HETEROGENEOUS FIRMS AND THE MICRO ORIGINS OF AGGREGATE FLUCTUATIONS*

Glenn Magerman

Karolien De Bruyne

Emmanuel Dhyne

Jan Van Hove

Abstract

This paper evaluates the impact of idiosyncratic productivity shocks to individual firms on the volatility of an economy's output. Two sources of firm-level heterogeneity contribute to aggregate fluctuations: (i) differences in input requirements and (ii) the skewed distribution of productivities. We first develop a static model in which monopolistic competitive firms use inputs from other firms. We derive a generalized centrality measure that takes these two sources of heterogeneity into account; more central firms can then contribute more to aggregate volatility. Using unique data on firm-to-firm relationships across all economic activities in Belgium, the model generates sizable aggregate volatility: even with many firms, the micro origins of aggregate fluctuations do not wash out. Importantly, the vast majority of volatility from micro origins remains after accounting for sector-level fluctuations. Moreover, the top 100 most influential firms contribute up to 90% of volatility generated by the model, underlining a strong granularity of the economy. Further counterfactual analysis shows that both sources of heterogeneity contribute to aggregate fluctuations, their relative contribution depending on the labor share of production. (*JEL Codes:* D24, E32, L23)

Keywords: Heterogeneous firms, networks, input-output linkages, aggregate volatility.

*Université libre de Bruxelles (ECARES). Avenue F.D. Roosevelt 42, 1050 Brussels, CP114/04. E-mail: glenn.magerman@ulb.ac.be. 13,281 words. We would like to thank Andrew B. Bernard, Vasco Carvalho, Thomas Chaney, John Cochrane, Paola Conconi, Ron B. Davies, Jan De Loecker, Xavier Gabaix, Jozef Konings, Kalina Manova, Catherine Schaumans, Alireza Tahbaz-Salehi, Joris Tielens and seminar participants at Stanford University, University of Chicago, Econometric Society Winter Meeting 2016, Santa Fe Institute, Université de Montreal, University of Nottingham, the National Bank of Belgium and ECARES for helpful comments and discussions. Glenn Magerman gratefully acknowledges the financial support of Fulbright USA, Methusalem funding at VIVES/KU Leuven, the PhD Research Fellowship at KU Leuven, the National Bank of Belgium Sponsorship program 2014 and the Otlet-La Fontaine scholarship. Part of this research has been performed while Glenn Magerman was visiting the Department of Economics at Stanford University and the National Bank of Belgium. The results presented fully respect confidentiality restrictions associated with the data sources used in our analysis. The views expressed are those of the authors and do not necessarily reflect the views of the National Bank of Belgium. All remaining errors are of course idiosyncratic.

I Introduction

Large fluctuations in aggregate output are costly for any economy: volatility has a negative impact on long-run growth (Ramey and Ramey (1995); Mobarak (2005)), it increases inequality and poverty (Jalan and Ravallion (1999)) and with imperfect capital markets, uncertainty increases assets' risk premia. Traditionally, macro models such as real business cycle models and New Keynesian models (e.g. Kydland and Prescott (1982); Prescott (1986); Christiano et al. (2005)) focus on large and common shocks to the economy (i.e. technology shocks, monetary and fiscal policies, government expenditures, aggregate demand shocks) as potential mechanisms to explain movements in aggregate variables such as GDP. However, while macroeconomic shocks are without doubt important, a substantial fraction of aggregate volatility remains unexplained by these models (Cochrane (1994b); Gabaix (2011)).

