FIRM SIZE DISTRIBUTION AND EMPLOYMENT FLUCTUATIONS: THEORY AND EVIDENCE

Holger Görg[a], Philipp Henze[a], Viroj Jienwatcharamongkhol[b], Daniel Kopasker[d],
Hassan Molana[c], Catia Montagna[d], Fredrik Sjöholm[b]

September 2017

Abstract

We show that the firm-size *distribution* is an important determinant of the relationship between an industry's employment and output. A theoretical model predicts that changes in demand for an industry's output have larger effects on employment, resulting from adjustments at both the intensive and extensive margin, in industries characterised by a distribution that has a lower density of large firms. Industry-specific shape parameters of the firm size distributions are estimated using firm-level data from Germany, Sweden and the UK, and used to augment a relationship between industry-level employment and output. The empirical results align with the predictions of the theory.

Corresponding Author: Catia Montagna, Economics, Business School, University of Aberdeen, Aberdeen AB24 3QY, UK.

_____

[a] University of Kiel and Kiel Institute for the World Economy; [b] Lund University; [c] University of Dundee and the Scottish Institute for Research in Economics; [d] University of Aberdeen and the Scottish Institute for Research in Economics; [e] Lund University and the Research Institute of Industrial Economics, Stockholm.

1. Introduction

Extensive empirical evidence documents differences in the relationship between employment and output across countries and industries. The large variability in how much a given change in output affects employment (e.g. Perman and Stephan, 2015; Hoffmann and Lemiueux, 2014) reflects country-specific and industry-specific productivity responses to output changes. At the country level, the dissimilarity in these responses have typically been explained by the existence of differences in labour market institutions, such as work-sharing agreements and employment protection laws, that affect the ease and time it takes to, e.g., lay off workers. While these country specific factors are important in explaining fluctuations in the aggregate employment/output ratio, they cannot account for the observed differences across industries within countries.

In this paper we argue that dissimilarities in the size distribution of firms across industries and countries can help to explain the varied response of aggregate employment to output change. There is ample evidence of significant intra-industry heterogeneity in firm characteristics and performance, both within narrowly defined industries and across countries (e.g., Syverson, 2011). A key implication of the vast literature that has developed in response to this evidence is that resource allocation across different firms matters to aggregate outcomes. Specifically, changes in aggregate performance (total factor productivity, employment, trade and foreign direct investment flows) result not only from changes within firms, but also from compositional changes of firms within and across industries via selection and reallocation effects.[1] It is then conceivable that differences in the distribution of firms across industries and countries will contribute to different aggregate employment responses to output shocks.

---

[1] See e.g. Olley and Pakes (1996) and Caves (1998) for some early contributions to this literature. More recently, several contributions have shown how misallocations across heterogeneous production units can affect aggregate productivity and the transmission of shocks (see, e.g., Restuccia and Rogerson, 2008; Foster *et al*. 2008; Bartelsman *et al*., 2013).

To explore this conjecture, we first provide a theoretical specification of the employment-output relationship at the industry level which captures the role of firm heterogeneity by considering an industry characterised by firm-specific productivity. To derive analytical results, we assume – as is done in a large part of the monopolistic competition literature[2] – a Pareto productivity distribution which is shown to provide a relatively good fit for the observed size distribution – for evidence see, e.g., Axtell (2001) and Bernard *et al.* (2014).[3] We show that exogenous changes in output have larger effects on employment when the productivity distribution is characterised by a lower concentration of high productivity firms, as reflected by a larger Pareto shape parameter – i.e. in industries that exhibit a lower average productivity. This suggests that a higher average industry productivity offers a greater 'insulation' of aggregate employment from output shocks. At the core of this result lies the fact that aggregate employment responses do not simply reflect changes in employment at the firm level, but also intra-industry reallocations ensuing from entry and exit of firms into the industry as a result of the output shocks. In this light, our findings suggest that even though larger and more productive firms might exhibit larger employment responses to aggregate output fluctuations at the intensive margin (consistent with the findings in Moscarini and Postel-Vinay, 2012), employment responses at the extensive margin are larger in industries characterised by a lower average productivity: by raising the minimum productivity required to survive in the industry, a negative shock will generate a larger shake out of firms when the firm productivity distribution has a higher concentration at low productivity levels.

---

[2] The Pareto distribution has been widely adopted in the international trade literature, see, e.g.: Helpman *et al.* (2004), Chaney (2008), Arkolakis *et al.* (2008, 2012), Bertoletti *et al.* (2017) among others.

[3] A number of recent contributions have discussed the adequacy of the Pareto distribution to represent the whole sample of firm size and have suggested that a Log-normal or a mixed Log-Normal-Pareto distribution can capture the distribution of firms more robustly (e.g., Head *et al.*, 2014; Nigai, 2017). One argument is that the Pareto distribution fails to reflect the wedge, observed in the data, between the distribution of firm productivities and that of exports/sales (which is not Pareto). However, the Pareto-to-Pareto 'cycle' depends on other aspects of the model, e.g. the nature of preferences and exports cost (see e.g., Anderson and De Palma, 2015; Mrázová *et al.*, 2015, Eaton, *et al.* 2011).

We then propose that, consistent with the theory, the aggregate industry-level effects of firm-level adjustments and reallocations can be synthetically captured empirically using a measure of the shape of the firms' size distribution. Using detailed firm-level data over the period 1999-2007, we proceed to estimate industry-level Pareto shape parameters for size distributions of firms in three countries: Germany, Sweden and the UK. Our theoretical prediction is then examined by using the industry-level data to estimate an employment-output relationship augmented with the estimated Pareto shape parameters as well as with additional variables to control for observable country differences. One advantage with our sample is that these three countries are characterised by different firm-size distributions, different sectoral structures, and different welfare state models and labour market institutions. Including countries with different characteristics strengthens the robustness of our results and enables us to draw more general conclusions.

Our data confirms that there is a large variability between industries and countries in the intra-industry distribution of firms. For instance, we show that the UK has a distribution which has a higher concentration of smaller firms in comparison to the distributions in Germany and Sweden. More importantly, we find that the distribution of firms plays a significant role in determining the effect of output changes on changes in employment at the industry level. As predicted by our theoretical model, employment responses are found to be larger in industries with higher shares of smaller firms. This result has implications for policy – e.g., to the extent that governments have in place employment creation and/or unemployment protection policies, such industries might require a greater intensity of intervention in recessions.

Although a number of recent contributions focus on the relationship between intra-industry reallocations and employment dynamics, the role of firms' distribution in determining fluctuations in employment remains relatively unexplored. As far as we are aware, the issue we address in this paper has not been dealt with in the existing literature.

A number of studies highlight the impact of firm characteristics on employment creation. Elsby and Michaels (2013) develop a search and matching model with endogenous job destruction adopting a Pareto distribution for idiosyncratic firm productivity and, aggregating up from the microeconomic behaviour, they show that the model can account for the key stylised facts concerning the amplitude and propagation of the dynamics of worker flows in and out of unemployment over the business cycle. In most cases, however, a focus on the firm-level makes it difficult to draw inferences about the consequences of reallocation across firms for aggregate employment, and about the relationship between employment and output, as we do in this paper. One strand of the literature focuses on the relationship between firm-level adjustments and employment dynamics and shows how different firms exhibit different cyclical patterns of net job creation.[4] For instance, using US private sector firm-level data, Neumark *et al*. (2011) find an inverse relationship between net job growth rates and firm size, with small firms contributing disproportionately to net job creation.[5] However, Haltiwanger *et al*. (2013) find that these results do not hold when firms' age is taken into account.

Moscarini and Postel-Vinay (2012) study the effect of firm characteristics on firm-level job creation. They show that large employers are more cyclically sensitive than smaller ones, shedding proportionally more jobs in recessions and creating more jobs during booms. Our result, that aggregate employment is more responsive to output changes in industries whose firm-size distribution has a higher density of smaller firms might, appear at first glance to contradict this evidence. However, these two results need not be inconsistent with each other. As we shall explain below, this is because aggregate employment responses to an output shock

---

[4] For an earlier theoretical contribution see Hopenhayn and Rogerson (1993).
[5] Although they are likely to experience higher job destruction rates, new firms tend to be small and grow more rapidly – which accounts for their significant contribution on job creation.

reflect the effect of intra-industry reallocations on employment adjustments at both the intensive and the extensive margin.[6]

Another strand of the literature explores the so-called 'granular hypotheses' of aggregate fluctuations. Gabaix (2011) argues that because firms' size distributions are very fat-tailed, idiosyncratic shocks to individual firms will not average out and will therefore be reflected in aggregate GDP fluctuations. Di Giovanni and Levchenko (2012, 2014) highlight the role of firm-to-firm linkages in aggregate fluctuations and show that the size composition of firms in industries interacts with trade openness in determining aggregate output volatility.[7] These papers focus on the effect of firm size distributions on output volatility.

