i-\frac{\nu-1}{\epsilon}\left[\sum_{m\neq i}\hat{\zeta}_{mr}\underbrace{\left(\sum_{o}\omega_{ior}\frac{L_{mor}}{L_{or}}\right)}_{S^c_{imr}}\right] \quad (12)$$

where $\omega_{ior} \equiv w_{ior}\Lambda_{ior}/(p_i y_{ir})$ is the cost share of occupation o in the industry-region pair ir, with $\sum_{i\in I}\omega_{ior} = 1$ from perfect competition. The index S^c_{imr} ($0 \le S^c_{imr} \le 1$) is similar to S^e_{imr} in that it captures the *similarity* in occupational use between industries i and m in region r, with the only difference that the weights are not based on the employment share of occupation o in industry i, but on its *cost share*. It shows that the value of

for an industry's expansion and therefore carry a higher weight. The fraction L_{-ir}/L_r on the other hand captures the fact that a relatively small industry should find it relatively easier to expand.

production of industry i is more reactive to changes in p_i if it can easily recruit workers from other industries in occupations which account for a large share of its costs. On the other hand, an expansion of the value of other industries tends to reduce the value of industry i. Industry i is particularly sensitive to the growth of industry m if m uses intensively occupations which account for a large share of i's costs.

Equation (12) must hold for each industry in a region, thereby establishing a system of I linear equations in I unknowns. Solving this system allows expressing each $\hat{\zeta}_{ir}$ as a linear combination of the price changes in all industries. By (11), it also implies that changes in sectoral employment at the regional level can be expressed as a linear combination of the change in industry prices. We denote the vector of employment growth in region r as $\hat{\boldsymbol{L}}_r \equiv \{\hat{L}_{ir}\}_{i\in I}$ and the vector of price growth as $\hat{\boldsymbol{p}} \equiv \{p_i\}_{i\in I}$. Furthermore, we define $\boldsymbol{S}_r^e$ and $\boldsymbol{S}_r^c$ as the respective matrices of S_{imr}^e and S_{imr}^c where i refers to the rows and m to the columns of the matrix. Since $S_{ir}^e = 1 - S_{iir}^e$, (11) and (12) can be combined to give:

$$\hat{\boldsymbol{L}}_r = \underbrace{\nu\,(\boldsymbol{I} - \boldsymbol{S}_r^e)\left(\boldsymbol{I} + \frac{\nu-1}{\epsilon}\boldsymbol{S}_r^c\right)^{-1}}_{\boldsymbol{E}_r}\hat{\boldsymbol{p}}. \tag{13}$$

The effect of price changes on regional employment in different industries is governed by the matrix $\boldsymbol{E}_r$, which captures the (scaled by ν) own and cross price *elasticity* of employment in region r. $\boldsymbol{E}_r$, which will be at the core of our empirical analysis, combines the mechanisms behind the two relationships (11) and (12). An increase in the price of a good raises the value of production in a region - the more so the more easily the industry can recruit the workers it needs (captured by $(\boldsymbol{I} + (\nu-1)/\epsilon\boldsymbol{S}_r^c)^{-1}$ - the index based on cost shares). A given increase in the value of production further translates into more employment in regions where the industry finds it easier to recruit workers (captured by $\boldsymbol{I} - \boldsymbol{S}_r^e$ - the index based on employment shares).

2.4 National growth of industries

By definition, the national growth of industry i ($\hat{L}_i$) is a weighted sum of the regional growth rates of that industry:

$$\hat{L}_i = \sum_r \chi_{ir}\hat{L}_{ir} \tag{14}$$

where $\chi_{ir} = L_{ir}/L_i$ represents the share of region r in the national employment of industry i. We denote $\boldsymbol{\chi}_r$ as the vector of χ_{ir} in region r and $\hat{\boldsymbol{L}}$ as the vector of national employment growth. Combining (13) and (14) yields:

$$\hat{\boldsymbol{L}} = \nu \underbrace{\left(\sum_r \boldsymbol{\chi}_r \circ \boldsymbol{E}_r \right)}_{\boldsymbol{E}} \hat{\boldsymbol{p}} \tag{15}$$

where ∘ is the Hadamard product of the two vectors (element by element multiplication). Equation (15) shows the matrix of price elasticity of employment at the national level. It proves that (i) the national growth rate of an industry's employment is a weighted sum of the growth rate of prices in all industries, (ii) an industry responds more to an aggregate shock in its own price if a larger share of its employment is located in regions where the employment elasticity is high. The national price elasticity of employment is a weighted sum of its regional counterparts.

3 Empirics

Our model predicts that the immobility of labour between regions and between occupations constitute two sources of frictions hampering the short-run responsiveness of industry-specific employment. In particular, we predict that an industry's employment will react more strongly to price shocks in regions where our flexibility index is larger, i.e in regions where the industry (i) accounts for a small share of regional employment and (ii) is close to neighboring industries in terms of occupational mix. Following these insights, we test our model using the cross-regional variation in employment growth within an industry.

In addition to being a natural choice considering the structure of our model (see (13)), using region-industry pairs as our unit of observation also provides a solution to the "degrees of freedom problem" which would plague an analysis using solely cross-industry variation in employment growth at the national level (i.e. an analysis based on an empirical counterpart to (14)). This problem has been recognised in recent years, and we follow a growing literature using regional variation to test the effect of nationwide shocks[14].

[14] See Chiquiar (2008), Hanson (2007), Topalova (2007), Topalova (2010), Kovak (2013) and Autor, Dorn, and Hanson (2012).

3.1 Empirical strategy

To test our model, we could use changes in output prices at the national level as our primitive shocks, predict the regional employment growth of an industry using equation (13) and test if the predicted value is in line with its observed counterpart in the data. Using price shocks is however problematic for three reasons. First, obtaining reliable price data for detailed industries is difficult. The lack of reliable prices for many tradable goods has been recognised in the literature (Autor, Dorn, and Hanson (2012)), and data on the prices of non-tradables are even more problematic. Second, an increase in prices can reflect either a decrease in U.S. productivity or an increase in U.S. demand, with opposite consequences for employment in the industry when the demand elasticity is larger than one. This is a particular concern for non-tradable goods, for which price changes are less likely to come from external forces to the U.S. economy. Third, the adjustment of employment to price changes can be sluggish in the presence of additional sources of frictions such as labour regulations, unionisation or search and matching, making it difficult to design an appropriate lag structure for our regression equation.

To circumvent these problems, we show that our model predicts a close connection between the national and the regional employment growth of industries[15], as evident from combining (13) and (15):

$$\hat{\boldsymbol{L}}_r = \boldsymbol{E}_r \left(\sum_r \boldsymbol{\chi}_r \boldsymbol{E}_r \right)^{-1} \hat{\boldsymbol{L}} = \boldsymbol{R}_r \hat{\boldsymbol{L}} \tag{16}$$

where $\boldsymbol{R}_r$ is the regional matrix of employment *responsiveness* to national employment growth that derives from the theory. It maps the vector of national employment growth in all industries to its regional counterpart in r and relates observable outcomes between which a contemporaneous relationship is likely to hold[16]. The diagonal entries of $\boldsymbol{R}_r$ are positive while the off-diagonal entries are typically negative, reflecting the competition for occupations between industries.

[15] We address at the end of this section the issue that regional and national growth are mechanically related since the second is a weighted average of the first over all regions.

[16] Using national employment changes to explain regional employment growth dates back to Blanchard, Katz, Hall, and Eichengreen (1992) or Bound and Holzer (2000). Blanchard, Katz, Hall, and Eichengreen (1992) study the heterogeneous response of states to national employment shocks and find evidence of a contemporaneous relationship. Lagged and led national growth are insignificant.

As shown in (16) and as discussed in section 2, our model predicts that the growth in national employment of industry i not only affects the regional growth of i (the "own-industry effect"), but also the regional growth of all industries $j \neq i$ (the "cross-industry effect"). In the following, we decompose the predicted growth of employment in industry i and region r into the effect of industry $i's$ national growth ($\hat{L}_{irt}^{own}$) and the effect of the national growth of all other industries ($\hat{L}_{irt}^{cross}$):

$$\hat{L}_{irt} = \underbrace{R_{iir}\hat{L}_{it}}_{\hat{L}_{irt}^{own}} + \underbrace{\sum_{j \neq i} R_{ijr}\hat{L}_{jt}}_{\hat{L}_{irt}^{cross}}, \tag{17}$$

where $\hat{L}_{irt}$ is the employment growth of industry i in region r in year t. We include the two components $\hat{L}_{irt}^{own}$ and $\hat{L}_{irt}^{cross}$ separately to allow for the own and cross-industry effects to have different explanatory power. As an empirical counterpart to (17), we use:

$$\hat{L}_{irt} = \beta_0 + \beta_1 \hat{L}_{irt}^{own} + \beta_2 R_{iir} + \beta_3 \hat{L}_{it} + \beta_4 \hat{L}_{irt}^{cross} + \boldsymbol{\gamma} \boldsymbol{X}_{irt} + \boldsymbol{\theta} \boldsymbol{X}_{rt} + \alpha_i + \alpha_r + \alpha_t + \varepsilon_{irt}, \tag{18}$$

where we discuss in the following our choice of controls. The coefficients of interest, β_1 and β_4, capture the average comovement between the actual employment growth and the two components of predicted growth: the own-industry and the cross-industry effects. We expect β_1 and β_4 to be positive. We include an industry dummy (α_i) to identify the cross-regional variation in employment growth and time fixed effects (α_t) to control for macroeconomic shocks common to all region-industry pairs. To control for the time-invariant heterogeneity in employment growth which may be correlated with the employment responsiveness across regions, we also include region (MSA) fixed effects (α_r). We replicate our analysis by controlling for time-varying industry specific effects ($\alpha_i * \alpha_t$) and present the corresponding results, which are in line with our baseline findings, in the Appendix 7.2.2. $\boldsymbol{X}_{irt}$ and $\boldsymbol{X}_{rt}$ include additional controls, which differ across specifications.

Controlling directly for our measure of responsiveness R_{iir} captures a possible source of bias in our estimation. Industries using similar occupations to neighboring industries may benefit from "thick labour market externalities" and see their productivity and employment grow more over time. This mechanism, in line with Marshallian externalities, would cause a positive correlation between an unobserved factor raising L_{irt} and the ease with which an industry can expand or shed labour (R_{iir}), thereby biasing our estimates[17].

[17]This argument is closely linked to the issue that industry location in the U.S. is endogenous, and that

Including R_{iir} in our estimating equation should to some extent control for this bias. To ensure that there is no issue of reverse causality from $\hat{L}_{irt}$ to our measure of responsiveness, we use a beginning of sample measure of R_{iir}, which is unlikely to be affected by subsequent region-industry shocks to employment. We assess the importance of including R_{iir} in (18) by replicating our analysis without controlling for R_{iir} and present the results, which yields larger estimates of β_1 and β_4, in the Empirical Appendix.

Our identification strategy also requires that unobserved regional shocks to an industry's employment are uncorrelated with the national growth of that (or any other) industry. Since national employment growth is a weighted sum of regional employment growth rates, this assumption appears mechanically violated. We approach this issue in two different ways. First, if a region only employs a small fraction of the national workforce of an industry, its effect on the national employment growth should be negligible. Since, in our sample, almost 93 percent of observations refer to region-industry pairs employing less than 1 percent of the national industry employment, the small industry assumption should be a valid approximation[18], giving us some confidence that national shocks can be considered exogenous from the perspective of such region-industry pairs. Second, we replicate all the analysis using the changes in the national employment net of the state employment in which an MSA-industry pair is located as the aggregate shock. This method ensures that the predicted employment growth for an MSA-industry does not incorporate the MSA or state level shocks. None of our qualitative results are affected by this alternative.

3.2 Data

To construct our matrix $\boldsymbol{R}_r$ in (16), which maps national to regional employment growth, we combine beginning of sample data on (i) the share of each industry's employment and total wage bill accounted for by each occupation at the national level (L_{io}/L_i and ω_{io}) with (ii) the industry employment at the regional level taken from the County Business Patterns (CBS) database of the U.S. Census Bureau.[19]

---

industries are likely to relocate towards regions with a large R_{iir} due to Marshallian externalities. This would create a spurious correlation between R_{iir} and $\hat{L}_{irt}$ which we address by controlling for R_{iir}.

[18]This observation also points out the importance of defining regional units as precisely as possible while investigating the regional evolutions given the national changes.

[19]While constructing $\boldsymbol{R}_r$, we implicitly assume that the employment and wage shares of occupations within an industry are constant across regions. These assumptions, motivated by data availability, are not

Employment data at the industry level are taken from the County Business Patterns (CBS) database of the U.S. Census Bureau. We use employment data on 4-digit NAICS industries in Micro- and Metropolitan Statistical Areas (henceforth, MSAs) between 2003 and 2008. Using MSAs rather than counties as our regional unit of observation has two important advantages. First, an MSA is defined as a collection of geographically close counties between which labour mobility is high whereas mobility across MSAs is relatively low. MSAs are therefore closer to the economic meaning of a region in our model than counties, between which there may be large short-run migrations. Second, employment data of county-industry pairs show a very large number of missing observation, due to imprecise estimations or to privacy reasons. Using MSAs, which are larger than counties, mitigates that concern. Even at the level of MSAs, however, industry-specific employment is not reported in many instances. For all missing observations, the CBP reports an approximate firm size distribution in the MSA-industry pair, with an upper and lower bound for employment. We use that information to reconstruct the MSA-industry employment in each year as explained in the data appendix. We show in section 4.2 that our results are not driven by the particular procedure in which we construct these approximations. In all our empirical exercises, we exclude the MSA-industry pairs employing less than 100 workers, which are quite sensitive to mis-measurement or idiosyncratic shocks.