In this paper, we appeal to a growing literature which argues that a sizable part of aggregate fluctuations can be attributed to idiosyncratic shocks to individual firms (Gabaix (2011); Acemoglu et al. (2012)). There is ample real-world evidence to support this view. First, some firms are large relative to the size of the economy; hence idiosyncratic shocks to these firms can have a directly visible impact on aggregate output. For instance, Nokia's value-added contributed on average 4% to Finland's GDP and 25% to GDP growth over the period 1998-2007, with a peak of 40% in 2000 (Seppala and Ali-Yrkko (2011)). After 2007, its share in GDP decimated to 0.4%, adding to a sluggish recovery of the Finnish economy after the Great Recession. Second, modern economies are intricate webs of connected firms through supplier-buyer relationships and value chains. Disruptions in these relationships can have a large impact on aggregate output through propagation effects across the production network. For example, the recent "Dieselgate" at Volkswagen Group has a yet unquantified impact on its own output (its value-added accounts for 1.3% of German GDP and almost 600,000 employees in 2014), that of downstream distribution systems and on sales to final consumers (Commission (2015)).

The main contribution of this paper lies in the empirical evaluation of firm-level idiosyncratic productivity shocks to volatility of GDP growth, using the observed network of firm-to-firm relations of the Belgian economy and a structural measure of firm-level TFP (Wooldridge (2009)). We show that, even in the presence of many firms, micro shocks matter in the aggregate, facilitated by asymmetries in the distributions of input requirements and firms' productivities.

To formalize these concepts, we first develop a static firm-level framework with monopolistically competitive firms that use labor and inputs from other firms to produce outputs. Firms then sell their output to other firms and to final demand. The model builds on the sectoral input-output presentation of Acemoglu et al. (2012, 2017), where the buyer-supplier interaction now is at the firm level and heterogeneous firms compete in monopolistic competition as in Melitz (2003); Melitz and Redding (2014).

Two channels then contribute to aggregate fluctuations from micro idiosyncratic shocks: (i) shocks to large firms can show up in the aggregate due to their contribution to aggregate output, and (ii) some firms are important suppliers to other firms in the economy, serving as hubs for shocks to propagate throughout the network of production. The model also provides micro-economic foundations for the existence of the observed skewed firm size distribution in reality: more productive firms charge lower prices, have higher profits and sell more to downstream firms and to final demand. If there are no intermediate goods, the model collapses back to Melitz (2003). In our more general setting however, firms' revenues and profits depend on the input-output structure of production and productivities across the whole network.

After realization of the equilibrium, independent productivity shocks hit firms in the economy. A positive shock to a firm leads to a decrease of its output price; this results in a cost reduction and lower output prices for its downstream buyers, affecting *their* buyers etc. up to final demand, after which the system returns to its equilibrium output. A productivity shock ultimately affects the real wage of consumers, generating aggregate effects in output. We derive the total contribution of a firm's value-added to GDP from a generalized Bonacich (1987) centrality measure, similar to the influence vector in Acemoglu et al. (2012, 2017). This measure quantifies the downstream impact of a productivity shock to a firm through all paths of all lengths from that firm to final demand. Aggregate volatility is then ultimately a weighted average of productivity shocks to individual firms, with weights given by this centrality measure. Centrality in turn, is governed by the two micro-economic sources of heterogeneity: input requirements and productivity.

Using unique data on the universe of firm-to-firm relationships within Belgium across all economic activities over the period 2002-2012, we present three key results. First, idiosyncratic firm-level volatility contributes substantially to observed aggregate volatility: the model predicts 1.11% aggregate volatility from idiosyncratic micro shocks, compared to observed volatility of GDP growth of 1.99% over the same period. This mechanism is complementary to existing macro shocks and other propagation mechanisms such as synchronization of investment (e.g. Gourio and Kashyap (2007)). Importantly, the bulk of aggregate fluctuations predicted by the model remains after controlling for comovement at the level of the economy or disaggregated sectors, up to around 500 4-digit NACE sectors. If sectoral, rather than firm-specific, idiosyncrasies would be the underlying source of aggregate volatility, this remaining variance should be zero.

Second, the model nests Acemoglu et al. (2012) and Gabaix (2011), and we show that both sources of heterogeneity contribute to aggregate fluctuations:[1] counterfactually shutting down either channel generates substantially lower predictions of aggregate volatility relative to our model. For Belgium, asymmetries in the input-output structure of production contribute relatively more to aggregate fluctuations than heterogeneity in sales to final demand.