The rest of the paper is organised as follows. Section 2 develops a theoretical model which highlights the role of the firm distribution. Section 3 carries out the empirical analysis and Section 4 concludes the paper.

2. The model

Consider an industry consisting of an upstream and a downstream sector. The latter is perfectly competitive and produces a final good by using as intermediate inputs horizontally differentiated varieties produced by the upstream sector. We postulate a constant elasticity of substitution technology,

$$Y = \left(M^{\frac{\lambda-1}{\sigma}} \int_{i \in M} \left(y(i) \right)^{1-1/\sigma} di \right)^{\frac{1}{1-1/\sigma}}, \ 0 \leq \lambda \leq 1, \ \sigma > 1, \tag{1}$$

---

where Y is the quantity of the final good, $y(i)$ is the quantity of input of variety i, M is the mass of available varieties, σ is the constant elasticity of substitution between any two varieties and λ captures the extent of 'variety effect': the larger is λ the larger is the increase in output resulting from a given increase in the mass of varieties, i.e. the stronger are industry-wide scale economies; thus, $\lambda=0$ and $\lambda=1$ correspond to the two extreme cases of 'no variety effect' and 'maximum variety effect', respectively.[8] Choosing $y(i)$ to maximise the aggregate profit of the sector, $PY - \int_{i \in M} p(i) y(i) di$, subject to (1) yields the demand function for individual varieties,

$$y(i) = M^{\lambda-1} Y \left(\frac{p(i)}{P} \right)^{-\sigma}, \qquad i \in M, \tag{2}$$

where $p(i)$ and P are the price of variety i and the aggregate price of the input basket, respectively. The zero profit condition, together with (2), then ensures that P is in fact the price index dual to the input basket in (1), hence

$$P = \left(M^{\lambda-1} \int_{i \in M} \left(p(i) \right)^{1-\sigma} di \right)^{\frac{1}{1-\sigma}}. \tag{3}$$

The upstream sector consists of a continuum mass of firms, M, where each variety $i \in M$ of the differentiated input is produced by one firm using a linear technology with increasing returns to scale that utilises a composite Cobb-Douglas basket of labour and the homogeneous final good. We denote the quantities of these inputs and the resulting composite input by l, z and v, respectively, and assume

$$v(i) = \left(\frac{l(i)}{1-\gamma} \right)^{1-\gamma} \left(\frac{z(i)}{\gamma} \right)^{\gamma}, \tag{4}$$

---

[8] See Montagna (2001) for details.

where $\gamma \in [0,1]$ measures the strength of vertical linkage within the industry.

Firms in the upstream sector are assumed to differ in their productivity. Henceforth, we drop the variety indicator i and distinguish firms by their productivity parameter $\varphi \in [1,\infty)$. To produce a quantity $y(\varphi)$, a firm with productivity φ requires the composite input level of

$$v(\varphi) = \alpha + \frac{y(\varphi)}{\varphi},$$ (5)

where α and $1/\varphi$ are the fixed and the variable input requirements, respectively. Denoting by P_v the unit composite input price, the input cost is

$$P_v v(\varphi) = Wl(\varphi) + Pz(\varphi),$$ (6)

which, upon minimisation subject to (4), yields the optimal unit cost

$$P_v = W^{1-\gamma} P^{\gamma}.$$ (7)

Applying Shepperd's lemma to $P_v v(\varphi)$ then yields the firm's demand for the two inputs which can be shown to imply

$$Wl(\varphi) = (1-\gamma) P_v v(\varphi),$$ (8)

$$Pz(\varphi) = \gamma P_v v(\varphi).$$ (9)

The firm chooses $p(\varphi)$ to maximise its profit, $\pi(\varphi) = p(\varphi) y(\varphi) - P_v v(\varphi)$, subject to its demand, input requirement and input price given by (2), (5) and (7) respectively. This yields the price setting rule

$$p(\varphi) = \frac{\sigma P_v}{(\sigma - 1)\varphi},$$ (10)

which, together with the definition of revenue, $r(\varphi) = p(\varphi) y(\varphi)$, can be used to rewrite profit as

$$\pi(\varphi) = r(\varphi)/\sigma - \alpha P_v.$$ (11)

Following Melitz (2003), we assume that there is a competitive pool, F, of firms that can enter the upstream sector by paying a sunk cost f measured in terms of the final good. This investment enables entrants to draw their technology as embodied in the specific value of the productivity parameter φ. The draw is from a common population with a known p.d.f. $g(\varphi)$ defined over the support $[1, \infty)$ with a continuous cumulative distribution $G(\varphi)$. A potential entrant's consequent decision to enter the industry depends on the magnitude of its $\varphi \in [1, \infty)$ in relation to the threshold productivity φ^c which yields $\pi(\varphi^c) = 0$; φ^c acts as a cut-off in that $\pi(\varphi) < 0$ for all $\varphi \in [1, \varphi^c)$ while $\pi(\varphi) > 0$ for all $\varphi \in (\varphi^c, \infty)$. Given (11), the zero profit condition for the marginal firm therefore is

$$r(\varphi^c) = \sigma \alpha P_v.$$ (12)

It is known, prior to entry, that only a fraction $M = (1 - G(\varphi^c))F$ of potential entrants will succeed to survive where, ex-post, M is the mass of varieties available in the market. We therefore redefine the p.d.f. of the surviving (incumbent) firms over $\varphi \in [\varphi^c, \infty)$ by $\mu(\varphi) = \dfrac{g(\varphi)}{1 - G(\varphi^c)}$, which can then be used to obtain a measure of the aggregate productivity of the surviving firms, denoted by $\tilde{\varphi}$, as the weighted average of their productivity levels (see Melitz, 2003, for details),

$$\tilde{\varphi} = \left(\int_{\varphi^c}^{\infty} \varphi^{\sigma-1} \mu(\varphi) \, d\varphi \right)^{\frac{1}{\sigma-1}}.$$ (13)

The demand function in (2) and the price rule in (10) respectively imply $p(\tilde{\varphi}) y(\tilde{\varphi}) / p(\varphi^c) y(\varphi^c) = (p(\tilde{\varphi}) / p(\varphi^c))^{1-\sigma}$ and $p(\tilde{\varphi}) / p(\varphi^c) = \varphi^c / \tilde{\varphi}$ which, together with the definition of revenue, yield

$$\frac{r(\tilde{\varphi})}{r(\varphi^c)} = \left(\frac{\tilde{\varphi}}{\varphi^c}\right)^{\sigma-1}. \tag{14}$$

Using this result, all the relevant aggregate variables can then be expressed in terms of $\tilde{\varphi}$: for any firm-level variable $x(\varphi)$, $x(\tilde{\varphi})$ is the value corresponding to the firm with average productivity and $Mx(\tilde{\varphi})$ is aggregate value. The average productivity $\tilde{\varphi}$ can then be defined in terms of the cut-off productivity φ^c using an appropriate specific p.d.f. for φ. Postulating a Pareto distribution

$$G(\varphi) = 1 - \varphi^{-\kappa} \quad \text{and} \quad g(\varphi) = \kappa \varphi^{-(1+\kappa)}, \quad \varphi \in [1, \infty), \tag{15}$$

where G and g are the cumulative distribution and probability density functions, the shape parameter κ provides an inverse measure of dispersion: the higher is κ the, the lower is the density of firms at lower productivity levels and the more homogeneous are the firms.[9] Then, (15) implies

$$M = \left(\varphi^c\right)^{-\kappa} F, \tag{16}$$

and using (13) we obtain

$$\tilde{\varphi} = \left(\frac{\kappa}{1+\kappa-\sigma}\right)^{1/(\sigma-1)} \varphi^c. \tag{17}$$

We assume that the entry process continues until the expected *net* profit of entry is zero at the industry level, hence,

$$M\pi(\tilde{\varphi}) - PFf = 0. \tag{18}$$

Finally, the industry-level labour demand is

$$L = Ml(\tilde{\varphi}). \tag{19}$$

---

[9] To obtain meaningful results we impose $\kappa > \sigma - 1$.