The time span of the empirical analysis is dictated by data comparability issues and the abrupt changes in macroeconomic conditions. Since the borders of MSAs change after every census (with a 3 year delay), we start our analysis from the last change in 2003 and stop it in 2008 due to the recent economic downturn. We also exclude outliers for the dependent variables and the variables of interest in all estimations presented below by trimming the lower and upper 1 percentile. None of our results depends on that particular threshold.

Data on occupations are taken from the beginning of sample version of the Occupational Employment Statistics (OES) of the Bureau of Labor Statistics. Occupations are defined at the 6-digit level of the standard occupational classification system (e.g. "economists" or "computer programmers"). The OES reports the share of the national

---

mutually consistent: a constant employment share is consistent with a Leontieff production function while a constant cost share is consistent with Cobb Douglas. Nevertheless, our model suggests a way to solve this issue which we discuss in detail in Section 4.2.3.

employment of an industry accounted for by each occupation (L_{io}/L_i) as well as their share of the national wage bill of the industry[20].

Finally, the computation of the matrix $\boldsymbol{R}_r$ requires making an assumption about the value of the ratio of parameters $(\nu - 1)/\epsilon$ (see (13)). Due to the lack of existing literature on the parameter ν, we decide to set this ratio to 1 in our baseline exercises[21] and test the sensitivity of our results to different values of this ratio in section 4.2. Further details of the exact computations of all variables, sources for additional controls and descriptive statistics of relevant variables are available in the Data Appendix.

After presenting the results of our baseline specification (18) in Section 4.1, we turn to a number of robustness tests to check the sensitivity of our results to the presence of alternative explanations, modeling assumptions and treatments of the data in Section 4.2.

4 Results

4.1 Main Results

Table 1 presents estimates of the determinants of the MSA-industry employment growth and shows the performance of our measure in explaining the cross-regional variation in industry employment growth given the national employment changes. Each column of Table 1 shows the estimation of a different version of equation (18), with all standard errors clustered at the industry-year level. As shown in Table 1, $\hat{L}_{irt}^{own}$ is strongly significant with a positive sign in all specifications. This finding implies that spatial variation in industry mix and the closeness of industries on the occupational space are successful in projecting the national shocks onto regional economic units. In our preferred specification (Column 5), the point estimate of β_1 is 0.914. This indicates that there is almost one-to-one relationship between the observed employment change of an MSA-industry and the one

[20] When labour is the only input and under perfect competition as we assumed, total industry output is equal to the total wage bill of the industry which justifies the computation of ω_{io} using total wage bill instead of total output. We test the robustness of our results to the labour share of industries by replicating our analysis for the industries with low and high labour share separately. All our qualitative results are robust to this additional control.

[21] In a contemporaneous work, Hsieh, Hurst, Jones, and Klenow (2013) arrive at a similar value using data from Decennial Censuses from 1960 to 2000 and 2006-2008 American Community Surveys.

predicted by our model in response to a national employment shock in this particular industry. Consider two MSA-industries which are at the 25th and 75th percentile of the R_{iir} distribution. When the industry employment changes 1 percent at the national level, the latter responds 0.97 percent while the former responds only 0.86 percent. The point estimate for the cross-industry effects, β_4, is 0.220 and significant at 1 percent. This shows that occupational similarity is important in predicting the impact of employment changes in other industries. In sum, we find strong evidence that the employment in the MSA-industries located closer to similar MSA-industries in terms of their occupational mix respond more to aggregate employment shocks.

In addition to our main finding, we observe that initial size is a strong predictor of the sub-sequent employment growth of an MSA-industry as initially larger industries grow significantly slower. The national employment growth is insignificant, which reflects the inability of aggregate employment shocks to explain the short-run changes in regional employment when the regional units are relatively small. Lastly, the cross-industry effects are very sensitive to the control for regional heterogeneity. Once we control for the different sources of time-varying heterogeneity in MSA-level employment growth, they appear to be important determinants of industry growth. We will elabourate on this point when discussing the sensitivity of our benchmark results to the modeling assumptions.

4.2 Robustness Tests

In this part, we test the sensitivity of our main results on three dimensions. First, we check whether our result is robust to controlling for alternative explanations, which could give rise to an omitted variable bias. Second, we relax three modeling assumptions and allow for a limited degree of geographic mobility, occupational mobility and unemployment. Third, we consider the robustness of our results to different ways of treating the data[22].

[22] We also conduct numerous additional robustness tests. Among others, we use 2-year averages instead of yearly data, split the sample between industries with high and low labour shares, exclude 2008 to control for the recent economic downturn or use a 2-digit occupation definition instead of the 6-digit one. These results are not reported but are available upon request. None of our qualitative results are affected.

Table 1: Main results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$
Resp. x Nat. growth ($\hat{L}_{irt}^{own}$)	0.852***	0.853***	0.888***	0.640***	0.914***	0.973***	1.433***
	[4.43]	[4.42]	[4.50]	[3.59]	[5.33]	[5.73]	[6.24]
Resp. (R_{iir})	0.261***	0.263***	0.261***	0.170***	0.018	0.015	0.013
	[26.48]	[26.08]	[26.08]	[19.11]	[1.40]	[1.17]	[0.97]
Nat. growth ($\hat{L}_{it}$)	-0.034	-0.035	-0.033	0.204	-0.078	-0.146	-
	[-0.18]	[-0.18]	[-0.17]	[1.16]	[-0.46]	[-0.87]	-
Cross-ind. effect ($\hat{L}_{irt}^{cross}$)		0.140	0.090	-0.251***	0.220**	0.362***	0.221**
		[1.81]	[1.18]	[-3.53]	[2.66]	[4.15]	[2.65]
Log init. size ($ln(L_{ir,2003})$)				-0.045***	-0.076***	-0.076***	-0.076***
				[-38.38]	[-34.21]	[-34.33]	[-34.16]
Industry FE's	Yes	Yes	Yes	Yes	Yes	Yes	No
Year FE's	No	No	Yes	Yes	Yes	No	No
MSA FE's	No	No	No	No	Yes	No	Yes
MSA*Year FE's	No	No	No	No	No	Yes	No
Industry*Year FE's	No	No	No	No	No	No	Yes
N	356470	356470	356470	356470	356470	356470	356470
R^2	0.034	0.034	0.035	0.065	0.074	0.087	0.049

The dependent variable is $\hat{L}_{irt}$, the MSA-industry growth rate of employment. "Resp." is the MSA-industry specific responsiveness measure, given by the corresponding diagonal entry of the R_r matrix as defined in (16). Standard errors are heteroscedasticity robust and clustered at the industry*year level. t-statistics in brackets * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All regressions include a constant.

4.2.1 Alternative Explanations

Mean Reversion: A potential source of bias for our estimates may come from the existence of mean reversion in employment levels. Some economic forces apart from labour availability may prevent industries from growing too large or becoming too small in a particular region (for example, the availability of industry-specific amenities may be a concern). In such a case, if an industry accounts for a large (small) share of regional employment compared to the national average, we would expect employment to decrease (increase) in that particular region-industry pair, or to increase less (more) quickly than in other regions if the national shock is positive. Under a positive national shock to industry i, our model also predicts that the industry's employment will expand comparatively less in regions where industry i accounts for a large share of employment, as it struggles to find the labour needed to grow. Such a mean reversion may therefore bias our estimate of β_1 upward[23]. To control for mean reversion, we include both the relative size of an MSA-industry with respect to the total MSA employment in 2003 and its interaction with the national employment growth. The inclusion of the interaction term serves two purposes. First, it controls for the aforementioned bias arising from mean reversion. Second, it tests whether the second dimension of our mechanism - closeness to other industries on the occupational space - provides relevant information about the cross-regional variation in employment growth which can not be explained solely by the relative size of the MSA-industry. Column 1 of Table 2 reports the estimates for the corresponding specification and confirms the robustness of our result. Furthermore, the occupational dimension of our measure of responsiveness remains an important factor explaining the regional responses to national employment shocks. The interaction term $((L_{ir,2003}/L_{r,2003}) * \hat{L}_{it})$ is negative and significant at the 10% level, showing that the MSA-industry pairs, which are larger relative to regional employment, face larger frictions in adjusting their labour input and respond less to aggregate shocks. This finding is consistent with our model where large industries are less responsive to aggregate shocks due to labour scarcity.[24] The point

[23]Note however that the prediction coming from the mean reversion argument is opposite to that of our model if the national shock to industry i is negative. The mean reversion argument would make an industry contract more in regions where it accounts for a comparatively large share of employment, while we predict the opposite.

[24]The initial share of the industry appears with a positive sign which might seem counter-intuitive. Note that the log size of the industry together with the MSA fixed effects controls for the log share of the

estimate of β_1 decreases in comparison to column (6) of Table 1. This is intuitive in the sense that the control of the labour share interacted with the national employment change captures the effect of the first dimension (relative size) of our measure. The cross-industry effect is positive but only significant at the 10 percent level.

Input output linkages: An alternative explanation for our result is that the employment growth of an industry depends on the geographic proximity to industries supplying its intermediate goods and not to industries using similar occupations. Such Input-Output (IO) links between industries may bias our results to the extent that industries with strong IO links use similar occupations and that some intermediate goods are not perfectly mobile. In this case, the occupational similarity between industries, which determines our measure of responsiveness, may capture the IO links between industries, giving rise to an omitted variable bias. To investigate the importance of these links, we define a new variable which captures the presence of an industry's suppliers in its region:

$$a_{ir} = \sum_{j \neq i} D_{ji} \frac{L_{jr}}{L_{ir}} \tag{19}$$

where D_{ji} designates the input share of industry j in the total output of industry i. Hence, a_{ir}, is a weighted sum of the size of the input suppliers relative to the size of industry i in region r, where the weights are given by the national IO matrix. A larger value of a_{ir} implies that the input factors are relatively abundant for industry i in region r, so that an industry with a large value of a_{ir} may be more responsive to national employment shocks. We control for a_{ir} and $a_{ir} * \hat{L}_{it}$ in column 2 of Table 2 and show that our results are robust to controlling for IO links with neighboring industries. The proximity to input industries is insignificant as a determinant of employment growth of an industry. While being close to input suppliers is one of the most important determinants of industry agglomeration (e.g. Rosenthal and Strange (2001) and Ellison, Glaeser, and Kerr (2010)), our analysis shows that the effect of collocation with input suppliers on short-run employment changes is rather weak.

industry. These two findings taken together imply that the effect of initial employment share on subsequent employment growth is negative for small MSA-industries and the initial size-growth relationship follows a U-shaped pattern.

4.2.2 Modeling Assumptions

Geographic Mobility: Although the definition of MSAs and the recent literature (Partridge, Rickman, Rose, and Kamar (2012)) suggest that the assumption of short-run geographic immobility is appropriate, we here consider the impact of a violation of this assumption on our empirical results. Such a violation might be problematic for two reasons. First, our estimates may be biased if the migration of workers are correlated with our variables of interest, $\hat{L}_{irt}^{own}$ and $\hat{L}_{irt}^{cross}$ although the direction of the bias is not known a priori. Second, the labour stock within an MSA becomes less relevant in predicting the employment growth of industries, i.e. our theoretical index becomes an imprecise predictor of the true employment responsiveness. Although the estimates are expected to be biased towards zero in the presence of noisy measurement, we address this issue formally by controlling for the growth of employed labour in the region, which captures the effects of migration flows, but also of changes in labour force participation and unemployment[25]. Column 3 of Table 2 reports the results and shows that our findings are robust to this additional control. Furthermore, the cross-industry effects ($\hat{L}_{irt}^{cross}$) of the growth in other industries turn out to be a strong predictor of MSA-industry growth once we control for the growth of the regional labour stock while other coefficient estimates change marginally, which implies that $\frac{1}{|I_r|}\sum_{i\in I_r}\hat{L}_{irt}^{cross} \propto \frac{1}{\hat{L}_{rt}}$ where I_r is the set of industries active in region r. This relationship is intuitive: a region with a dispersed (on the occupational space) industry mix has a low $\hat{L}_{irt}^{cross}$ reflecting the limited interaction between dissimilar industries. In such a region, the changes in labour demand should be satisfied (or absorbed) by migration flows or adjustments in the labour force participation or unemployment rates which lead to larger time-varying regional employment shocks, $\hat{L}_{rt}$. Namely, there is a negative correlation between the mean of our measure of cross-industry effects and the time-varying employment shocks at the regional level. It is therefore essential to control for the time-varying regional employment shocks to fully understand the importance of cross-industry effects, a result which is already apparent in Column 6 of Table 1.

Occupational Mobility: Our response matrix ($\boldsymbol{R_r}$) relating national to regional employment growth is based on the assumption that there is no mobility between occupations. While a growing literature points to the substantial costs of switching occupations

[25] We discuss the effects of relaxing the full employment assumption from a different perspective in the following subsections.