Third, the influence of individual firms is very skewed: the 100 most influential firms explain around 90% of the variance in aggregate volatility generated by the model. A variance decomposition across industries shows that firms in Services contribute most to aggregate volatility (40%), followed by Utilities (36%), Manufacturing (24%) and virtually 0% for the Primary industry. Considering that Manufacturing and Services contribute respectively 15% and 47% to Belgian GDP over our sample period, this suggests that Manufacturing is relatively more volatile than the Services industry. Additional sensitivity analyses, accounting for exports, imports and alternative measures for firm-level growth in the data confirm our main results. Moreover, extensions to the baseline model to account for investment goods and firm-level heterogenous labor shares reinforce our main findings.

Finally, we present two additional results. First, performing simple counterfactuals across a sequence of economies, we show that the relative contribution of each channel crucially depends on the labor share of production: in the limit, labor is the only factor of production, all sales are directly to final demand and any network propagation mechanisms are mute. As the labor share decreases, relatively more intermediate inputs are used in production and the fragmentation of the production process generates network multipliers (Carvalho (2014)) as in Acemoglu et al. (2012, 2017). Second, Gabaix (2011) and Acemoglu et al. (2012) provide sufficient conditions for the influence vector under which idiosyncrasies can surface in the economy. We empirically confirm that these conditions are satisfied in our setting.

This paper contributes to three strands of literature that posit micro origins for aggregate

[1] Identical to Gabaix (2011); Acemoglu et al. (2012, 2017), a sufficient statistic for the aggregate impact of firm-level idiosyncratic shocks in our model is given by firms' sales shares in the economy (Hulten (1978)). However, we are specifically interested in the relative contribution of the two underlying micro channels to aggregate fluctuations and decompose sales in both the model and the data.

fluctuations. First, a large literature focuses on the aggregate effects of sectoral shocks via the input-output structure of the economy, using the standard deviation of aggregate output growth as a statistic for aggregate volatility (Long and Plosser (1983); Horvath (1998, 2000); Dupor (1999); Shea (2002); Foerster et al. (2011); Carvalho and Gabaix (2013); Acemoglu et al. (2015); Atalay (2015)). Shocks to sectors that are important suppliers to other sectors can propagate throughout the network of input-output linkages and show up in aggregate output. Acemoglu et al. (2012) present a tractable theoretical framework and derive conditions for when shocks to key sectors can show up in the aggregate. We build on this literature by bringing the network structure to the firm level, while we provide micro-economic foundations for the skewed distribution of firm sizes in reality. Importantly, the mechanisms in this paper rely on observable quantities, allowing us to directly estimate the model in the data.

A second strand of literature focuses on the impact of individual firms on aggregate output (e.g. Durlauf (1993); Carvalho and Grassi (2016)): some firms are large relative to the size of the economy and Gabaix (2011) shows that the skewed distribution of firm sizes, measured by the economy's sales Herfindahl index, can generate aggregate volatility from firm-level shocks. Moreover, even within narrowly defined industries there is a large amount of heterogeneity in firm sizes and productivities (see Syverson (2011) for an excellent survey), suggesting that some firms dominate sector-level movements. We contribute to this literature by developing a model that accounts for heterogeneity in productivity, while explicitly allowing for intermediate input consumption. While we abstract from changes in extensive margins, a few other papers evaluate the effects of non-linearities and cascades in production networks, such as Elliot et al. (2014); Baqaee (2015) or that of sticky network formation (Lim (2016)).

Third, we add to a growing literature on the empirical validation of production networks. Applications at the sector level have identified the network structure of production to be the major channel of contribution to aggregate fluctuations (Atalay (2015)). Some first evidence at the firm-level however, shows that the firm-specific component dominates sector-level sources of aggregate fluctuations (di Giovanni and Levchenko (2012)). Empirical evidence from natural experiments such as the Tohoku earthquake in 2011 also point to sizable supply chain distortions at the firm level (Carvalho et al. (2015); Boehm et al. (2015)). Only few other papers have substantial information on domestic firm-to-firm linkages, such as Bernard et al. (2015) for Japan. We add to this literature by providing a structural firm-to-firm framework at this level of detail.