2.1. The relationship between employment and output

Our main purpose in this paper is to examine how the relationship between industry-level labour demand and output is affected by the shape of the distribution of firms captured by κ. To this end, we reduce the model to the following 3 equations (see Appendix A.i) that determine $l(\tilde{\varphi})$, W/P and M which, for convenience, are written in logarithmic form as,

$$\ln l(\tilde{\varphi}) = \bar{\eta} - \gamma \ln\left(\frac{W}{P}\right), \tag{20}$$

$$\ln M - \frac{(1-\gamma)(\sigma-1)(\kappa-1)}{\lambda\kappa} \ln\left(\frac{W}{P}\right) = \bar{\mu}_u, \tag{21}$$

$$\ln M + (1-\gamma)\ln\left(\frac{W}{P}\right) = \bar{\mu}_d + \ln Y, \tag{22}$$

where $\bar{\eta}$, $\bar{\mu}_d$ and $\bar{\mu}_u$ are constant parameters that depend on $(\alpha, f, \gamma, \lambda, \sigma, \kappa)$. Equation (20) gives the labour demand for the firm with average productivity. Equations (21) and (22) are respectively derived from the zero-profit conditions for the upstream and downstream industries and represent the combinations of $\ln W/P$ and $\ln M$ which ensure these equilibrium conditions hold; for any given Y, their solution yields

$$\ln\left(\frac{W}{P}\right) = \bar{\omega} + \frac{\lambda\kappa}{(1-\gamma)(\lambda\kappa + (\sigma-1)(\kappa-1))} \ln Y \tag{23}$$

and

$$\ln M = \bar{\mu} + \frac{(\sigma-1)(\kappa-1)}{\lambda\kappa + (\sigma-1)(\kappa-1)} \ln Y \tag{24}$$

where $\bar{\omega}$ and $\bar{\mu}$ are constant parameters that depend on $(\alpha, f, \gamma, \lambda, \sigma, \kappa)$. Thus, while both W/P and M are increasing in Y, the impact of Y depends on the distribution parameter κ. A ceteris paribus increase in κ reduces the impact of Y on W/P. Therefore, as equation (20) shows, a change in Y affects the firm-level employment via its impact on W/P and this impact will be smaller the more homogenous are the firms. In other words, the response of the average firm-

11

level labour demand to a change in the industry-level output demand in an industry which is populated with a relatively bigger number of smaller (or less productive) firms is smaller. This is in line with the evidence provided by Moscarini and Postel-Vinay (2012) − as shown in Appendix A.ii − and reflects employment adjustments at the intensive margin. However, the rise in κ has the opposite effect on the adjustments at the extensive margin: as equation (24) shows the larger is κ, the larger is the impact of Y on M. Intuitively, a fall in Y that increases the minimum productivity required to survive in the industry will trigger a larger shake out of firms when the industry has a higher density of low productivity firms. As a result, unless the effects at intensive and extensive margins exactly cancel each other out, κ will influence the impact of Y on industry employment, $L = Ml(\tilde{\varphi})$, and the influence will be positive if the extensive margin effect dominates. To see this, we substitute from (20), (23) and (24) into equation (19) to obtain our main result (see Appendix A.i) which, expressed in the form of elasticity, is

$$\partial \ln L = \beta(\kappa) \partial \ln Y, \tag{25}$$

where $\beta(\kappa) = 1 - \dfrac{\lambda\kappa}{(1-\gamma)\left[\lambda\kappa + (\sigma-1)(\kappa-1)\right]}$. It follows that $0 < \beta(\kappa) < 1$, provided that

$0 \le \gamma < \dfrac{1}{1 + \lambda\kappa/(\kappa-1)(\sigma-1)}$ and $0 < \lambda \le 1$ hold. In other words, as long as the extent of vertical linkages in the upstream sector (captured by γ) is not too large and there are some industry-wide scale economies (captured by a positive value of λ), the model predicts a positive relationship between employment and output changes at the industry level. Specifically, the effect of a change in output on employment depends on the productivity distribution of firms in the industry, captured by the shape parameter κ: the larger is κ, the higher the density of low productivity firms, and the bigger is the effect of an exogenous change in the industry's output

demand on its employment. As explained above, this reflects the dominance of adjustments of employment at the extensive margin over those at the intensive margin.

In sum, the model outlined in this section provides theoretical support for the conjecture that inter-industry differences in productivity distributions affect the responsiveness of employment to output changes. Specifically, it suggests that in industries characterised by, on average, larger and more productive firms and by a higher degree of firm heterogeneity, aggregate employment is more 'insulated' from output shocks. This is the main prediction of the model which motivates our empirical analysis in the next section.

3. Empirical estimations

3.1. Data and descriptive statistics

Firm-level data are required to investigate the effects of the firm size distribution empirically. We are able to use firm data for three countries, Germany, Sweden and the UK, which include information on employment and on the industry of the establishment.

German firm-level data is available from the Establishment History Panel (EHP). It provides information on the population of establishments in Germany. We have access to a randomly drawn sample covering 50 percent of the population, yielding information for about 800,000 plants per year over the period 1997 - 2011. The EHP is made available by the Institute for Employment Research (IAB).

Swedish data is provided by Statistics Sweden and covers the population of all registered firms. From 1997-2012, around 170,000 unique companies are included in the data. Unlike data from Germany and the UK, the Swedish data is on firms rather than on plants. However, 78 percent of Swedish firms are single-plant firms.

Firm-level data for the UK is from the Annual Respondents Database (ARD) and includes data from the Inter-Departmental Business Register (IDBR), which is the key sampling frame

for UK business statistics used by the Office for National Statistics (ONS). As the German EHP, the ARD is essentially a census of UK businesses, it contains about 70 thousand enterprises.[10] With the addition of data from the IDBR, the dataset contains around 3.7 million enterprises. The sample used covers the period 1997-2007.

Firm-level data from Germany, Sweden and the UK are used to construct a measure of the size distribution of firms within industries. The unit of observations is country specific 3-digit NACE manufacturing industries over the period 1999-2007.[11] We have chosen these three countries because they are characterised by different welfare state regimes and labour market institutional settings as well as by different sectoral structures – factors that can influence the responsiveness of employment to output changes.

One aspect of the country differences can be seen in Figure 1, which illustrates the relative size of different aggregated 2-digit industries in the three countries.[12] Food, Beverage, and Tobacco (DA) is the largest industry in the UK accounting for around 17 percent of total manufacturing employment, substantially more than the shares in Germany and Sweden. Another large industry in the UK is Paper products (DE), which is large also in Sweden but not in Germany. Germany and Sweden have both their largest employment shares in Machinery (DK) with around 18 percent of total manufacturing employment.

-- Figure 1 around here --

Another difference between these countries is in the type of welfare states and labour market regulations, which are prominent among the explanations advanced for the existing

---

[10] The ARDis described by Barnes (2002).
[11] There are 97 industries included in our sample. See the appendix for more information on the data and variables. Since the available firm level data for the UK currently covers only up to 2008 (data for 2008 is incomplete), a consistent sample for the three countries cannot be extended beyond 2007.
[12] Again, our econometric analysis is conducted at the more disaggregated 3-digit level of NACE.

inter-country differences in employment performance (e.g., Perman and Stephan, 2015; Hoffmann and Lemiueux, 2014). As noted, these three countries correspond to three types of social models within the European Union: the Anglo-Saxon (UK), the Central European (Germany) and the Scandinavian (Sweden). These social models substantially differ in terms of labour market institutions and legislation, particularly with respect to employment protection, unemployment benefits, minimum wages or the role of unions. Another key difference is their reliance on active labour market policies. For instance, expenditures on active labour market policies during 1997 to 2007 averaged 3.1 percent of GDP in Germany, 2.6 percent in Sweden and 0.6 percent in the UK.[13]

For each industry, in each country, and for each year, we estimate a Pareto shape parameter for firm size distribution which we then use as a proxy for the parameter κ that features in the above theoretical analysis.[14] We obtain the maximum likelihood estimates of κ based on the C.D.F. $F(s) = 1 - (s/s_0)^{-\kappa}$, $0 < s_0 \leq s$, where s denotes the firm size (number of employees) and s_0 is the lower bound. We define $x_i = \dfrac{s_i}{s_0}$ and use it to write the joint likelihood function $L(\kappa, s_0) = \prod_{i=1}^{N} \kappa s_0^{\kappa} x_i^{-(1+\kappa)} = \kappa^N s_0^{N\kappa} \prod_{i=1}^{N} x_i^{-(1+\kappa)}$ which yields the log-likelihood function $\ln L(\kappa, s_0) = N \ln \kappa + N\kappa \ln s_0 - (1+\kappa) \sum_{i=1}^{N} \ln x_i$. It follows that $\hat{s}_0 = \min(x_i)$ and the solution to the first order condition, $\dfrac{\partial \ln L}{\partial \kappa} = N/\kappa + N \ln s_0 - \sum_{i=1}^{N} \ln x_i = 0$, yields $\hat{\kappa} = N \left[\sum_{i=1}^{N} \ln (x_i / x_1) \right]^{-1}$.

---

[14] We use firm size as a proxy for productivity, consistent with both the one-to-one correspondence between size (in terms of number of employees) and productivity emerging from the theory – see A.ii in the Appendix – and the high correlation observed empirically between firm size and productivity. It is observed in the literature that firms' distributions of employees and sales can be closely approximated by a Pareto distribution, except in the region of very small firms. Accordingly, the Pareto distribution of firm size is subject to bias if entrants are disproportionately small and their share is large. Bernard *et al.* (2015) show that the Pareto fit becomes more robust once the data are corrected for the 'partial year' effect, as this decreases the share of small firms.