(Kambourov and Manovskii (2009)), infinite costs of occupational mobility is a strong assumption. In the theory appendix 7.3.3, we extend our model to allow for some degree of occupational mobility, meaning that workers endogenously choose their occupation as a function of the relative wages offered by these different occupations. We show that controlling for the share of industries in regional employment, interacted with their national employment growth is sufficient to control for the possibility of occupational mobility. In other words, the same controls as we introduced to capture the potential mean reversion (L_{ir}/L_r and $(L_{ir}/L_r) * \hat{L}_{it}$), to which we add the cross-industry counterpart ($\sum_{j \neq i} (L_{jr}/L_r) * \hat{L}_{jt}$), also capture the effect of mobility between occupations. The intuition is as follows: if occupational mobility was perfect, different occupations would boil down to a single input factor: labour. In such a case, the occupational dimension of our responsiveness measure would become irrelevant, and the only remaining factor affecting the responsiveness of an industry's employment would be its share of regional employment. By adding the industry's share of regional employment separately to our index - both its level and its interaction with national industry growth - we effectively allow the impact of occupations to differ from the one predicted by our theory in a way which is consistent with occupational mobility. Column 4 of table 2 presents the results of the corresponding estimation. The point estimates of $\hat{L}_{irt}^{own}$ and $\hat{L}_{irt}^{cross}$ are almost unaffected and the additional control is itself insignificant.

Unemployment: Although our model assumes that the labour force is fully employed, allowing for unemployment can be important for two reasons. First, the presence of unemployment may give rise to measurement error, causing a downward bias in our estimates. Although the inclusion of MSA fixed effects and of the growth rate of employed labour capture the time invariant part as well as some of the time variation of unemployment rates, the lack of information on the composition of the unemployed labour stock in terms of occupations gives rise to classical measurement error, which should go against our results. Second, the possibility for workers to be unemployed creates an asymmetry between expanding and contracting industries. While a growing industry needs to recruit workers in the particular occupations that it uses, a contracting industry can simply lay off workers regardless of the scarcity of occupations in its region. In this case, our measure of employment responsiveness would only be a useful predictor of regional employment growth for expanding industries, while it would have no explanatory power for contracting industries.

In our model with full employment, the occupational dimension matters for contracting industries as workers with lower chances of being employed in other industries take a large wage cut to avoid being laid off. This guarantees that employment in region-industry pairs with a low index of employment responsiveness are also less reactive to negative shocks. In reality, a similar mechanism may take place even in the presence of unemployment as workers in contracting industries may be more willing to negotiate wage cuts to remain employed if their occupation is used only by few other industries in the region. Still, the above reasoning suggests that our theoretical mechanism matters more for expanding than contracting industries. We test for a possible heterogeneity between expanding and contracting industries by estimating our preferred specification separately for both groups of industries. To distinguish these groups, we compute the average growth for each industry over our sample period and use the first and the third terciles of the industry growth distribution for contracting and expanding industries, respectively.[26] Columns 5 and 6 of Table 2 show the results of the estimations for contracting and expanding industries, respectively. In line with our expectations, our measure performs better in explaining the regional responses for expanding industries.

Small Regions Assumption Our identification strategy relies on the assumption that national employment growth is not affected by shocks to employment in a particular region. However, national employment growth is a weighted sum of regional employment responses. Nevertheless, as long as regions are small enough from the perspective of the national economy, a regional shock should only have a negligible effect on the national growth of employment. In very large regions or for very large MSA-industries, however, there might be a mechanical comovement between the national and regional growth rates. Hence we expect that our measure ($\hat{L}_{irt}^{own}$) will lose some of its explanatory power in large MSAs at the expense of the national growth rate of the industry ($\hat{L}_{it}$). To test whether the performance of our measure indeed differs for MSAs of different sizes, we estimate our preferred specification for small and large regions separately. In doing so, we use the total employed labour in an MSA in 2003 and identify those in the first and the in third terciles as small and large respectively[27]. Column 1 and 2 of Table 3 present the results. The coefficient estimate of $\hat{L}_{irt}^{own}$ - the own effect predicted by our theory - is close to one

[26] The thresholds are -4 and +2 percent growth for the first and the third terciles.

[27] The thresholds are 41919 and 148738 employees for the first and the third terciles.

and very significant for smaller regions, while it is insignificant for the large MSAs. As expected, in large MSAs, the coefficient on the national growth rate of the industry turns positive and significant. Interestingly, the cross-industry effects ($\hat{L}_{irt}^{cross}$) become large and significant. In large regions, the number of active industries and level of diversification is generally high. In such cases, industry specific aggregate shocks capture only a small part of the relevant information to explain the MSA-industry growth. This finding shows that incorporating cross-industry effects is important to understand short-run employment changes of an industry, especially in large regions.

Another approach which we use to address the mechanical relationship between the national and regional employment growth is to exclude the employment of the state where the MSA-industry is located while constructing the national shocks. This strategy ensures that the predicted employment growth does not incorporate MSA- or state-specific shocks and is not mechanically related to the observed employment growth of an MSA-industry. We replicate all the analysis using this strategy and find that none of our benchmark results are affected.

4.2.3 Data Construction

Weight of the cost shares of occupations: To conduct all previous regressions, we constructed the matrix E_r (see equation (13)) under the assumption that $\frac{\nu-1}{\varepsilon} = 1$, which is the weight of the cost shares of occupations S_r^c in our measure of employment responsiveness. Although in line with Hsieh, Hurst, Jones, and Klenow (2013), this choice is somewhat arbitrary. We therefore reconstruct our measure using different values of $\frac{\nu-1}{\varepsilon}$ between 0.25 and 4 to test the sensitivity of our baseline results. Columns 3 and 4 of Table 3 present the estimates for $\frac{\nu-1}{\varepsilon} = 0.25$ and $\frac{\nu-1}{\varepsilon} = 4$, respectively. Our results are robust to this additional test.

Treatment of missing data: Employment data at the MSA-industry level are often not reported due to imprecise estimates or privacy reasons in the CBP database. As described in the data appendix, we use the information on the approximate size distribution of firms and on the intervals provided by the CBP to approximate the actual employment level for each MSA-industry. Approximated values are available for all MSA-industry-year tuples, while exact data account for only 17% of all observations. In all our regressions, we used exact employment data only for the MSA-industry pairs for which they are available for all years. For all other MSA-industry pairs, i.e. if exact data are missing for at least one year, we use approximate values for all years. The reason for this procedure is to avoid computing growth rates based on approximate values for one year and actual ones for the next, as this would introduce additional noise in the growth rates. To check the robustness of our results, we replicate all regressions using the actual values whenever they are available in two consecutive years and use the growth based on approximate values otherwise. This method increases the share of actual values substantially but does not affect our results as can be seen in Column 5 of Table 3. Unreported results show that using regressions based solely on the subsample of MSA-industry-year observations for which growth rates are computed with non-approximated data does not affect our estimate of β_1. The cross-industry effect β_4 however becomes insignificant.

Wage and employment shares of occupations: Due to the unavailability of data on wage and employment shares of occupations at the MSA-industry level (ω_{ior} and L_{ior}/L_{ir}, respectively), we assume that they do not exhibit any regional variation and we use their national counterpart to construct the data. As argued earlier however, these two assumptions are mutually inconsistent[28]. To check the sensitivity of our results to these assumptions, first note that the "expected" wage for an occupation in our model, w_{or}, should be the same across all industries in a region. Indeed, multiplying both sides of equation (7) by w_{ior} and using equation (6), we find that $w_{or} = \Delta(\sum_i w_{ior}^{\nu})^{1/\nu}$, which is constant for an occupation in a particular region. This modified no-arbitrage condition together with the assumption that the employment share of an occupation within an industry is constant across regions, i.e. $L_{ior}/L_{ir} = L_{io}/L_i \;\; \forall r$, leads to the following expression for the wage share of an occupation o in industry i in region r:

$$\omega_{ior} = \frac{w_{or}\frac{L_{ior}}{L_{ir}}}{\sum_{o'} w_{o'r}\frac{L_{io'r}}{L_{ir}}} \tag{20}$$

To compute these wage shares, we compile data on the annual average wage share of occupations at the regional level using regional OES data (see Section 7.1). We reestimate our model using ω_{ior} as given by (20) to compute the matrix $\boldsymbol{R}_r$ and show that our results are unaffected (Column 6 of Table 3).

[28]The first would be correct if the production function was Cobb Douglas, while the second would hold with a Leontieff production function.

Table 2: Robustness tests

	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$
Resp. x Nat. growth ($\hat{L}_{irt}^{own}$)	0.777***	0.778***	0.804***	0.805***	0.677*	1.065***
	[4.53]	[4.53]	[4.73]	[4.74]	[2.55]	[3.33]
Resp. (R_{iir})	0.041**	0.043***	0.041**	0.041**	0.023	-0.025
	[3.24]	[3.38]	[3.30]	[3.28]	[1.06]	[-1.34]
Nat. growth ($\hat{L}_{it}$)	0.098	0.097	0.068	0.066	0.167	-0.142
	[0.56]	[0.56]	[0.39]	[0.38]	[0.60]	[-0.47]
Cross-ind. effect ($\hat{L}_{irt}^{cross}$)	0.212*	0.211*	0.329***	0.345***	0.393**	0.006
	[2.57]	[2.55]	[4.02]	[4.02]	[3.26]	[0.03]
Log init. size ($ln(L_{ir,2003})$)	-0.080***	-0.079***	-0.079***	-0.079***	-0.080***	-0.096***
	[-33.54]	[-31.43]	[-31.65]	[-31.65]	[-19.77]	[-18.73]
$L_{ir,2003}/L_{r,2003}$	0.893***	0.842***	0.845***	0.845***	0.260	1.181**
	[6.56]	[6.18]	[6.25]	[6.25]	[1.12]	[3.30]
$(L_{ir,2003}/L_{r,2003}) * \hat{L}_{it}$	-6.106**	-6.096**	-6.454**	-6.501**	-10.960**	-8.428
	[-2.69]	[-2.69]	[-2.87]	[-2.89]	[-3.13]	[-1.78]
$\alpha_{ir} * \hat{L}_{it}$		-0.001	-0.001	-0.001	0.001	-0.003*
		[-0.58]	[-0.59]	[-0.59]	[0.36]	[-2.20]
α_{ir}		0.000	0.000	0.000	0.000	0.000
		[1.33]	[1.36]	[1.36]	[0.51]	[0.54]
$\hat{L}_{rt}$			0.632***	0.630***	0.567***	0.652***
			[25.95]	[25.46]	[14.89]	[14.86]
$\sum_{j \neq i} (L_{jrt}/L_{rt}) * \hat{L}_{jt}$				0.110	0.177	-0.165
				[0.69]	[0.70]	[-0.58]
Industry FE's	Yes	Yes	Yes	Yes	Yes	Yes
Year FE's	Yes	Yes	Yes	Yes	Yes	Yes
MSA FE's	Yes	Yes	Yes	Yes	Yes	Yes
N	356470	356470	356470	356470	115259	119534
R^2	0.074	0.074	0.078	0.078	0.075	0.089

The dependent variable is $\hat{L}_{irt}$, the MSA-industry growth rate of employment. "Resp." is the MSA-industry specific responsiveness measure, given by the corresponding diagonal entry of the R_r matrix as defined in (16). Standard errors are heteroscedasticity robust and clustered at the industry*year level. t-statistics in brackets * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All regressions include a constant.

Table 3: Robustness tests, cont'd

	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$
Resp. x Nat. growth ($\hat{L}^{own}_{irt}$)	1.084***	0.031	0.866***	0.681***	0.619***	0.814***
	[3.42]	[0.13]	[4.54]	[5.43]	[4.25]	[4.33]
Resp. (R_{iir})	0.012	-0.011	0.048**	0.034***	0.035***	0.050**
	[0.61]	[-0.53]	[2.91]	[4.03]	[3.39]	[2.86]
Nat. growth ($\hat{L}_{it}$)	-0.281	0.848***	0.021	0.176	0.171	0.065
	[-0.87]	[3.49]	[0.11]	[1.36]	[1.15]	[0.35]
Cross-ind. effect ($\hat{L}^{cross}_{irt}$)	-0.104	0.425***	0.404***	0.248***	0.241***	0.235***
	[-0.77]	[3.42]	[3.91]	[3.84]	[3.31]	[3.47]
Log init. size ($ln(L_{ir,2003})$)	-0.207***	-0.071***	-0.080***	-0.077***	-0.062***	-0.0790***
	[-27.40]	[-20.12]	[-32.21]	[-30.00]	[-31.44]	[-32.38]
$L_{ir,2003}/L_{r,2003}$	3.994***	3.660***	0.879***	0.777***	0.524***	0.935***
	[16.52]	[8.29]	[6.16]	[6.25]	[7.73]	[6.82]
$(L_{ir,2003}/L_{r,2003}) * \hat{L}_{it}$	-1.194	-4.595	-7.975***	-4.876*	-3.340*	-6.695**
	[-0.37]	[-1.00]	[-3.46]	[-2.30]	[-2.46]	[-3.21]
$\hat{L}_{rt}$	0.608***	0.593***	0.630***	0.632***	0.554***	0.626***
	[18.15]	[14.22]	[25.55]	[25.32]	[27.29]	[25.58]
$\sum_{j \neq i} (L_{jrt}/L_{rt}) * \hat{L}_{jt}$	-0.280	0.614*	0.198	-0.018	0.054	0.073
	[-1.11]	[2.35]	[1.19]	[-0.12]	[0.39]	[0.46]
$\alpha_{ir} * \hat{L}_{it}$	0.005	-0.001	-0.001	0.000	-0.001	-0.001
	[0.10]	[-0.96]	[-0.80]	[0.26]	[-1.26]	[-1.07]
α_{ir}	-0.007***	0.001***	0.000	0.000	0.000	0.000
	[-4.47]	[4.59]	[1.47]	[1.86]	[1.27]	[1.39]
Industry FE's	Yes	Yes	Yes	Yes	Yes	Yes
Year FE's	Yes	Yes	Yes	Yes	Yes	Yes
MSA FE's	Yes	Yes	Yes	Yes	Yes	Yes
N	114680	122257	355904	357577	354081	354595
R^2	0.125	0.074	0.078	0.078	0.077	0.077

The dependent variable is $\hat{L}_{irt}$, the MSA-industry growth rate of employment. "Resp." is the MSA-industry specific responsiveness measure, given by the corresponding diagonal entry of the R_r matrix as defined in (16). Standard errors are heteroscedasticity robust and clustered at the industry*year level. t-statistics in brackets * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All regressions include a constant.