Closest to us is di Giovanni et al. (2014), who empirically evaluate the impact of micro shocks on aggregate volatility in the presence of production networks. We differ in three key aspects. First, while these authors start from a decomposition of Melitz (2003) into variance and covariance components, we develop a structural model of production networks. Second, while the authors assume comovement in growth rates is correlated with unobserved firm-level input-output linkages, we observe these linkages at the firm level. Finally, we evaluate the propagation of idiosyncratic productivity shocks, rather than movements in the outcome variable of sales.

II The Model

In this section, we present a static model of a production economy with intermediate goods. The model largely follows Long and Plosser (1983); Acemoglu et al. (2012, 2017), where we allow for monopolistic competitive firms with heterogeneous productivities as in Melitz (2003); Melitz and Redding (2014).

A Environment

The representative household has constant elasticity of substitution (CES) preferences over n goods and maximizes utility

$$U = \left(\sum_{i=1}^{n} q_i^{\rho} \right)^{1/\rho}$$

where q_i is the quantity consumed of good i and $\eta = \frac{1}{1-\rho} > 1$ is the elasticity of substitution, common across goods. Each household is endowed with one unit of labor, supplied inelastically and paid in wages w. The size of the economy is normalized to 1 and labor is the only source of value-added, so that $w = Y$, where Y represents total spending on final consumption. Residual final demand for good i is thus given by $q_i = \frac{p_i^{-\eta}}{P^{1-\eta}} Y$, where p_i is the price of good i and the aggregate price index is given by $P^{1-\eta} = \sum_{i=1}^{n} p_i^{1-\eta}$ (see Appendix A for the model derivation).

On the supply side, a two-stage game governs production. In the first stage, there is a large pool of potentially entering firms. After payment of an investment cost $f_e > 0$, each firm i observes its own productivity ϕ_i (drawn from a Pareto distribution) and it receives a blueprint for production, stipulated by a set of particular input requirements $\omega_{ji} \in \Omega$. The existence of an *ex ante* investment cost dictates that there is a cutoff productivity ϕ_i^*, below which firm i does not make positive expected profits. Firms with expected profits lower than the sunk investment cost do not enter the market. In the second stage, sufficiently productive firms enter and search for input suppliers. The production blueprint is contingent, so that firms know how to produce an input in-house if it is not available on the market (due to non-producing firms with below-threshold productivities). We assume that in-house production is more expensive than sourcing these inputs, for example due to diseconomies of scope.

Ex post, the economy consists of n firms, competing in monopolistic competition. Each firm produces a differentiated good, which can be used for final consumption and as an input by other firms. Output of firm i is thus given by $x_i = \sum_j^n x_{ij} + q_i$, where x_{ij} is output produced by i and used as an input for production by firm j. Production entails fixed costs f, paid in terms of labor and inputs and is common across all firms, so that x_i output requires a total cost of $\Gamma_i = \frac{x_i}{\phi_i} + f$ of inputs in terms of labor and intermediate goods.[2] After payment of the fixed costs, output x_i follows a Cobb-Douglas production technology with constant returns to scale

$$x_i = (z_i l_i)^\alpha \prod_{j=1}^{n} x_{ji}^{(1-\alpha)\omega_{ji}} \tag{1}$$

where z_i is total factor productivity (TFP) of firm i, l_i is the amount of labor needed to produce x_i, $\alpha \in (0, 1]$ is the share of labor (common across firms), x_{ji} is the amount of inputs j used in production of i and $\omega_{ji} \in [0, 1]$ is the share of input j in total intermediate input use of firm i.[3] If firm i does not use input j in its production, $\omega_{ji} = 0$ and $\sum_j^n \omega_{ji} = 1$ so that constant returns to scale hold between labor and inputs. Denote $\varepsilon_i \equiv \ln z_i$ and the distribution of ε_i by $F(\varepsilon)$. Furthermore assume $E(\varepsilon_i) = 0$ and $\text{var}(\varepsilon_i) = \sigma_i^2 \in (\underline{\sigma}^2, \overline{\sigma}^2)$ with $0 < \underline{\sigma}^2 < \overline{\sigma}^2 < \infty$.