The asymptotic standard errors for $\hat{s}_0$ and $\hat{\kappa}$ are obtained using a standard bootstrapping approach.

The distribution of firms is more homogenous and has a higher concentration of small firms if the shape parameter is larger. Accordingly, a smaller shape parameter implies a relatively more heterogeneous distribution of firms, i.e. with a thicker tail of large firms. In Figure 2 we show the aggregated annual Pareto shape parameter estimates for our three countries over the period 2000-2007. The series are normalized with the Pareto distribution in the UK in 2000 set to unity. A few conclusions can be drawn. Firstly, there is a large difference between the UK on the one hand, and Sweden and Germany on the other: the Pareto shape parameter is substantially higher in the UK. This means that the UK firms are relatively more homogeneous with a relatively large presence of small firms. The distributions in Sweden and Germany are similar in the first years of observations but there is a divergence over time with the German distribution shifting towards larger firms (i.e. its Pareto parameter becomes smaller). Moreover, the distributions are relatively stable over time with only a modest change towards less heterogeneity in the UK, and the above mentioned change towards more heterogeneity in Germany. There is practically no change in the aggregate distribution for Sweden.

-- Figure 2 about here --

Figure 3 highlights the differences for three industries, selected on the basis that they represent different types of activities. The jewellery industry involves very diverse firms including ones with mining activities at one end of the spectrum and small retailers at the other. The medical equipment industry consists primarily of high precision manufacturing firms with diverse technological features. Paper is a more conventional capital intensive industry that

16

comprises firms with large forestry activities, which are specialised in the production of pulp, as well as smaller firms which focus on producing highly specific types of paper, e.g. for cutting or packaging. We include two measures in Figure 3: the Employment/Output ratios and the Pareto parameters.

A few conclusions can be drawn from the above discussion. Firstly, the Employment/Output ratios differ between countries: it tends to be lowest in Germany and highest in the UK. It also differs between industries with the capital-intensive paper industry having the lowest ratio. Secondly, the Employment/Output ratios are relatively stable in Germany and relatively volatile in the UK, with Sweden as an intermediate case. The size distributions differ substantially across industries and countries, with the highest value of κ in the jewellery industry. Germany has the lowest concentration at small firm sizes in two out of three industries. The distributions are relatively stable over time in most industries and countries. One exception is the jewellery industry in the UK and Sweden where we see a higher density of small firms in later years.

-- Figure 3 about here --

3.2. Testing the key theoretical prediction

The key prediction of our theoretical model, as summarised in equation (25), is that the impact of output on employment depends on the shape of the firm size distribution, with the output elasticity of employment being larger the larger is κ. In its most basic form, this relationship can be represented by the reduced form panel regression equation

$$\Delta \ln L_{ict} = \left(\beta_1 + \beta_2 \hat{\kappa}_{ict} \right) \Delta \ln Y_{ict} + \varepsilon_{ict} , \tag{26}$$

where L_{ict} and Y_{ict} are employment and output in industry i, country c, in year t; $\hat{\kappa}_{ict}$ is an estimate of the Pareto shape parameter for firm size distribution; and ε_{ict} is an idiosyncratic disturbance term (once it is corrected for the industry, country and time specific fixed effects).

As a generalisation, we add a vector of regressors to control for country-specific factors. Specifically, to capture the effects of labour market institutional factors and the role of government in the economy, we include measures of: strictness of employment protection policies, union density, active and passive labour market policies, and government size. We also include trade openness, since existing evidence suggests that it affects competitive forces and has selection and reallocation effects.[15] Finally, we also include $\hat{\kappa}_{ict}$ as an additional regressor to capture the 'independent effect' of any shifts in firm heterogeneity beyond that exerted via its interaction with $\Delta \ln Y_{ict}$. We therefore estimate the more general regression equation

$$\Delta \ln L_{ict} = \beta_0 + \beta_1 \Delta \ln Y_{ict} + \beta_2 \hat{\kappa}_{ict} \Delta \ln Y_{ict} + \beta_3 \hat{\kappa}_{ict} + \sum_j \beta_{4+j} z_{jct} + \mu_i^F + \mu_c^F + \mu_t^F + \varepsilon_{ict} , \quad (27)$$

where z_{jct} refers to the jth country-specific control variable and μ_i^F, μ_c^F and μ_t^F are industry, country and time fixed effects to control for unobservable along these dimensions. We expect to find $\hat{\beta}_1 > 0$ and $\hat{\beta}_2 > 0$. The coefficient β_3, which captures any direct and independent effect that the firms' size distribution might exert on employment, cannot be analytically signed in our model. Nevertheless, it would not be implausible to expect a priori to find $\beta_3 < 0$ since, as the literature suggests, an industry with a higher degree of heterogeneity is characterised *ceteris paribus* by higher entry and a larger mass of surviving firms, and hence higher aggregate employment.

---

[15] See, e.g., Tybout (2003) for a survey and Corcos *et al.* (2012) among others.

3.3. Results

The variables and datasets we use in our regressions are explained in detail in Appendix A.iii.   Our estimation results are reported in Table 1. The first column of Table 1 includes output growth as the only explanatory variable together with the fixed effect (industry, time and country) dummies, and shows that the elasticity of employment with respect to output is 0.39: a ten percent increase in industry-level output increases employment by about 4 percent. The Pareto shape parameter is added as an independent regressor in column 2 and the negative and significant sign of its coefficient estimate confirms our conjecture that industries populated by more homogenous (and smaller) firms exhibit lower employment growth. Column 3 adds our main explanatory variable of interest, the interaction term between output growth and the firm-size distributions parameter $\hat{\kappa}_{ict} \Delta \ln Y_{ict}$. Its positive and statistically significant effect confirms our main theoretical conjecture that a change in output has a larger effect on employment in industries with a more homogenous distribution of firms. Moreover, while including the interaction term does not alter the independent impact of $\hat{\kappa}$, it reduces the size of the direct effect of output growth; when measured directly, the elasticity of employment now falls from 0.38 to 0.25. The total elasticity, however, is given by $0.25 + 0.21\overline{\hat{\kappa}}$, where $\overline{\hat{\kappa}}$ is a sample-based measure of the shape parameter. Using the sample mean of $\overline{\hat{\kappa}} = 0.556$ implies the total elasticity of around 0.367 which is of the same magnitude as the estimates in the first and second columns of Table 1. The sub-sample means for Germany, Sweden and the UK respectively are $\overline{\hat{\kappa}}_G = 0.425$, $\overline{\hat{\kappa}}_S = 0.628$ and $\overline{\hat{\kappa}}_U = 0.628$ which indicate that German industries are likely to show, on average, a smaller response of employment to output compared with their Swedish and British counterparts.

Finally, we include as additional regressors the lagged values of country-specific control variables in the last column. We have included the lagged, rather than the current, values of country-specific control variables since it is likely that it takes time for the policies to have an

impact on employment. In addition, albeit in a crude way, it helps avoiding the simultaneity bias problem. The main results do not change, except for a drop in the statistical significance of the coefficient capturing the direct impact of the Pareto shape parameter whose point estimate remains the same but becomes statistically insignificant. Most of the control variables have statistically significant effects with the expected signs. More precisely, labour protection regulations and active labour market policies are associated with higher employment growth, and passive labour market policies and large public sector with decreases in employment. However, trade openness and the degree of unionisation of the labour force do not significantly affect employment growth. On the whole, the regressions have reasonable fits with the adjusted goodness of fit measure between 27% and 29% and both the log-likelihood values and the AIC support the most general specification as proposed by equation (27).

-- Table 1 about here --

We continue with some robustness checks by estimating alternative specifications of equation (27). These are reported in Table 2. In the first column we use the current instead of the lagged values of the control variables and find these not to alter the main results; the estimates and statistical significance of β_1, β_2 and β_3 are almost identical to the ones reported in the last column of Table 1. Next, we experiment with a different measure of trade openness. As previously discussed, openness has been suggested to affect competitive forces and employment but in the results reported in Table 1 our country-level measure on trade openness has no statistically significant effect. We therefore re-examine the role of openness further in the last two columns by including an industry-specific share of trade in production in the second column of Table 2, and including both industry-specific and aggregate openness in the last column. While neither of the openness variables have a statistically significant effect, including

20

the industry-specific measure eliminates the direct effect of output growth while inflating the effect output exerts via its interaction with the Pareto shape parameter. On the grounds that output is likely to be the most crucial determinant of employment, we can only interpret this as a spurious result.

4. Summary and conclusions

In this paper we have shown that the degree of firm heterogeneity is a channel that can contribute to explaining the observed differences in the output-employment relationship at the industry level. Within a theoretical model in which firms are characterised by heterogeneous productivities that follow a Pareto distribution, we have shown that employment adjustments at both the intensive and extensive margins depend on the shape parameter of the distribution. Specifically, the effect of output changes on employment changes are shown to depend negatively on the extent of firm-dispersion in the industry: the larger is the Pareto coefficient and the more homogenous is the distribution of firms, the larger is the impact of output changes on employment changes. Accordingly, employment is perceived to be more 'insulated' from output shocks in those industries that have (on average) larger and more productive firms.