5 Aggregate consequences

In this section, we draw the aggregate consequences of our model and derive a measure of the sector-specificity of workers in each industry at the national level. For this, we use equation (15) of our model and the data described in section 3 to compute the elasticity of an industry i's national employment to its output price, which is given by νE_{ii} in (15), where E_{ii} denotes the i'th diagonal entry of the matrix $\boldsymbol{E}$. The vector of elasticities is important for two reasons. First, it determines the adaptability of an economy to changes in economic conditions. The ability of an economy to quickly reallocate labour away from ailing sectors towards growing ones is a determinant of its short-term performance, notably its GDP growth and unemployment rate. Second, it quantifies the degree to which workers in i are "specific" to that industry, where a higher E_{ii} means a lower specificity. The specific factors literature (Mussa (1974), Neary (1978)) shows that the wage of workers which are tied to an industry is very dependent on the output price of that industry, while the opposite holds for workers who can move between industries. Our model builds on a similar structure, but makes explicit that the *degree of industry specificity*, captured by E_{ii}, depends on the geographic proximity of similar occupations[29]. From (15), an industry has a high E_{ii} if it is mostly located in MSAs where the occupations it uses intensively are abundantly present, i.e. in MSAs where other industries using similar occupations are located. This is more likely to happen if (i) the occupations that an industry uses are commonly used by other industries in the economy, (ii) the industry is mostly located in MSAs where it does not account for a large share of employment, and (iii) the industry collocates with the other industries using similar occupations.

Before we proceed, two caveats are in order. First, we want to stress that the elasticities reported in this section are those implied by our model, and therefore solely capture the effect of our two supply side frictions, namely the lack of short run mobility between MSAs and between occupations. Although other factors certainly matter in reality, e.g. the demand elasticity, the aim of the present exercise is to isolate the contribution of our two supply side frictions. Second, the vector of employment elasticity is given by the diagonal entry of the matrix $\boldsymbol{E}$ times ν, which we do not know. Since ν scales the employment

[29] The lower the E_{ii} (the more specific the workers), the more the average wage in i will react to a change in p_i. Obtaining a measure of an industry's worker specificity is therefore essential to understand and quantify the redistributive consequences of industry-specific shocks or policies.

elasticity for all sectors, we decide to ignore this common factor in the following and to report a *normalised* elasticity for all industries i, equal to E_{ii} and with a maximum value of one[30]. Since the normalised elasticity is only equal to one in a frictionless world (with neither occupations nor geography are a source of friction) where all industries are small, the distance from one reflects the reduction in elasticity due to the presence of our two frictions combined.

Table 4 reports the value of E_{ii} for all 2-digit sectors and for all 3-digit manufacturing industries[31]. They show a substantial variation between industries. Agriculture exhibits the highest degree of workforce specificity (the lowest elasticity), suggesting that employment in agriculture should be little responsive to changes in output price, but that labour compensation should take the bulk part of the adjustment. Similarly, mining and education have relatively specific workers. On the other hand, wholesale trade has the highest employment elasticity. Within manufacturing, textile-related industries generally feature a high degree of workforce specificity, while metal and machinery manufacturing do not. The fact that workers in textile and agriculture exhibit a high degree of sector specificity is of particular interest considering that these sectors are among the most protected in the United States. The argument that workers in these branches find it difficult to move to other industries is confirmed by our model.

To obtain a better understanding of the driving forces behind the values of E_{ii} in column (1), we compute two benchmark cases, corresponding to our two types of frictions. The first, reported in column (2), shows the value that E_{ii} would take if workers could change occupation costlessly but still could not migrate between MSAs. As shown in Table 4, column (2) is very close to 1, meaning that no industry is primarily located in MSAs where it accounts for a large share of employment. The geographic frictions alone can therefore

[30] Note that our calculations are not totally free of assumptions about the parameters ν and ϵ as these enter the matrix A . All computations presented in this section are based on the assumption that $(\nu-1)/\epsilon = 1$. We have recomputed all results of this section for $(\nu-1)/\epsilon = 0.25$ and $(\nu-1)/\epsilon = 4$. The qualitative conclusions, and in particular the ranking of sectors are left unaffected.

[31] For aggregation at the 2 or 3 digit level, we compute the weighted average of all own price elasticities at the 4-digit level, where the weights are determined by the share of employment that a given 4-digit industry accounts for in its corresponding 2 or 3 digit sector. Using an unweighted average instead of a weighted average does not change the results much. A full set or results for each 4-digit industry is available from the authors upon request.

TWO DIGIT SECTORS

Code	Naics description	(1)	(2)	(3)	(4)
11	Agriculture, Forestry, Fishing and Hunting	0.41	0.99	0.55	0.76
72	Accommodation and Food Services	0.42	0.96	0.43	0.97
61	Educational Services	0.47	0.98	0.49	0.95
21	Mining	0.56	0.98	0.70	0.67
23	Construction	0.59	0.99	0.61	0.95
54	Professional, Scientific, and Technical Services	0.59	0.99	0.62	0.93
81	Other Services (except Public Administration)	0.60	0.99	0.61	0.96
62	Health Care and Social Assistance	0.60	0.98	0.62	0.96
52	Finance and Insurance	0.64	0.98	0.66	0.94
22	Utilities	0.67	0.99	0.71	0.90
56	Administrative and Support and Waste Management and Remediation Services	0.69	0.98	0.71	0.96
48-49	Transportation and Warehousing	0.70	0.99	0.75	0.85
71	Arts, Entertainment, and Recreation	0.71	0.99	0.76	0.82
51	Information	0.77	0.99	0.79	0.88
31-33	Manufacturing	0.79	0.98	0.87	0.61
44-45	Retail Trade	0.83	0.99	0.84	0.93
53	Real Estate and Rental and Leasing	0.84	1.00	0.86	0.90
55	Management of Companies and Enterprises	0.87	0.97	0.90	0.81
42	Wholesale Trade	0.93	0.99	0.94	0.78

THREE DIGIT INDUSTRIES IN MANUFACTURING

Code	Naics description	(1)	(2)	(3)	(4)
323	Printing and Related Support Activities	0.60	0.99	0.62	0.94
315	Apparel Manufacturing	0.63	0.99	0.76	0.65
313	Textile Mills	0.65	0.98	0.79	0.61
311	Food Manufacturing	0.68	0.97	0.77	0.72
322	Paper Manufacturing	0.71	0.99	0.83	0.59
337	Furniture and Related Product Manufacturing	0.73	0.98	0.81	0.71
336	Transportation Equipment Manufacturing	0.74	0.97	0.88	0.45
316	Leather and Allied Product Manufacturing	0.75	1.00	0.85	0.60
324	Petroleum and Coal Products Manufacturing	0.75	0.99	0.84	0.64
331	Primary Metal Manufacturing	0.78	0.98	0.90	0.45
326	Plastics and Rubber Products Manufacturing	0.78	0.98	0.84	0.74
314	Textile Product Mills	0.78	0.96	0.95	0.25
321	Wood Product Manufacturing	0.82	0.99	0.91	0.50
325	Chemical Manufacturing	0.82	0.99	0.91	0.48
334	Computer and Electronic Product Manufacturing	0.83	0.99	0.90	0.57
339	Miscellaneous Manufacturing	0.85	0.99	0.88	0.77
335	Electrical Equip., Appliance, and Component Manuf.	0.86	0.98	0.97	0.25
332	Fabricated Metal Product Manufacturing	0.88	0.99	0.92	0.68
333	Machinery Manufacturing	0.89	0.99	0.97	0.32
327	Nonmetallic Mineral Product Manufacturing	0.90	0.99	0.96	0.36
312	Beverage and Tobacco Product Manufacturing	0.92	0.99	0.98	0.21

Table 4: Measure and decomposition of average factor specificity per industry.

not explain much of the heterogeneity in elasticities highlighted in column (1)[32]. The second benchmark, shown in column (3) of both tables, is the normalised elasticity which would result if workers could migrate at no cost but were unable to change occupations. At the 2-digit level, column (3) tracks column (1) very closely. Column (4) makes this close fit explicit by showing the extent to which the occupational dimension alone can account for the deviation of the normalised elasticity from the frictionless case[33]. For all 2-digit sectors, the occupational dimension accounts for at least 60% of this deviation, and for most sectors at least 90%. Interestingly, the two sectors for which this fraction is lowest are mining and manufacturing, for which we would expect a higher degree of geographical concentration, respectively due to the location of natural resources and to the importance of spillovers (Ellison, Glaeser, and Kerr (2010)). These results suggest that the occupational dimension contributes more than the spatial dimension to the cross-industry heterogeneity in worker specificity.[34]

In the appendix 7.4, we provide an additional perspective on the aggregate consequences of our model for the U.S. and assess the costs of occupational and geographic frictions in terms of foregone GDP growth at the national level. In a frictionless world, i.e. where workers can change occupation and migrate, short-run exogenous shocks to output prices would give rise to a substantial reallocation of labour from industries with a negative shock to those with a positive one. As we have just shown, however, the elasticity of labour employment at the national level substantially differs from the frictionless case and is heterogeneous between industries. The lack of short run labour reallocation has a cost in terms of GDP growth which we quantify. For this, we consider the large increase in the dollar price of tradable goods in the period 2005-2008 - concomitant with a strong

[32]Note that the values for the 2 and 3 digit sectors in column (2) are also weighted averages over those of the corresponding 4-digit sectors. Even if a 2-digit sector represents an important part of total employment in the MSAs where it is present, it may be that each of the 4-digit industries composing it accounts for a small share of local employment, in which case the value of column (2) is close to 1. The measured importance of geography is not independent of the level of aggregation chosen.

[33]Column (4) is computed as $(1 - Column(3))/(1 - Column(1))$

[34]We should note that there is also a third component of the frictions due to the interaction between geographic and occupational immobility of labour. For instance, in the first row of the upper panel of Table 4 geography alone explains virtually none of the total fricitons for the 2-digit NAICS industry, "Agriculture, Forestry, Fishing and Hunting", whereas occupational immobility alone explains only 76 percent as reported in Column 4. This suggests that almost 23 percent of the deviation from one is explained by the interaction between the occupational and geographic immobility of workers.

depreciation of the dollar - and compare the GDP growth implied by our model to the GDP growth which would obtain in a model with no frictions. Assuming no change in the price of non-tradables, we find that the geographic and occupational frictions cost 1.15 percentage points of GDP growth, thereby halving the growth rate which would obtain with no frictions.

6 Conclusion

This paper shows that the geographical distribution of occupational employment within a country affects the sensitivity of an industry's employment to external shocks. We derive a theory-based measure capturing the similarity of occupational employment between industries and show that the geographic proximity to similar industries matters for the response of employment to short-run demand shocks. We assess the importance of our theoretical mechanisms in two ways.

First, we test our theory empirically at the regional level. Using data on the employment growth of MSA-industry pairs in the U.S., we show that the employment of an industry responds more to national shocks in regions where other industries using similar occupations are located. Second, we aggregate the regional U.S. data on industry composition to compute our model's prediction for the price elasticity of employment and output of each industry at the national level. We draw two conclusions from this exercise. First, we show that our model predicts substantial cross-industry differences in the degree to which workers are specific to their industry. Our model implies that workers in agricultural sectors and in textile manufacturing are among the most specific, meaning that they are most vulnerable to demand shocks to their own industry. Second, we assess the efficiency loss coming from the short-run immobility of workers between MSAs and occupations by computing the change in real GDP growth implied by our model under two sets of assumptions: (i) geographic and occupational immobility and (ii) perfect mobility between regions and occupations. We find that, for a price shock similar to the increase in tradable prices between 2005 and 2008, U.S. short-run real GDP would grow twice as quickly if workers could costlessly move between MSAs and between occupations.

References

Artuç, E., S. Chaudhuri, and J. McLaren (2010): "Trade shocks and labor adjustment: a structural empirical approach," *American Economic Review*, 100, 1008–1045.