Next, the unit cost function associated with (1) is given by

$$c_i = B_i \left(\frac{w}{z_i}\right)^\alpha \prod_{j=1}^{n} p_{ji}^{(1-\alpha)\omega_{ji}} \tag{2}$$

where $B_i = \alpha^{-\alpha} \prod_j^n ((1-\alpha)\omega_{ji})^{-(1-\alpha)\omega_{ji}}$ is a constant and p_{ji} is the price of input j to firm i. Assuming no arbitrage (or equivalently no price discrimination), we can set $p_{ji} = p_j$. From monopolistic competition and CES preferences, prices are set as a constant markup over marginal

[2] In this model, aggregate profits are absorbed by fixed investment costs similar to Melitz (2003). Alternatively, one can imagine a setting with no fixed entry costs and an exogenous set of potential entrants as in Chaney (2008). Profits would then be redistributed across consumers through a fund in which all households have equal shares, so that $Y = w + \sum_i \pi_i$.

[3] In Appendix B, we derive an extension of the model with firm-specific labor shares. We then re-estimate the model predictions for aggregate volatility with this extra source of heterogeneity. Results are almost identical to our baseline results, underlining that heterogeneity in labor shares is not a main driver for aggregate fluctuations in the light of our model.

cost, so that $p_i = \frac{c_i}{\rho\phi_i}$.[4]

B Competitive equilibrium

Firms maximize profits given optimal amounts of labor and inputs, l_i and x_{ji}, and obtain output prices p_i, taking as given the prices of wages and inputs, w and p_j, and the total cost function Γ_i. The firm's problem can thus be written as

$$\pi_i = p_i x_i - wl_i - \sum_{j=1}^{n} x_{ji} p_j - wf - f \sum_{j \in \mathcal{S}_i} p_j$$

where $\mathcal{S}_i$ represents the set of suppliers of firm i. Denoting revenues by $r_i \equiv p_i x_i$, we can write equilibrium revenues as

$$r_i(\phi) = \underbrace{(1-\alpha)\sum_{j=1}^{n} \omega_{ij} r_j(\phi)}_{\text{downstream revenue}} + \underbrace{p_i q_i(\phi_i)}_{\text{final demand revenue}} \tag{3}$$

Relative revenues can be written as

$$\frac{r_1(\phi)}{r_2(\phi)} = \frac{(1-\alpha)\sum_{j=1}^{n} \omega_{1j} r_j(\phi) + Y\left(\frac{\rho P \phi_1}{c_1}\right)^{\eta-1}}{(1-\alpha)\sum_{j=1}^{n} \omega_{2j} r_j(\phi) + Y\left(\frac{\rho P \phi_2}{c_2}\right)^{\eta-1}} \tag{4}$$

Firm profits are given by

$$\pi_i = \frac{r_i}{\eta} - c_i f = \underbrace{\frac{1-\alpha}{\eta}\sum_{j=1}^{n} \omega_{ij} r_j(\phi)}_{\text{downstream}} + \underbrace{\left(\frac{\rho \phi_i P}{c_i}\right)^{\eta-1} \frac{Y}{\eta}}_{\text{final demand}} - c_i f \tag{5}$$

Some things are worth mentioning at this point. First, (3) shows the relationship between revenues, productivities and the network structure of production. Firms are larger if they sell more to other firms in the economy or to final demand. All other things equal, if the intermediate input share $1-\alpha$ is higher, if i sells to more downstream firms j, if buyers j have higher input requirements ω_{ij} from i or if any j is larger, firm i generates higher revenues.