To examine the empirical validity of these theoretical priors, we have estimated industry-specific Pareto shape parameters of the firm size distributions using firm-level data from Germany, Sweden and the UK, and then used these estimates to augment a relationship between industry-level employment and output. Estimates based on the available cross-country cross-industry data support our theoretical conjecture and confirm that the firm-size distribution provides a channel for the transmission of output shocks and that intra-industry reallocations.

Previous literature has pointed to cross-country differences in labour market institutions as an important reason for differences in employment volatility. This paper complements this

literature by highlighting the role of one particular industry characteristic: the size distribution of firms. Our results suggest that a deeper understanding of employment volatility needs to take both country and industry specific factors into account; in so doing, they point to the importance of making a more nuanced distinction between labour market and industrial policies. Our paper provides a first step in this direction and hopefully offers important insights for further research and to policymakers alike.

References

Anderson, S.P. and De Palma, A., 2015, Economic distributions and primitive distributions in monopolistic competition, CEPR Discussion Paper No. 10748.

Arkolakis, C., Esposito, F. 2015, Endongenous labour supply and the gains from international trade, mimeo, Yale University.

Arkolakis C., Klenow P., Demidova S., Rodriguez-Clare A., 2008, Endogenous variety and the gains from trade. American Economic Review. Papers and Proceedings. 98(2): 444-50.

Arkolakis, C., Costinot, A., Rodriguez-Clare, A, (2012), New trade models, same old gains? American Economic Review 102 (1), 94-130.

Axtell R.L., 2001, Zipf distribution of US firm sizes, Science, 293:1818-20.

Barnes, M. (2002), Annual Respondents Database User Guide, Office of National Statistics, London.

Bartelsman, E., Haltiwanger, J., Scarpetta, S., 2013, Cross-country differences in productivity: the role of allocative efficiency, American Economic Review, 103(1): 305-334.

Bernard, A.B., Massari, R., Reyes, J.-D., Taglioni, D., 2014, Exporter dynamics, firm size and growth, and partial year effects, NBER Working Paper No. 19865.

Bertoletti, P., Etro, F., Simonovska, I., 2017, International trade with indirect additivity, American Economic Journal: Microeconomics (Forthcoming).

Caves, R., 1998, Industrial organization and new findings on the turnover and mobility of firms, Journal of Economic Literature, Vol.36, pp.1947–82.

Chaney T., 2008, Distorted gravity: the intensive and extensive margins of international trade. American Economic Review. 98(4): 1707-1721.

Corcos, G., Del Gatto, M., Mion, G., Ottaviano, G.I.P., 2012, Productivity and firm selection: quantifying the 'New' Gains from Trade, The Economic Journal, Vol.122, 561, 754-798.

Di Giovanni J., Levchenko, A.A., 2012, Country size, international trade, and aggregate fluctuations in granular economies, Journal of Political Economy, Vol.120, No.6, pp. 1083-1132.

Di Giovanni, J., Levchenko, A.A., 2014, Firms, destinations and aggregate fluctuations, Econometrica, Vol.82, No.4, pp. 1303-1340.

Elsby M.W.L., Michaels, R., 2013, Marginal jobs, heterogeneous firms, and unemployment flows, American Economic Journal: Macroeconomics, 5(1): 1-48.

Felbermayr G., Prat J., Schmerer H., 2011, Globalization and labor market outcomes: wage bargaining, search frictions, and firm heterogeneity, Journal of Economic Theory, 146(1): 39-73.

Foster, L., Haltiwanger, J., Syverson, C., 2008, Reallocation, firm turnover, and efficiency: selection on productivity or profitability?, American Economic Review, 98(1), 394-425, March.

Gabaix, X, 2011, The granular origin of aggregate fluctuations, Econometrica, Vol. 79, No. 3, 733-772.

Haltiwanger, J., Jarmin, R.S., Miranda, J., 2013, Who creates jobs? Small versus large versus young, The Review of Economics and Statistics, Vol XCV, No.2, 347-361.

Head, K., Mayer, T., Thoenig, M., 2015, Welfare and trade without Pareto, American Economic Review: Papers and Proceedings 2014, 104(5): 310-316.

Helpman E., Itskhoki, O., 2010, Labor market rigidities, trade and unemployment, Review of Economic Studies, 77 (3): 1100-1137.

Helpman E., Melitz M.J., Yeaple S.R., 2004, Export versus FDI with heterogeneous firms, American Economic Review, 94(1): 300-16.

Hoffmann, F., Lemieux, T., 2014, Unemployment in the Great Recession: a comparison of Germany, Canada and the United States, National Bureau of Economic Research Working Paper No. 20694.

Hopenhayn, H., Rogerson, R., 1993, Job turnover and policy evaluation: a general equilibrium analysis, Journal of Political Economy, 101(5): 915-38.

Mrázová, M., Neary, J.P., Parenti, M., 2105, Technology, demand, and the size distribution of firms, Department of Economics Discussion Paper Number 774.

Melitz, M.J., 2003, The impact of trade on intra-industry reallocations and aggregate industry productivity, Econometrica, 71, Issue 6, 1695-1725.

Molana, H., Montagna, C., 2017, Selection, trade and employment: the strategic use of subsidies, mimeo, University of Aberdeen.

Montagna, C., 2001, Efficiency gaps, love of variety and international trade, Economica, Vol 68, no. 269, pp. 27-44.

Montagna, C., Nocco, A., 2013, Unionization, international integration, and selection, Canadian Journal of Economics, Vol. 46, No1, 23-45.

Montagna, C., Nocco, A., 2015, (De)Unionization, trade liberalisation and selection, Metroeconomica, 66:1, 91-122.

Moscarini, G., Postel-Vinay, F., 2012, The contribution of large and small employers to job creation in times of high and low unemployment, American Economic Review, 102, 2509-2539.

Neumark, D., Wall, B., Zhang, J., 2011, Do small businesses create more jobs? New evidence for the United States from the National Establishment Time Series, The Review of Economics and Statistics, 93, 16-29.

Nigai, S., 2017, A tale of two tails: productivity distribution and the gains from trade, Journal of International Economics, Volume 104, 44-62.

Olley, S., Pakes, A., 1996, The dynamics of productivity in the telecommunications equipment industry, Econometrica, Vol.64, pp.1263–97.

Perman, R., Stephan, G., Tavéra, C., 2015, Okun's Law – a meta-analysis, The Manchester School, vol. 83(1): 101-126.

Restuccia, D., Rogerson, R., 2008, Policy distortions and aggregate productivity with heterogeneous establishments, Review of Economic Dynamics, vol. 11, Issue 4, 707-720.

Stella, A., 2015, Firm dynamics and the origins of aggregate fluctuations, Journal of Economic Dynamics and Control, 55, 71-88.

Syverson, C., 2011, What determines productivity?, Journal of Economic Literature, 49, 326-365

Tybout, J.R., 2003, "Plant and firm level evidence on the 'New' Trade Theories", in E.K. Choi and J. Harrigan, eds. *Handbook of International Trade*, Blackwell, Malden, 388-415.

Appendix

A.i Equations of the model and derivation of employment-output relationship

The equations below are repeated for convenience and the numbers in square brackets after an equation's description refers to the corresponding number in the main text.

$$r(\tilde{\varphi}) = p(\tilde{\varphi}) y(\tilde{\varphi}) \qquad \text{definition of revenue of firms with average productivity} \qquad \text{(A1)}$$

$$Mr(\tilde{\varphi}) = PY \qquad \text{zero profit condition in downstream sector} \qquad \text{(A2)}$$

$$P = M^{-\frac{\lambda}{\sigma-1}} p(\tilde{\varphi}) \qquad \text{aggregate CES price index [3]} \qquad \text{(A3)}$$

$$v(\tilde{\varphi}) = \alpha + \frac{y(\tilde{\varphi})}{\tilde{\varphi}} \qquad \text{input requirement of firms with average productivity [5]} \qquad \text{(A4)}$$

$$P_v = P^\gamma W^{1-\gamma} \qquad \text{unit input cost [7]} \qquad \text{(A5)}$$

$$Wl(\tilde{\varphi}) = (1-\gamma) P_v v(\tilde{\varphi}) \qquad \text{labour demand of firms with average productivity [8]} \qquad \text{(A6)}$$

$$p(\tilde{\varphi}) = \frac{\sigma P^v}{(\sigma-1)\tilde{\varphi}} \qquad \text{price-markup rule of firms with average productivity [10]} \qquad \text{(A7)}$$

$$\pi(\tilde{\varphi}) = r(\tilde{\varphi})/\sigma - \alpha P_v \qquad \text{profit of firms with average productivity [11]} \qquad \text{(A8)}$$

$$r(\varphi^c) = \sigma \alpha P_v \qquad \text{revenue of marginal firms [12]} \qquad \text{(A9)}$$

$$\frac{r(\tilde{\varphi})}{r(\varphi^c)} = \left(\frac{\tilde{\varphi}}{\varphi^c}\right)^{\sigma-1} \qquad \text{relationship between the marginal and average firms' revenue [14]} \qquad \text{(A10)}$$

$$M = (\varphi^c)^{-\kappa} F \qquad \text{mass of successful entrants [16]} \qquad \text{(A11)}$$

$$\tilde{\varphi} = \left(\frac{\kappa}{1+\kappa-\sigma}\right)^{1/(\sigma-1)} \varphi^c \qquad \text{relationship between the marginal and average productivities [17]} \qquad \text{(A12)}$$

$$M\pi(\tilde{\varphi}) = PFf \qquad \text{zero profits condition in upstream sector [18]} \qquad \text{(A13)}$$

$$L = Ml(\tilde{\varphi}) \qquad \text{aggregate labour demand [19]} \qquad \text{(A14)}$$