Artuç, E., and J. McLaren (2012): "Trade Policy and wge inequality: a structural analysis with occupational and sectoral mobility," mimeo.

Autor, D., D. Dorn, and G. Hanson (2012): "The China syndrome: local labor market effects of import competition in the United States," American Economic Review, forthcoming.

Blanchard, O., L. Katz, R. Hall, and B. Eichengreen (1992): "Regional evolutions," *Brookings papers on economic activity*, 1992(1), 1–75.

Bound, J., and H. Holzer (2000): "Demand shifts, population adjustments, and labor market outcomes during the 1980s," *Journal of Labor Economics*, 18(1), 20–54.

Chiquiar, D. (2008): "Globalization, regional wage differentials and the Stolper-Samuelson Theorem: Evidence from Mexico," *Journal of International Economics*, 74, 70–93.

Cosar, K. (2013): "Adjusting to trade liberalization: reallocation and labor market policies," Chicago Booth, mimeo.

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

Ebenstein, A., A. Harrison, M. McMillan, and S. Phillips (2011): "Estimating the impact of trade and offshoring on American workers using the Current Population Surveys," World Bank Policy Research paper 5750.

Ellison, G., E. Glaeser, and W. Kerr (2010): "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," *American Economic Review*, 100, 1195–1213.

Freeman, R., and L. Katz (1991): "Industrial wage and employment determination in an open economy," in *Immigration, Trade, and the Labor Market*, ed. by J. Abowd, and R. Freeman, pp. 335–359. University of Chicago Press.

Gaston, N., and D. Trefler (1997): "The labor market consequences of the Canada-US free trade agreement," *Canadian Journal of Economics*, 30, 18–41.

Godlberg, P., and N. Pavcnik (2007): "Distributional effects of globalization in developing countries," *Journal of Economic Literature*, 45, 39–82.

Grossman, G. (1986): "Imports as a cause of injury: the case of the U.S. steel industry," *Journal of International Economics*, 20, 201–223.

Hanson, G. (2007): "Globalization, labor income, and poverty in Mexico," in *Globalization and poverty*, pp. 417–456. University of Chicago Press.

Hsieh, C.-T., E. Hurst, C. Jones, and P. Klenow (2013): "The allocation of talent and U.S. economic growth," mimeo.

Kambourov, G. (2009): "Labor market regulations and the sectoral reallocation of workers: the case of trade reforms," *Review of Economic Studies*, 76, 1321–1358.

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

Kovak, B. (2013): "Regional effects of trade reform: what is the correct measure of liberalization," *American Economic Review*, 103(5), 1960–1976.

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

McLaren, J., and S. Hakobyan (2012): "Looking for local labor market effects of NAFTA," NBER Working Paper 16535.

McMillan, M., and D. Rodrik (2011): "Globalization, structural change and productivity growth," NBER Working Paper 17143.

Menezes-Filho, N., and M.-A. Muendler (2011): "Labor reallocation in response to trade reforms," mimeo.

Mussa, M. (1974): "Tariffs and the distribution of income: the importance of factor specificity," *Journal of Political Economy*, 82(6), 1191–1203.

Neary, P. (1978): "Short-run capital specificity and the pure theory of international trade," *Economic Journal*, 88, 488–510.

Overman, H., and D. Puga (2010): "Labor pooling as a source of agglomeration. An empirical investigation.," in *Agglomeration Economics*, pp. 133–150. University of Chicago Press.

Partridge, M., D. Rickman, O. Rose, and A. Kamar (2012): "Dwindling U.S. internal migration: evidence of spatial equilibrium or structural shifts in local labor markets?," *Regional Science and Urban Economics*, 42(1-2), 375–388.

Revenga, A. (1992): "Exporting jobs? The impact of import competition on employment and wages in U.S. manufacturing," *Quarterly Journal of Economics*, 107, 255–284.

——— (1997): "Employment and wage effects of trade liberalization: the case of Mexican manufacturing," *Journal of Labor Economics*, 15, 20–43.

Rosenthal, S., and W. Strange (2001): "The determinants of agglomeration," *Journal of Urban Economics*, 50(2), 191–229.

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

Topalova, P. (2007): "Trade liberalization, poverty and inequality: evidence from Indian districts," in *Globalization and Poverty*, pp. 291–336. University of Chicago Press.

——— (2010): "Factor immobility and regional impacts of trade liberalization: evidence from poverty in India," *American Economic Journal: Applied Economics*, 2, 1–41.

Trefler, D. (2004): "The long and short of the Canadian-U.S. free trade agreement," *American Economic Review*, 4, 870–895.

Vannoorenberghe, G., and E. Janeba (2013): "Trade and the political economy of redistribution," CESifo Working Paper 4062.

Wacziarg, R., and J. Wallack (2004): "Trade liberalization and intersectoral labor movements," *Journal of International Economics*, 64, 411–439.

7 Appendix (Not for publication)

7.1 Data Appendix

This appendix gives the sources of the data and the detailed procedure we followed to treat them.

7.1.1 Data on MSA-industry employment and employment growth

We use employment data from 2003-2008 for MSAs in 4-digit NAICS industries from the County Business Patterns data (CBP) of the Census Bureau. MSAs include both Micropolitan Statistical Areas, with an urban core of 10.000 to 49.999 persons and Metropolitan Statistical Areas, with an urban core of more than 50.000 persons. The main concern with the CBP is the prevalence of missing data, which are either not reported due to confidentiality issues or due to poor quality. Considering the high level of disaggregation that we use (4-digit NAICS employment per MSA), a large fraction of the observations is missing (only 17% of all MSA-industry observations have non-missing data in all years). For all observations, however, the CBP also reports an approximation of the firm size distribution, which consists for each MSA-industry-year tuple in the number of firms in each size bin (e.g. 2 firms between 1-4 employees and 3 firms between 10-19). The CBP also reports for each observation an upper bound and lower bound for the number of employees in an MSA-industry-year. For each observation, we compute an approximation of employment by (i) assuming that all firms within a size bin have the average of the upper and lower bound of the bin, (ii) multiplying the obtained size by the number of firms in the bin, and adding up over bins, (iii) truncating the obtained value at the lower or upper bound provided by the CBP if the computed value lies outside. In our example, that means 2*2.5+3*14.5=48.5 employees if this lies between the bounds of the CBP. If the CBP reports that employment lies between 30 and 45 for that MSA-industry-year observation, we would then use 45 as the number of employees.

In the main part of the analysis, we compute the growth rate of employment as the growth rate of the approximation, except for the MSA-industry pairs which report exact employment data *every year*. This procedure ensures that the change from approximated to exact data does not cause additional noise. For example, consider an MSA-industry pair which in reality has 100 employees from 2003 to 2005, and for which we approximate 110 employees each year by using the size distribution of firms. Assume that the CBP only reports the actual data in 2004. If we were to use actual data whenever available and approximated data otherwise, we would measure in this situation a contraction of employment by 10 employees in 2004 and a growth of 10 employees in 2005. To avoid generating such noise, we use either only approximations or only actual data but never mix the two. In section 4.2.3, we experiment with an alternative method: taking actual data to compute the growth rate of employment whenever actual data are available two years in a row for an MSA-industry pair. This procedure relies more on actual data without being subject to the noise previously described.

7.1.2 Data on occupations and the computation of $\boldsymbol{R}_r$

Data on occupations are from the Occupational Employment Statistics (OES) of the Bureau of Labor Statistics (BLS). The OES reports data on the employment and wage of each occupation and each industry at the national level. We use these data to compute the matrices $\boldsymbol{S}_r^e$ and $\boldsymbol{S}_r^c$. We compute L_{io}/L_i as the beginning of sample share of industry i's employment accounted for by occupation o and ω_{io} as the total wage paid by industry i to occupation o as a fraction of total wage payments of industry i at the beginning of sample.

The OES data reports some missing observations at the detailed 4-digit NAICS/ 6-digit SOC level. In particular, data on employment and/or wages of an occupation may be missing for two reasons: privacy concerns and poor quality of the data. Observations are typically only missing for occupations which account only for a small fraction of an industry's employment.

To account for missing data in an efficient way, we use the data on occupations from the OES surveys for May 2003, November 2003 and May 2004[35] and proceed in two steps for

[35]There is here a tradeoff between using data on beginning of the sample only, which is warranted for exogeneity concerns, and the fact that May and November 2003 have less precise information than May 2004, in which there are many less missings. There are two reasons why we think that this is not a concern to use May 2004. First, the occupational data are collected on a rolling basis, where only 1/3 of the observations are replaced each year. The data for May 2004 is therefore based for 2/3 on earlier observations. Second, we ran the whole analysis based on the May 2003 data only and did not find any

each industry. First, for each occupation with missing data in one or two out of the three surveys, we assign the share of that occupation in the industry in the other surveys, scaled by the share of the 2-digit occupation in the industry. For example, if the occupation "economists" is missing in industry "Universities" in 2003 but not in 2004, we impute to the 2003 data a share of economists in universities given by (i) the share of economists in all the technical occupations used by the Universities in 2004 times (ii) the share of technical occupations in universities in 2003. Second, we add up the (imputed and original) shares for all 6-digit occupations of an industry in each survey and use for that industry the data on the survey for which the sum of shares is closest to 1. Combining the occupation data with the employment data from the CBP, we have a total of 289 industries and 922 MSAs.

7.1.3 Industry Correspondence

The industry classification system used by CBP changes over time. While the data from 2003 to 2007 are in NAICS 2002, the data in 2008 are in NAICS 2007. To get a consistent industry classification over time, we use a 6-digit concordance table to match 2008 data to NAICS 2002. We keep the data as they are if the change in a 6-digit code does not result in a change at the 4-digit and do the necessary updates otherwise. The only non-trivial update relates the 6-digit NAICS 2007 industries falling under more than one 4-digit NAICS 2002 codes. In these rare cases, we assign the 6-digit industry to its first match.

7.1.4 Additional data sources

National growth rates per 4-digit industry come from the *Statistics of U.S. Businesses (SUSB)* of the Census bureau.

Input Output data come from the Bureau of Economic Analysis standard make and use tables for 2002. To compute the matrix of D_{ji} in (19), we multiply the transpose of the make matrix by the use matrix.

Regional wages for occupations (w_{or}) are needed to compute ω_{ior} as defined in (20). The OES publishes estimates of these for *metropolitan* state areas and some micropolitan regions, the latter are however not defined along the lines of the micropolitan state areas reported in the CBP data. The OES also reports state estimates of wages per occupation,

qualitative difference.

which we use as a proxy for the wage of occupations in the micropolitan areas located in that state[36]. We use the metropolitan and the state estimates of the OES from 2005. Although the new definition of MSAs is based on the 2000 census, it is only implemented as of 2003 in the CBP and as of 2005 in OES[37].

7.1.5 Descriptive statistics

Table 5: Descriptive statistics

Variable	Min	Median	Average	Max	Obs
Employees per MSA	1972	79039	330732	8378058	356470
Employees per industry	2086	526946	859372	5230878	356470
R_{iir}	0.686	0.990	1.012	2.056	356470
$L_{ir,2003}/L_{r,2003}$	4.55e-06	0.005	.009	0.209	356470
$\hat{L}_{it}$	-0.143	0.010	0.010	0.224	356470

The figures come from the observations which are used in our empirical analysis, hence outliers and the MSA-industries with fewer than 100 employees are omitted. For 2003, the number of MSAs per industry is 620 on average with a minimum of 28 and a maximum of 922 while the number of industries per MSA is 194 is on average with a minimum of 78 and a maximum of 288.

7.2 Empirical Appendix

7.2.1 Without controlling for R_{iir}

In this section, we replicate Tables 1 and 2 in the main text without controlling for R_{iir} to show that our results are not driven by its inclusion in the estimating equations as a control variable.

[36]Partialing out the metropolitan state areas from the state data to obtain the is complicated by the fact that many MSA's are defined across state borders. Since this would create additional noise, we prefer assigning the state average wage to the Micropolitan state areas in that state.

[37]The coding of Metropolitan state slightly differs between the OES and the CBP data in New England, for which the OES does not report the MSA but the "NECTA" (New England City and Town Areas) only. We match the NECTA (used by OES) to the MSA (used by CBP) based on their name.