Second, (3) to (5) provide micro foundations for the skewed distribution of firm sizes ob-

[4] Firms that are large relative to the size of the economy could internalize the impact of their pricing strategy on the aggregate price index, leading to heterogeneous markups, even with CES preferences (Melitz and Redding (2014)). However, this is a general critique to heterogeneous firms models with CES preferences (Dixit and Stiglitz (1993); di Giovanni and Levchenko (2012)) and not specific to our setup. With non-CES preferences, pass-through can be incomplete (see for instance Amiti and Konings (2007)).

served in reality: all other things equal, more productive firms charge lower prices through $p_i = \frac{c_i}{\rho\phi_i}$, generate higher revenues and earn higher profits. When there are no intermediate inputs ($\omega_{ij} = 0$ for all j and $c_i = w$), (4) collapses to Melitz (2003), with $\frac{r_1(\phi)}{r_2(\phi)} = \left(\frac{\phi_1}{\phi_2}\right)^{\eta-1}$. Then, relative revenues only depend on relative productivities and the elasticity of substitution: the closer substitutes goods are, the more small differences in productivity result in higher revenues. In our more general setting, firms' cost functions c_i are also heterogeneous in terms of input requirements, so that prices, revenues and profits subsequently depend on the whole input-output structure of production.

Third, revenues have a recursive structure: sales of i depend on sales of j, which in turn depend on sales of buyers of j etc. (3) provides a mechanism for idiosyncratic shocks to propagate over the network of production: a positive shock to a firm z_i leads to a decrease of its output price through c_i. This leads to lower cost functions c_j for downstream buyers j, resulting in lower output prices for their goods, affecting *their* buyers, and so on up to final demand. A productivity shock ultimately affects the real wage of consumers, resulting in changes in aggregate welfare. Finally, while we do not model the endogenous matching of supplier-buyer relationships in this paper, $r_i(\phi, r_j(\phi))$ in (3) can be seen as a supplier-buyer matching function similar to Lim (2016).

C The network structure of production

The production economy can also be seen as a weighted directed graph, in which vertices represent firms and directed edges from i to j are given by the non-negative elements of the adjacency matrix $\omega_{ij} \in \mathbf{\Omega}$. Writing (3) in matrix form:

$$\mathbf{r} = \left[\mathbf{I} - (1-\alpha)\mathbf{\Omega}\right]^{-1} \mathbf{b} \tag{6}$$

where $p_i q_i \equiv b_i \in \mathbf{b}$. $\mathbf{b}$ represents the vector of equilibrium revenues from final demand and is of dimension $n \times 1$. $\mathbf{I}$ is the identity matrix and $\mathbf{\Omega}$ is a square matrix with elements $\omega_{ij} \in [0, 1]$, both of dimension $n \times n$. Hence, aggregate output depends on the network structure of production through a Leontief inverse $[\mathbf{I} - (1-\alpha)\mathbf{\Omega}]^{-1}$, on the share of intermediate inputs $(1-\alpha)$, on input requirements $\omega_{ij} \in \mathbf{\Omega}$ and on final demand revenues b_i. Additionally, since $\mathbf{\Omega}$ is row stochastic, $[\mathbf{I} - (1-\alpha)\mathbf{\Omega}]$ is invertible for values $\alpha \in (0, 1]$ establishing existence and uniqueness of equilibrium.