For any given value of Y and the parameters $(\alpha, f, \gamma, \lambda, \sigma, \kappa)$, equations (A1)-(A14) determine the values of 14 endogenous variables: L, W/P, F, M, $l(\tilde{\varphi}), v(\tilde{\varphi}), y(\tilde{\varphi}), p(\varphi^c)$, $\pi(\tilde{\varphi}), r(\tilde{\varphi}), r(\varphi^c), \tilde{\varphi}, \varphi^c, P_v$. It is however more informative to reduce the above equations to 4 equations determining $l(\tilde{\varphi})$, W/P, M and L given by equations (A18), (A22), (A23) and (A24)

below. To obtain these, first we use (A1), (A4), (A5), (A7), (A9) and (A11) to eliminate $r(\tilde{\varphi})$, $v(\tilde{\varphi})$, P_v, $p(\tilde{\varphi})$, $r(\varphi^c)$ and F. Then we use (A12) to substitute out φ^c. The resulting substitutions in (A3), (A6), (A8) and (A10) can then be shown to yield the following solutions for $\tilde{\varphi}$, $y(\tilde{\varphi})$, $\pi(\tilde{\varphi})$ and $l(\tilde{\varphi})$ in terms of W/P and M:

$$\tilde{\varphi} = \frac{\sigma}{\sigma-1} M^{-\frac{\lambda}{\sigma-1}} \left(\frac{W}{P}\right)^{1-\gamma}, \tag{A15}$$

$$y(\tilde{\varphi}) = \frac{\alpha\sigma\kappa}{\kappa+1-\sigma} M^{-\frac{\lambda}{\sigma-1}} \left(\frac{W}{P}\right)^{1-\gamma}, \tag{A16}$$

$$\pi(\tilde{\varphi}) = \frac{\alpha(\sigma-1)}{\kappa+1-\sigma} \left(\frac{W}{P}\right)^{1-\gamma}, \tag{A17}$$

$$l(\tilde{\varphi}) = \frac{\alpha(1-\gamma)(\kappa\sigma+1-\sigma)}{\kappa+1-\sigma} \left(\frac{W}{P}\right)^{-\gamma}. \tag{A18}$$

Equation (20) in the paper corresponds to (A18).

Then we substitute from the above solutions in the two remaining equations, A(2) and (A13), which are the zero profit conditions in the downstream and upstream sectors respectively, to obtain two equations in terms of W/P and M:

$$M = \bar{M}_u \left(\frac{W}{P}\right)^{\frac{(1-\gamma)(\sigma-1)(\kappa-1)}{\lambda\kappa}}, \tag{A20}$$

$$M = \bar{M}_d Y \left(\frac{W}{P}\right)^{-(1-\gamma)}, \tag{A21}$$

where $\quad \bar{M}_u = \dfrac{\kappa+1-\sigma}{\alpha\sigma\kappa} \quad$ and $\quad \bar{M}_d = \left(\dfrac{f}{\alpha}\right)^{\frac{\sigma-1}{\lambda\kappa}} (\kappa)^{-\frac{1}{\lambda}} (\sigma)^{\frac{\sigma-1}{\lambda}} (\sigma-1)^{-\frac{(\kappa+1)(\sigma-1)}{\lambda\kappa}} (\kappa+1-\sigma)^{\frac{\kappa+\sigma-1}{\lambda\kappa}}.$

Equations (21) and (22) in the paper correspond to (A20) and (A21), respectively.

Next, solving (A20) and (A21) determines W/P and M for any given level of Y:

$$\frac{W}{P} = \left(\frac{\bar{M}_d Y}{\bar{M}_u}\right)^{\frac{\lambda\kappa}{(1-\gamma)[\lambda\kappa+(\sigma-1)(\kappa-1)]}}, \tag{A22}$$

$$M = \left(\bar{M}_u\right)^{\frac{\lambda\kappa}{\lambda\kappa+(\sigma-1)(\kappa-1)}} \left(\bar{M}_d\right)^{\frac{(\sigma-1)(\kappa-1)}{\lambda\kappa+(\sigma-1)(\kappa-1)}} Y^{\frac{(\sigma-1)(\kappa-1)}{\lambda\kappa+(\sigma-1)(\kappa-1)}}. \tag{A23}$$

Equations (23) and (24) in the paper correspond to (A22) and A(23) respectively.

Finally, we then use (A18), (A22) and (A23) to eliminate $l(\tilde{\varphi})$, W/P and M from (A14) to obtain

$$L = \frac{\alpha(1-\gamma)(\kappa\sigma+1-\sigma)}{\kappa+1-\sigma}(\bar{M}_u)^{\frac{\lambda\kappa}{(1-\gamma)[\lambda\kappa+(\sigma-1)(\kappa-1)]}}(\bar{M}_d)^{1-\frac{\lambda\kappa}{(1-\gamma)[\lambda\kappa+(\sigma-1)(\kappa-1)]}}Y^{1-\frac{\lambda\kappa}{(1-\gamma)[\lambda\kappa+(\sigma-1)(\kappa-1)]}}. \qquad (A24)$$

The elasticity expression in equation (25) in the paper is based on (A24).

A.ii The relationship between employment of firms with average and marginal productivity

In the absence of data on firm-level productivity, in our empirical investigation we have approximated firms' productivity distribution by their size distribution. We have justified this on the basis of the evidence in the literature that larger firms are found to be more productive. To see that this in fact holds in our model, we compare the employment of firms with average and marginal productivity. Given that in general a firm's employment in the model is given by

$l(\varphi) = (1-\gamma)v(\varphi)\left(\dfrac{W}{P}\right)^{-\gamma}$, where $v(\varphi) = \alpha + \dfrac{y(\varphi)}{\varphi}$ (see equations (5), (7) and (8) in the paper), we can write,

$$l(\varphi) = (1-\gamma)\left[\alpha + \frac{y(\tilde{\varphi})}{\tilde{\varphi}}\right]\left(\frac{W}{P}\right)^{-\gamma} = (1-\gamma)\left[\alpha + \left(\frac{y(\tilde{\varphi})}{y(\varphi^c)}\frac{\varphi^c}{\tilde{\varphi}}\right)\frac{y(\varphi^c)}{\varphi^c}\right]\left(\frac{W}{P}\right)^{-\gamma}$$

which, using (A10), (A12), $p(\tilde{\varphi})/p(\varphi^c) = \varphi^c/\tilde{\varphi}$ and $v(\varphi^c) = \alpha + \dfrac{y(\varphi^c)}{\varphi^c}$, can be rewritten as

$l(\tilde{\varphi}) = (1-\gamma)\left[\alpha + \left(\dfrac{\kappa}{\kappa-(\sigma-1)}\right)\left(v(\varphi^c)-\alpha\right)\right]\left(\dfrac{W}{P}\right)^{-\gamma}$. Next, $l(\varphi^c) = (1-\gamma)v(\varphi^c)\left(\dfrac{W}{P}\right)^{-\gamma}$ can

be used to obtain $l(\tilde{\varphi}) = \left(\dfrac{\kappa}{\kappa-(\sigma-1)}\right)l(\varphi^c) - \alpha\left(\dfrac{\kappa}{\kappa-(\sigma-1)}-1\right)(1-\gamma)\left(\dfrac{W}{P}\right)^{-\gamma}$. However, we

also know that, due to the zero profit condition, the marginal firms' input requirement is

constant, i.e., $v(\varphi^c) = \alpha + \dfrac{y(\varphi^c)}{\varphi^c} = \alpha + \dfrac{p(\varphi^c)y(\varphi^c)}{\varphi^c p(\varphi^c)}$. This, upon substitution from (A9) and

taking into account $p(\varphi^c) = \dfrac{\sigma P^v}{(\sigma-1)\varphi^c}$, implies $v(\varphi^c) = \alpha\sigma$ and thus