Table 6: Main results without R_{iir}

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$
Resp. x Nat. growth ($\hat{L}_{irt}^{own}$)	1.590***	1.584***	1.617***	1.088***	0.950***	1.003***	1.473***
	[6.24]	[6.23]	[6.33]	[5.40]	[5.59]	[5.95]	[6.52]
Nat. growth ($\hat{L}_{it}$)	-0.776**	-0.771**	-0.762**	-0.245	-0.114	-0.176	-
	[-3.07]	[-3.05]	[-3.00]	[-1.23]	[-0.68]	[-1.06]	-
Cross-ind. effect ($\hat{L}_{irt}^{cross}$)		-0.130	-0.186*	-0.445***	0.228**	0.370***	0.227***
		[-1.70]	[-2.48]	[-6.33]	[2.77]	[4.26]	[2.73]
Log init. size ($ln(L_{ir,2003})$)				-0.0480***	-0.0778***	-0.0779***	-0.0772***
				[-39.02]	[-43.42]	[-43.67]	[-43.49]
Industry FE's	Yes	Yes	Yes	Yes	Yes	Yes	No
Year FE's	No	No	Yes	Yes	Yes	No	No
MSA FE's	No	No	No	No	Yes	No	Yes
MSA*Year FE's	No	No	No	No	No	Yes	No
Industry*Year FE's	No	No	No	No	No	No	Yes
Observations	356470	356470	356470	356470	356470	356470	356470
R^2	0.027	0.027	0.027	0.061	0.074	0.087	0.049

The dependent variable is $\hat{L}_{irt}$, the MSA-industry growth rate of employment. Standard errors are heteroscedasticity robust and clustered at the industry*year level. t-statistics in brackets * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Table 2 without R_{iir}

	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$
Resp. x Nat. growth ($\hat{L}^{own}_{irt}$)	0.852***	0.856***	0.880***	0.881***	0.609*	0.944**
	[4.98]	[4.99]	[5.19]	[5.19]	[2.38]	[3.19]
Nat. growth ($\hat{L}_{it}$)	0.022	0.0181	-0.00894	-0.0105	0.239	-0.0199
	[0.13]	[0.10]	[-0.05]	[-0.06]	[0.88]	[-0.07]
Cross-ind. effect ($\hat{L}^{cross}_{irt}$)	0.229**	0.228**	0.346***	0.364***	0.411***	0.00361
	[2.79]	[2.78]	[4.25]	[4.27]	[3.48]	[0.02]
Log init. size ($ln(L_{ir,2003})$)	-0.083***	-0.082***	-0.082***	-0.082***	-0.081***	-0.094***
	[-39.93]	[-37.41]	[-37.63]	[-37.65]	[-22.40]	[-20.89]
$L_{ir,2003}/L_{r,2003}$	0.731***	0.682***	0.690***	0.691***	0.162	1.264***
	[5.19]	[4.85]	[4.94]	[4.95]	[0.66]	[3.44]
$(L_{ir,2003}/L_{r,2003}) * \hat{L}_{it}$	-6.098**	-6.089**	-6.447**	-6.501**	-11.54***	-8.414
	[-2.68]	[-2.68]	[-2.87]	[-2.89]	[-3.36]	[-1.79]
$\alpha_{ir} * \hat{L}_{it}$		-0.001	-0.001	-0.001	0.001	-0.003*
		[-0.54]	[-0.56]	[-0.55]	[0.35]	[-2.22]
α_{ir}		0.000	0.000	0.000	0.000	0.000
		[1.13]	[1.16]	[1.16]	[0.43]	[0.64]
$\hat{L}_{rt}$			0.632***	0.630***	0.568***	0.652***
			[25.96]	[25.46]	[14.89]	[14.85]
$\sum_{j\neq i}(L_{jrt}/L_{rt}) * \hat{L}_{jt}$				0.125	0.193	-0.168
				[0.78]	[0.77]	[-0.59]
Industry FE's	Yes	Yes	Yes	Yes	Yes	Yes
Year FE's	Yes	Yes	Yes	Yes	Yes	Yes
MSA FE's	Yes	Yes	Yes	Yes	Yes	Yes
Observations	356470	356470	356470	356470	115259	119534
R^2	0.074	0.074	0.078	0.078	0.075	0.088

The dependent variable is $\hat{L}_{irt}$, the MSA-industry growth rate of employment. "Resp." is the MSA-industry specific responsiveness measure, given by the corresponding diagonal entry of the R_r matrix as defined in (16). Standard errors are heteroscedasticity robust and clustered at the industry*year level. t-statistics in brackets * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

7.2.2 Industry*Year Fixed Effects

In the benchmark empirical analyses, we control for year and industry fixed effects to capture the macroeconomic shocks common to all MSA-industries and time-invariant industry characteristics, respectively. In this section, we will replicate Tables 1 and 2 in the main text to show that our results are robust to the control of industry*year fixed effects which capture the time-varying industry specific characteristics. Column 4 of Table (8) corresponds to the column 7 of Table (1).

Table 8: Main results with Time-varying Industry Specific Effects

	(1)	(2)	(3)	(4)
	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$
Resp. x Nat. growth ($\hat{L}_{irt}^{own}$)	1.396***	1.395***	0.991***	1.433***
	[5.06]	[5.06]	[4.08]	[6.24]
Resp. (R_{iir})	0.255***	0.256***	0.167***	0.0127
	[26.13]	[25.67]	[18.79]	[0.97]
Cross-ind. effect ($\hat{L}_{irt}^{cross}$)		0.0721	-0.284***	0.221**
		[0.93]	[-3.98]	[2.65]
Log init. size ($ln(L_{ir,2003})$)			-0.0453***	-0.0759***
			[-38.44]	[-34.16]
Industry FE's	No	No	No	No
Year FE's	No	No	No	No
MSA FE's	No	No	No	Yes
Industry*Year FE's	Yes	Yes	Yes	Yes
Observations	356470	356470	356470	356470
R^2	0.009	0.009	0.039	0.049

The dependent variable is $\hat{L}_{irt}$, the MSA-industry growth rate of employment. Standard errors are heteroscedasticity robust and clustered at the industry*year level. t-statistics in brackets * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Table 2 with Time-varying Industry Specific Effects

	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$	$\hat{L}_{irt}$
Resp. x Nat. growth ($\hat{L}^{own}_{irt}$)	1.217***	1.214***	1.240***	1.241***	0.988*	2.345***
	[4.62]	[4.61]	[4.72]	[4.72]	[2.32]	[3.57]
Resp. (R_{iir})	0.0369**	0.0387**	0.0373**	0.0372**	0.031	-0.0635**
	[2.89]	[3.03]	[2.93]	[2.92]	[1.28]	[-2.69]
Cross-ind. effect ($\hat{L}^{cross}_{irt}$)	0.210*	0.208*	0.335***	0.346***	0.386**	0.046
	[2.52]	[2.50]	[4.04]	[3.98]	[3.12]	[0.25]
Log init. size ($ln(L_{ir,2003})$)	-0.0796***	-0.0783***	-0.0788***	-0.0788***	-0.0792***	-0.0955***
	[-33.54]	[-31.41]	[-31.65]	[-31.64]	[-19.70]	[-18.89]
$L_{ir,2003}/L_{r,2003}$	0.870***	0.823***	0.827***	0.827***	0.325	1.058**
	[6.50]	[6.14]	[6.21]	[6.21]	[1.25]	[2.65]
$(L_{ir,2003}/L_{r,2003}) * \hat{L}_{it}$	-3.894	-3.933	-4.354	-4.391	-8.763	-4.476
	[-1.24]	[-1.25]	[-1.39]	[-1.40]	[-1.85]	[-0.59]
$\alpha_{ir} * \hat{L}_{it}$		-0.00071	-0.000699	-0.000698	0.000424	-0.00206
		[-0.54]	[-0.53]	[-0.53]	[0.16]	[-1.31]
α_{ir}		0.000138	0.000141	0.000141	0.0000673	0.0000398
		[1.26]	[1.29]	[1.29]	[0.45]	[0.18]
$\hat{L}_{rt}$			0.630***	0.628***	0.566***	0.649***
			[25.78]	[25.34]	[14.86]	[14.77]
$\sum_{j \neq i} (L_{jrt}/L_{rt}) * \hat{L}_{jt}$				0.074	0.0957	-0.099
				[0.46]	[0.37]	[-0.34]
Industry FE's	No	No	No	No	No	No
Year FE's	No	No	No	No	No	No
MSA FE's	Yes	Yes	Yes	Yes	Yes	Yes
Industry*Year FE's	Yes	Yes	Yes	Yes	Yes	Yes
Observations	356470	356470	356470	356470	115259	119534
R^2	0.049	0.049	0.053	0.053	0.057	0.064

The dependent variable is $\hat{L}_{irt}$, the MSA-industry growth rate of employment. "Resp." is the MSA-industry specific responsiveness measure, given by the corresponding diagonal entry of the R_r matrix as defined in (16). Standard errors are heteroscedasticity robust and clustered at the industry*year level. t-statistics in brackets * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

7.3 Theory Appendix

7.3.1 Derivation of equation (8)

Equating the demand and supply of effective labour in each industry-occupation-region tuple (equations (4) and (7)) gives:

$$w_{ior}^{\nu} = \left(\alpha_{io}\Delta^{-1}\zeta_{ir}\right)^{\Omega} \left(\sum_{i'} w_{i'oc}^{\nu}\right)^{\frac{\nu-1}{\nu}\Omega} L_{or}^{-\Omega}. \tag{21}$$

Summing up over all industries and rearranging, the geometric average of wages in an occupation-region pair is:

$$\sum_{i} w_{ior}^{\nu} = (\Delta L_{or})^{-\frac{\nu}{\epsilon}} \left(\sum_{i} (\alpha_{io}\zeta_{ir})^{\Omega}\right)^{\frac{\nu+\epsilon-1}{\epsilon}}. \tag{22}$$

Plugging the above back in (21) and rearranging yields (8). Combining (6), (7), (21) and (22) gives:

$$L_{ior} = \frac{\left(\alpha_{io}\zeta_{ir}\right)^{\Omega}}{\sum_{i'}\left(\alpha_{i'o}\zeta_{i'c}\right)^{\Omega}} L_{or} \tag{23}$$

$$\Lambda_{ior} = \Delta L_{or}(\alpha_{io}\zeta_{ir})^{\frac{\nu-1}{\nu+\epsilon-1}} \left(\sum_{j} (\alpha_{jo}\zeta_{jr})^{\Omega}\right)^{\frac{1-\nu}{\nu}} \tag{24}$$

7.3.2 Comparative statics

From (21), it can easily be shown that:

$$\sum_{i} w_{ior}^{\nu} = (\Delta L_{or})^{-\frac{\nu}{\epsilon}} \left(\sum_{i} (\alpha_{io}\zeta_{ir})^{\Omega}\right)^{\frac{\nu+\epsilon-1}{\epsilon}}. \tag{25}$$

Plugging the above and (21) in (7) and rearranging, the total payment to occupation o in the industry region pair ir is:

$$w_{ior}\Lambda_{ior} = (\Delta L_{or})^{\frac{\epsilon-1}{\epsilon}} (\alpha_{io}\zeta_{ir})^{\Omega} \left(\sum_{j} (\alpha_{j}\zeta_{jr})^{\Omega}\right)^{\frac{1-\nu}{\nu}\frac{\epsilon-1}{\epsilon}} \tag{26}$$

To derive (12), we first differentiate the right hand side of (9) (rhs) with respect to ζ_{mr}

for $m \neq i$:

$$\begin{aligned}\frac{\partial rhs}{\partial \zeta_{mr}} &= \Delta p_i^{\epsilon-1}\Omega\frac{1-\nu}{\nu}\frac{\epsilon-1}{\epsilon}\left[\sum_o \alpha_{io}^{\Omega} L_{or}^{\frac{\epsilon-1}{\epsilon}}\left(\sum_j (\alpha_{jo}\zeta_{jr})^{\Omega}\right)^{\frac{1-\nu}{\nu}\frac{\epsilon-1}{\epsilon}-1}\alpha_{mo}^{\Omega}\zeta_{mr}^{\Omega-1}\right] \\ &= \frac{1-\nu}{\epsilon}(1-\Omega)\zeta_{mr}^{-1}p_i^{\epsilon-1}\zeta_{ir}^{-\Omega}\left[\sum_o w_{ior}\Lambda_{ior}\frac{L_{mor}}{L_{or}}\right] \qquad (27)\end{aligned}$$

where we use (23) and (26) to derive the second equality. Using the definition of ω_{ior} as the cost share of occupation o in the industry-region pair ir, the above becomes:

$$\frac{\partial rhs}{\partial \zeta_{mr}} = \frac{1-\nu}{\epsilon}(1-\Omega)\zeta_{mr}^{-1}\zeta_{ir}^{1-\Omega}\left(\sum_o \omega_{ior}\frac{L_{mor}}{L_{or}}\right). \qquad (28)$$

Using the above equation, the total differentiation of (9) gives:

$$\hat{\zeta}_{ir} = (\nu+\epsilon-1)\hat{p}_i + \frac{1-\nu}{\epsilon}\left[\sum_m \hat{\zeta}_{mr}\left(\sum_o \omega_{ior}\frac{L_{mor}}{L_{or}}\right)\right] \qquad (29)$$

7.3.3 Imperfect mobility between occupations

In this section, we relax the assumption that workers are fully immobile between occupations. To model imperfect mobility parsimoniously, we now assume that workers within an occupation are perfectly mobile between industries, but that each worker draws independently for each occupation a productivity parameter from the Fréchet distribution (2). To avoid confusion with the parameter ν, which refers to the mobility between industries, we here denote the parameter of the Fréchet distribution for occupations as $\tilde{\nu}$. Similarly, we define $\tilde{\Omega} \equiv \tilde{\nu}/(\tilde{\nu}+\epsilon-1)$. Workers choose the occupation in which they obtain the highest income. The number of workers in region r choosing to work for occupation o and the number of effective labour units in occupation o are respectively given by:

$$\begin{aligned}L_{or} &= \frac{w_{or}^{\tilde{\nu}}}{\sum_o w_{or}^{\tilde{\nu}}}L_r \qquad (30)\\ \Lambda_{or} &= \Delta w_{or}^{\tilde{\nu}-1}\left(\sum_o w_{or}^{\tilde{\nu}}\right)^{\frac{1-\tilde{\nu}}{\tilde{\nu}}} L_r \qquad (31)\end{aligned}$$

where L_r is the exogenous number of workers in region r.

From (4), the demand for effective labour units of occupation o in industry i is given by:

$$\Lambda_{ior} = \alpha_{io}w_{or}^{-\epsilon}\zeta_{ir}. \qquad (32)$$

Equating the effective supply of occupation o in (31) with the demand for the occupation, given by the sum of (32) across all industries, pins down the wage of an occupation:

$$w_{or}^{\tilde{\nu}} = \left(\sum_i \alpha_{io}\zeta_{ir}\right)^{\tilde{\Omega}} \left(\sum_{o'} w_{o'}^{\tilde{\nu}}\right)^{\frac{\tilde{\nu}-1}{\nu}\tilde{\Omega}} L_r^{-\tilde{\Omega}}. \tag{33}$$

Summing up over all o, solving for $\sum_o w_{or}^{\tilde{\nu}}$ and plugging back in the above equation gives:

$$w_{or}^{\tilde{\nu}} = \left(\sum_i \alpha_{io}\zeta_{ir}\right)^{\tilde{\Omega}} \left[\sum_{o'} \left(\sum_i \alpha_{io'}\zeta_{ir}\right)^{\tilde{\Omega}}\right]^{\frac{\tilde{\nu}-1}{\epsilon}} L_r^{-\frac{\tilde{\nu}}{\epsilon}}. \tag{34}$$

Using the equilibrium w_{or} in (30), (31), and (32) gives the equilibrium L_{or}, Λ_{or}, and L_{ior} respectively. Note that L_{ior} is not determined as a worker in occupation o has the same productivity in all industries, and all that matters for an industry is the number of effective units of labour it employs. It can for example employ many workers with a low productivity in the occupation and few workers with a high productivity. Workers are also indifferent as they receive the same wage per effective labour unit in all industries. A natural assumption is to equate the fraction of occupation o workers in industry i to the share of effective labour units of o in i: $L_{ior} = L_{or}\Lambda_{ior}/\Lambda_{or}$. From this assumption, the size of employment in industry i is given by $L_{ir} = \sum_o L_{ior}$:

$$L_{ir} = \sum_o \frac{\alpha_{io}\zeta_{ir}\left(\sum_j \alpha_{jo}\zeta_{jr}\right)^{\tilde{\Omega}-1}}{\sum_{o'}\left(\sum_j \alpha_{jo'}\zeta_{jr}\right)^{\tilde{\Omega}}} L_r \tag{35}$$

which is the counterpart to (10) when workers are mobile between occupations. On the other hand, plugging the solution for w_{or} obtained in (34) into (5) gives the counterpart to (9):

$$p_i^{1-\epsilon} = \left[\sum_o \alpha_{io}\left(\sum_j \alpha_{jo}\zeta_{jr}\right)^{\frac{1-\epsilon}{\tilde{\nu}+\epsilon-1}}\right]\left[\sum_o\left(\sum_j \alpha_{jo}\zeta_{jr}\right)^{\tilde{\Omega}}\right]^{\frac{\tilde{\nu}-1}{\tilde{\nu}}\frac{1-\epsilon}{\epsilon}} L_r^{\frac{\epsilon-1}{\epsilon}}. \tag{36}$$

We now turn to the comparative statics exercise. Differentiating (35) with respect to ζ_{mr} for $m \neq i$ gives:

$$\frac{\partial L_{ir}}{\partial \zeta_{mr}}\frac{\zeta_{mr}}{L_{ir}} = -\sum_o \frac{L_{ior}}{L_{ir}}\frac{L_{mor}}{L_{or}} + \tilde{\Omega}\left(\sum_o \frac{L_{ior}}{L_{ir}}\left(\frac{L_{mor}}{L_{or}} - \frac{L_{mr}}{L_r}\right)\right), \tag{37}$$

$$\frac{\partial L_{ir}}{\partial \zeta_{ir}}\frac{\zeta_{ir}}{L_{ir}} = -\sum_o \frac{L_{ior}}{L_{ir}}\frac{L_{-ior}}{L_{or}} + \tilde{\Omega}\left(\sum_o \frac{L_{ior}}{L_{ir}}\left(\frac{L_{ior}}{L_{or}} - \frac{L_{ir}}{L_r}\right)\right). \tag{38}$$

The total derivative of L_{ir} is therefore given by:

$$\hat{L}_{ir} = \hat{\zeta}_{ir}\left(\sum_o \frac{L_{ior}}{L_{ir}}\frac{L_{-ior}}{L_{or}}\right) - \sum_{m\neq i}\hat{\zeta}_{mr}\left(\sum_o \frac{L_{ior}}{L_{ir}}\frac{L_{mor}}{L_{or}}\right) + \tilde{\Omega}\left[\sum_m \hat{\zeta}_{mr}\left(\sum_o \frac{L_{ior}}{L_{ir}}\left(\frac{L_{mor}}{L_{or}} - \frac{L_{mr}}{L_r}\right)\right)\right] \quad (39)$$

Equation (39) consists of two parts. The first line is similar to (11) with the only difference that the coefficient Ω in (11) has dropped out. The reason is that we assume perfect mobility between industries within an occupation, so that the Ω in (11) is now equal to 1. The second line shows an important difference of allowing for imperfect mobility between occupations. The growth of an industry m has now an additional impact on the employment in other industries through its effect on the equilibrium supply of given occupations.Consider the case where workers cannot switch occupations. In this case, growth in industry m reduces the supply of occupations used by m in other industries, an effect captured by the second part of the first line in (39). With occupational mobility however, the occupations which m uses intensively (i.e. the occupations for which $L_{mor}/L_{or} - L_{mr}/L_r > 0$) will see their supply increase, thereby dampening the direct effect of growth in m on the other industries using these occupations. For the occupations that m does not use intensively however, the effect goes in the opposite direction, as some workers in these occupations will now switch to occupations used intensively by industry m. The supply of occupations which are not used intensively by the growing m industry therefore contracts, and - compared to the baseline case of no occupational mobility - affects negatively the employment of other industries which use these occupations. Equation (39) can be rewritten as:

$$\hat{L}_{ir} = \hat{\zeta}_{ir}\left(\sum_o \frac{L_{ior}}{L_{ir}}\left(\frac{L_{-ior}}{L_{or}} + \tilde{\Omega}\frac{L_{ior}}{L_{or}}\right)\right) - (1-\tilde{\Omega})\sum_{m\neq i}\hat{\zeta}_{mr}\left(\sum_o \frac{L_{ior}}{L_{ir}}\frac{L_{mor}}{L_{or}}\right) - \tilde{\Omega}\sum_m \hat{\zeta}_{mr}\frac{L_{mr}}{L_r} \quad (40)$$

This formulation shows that, if $\tilde{\Omega} = 0$ and workers cannot move between occupations, the indices of occupational closeness S^e_{ir} and S^e_{imr} defined in (39) are the relevant measures of the marginal effect of ζ_{mr} on $\hat{L}_{ir}$. If $\tilde{\Omega} > 0$ on the other hand, the last term above shows that $\hat{\zeta}_{mr}$ not only affects $\hat{L}_{ir}$ through the indices of occupational closeness but also directly through the share of employment that m accounts for in the region. If $\tilde{\Omega} \to 1$, workers are perfectly mobile between occupations and the occupational structure becomes irrelevant. In this case, the share of an industry in the region's total employment replaces our measure of occupational closeness as the relevant index.

We now turn to the total derivative of (36). It can easily be shown that:

$$(\nu + \epsilon - 1)\hat{p}_i = \sum_j \hat{\zeta}_{jr} \left(\sum_o \omega_{ior} \frac{L_{jor}}{L_{or}} \right) + \frac{\tilde{\nu} - 1}{\epsilon} \left(\sum_j \hat{\zeta}_{jr} \frac{L_{jr}}{L_r} \right). \tag{41}$$

There are two main differences between the above equation and its counterpart with occupational immobility and partial cross-industry mobility in (12). First, $\hat{\zeta}_{ir}$ enters the equation in a similar way as all other $\hat{\zeta}_{jr}$ when $\hat{p}_i$ is on the left hand side. This is due to the assumption of perfect mobility between industries. If, in (12), $\nu \to \infty$, $\hat{\zeta}_{ir}$ would also enter in the same way as all other $\hat{\zeta}_{jr}$ in (12). Second, (41) shows that the share of an industry enters the equation separately from the measure S^c_{imr} as defined in (12). The intuition is similar to the one in (39). If workers are very mobile between occupations, the occupational closeness becomes a less important determinant of the effect of growth in an industry on employment growth of other industries.

In a similar way to (13), we can express the vector of the growth rates of regional employment across industries as:

$$\hat{\boldsymbol{L}}_r = (\tilde{\nu} + \epsilon - 1) \left(\boldsymbol{I} - (1 - \tilde{\Omega}) \boldsymbol{S}^e_r - \tilde{\Omega} \boldsymbol{X}_r \right) \left(\boldsymbol{S}^c_r + \frac{\tilde{\nu} - 1}{\epsilon} \boldsymbol{X}_r \right)^{-1} \hat{\boldsymbol{p}}, \tag{42}$$

where the matrix $\boldsymbol{X}_r$ is a matrix of the share of employment of each industry in the region:

$$\boldsymbol{X}_r = \begin{bmatrix} L_{1r}/L_r & L_{2r}/L_r & ... & L_{Nr}/L_r \\ L_{1r}/L_r & L_{2r}/L_r & ... & L_{Nr}/L_r \\ ... & ... & ... & ... \\ L_{1r}/L_r & L_{2r}/L_r & ... & L_{Nr}/L_r \end{bmatrix}.$$

As shown by (42), and in line with the explanations above, the growth of employment is less sensitive to our measures of the occupational mix (S^e_r and S^c_r), but more to the share of regional employment in the industry. Controlling directly for the share of employment in our estimating equation (as in Table 2) therefore is similar to allowing for occupational mobility. Since the share of employment of industries enter equation (42) in a highly non-linear way, we show in additional (unreported) specifications that controlling for employment shares in a non-linear way does not affect our results.

7.4 Quantitative impact of frictions on GDP growth

As previously argued, we expect the ability of an economy to reallocate resources across sectors to be an important determinant of short run GDP growth. The present section takes a step further in this direction and quantifies the impact of our spatial and occupational frictions on short run growth. For this, we compute the predicted increase in real GDP growth stemming from an exogenous shock to the price vector under two scenarios. The first assumes, as in our theory, the impossibility for workers to move between MSAs and occupations. The second assumes perfect mobility of workers between MSAs and occupations. The difference in real GDP growth between the two scenarios provides an estimate of the short-run costs of occupational and spatial immobility in the U.S. economy.

7.4.1 Computing real GDP growth

First, we compute the own and cross price elasticities of the national output of each industry as predicted by our model, and denote the matrix elements as $\epsilon_{in} = \partial y_i / \partial p_n * p_n / y_i$. The matrix ϵ_{in} is, up to a constant, very similar to the matrix of price elasticities of employment since the ability to scale output up or down depends on the ability to recruit or lay off workers[38]. Second we compute the matrix of own and cross price elasticities which would obtain if workers were perfectly mobile between occupations and MSAs. We denote the elements of the elasticity matrix of this "no-frictions" case as ϵ_{in}^{NF}.

In a third step, detailed in section 7.4.3, we compute an approximation of the growth rate of real GDP based on the matrix of cross price elasticities of output at the national level (the matrix of ϵ_{in} or ϵ_{in}^{NF}) and on the price vector. The difference between the real GDP growth[39] with and without our two supply side frictions is given by:

$$\widehat{RGDP} - \widehat{RGDP}^{NF} = \sum_n \sum_i \left(s_i \epsilon_{in} - s_i^{NF} \epsilon_{in}^{NF} \right) \hat{p}_n \hat{p}_i \tag{43}$$

where s_i and s_i^{NF} are the share of national GDP accounted for by industry i in the cases with and without frictions respectively. $\widehat{RGDP}$ and $\widehat{RGDP}^{NF}$ denote the real GDP growth with and without frictions. Equation (43) has three interesting properties. First, for a marginal change in the price vector, the product $\hat{p}_n \hat{p}_i$ is negligible and the above expression is zero. This is an application of the envelope theorem: if the output vector

[38] The strength of this connection is reinforced by the assumption that labour is the only factor of production in the model. Formally, we derive the values of ϵ_{in} from (12) by noting that $\hat{\zeta}_{ir} = \epsilon \hat{p}_i + \hat{y}_{ir}$.

[39] We assume that each good accounts for the same share of production and consumption to concentrate purely on the effect of our occupational and spatial frictions. Terms of trade effects, which would decrease the real value of GDP if the U.S. imports goods of which the price goes up more strongly, are neglected in this exercise as they do not directly relate to the the reallocation of labour between sectors.

initially maximises real GDP, the change in the output vector following a marginal change in prices does not affect real GDP. Since the effect of the spatial and occupational frictions solely come through the price elasticity of output, the costs of frictions in terms of foregone real GDP growth is zero for small price changes. Second, consider for simplicity the case where only the price of good i changes and where $s_i = s_i^{NF}$. Since $\epsilon_{ii}^{NF} > \epsilon_{ii}$ (the reaction of output to a price change is more important when there are less frictions), the real GDP growth must be larger in the case with no frictions. The extent to which frictions affect the elasticity of output is key to determining the foregone real GDP growth. Third, as evident from (49), if all prices change in the same proportion, $\widehat{RGDP} - \widehat{RGDP}^{NF} = 0$. A proportional change of all prices does not affect the optimal allocation of resources between sectors, and therefore makes frictions irrelevant.