$l(\varphi^c) = \alpha\sigma(1-\gamma)\left(\dfrac{W}{P}\right)^{-\gamma}$. Using this and the previous equation, we have the required result, namely,

$$\frac{l(\tilde{\varphi})}{l(\varphi^c)} = 1 + \frac{(\sigma-1)^2}{\sigma(\kappa-(\sigma-1))} > 1 . \tag{A25}$$

The above result, which implies $dl(\tilde{\varphi}) > dl(\varphi^c)$, is consistent with the evidence reported in Moscarini and Postel-Vinay (2012), that the response over the business cycle of employment to shocks in larger firms is relatively higher. This result also holds for the variable input requirement. To see this, we compare how employment responds to changes in variable input requirement in firms with average and marginal productivity. Denoting the employment associated with the variable input requirement of a firm by $l^{VC}(\varphi)$, (A4) and (A6) imply

$l^{VC}(\tilde{\varphi}) = (1-\gamma)\left(\dfrac{y(\tilde{\varphi})}{\tilde{\varphi}}\right)\left(\dfrac{W}{P}\right)^{-\gamma}$, which can be written in terms of the employment associated with the variable input requirement of the marginal firm by noting that

$l^{VC}(\tilde{\varphi}) = (1-\gamma)\left(\dfrac{y(\tilde{\varphi})}{y(\varphi^c)}\dfrac{\varphi^c}{\tilde{\varphi}}\right)\dfrac{y(\varphi^c)}{\varphi^c}\left(\dfrac{W}{P}\right)^{-\gamma}$. As above, equations (A10), (A12) and

$p(\tilde{\varphi})/p(\varphi^c) = \varphi^c/\tilde{\varphi}$ can be used to write $\dfrac{y(\tilde{\varphi})}{y(\varphi^c)}\dfrac{\varphi^c}{\tilde{\varphi}} = \dfrac{\kappa}{\kappa-(\sigma-1)}$ which, when used together

with $l^{VC}(\varphi^c) = (1-\gamma)\left(\dfrac{y(\varphi^c)}{\varphi^c}\right)\left(\dfrac{W}{P}\right)^{-\gamma}$, yields

$$\frac{l^{VC}(\tilde{\varphi})}{l^{VC}(\varphi^c)} = \frac{\kappa}{\kappa-(\sigma-1)} > 1 , \tag{A26}$$

which implies that $dl^{VC}(\tilde{\varphi}) > dl^{VC}(\varphi^c)$, i.e.: a larger firm has a larger response to an exogenous aggregate output shock.

A.iii Data: definitions and sources

Data on the main variables are as follows:

Y: *Measure of industry output*

> For each country, the observations were constructed by deflating the nominal annual production values (sbs_na_2a_dade) from Eurostat for industry aggregates by the Eurostat GDP (Euro) deflator (nama_05_gdp_p).

L: *Measure of employment, total industry employment*

For Sweden and the UK, this was obtained by aggregating the firm-level employment for each industry. For Germany, this was obtained by aggregating full-time equivalent employees for each industry. The minimum firm size is 1 employee, and the minimum industry size is 10 firms.

UK source: Office for National Statistics. (2012). Annual Respondents Database, 1973-2008: Secure Access. [data collection]. 3rd Edition. UK Data Service. SN: 6644.
Sweden source: Statistics Sweden. (2015). Business Statistics Database, 1986-2012: Secure Access. [data collection]. Microdata Online Access.
Germany source: Institute for Employment Research,
http://fdz.iab.de/en/FDZ_Establishment_Data/Establishment_History_Panel.aspx

κ: *Measure of firms' size distribution*

Approximated by the Pareto shape parameter and estimated using the annual firm-level employment for each industry.

The definition, measurement and data source for the additional explanatory variables are as follows:

Employment Protection Policies: Measure of overall strictness of employment protection. Scale 0 to 6 representing least to most stringent (source: OECD).

Union Density: Measure of union density. Ratio of wage and salary earners that are trade union members, divided by the total number of wage and salary earners (source: OECD).

Active Labour Market Policies: Measure of expenditure on active labour market policies. Total annual expenditure as a percentage of GDP (source: OECD).

Passive Labour Market Policies: Measure of expenditure on passive labour market policies. Total annual expenditure as a percentage of GDP (source: OECD).

Government Size: Measure of size of the public sector. Total annual government expenditure as a percentage of GDP (source: OECD).

Trade Openness: Measure of trade openness. Total annual value of imports and export as a percentage of GDP (source: OECD).

Trade Openness at the Industry Level: Measure of trade openness. Total annual value of industry-level imports and exports as a percentage of total industry-level production (source: Eurostat PRODCOM).

FIGURES AND TABLES

Figure 1. Average Share of Total Employment by Country and NACE rev.1
Manufacturing Subsections (2000-2007)
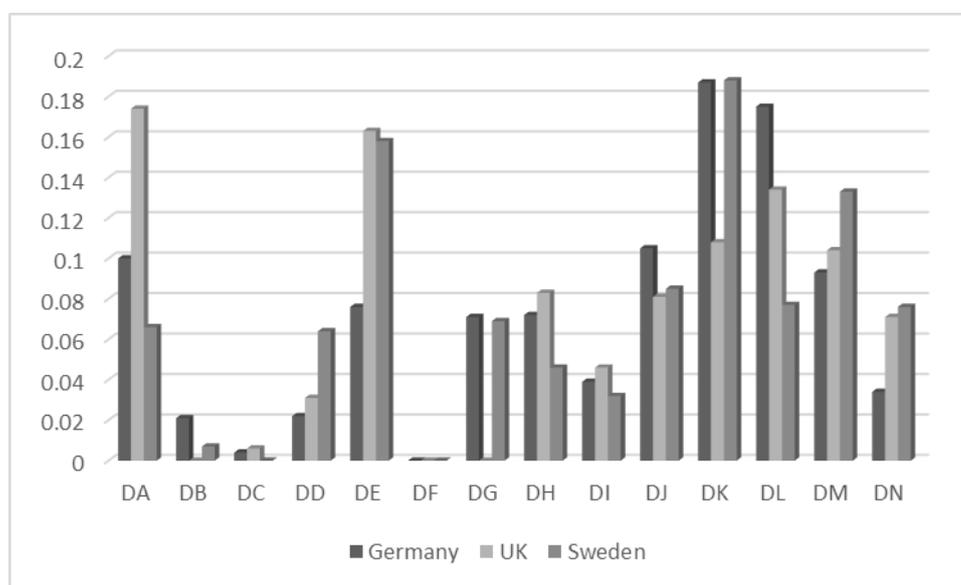
Note: Each share is calculated by dividing the 2-digit industry total employment by the total manufacturing sector employment. The sample is balanced within countries and shares for each country add one.

DA: Manufacture of food products, beverages and tobacco

DB: Manufacture of textiles and textile products

DC: Manufacture of leather and leather products

DD: Manufacture of wood and wood products

DE: Manufacture of pulp, paper and paper products; publishing and printing

DF: Manufacture of coke, refined petroleum products and nuclear fuel

DG: Manufacture of chemicals, chemical products and man-made fibres

DH: Manufacture of rubber and plastic products

DI: Manufacture of other non-metallic mineral products

DJ: Manufacture of basic metals and fabricated metal products
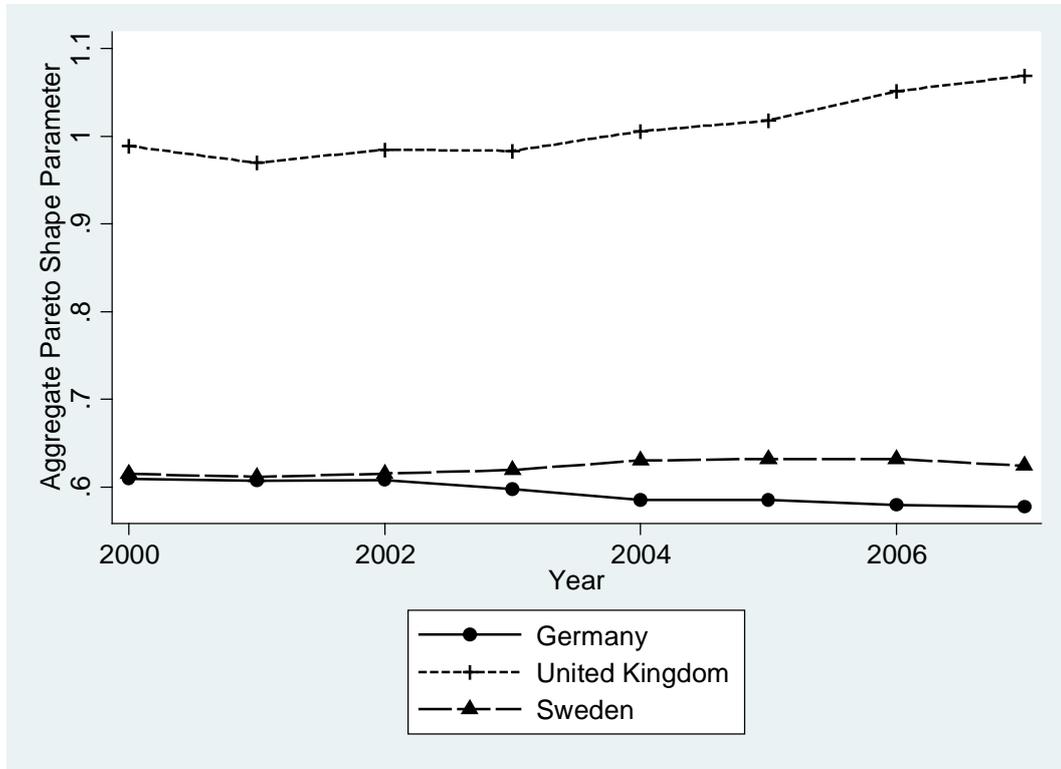
DK: Manufacture of machinery and equipment n.e.c.