7.4.2 The 2005-2008 shock to tradable prices

We consider as an external shock to the economy the large change in the prices of tradable goods which followed the strong depreciation of the dollar and the boom in world commodity prices[40] between 2005 and 2008. We use the data on import prices given by the import/export price data (MXP) of the Bureau of Labor Statistics at the 3-digit naics level and report the full list of price changes in Table 12. Table 10 summarises the price changes at the 2-digit level. We impute the vector of price changes described in Table

Table 10: Change in tradable prices, 2005-2008

2-digit naics	2005-08 %Δ in import prices
11 Agriculture, Forestry, Fishing and Hunting	+34.8%
21 Mining	+62.5%
31 Food, Beverages and Textile manufacturing	+8.5%
32 Wood, Paper, Chemical, Plastics, Coal manufacturing	+25.5%
33 Metal, machinery, electric and electronic equipment manufacturing	+8.4%

12 into equation (43) and compute the difference in GDP growth rates implied by our model between a case with and a case without spatial and occupational mobility. Following Hsieh, Hurst, Jones, and Klenow (2013), we set $\nu - 1 = 3$ and assume that, in

[40] Although these changes were driven to a large extent by factors external to the U.S., we do not require that these were absolutely exogenous to the U.S. economy. We think of these numbers as providing an interesting example of a realistic price shock that may hit the U.S. economy.

the scenario where workers are immobile between occupations, $\epsilon = \nu - 1$. In response to the sole change in tradable prices over the 2005-08 period, our model predicts a real GDP growth of 1.23% between 2005 and 2008 for the case where workers are immobile between MSAs and occupations (i.e. using ϵ_{in}) but a growth of 2.37% if workers are freely mobile between occupations and MSAs (using ϵ_{in}^{NF}), implying a growth differential of 1.14 percentage points. Our model therefore predicts that the short-run GDP growth following of a price shock comparable to the one observed in 2005-2008 would be twice as large if there were no spatial and occupational frictions in the U.S. Table 11 reports the GDP growth with and without frictions for different values of ϵ and shows that the foregone GDP growth remains large for a wide range of ϵ.

The above results relied partly on the assumption that the price of non-tradables did not change over the 2005-2008 period. This assumption may however matter as the frictions emphasised in the present paper arise when different sectors face asymmetric shocks. A contemporaneous rise in the price of the non-tradable sectors would therefore decrease the heterogeneity in price changes between industries and dampen the impact of the occupational and spatial frictions. Assuming that the price of all non-tradable sectors grew at the same rate as the CPI[41], our model still predicts a 0.61 percentage point differential growth between the scenarios with and without frictions (from 0.94 to 1.55%). Although we consider only shocks to tradable sectors, which account for a relatively small fraction of total U.S. employment (around 12% in 2005), and although we assume a homogeneous change in the price of all non-tradable sectors, the effect of our two frictions on real GDP growth remains substantial.

7.4.3 Calibration

Consider two price vectors, $\boldsymbol{p} = \{p_i\}_{i \in I}$ and $\boldsymbol{p'} = \{p'_i\}_{i \in I}$. If the price vector changes from p to p', the associated change in nominal GDP is:

$$\Delta GDP = \sum_i p'_i y_i(\boldsymbol{p'}) - \sum_i p_i y_i(\boldsymbol{p}) = \sum_i (p'_i - p_i) y_i(\boldsymbol{p}) + p'_i (y_i(\boldsymbol{p'}) - y_i(\boldsymbol{p})). \quad (44)$$

[41] The average CPI grew by 10.1% over the three years, see BLS. Note that part of the rise in CPI should be accounted for by the very large increase in import prices. CPI growth is therefore likely to be larger than the growth of the price of non-tradables. Since a large increase in the price of non-tradables (of the same magnitude as that of tradables) would reduce the quantitative effect of the frictions, using the CPI for non-tradable prices provides a conservative estimate of the impact of frictions.

Table 11: Predicted real GDP growth following the 2005-2008 change in tradable prices

	Constant non-tradable prices	Non-tradable prices follow CPI
Occupational and Spatial frictions		
$(\nu-1)/\epsilon = 4$	0.84%	0.67%
$(\nu-1)/\epsilon = 1$	1.23%	0.94%
$(\nu-1)/\epsilon = 0.25$	1.44%	1.08%
No frictions	2.37%	1.55%

The first column reports the simulated real GDP growth following the 2005-2008 change in tradable prices under the assumption that non-tradable prices did not change. The second column assumes that non-tradable prices grew at the CPI rate. The first three lines assume that there are occupational and spatial frictions, varying the value of ϵ. The last line reports simulated GDP growth with no frictions. All simulations assume that $\nu - 1 = 3$. To compute the results for $\nu' \neq \nu$, all numbers in the table should be multiplied by $(\nu' - 1)/3$.

The growth of GDP following the price change can easily be shown to equal:

$$\widehat{GDP} = \sum_i \hat{p}_i s_i + \sum_i (1 + \hat{p}_i) s_i \hat{y}_i \tag{45}$$

where $\hat{y}_i$ and $\hat{p}_i$ respectively denote the growth rate of output and of prices when $\boldsymbol{p}$ changes to $\boldsymbol{p'}$, and where s_i is the share of GDP accounted for by industry i at prices $\boldsymbol{p}$ ($s_i \equiv p_i y_i(\boldsymbol{p})/(\sum_n p_n y_n(\boldsymbol{p}))$). The first sum in (45) reflects a pure valuation effect: if the price of good i increases, it causes a mechanical increase in the value of nominal GDP. In terms of real GDP however, the effect is unclear and depends on whether the fraction of good i in total spending on consumption is larger or smaller than s_i (a terms of trade effect). Since our focus is to quantify the effect of the spatial and occupational frictions on the efficiency of the allocation of labour between sectors, we want to abstract from these valuation effects and assume that the share of good i in GDP (s_i) is equal to the share of consumption spending on good i. Under this assumption, we have:

$$\widehat{RGDP} = \sum_i (1 + \hat{p}_i) s_i \hat{y}_i. \tag{46}$$

To compute the growth of real GDP as a function of the change in the price vector only, we make the following approximation:

$$\hat{y}_i = \sum_n \epsilon_{in}(\boldsymbol{p}) \hat{p}_n, \tag{47}$$

where:

$$\epsilon_{in}(\boldsymbol{p}) = \frac{\partial y_i(\boldsymbol{p})}{\partial p_n} \frac{p_n}{y_i(\boldsymbol{p})}. \tag{48}$$

Equation (47) is an approximation in the sense that we abstract from the effect of changes of $\boldsymbol{p}$ on the matrix of cross price elasticities. At the initial prices, the production vector must maximise the value of GDP, i.e. the following must hold for all n:

$$\sum_i p_i \frac{\partial y_n}{\partial p_n} = 0 \Rightarrow \sum_i s_i \epsilon_{in} = 0 \Rightarrow \sum_i s_i \hat{y}_i = 0. \tag{49}$$

The middle equality in (49) shows how to obtain the equilibrium vector of GDP share for each industry (s_i): it is the (normalised) eigenvector of the elasticity matrix which has eigenvalue zero. Under the approximation, the growth of real GDP in (46) therefore collapses to:

$$\widehat{RGDP} = \sum_i \hat{p}_i s_i \hat{y}_i = \sum_i \sum_n s_i \epsilon_{in} \hat{p}_n \hat{p}_i. \tag{50}$$

Repeating the same exercise using the cross elasticity matrix without frictions ϵ_{in}^{NF} and the corresponding equilibrium shares s_i^{NF} gives (43). From (49), it is immediate that if all prices change in the same proportion, real GDP does not change.

To see that the growth of real GDP based on the national matrix of cross price elasticity is identical to the one which would obtain when solving for each industry-MSA pair, first note that (46) is equivalent to:

$$\widehat{RGDP} = \sum_r \sum_n (1 + \hat{p}_i) \kappa_{ir} s_i \hat{y}_{ir} \tag{51}$$

where $\hat{y}_{ir}$ is the growth of output in the industry-MSA pair ir and $\kappa_{ir} \equiv y_{ir}/y_i$. Defining $\epsilon_{inc}(\boldsymbol{p}) \equiv \frac{\partial y_{ir}(\boldsymbol{p})}{\partial p_n} \frac{p_n}{y_{ir}(\boldsymbol{p})}$ and using the fact that:

$$\hat{y}_i = \sum_r \kappa_{ir} \hat{y}_{ir} = \sum_r \sum_n \kappa_{ir} \epsilon_{inc}(\boldsymbol{p}) \hat{p}_n \tag{52}$$

in combination with (47) shows that:

$$\epsilon_{in}(\boldsymbol{p}) = \sum_r \kappa_{ir} \epsilon_{inr}(\boldsymbol{p}) \tag{53}$$

and that (51) and (46) are equivalent[42].

To compute the elasticity matrix $\epsilon_{inr}(\boldsymbol{p})$, we use the system of equations defined by (12):

$$\left(\boldsymbol{I} + \frac{\nu - 1}{\epsilon} \boldsymbol{S}_{\boldsymbol{r}}^{\boldsymbol{c}} \right) \hat{\boldsymbol{\zeta}}_r = (\nu + \epsilon - 1) \hat{\boldsymbol{p}} \tag{54}$$

[42] Due to the lack of data on output at the MSA level, we use the share of employment L_{ir}/L_i as a proxy for the share of output y_{ir}/y_i.

and the definition of ζ, which implies that $\hat{\boldsymbol{\zeta}}_r = \epsilon\hat{\boldsymbol{p}} + \hat{\boldsymbol{y}}_r$. Rearranging gives:

$$\hat{\boldsymbol{y}}_r = \underbrace{(\nu - 1)\left(\boldsymbol{I} + \frac{\nu - 1}{\epsilon}\boldsymbol{S}_{\boldsymbol{r}}^{\boldsymbol{c}}\right)^{-1}(\boldsymbol{I} - \boldsymbol{S}_{\boldsymbol{r}}^{\boldsymbol{c}})}_{\boldsymbol{\epsilon}_r(\boldsymbol{p})}\hat{\boldsymbol{p}} \tag{55}$$

where the nationwide matrix of the price elasticity of output is given by aggregating the regional matrix $\boldsymbol{\epsilon}_r(\boldsymbol{p})$ across MSAs using (53).

To obtain the elasticity matrix with no frictions $\boldsymbol{\epsilon}_r^{NF}(\boldsymbol{p})$, note that if there is only one occupation and one region, (12) becomes:

$$\hat{\zeta}_i = -\frac{\nu - 1}{\epsilon}\left(\sum_m \frac{L_m}{L}\hat{\zeta}_m\right) + (\nu + \epsilon - 1)\hat{p}_i \tag{56}$$

where L_m/L is the share of national employment in industry m. Multiplying both sides by L_i/L, adding up over all i and plugging back in the above equation gives:

$$\hat{\zeta}_i = (\nu + \epsilon - 1)\hat{p}_i - (\nu - 1)\left(\sum_m \frac{L_m}{L}\hat{p}_m\right). \tag{57}$$

From the definition of ζ and using matrix notation, the vector of output growth can be expressed as:

$$\hat{\boldsymbol{y}} = \underbrace{(\nu - 1)(\boldsymbol{I} - \boldsymbol{X})}_{\boldsymbol{\epsilon}^{NF}(\boldsymbol{p})}\hat{\boldsymbol{p}} \tag{58}$$

where $\boldsymbol{X}$ is a matrix of employment shares of industries defined as the national equivalent of $\boldsymbol{X}_r$ in the previous section.

3-digit naics	2005-08 %Δ import price
113 Forestry and Logging	35%
115 Support Activities for Agriculture and Forestry	35%
211 Oil and Gas Extraction	63%
212 Mining (except Oil and Gas)	63%
213 Support Activities for Mining	63%
311 Food Manufacturing	26%
312 Beverage and Tobacco Product Manufacturing	6%
313 Textile Mills	9%
314 Textile Product Mills	2%
315 Apparel Manufacturing	2%
316 Leather and Allied Product Manufacturing	5%
321 Wood Product Manufacturing	-6%
322 Paper Manufacturing	11%
323 Printing and Related Support Activities	26%
324 Petroleum and Coal Products Manufacturing	51%
325 Chemical Manufacturing	25%
326 Plastics and Rubber Products Manufacturing	11%
327 Nonmetallic Mineral Product Manufacturing	14%
331 Primary Metal Manufacturing	67%
332 Fabricated Metal Product Manufacturing	19%
333 Machinery Manufacturing	10%
334 Computer and Electronic Product Manufacturing	-7%
335 Electrical Equipment, Appliance, and Component Manufacturing	12%
336 Transportation Equipment Manufacturing	8%
337 Furniture and Related Product Manufacturing	8%
339 Miscellaneous Manufacturing	8%

Table 12: Change in 3-digit naics MXP import prices between 2005 and 2008