DL: Manufacture of electrical and optical equipment

DM: Manufacture of transport equipment

DN: Manufacturing n.e.c.

Figure 2. Comparison of the Pareto Shape Parameter Estimates
(2000-2007, UK 2000 =1)

Maximum likelihood estimates. The sample for each year for each country includes all industries.

Figure 3. Inter-Industry and Inter-Country Differences in Employment/Output Fluctuations and Firm Size Distribution
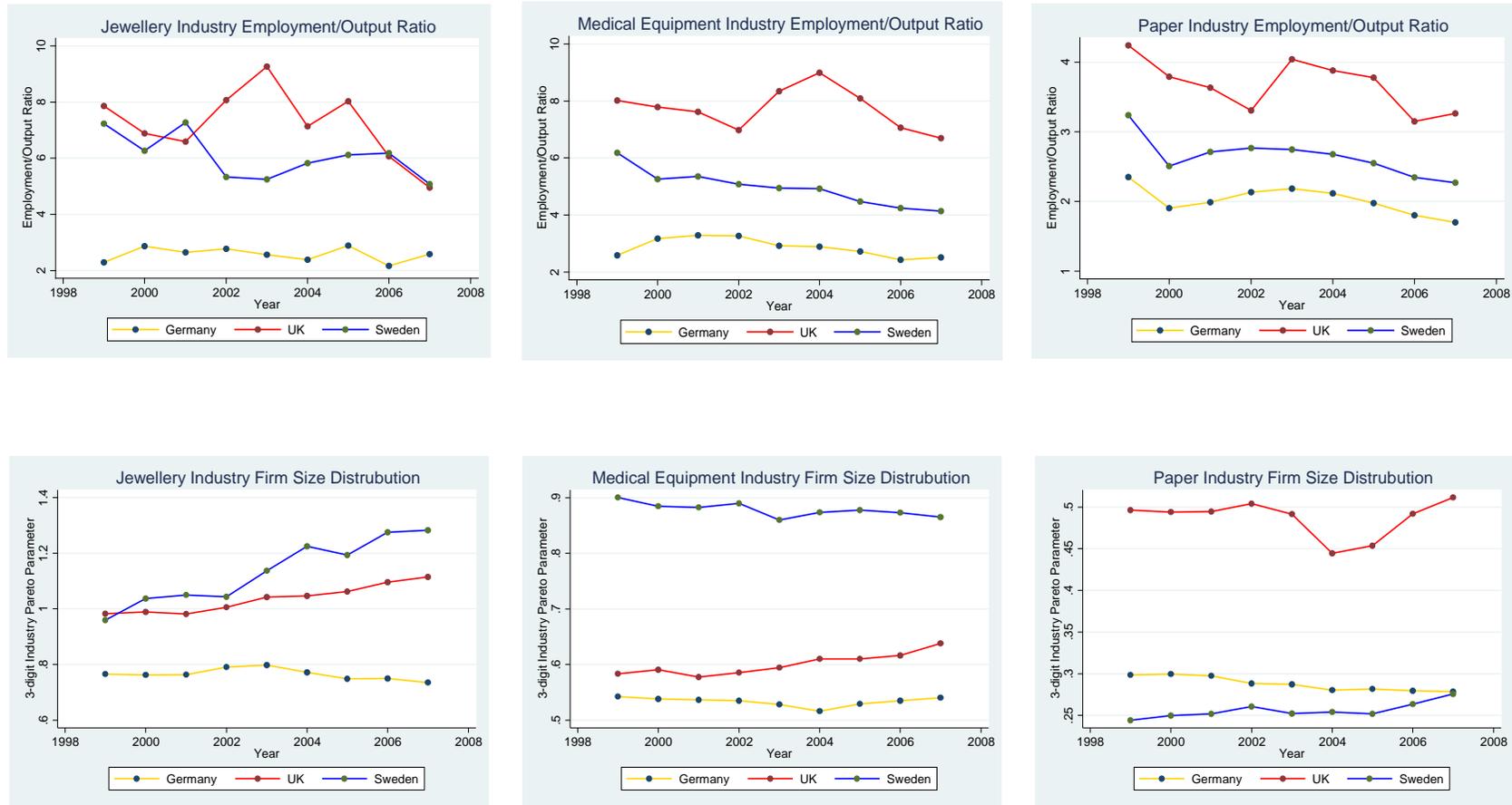
Table 1. LSDV estimates of equation (27) with different nested specifications
Dependent Variable: Employment Growth $\Delta \ln L$

Regressors	(I)	(II)	(III)	(IV)
Output Growth $\Delta \ln Y$	0.39*** (21.69)	0.38*** (21.63)	0.25*** (5.31)	0.23*** (4.92)
$\hat{\kappa} \cdot \Delta \ln Y$	--	--	0.21*** (3.10)	0.22*** (3.38)
Pareto Shape Parameter $\hat{\kappa}$	--	-0.06** (2.41)	-0.05** (2.11)	-0.04 (1.38)
Employment Protection	--	--	--	0.13** (2.09)
Union Density	--	--	--	-0.01 (1.49)
Active Labour Market Policies	--	--	--	0.12*** (3.76)
Passive Labour Market Policies	--	--	--	-0.11*** (4.49)
Government Size	--	--	--	-1.61*** (3.37)
Trade Openness	--	--	--	-0.23 (1.31)
UK dummy	-0.03*** (5.81)	-0.02*** (2.84)	-0.02*** (3.16)	0.03 (0.24)
Sweden dummy	0.00 (0.35)	0.01* (1.65)	0.01 (1.43)	0.84 (1.57)
Intercept	-0.05** (2.00)	-0.02 (0.95)	-0.02 (1.01)	0.96*** (3.36)
Log-likelihood	1731.63	1734.69	1739.76	1769.28
$\bar{R}^2$	0.27	0.27	0.27	0.29
AIC	-3249.27	-3253.38	-3261.52	-3308.56

- All the control variables are at the aggregate country level and are lagged once.
- The sample size is 2032, consisting of unbalanced annual observations for 90 German, 91 UK, and 73 Swedish 3-digit manufacturing industries for the period 2000-2007.
- All regressions include industry and time dummies which are found to be jointly significant.
- The numbers in parentheses are the t-ratios and *, ** and *** respectively denote significance at 10%, 5% and 1%. These are based on 'un-clustered' standard errors.

Table 2. LSDV estimates of equation (27) with alternative specifications
Dependent Variable: Employment Growth $\Delta \ln L$

Regressors	(I)	(II)	(III)
Output Growth $\Delta \ln Y$	0.23*** (4.92)	0.02 (0.31)	-0.02 (0.28)
$\hat{\kappa} \cdot \Delta \ln Y$	0.23*** (3.47)	0.56*** (5.27)	0.61*** (5.76)
Pareto Shape Parameter $\hat{\kappa}$	-0.03 (1.14)	-0.07** (2.09)	-0.04 (1.13)
Trade Openness (*Industry Level*)	--	0.00 (0.07)	0.00 (0.07)
Trade Openness (*Aggregate Level*)	--	--	-0.05 (0.30)
Other Control Variables	Included (*not lagged*)	included	included
Time, Country and Industry Dummies	Included	included	included
Log-likelihood	1782.43	1509.92	1509.97
$\bar{R}^2$	0.30	0.27	0.27
AIC	-3334.86	-2817.85	-2817.85

- For Other Control Variables see those included in Table 1.
- The sample size is 2032, consisting of unbalanced annual observations for 90 German, 91 UK, and 73 Swedish 3-digit manufacturing industries for the period 2000-2007.
- The numbers in parentheses are the t-ratios and *, ** and *** respectively denote significance at 10%, 5% and 1%. These are based on 'un-clustered' standard errors.