From Hulten (1978) and in the presence of Harrod-neutral productivity shocks, we define the *influence vector* $\mathbf{v}$, similar to Acemoglu et al. (2012, 2017), as

$$\mathbf{v} \equiv \frac{\alpha}{Y}\mathbf{r} = \frac{\alpha}{Y}\left[\mathbf{I} - (1-\alpha)\boldsymbol{\Omega}\right]^{-1}\mathbf{b} \tag{7}$$

The influence vector captures how important each firm is in contributing to aggregate value-added Y. $\mathbf{v}$ can also be seen as a generalized Bonacich (1987) centrality, where $\mathbf{b}$ contains firm-specific elements outside the network (Newman (2010)). Additionally, note that α is the share of value-added in output, so we can write $v_i = \frac{\alpha r_i}{\alpha R}$, where v_i is the i^{th} element of the influence vector $\mathbf{v}$ and $R = \sum_{i=1}^{n} r_i$ represents gross output of the economy. The influence vector thus equivalently reflects the vector of equilibrium market shares in the economy.

This framework nests several existing contributions on the micro origins of aggregate fluctuations. First, if firms source equally from all other firms in the economy and also sell in equal proportions to final demand, (7) collapses to $\mathbf{v} = \frac{1}{n}$. This is a restatement of the classical diversification argument by Lucas (1977): aggregate volatility is then proportional to $1/\sqrt{n}$ and shocks to individual firms wash out in the aggregate from a law of large numbers argument, as shown below.

Second, with homogeneous input shares ($\omega_{ij} = \frac{1}{n}, \forall i, j$) and heterogeneity in final demand, (7) collapses to $\mathbf{v} = \frac{\alpha}{Y}\left[\mathbf{I} - \frac{(1-\alpha)}{n}\right]^{-1}\mathbf{b}$. For large n, this can be approximated by $\mathbf{v} \simeq \frac{\alpha}{Y}\mathbf{b}$. Hence the influence of individual firms is proportional to their sales to final demand. If additionally $\alpha = 1$, (7) becomes $\mathbf{v} = \frac{1}{Y}\mathbf{b} = \frac{1}{R}\mathbf{r}$. This is a direct statement of the granular hypothesis presented by Gabaix (2011): in the presence of a sufficiently skewed sales Herfindahl index that is consistent with a power law with fat tails, shocks to large firms can contribute significantly to aggregate fluctuations.[5]

Third, with heterogeneous input shares and homogeneous sales to final demand ($\frac{1}{Y}b_i = \frac{1}{n}$), (7) collapses to $\mathbf{v} = \frac{\alpha}{n}\left[\mathbf{I} - (1-\alpha)\boldsymbol{\Omega}\right]^{-1}\mathbf{1}$, where $\mathbf{1}$ is a vector of ones. This specification is the influence vector in Acemoglu et al. (2012): if the distribution of $\mathbf{v}$ is consistent with a power law distribution with fat tails, idiosyncratic shocks to important suppliers can propagate through the network of production and show up in the aggregate of the economy. Acemoglu et al. (2017) derive a similar influence vector as in (7), where $\mathbf{b}$ is derived from heterogeneity in final demand preference shares rather than from productivity differences.

[5] The case of $\alpha = 1$ is for expositional clarity. Gabaix (2011) also holds more generally in the presence of input-output linkages (i.e. $\alpha < 1$). Hulten (1978) shows that a firm's sales are a sufficient statistic to capture the total effects of productivity shocks on changes in aggregate welfare.

D Aggregate volatility

Log-linearizing (1) and summing over all firms allows us to write (7) as

$$\ln Y \equiv y = \mu + \mathbf{v}'\varepsilon \tag{8}$$

where μ represents the mean output of the economy, independent of the vector of idiosyncratic productivity shocks ε. The log of aggregate value-added, y, is thus a random variable and from $E(\varepsilon_i) = 0$, it follows that $E(y) = \mu$; i.e. any deviation from equilibrium output is only the result of idiosyncratic shocks ε, weighted by their impact through the influence vector $\mathbf{v}$. (8) is the link between the production network of the economy, idiosyncratic shocks and the resulting aggregate volatility. Given i.i.d. shocks, we can then write the standard deviation of aggregate output, σ_y, as

$$\sigma_y = \sqrt{\sum_{i=1}^{n} v_i^2 \sigma_i^2} \tag{9}$$

Consider for simplicity of exposition $\sigma_i = \bar{\sigma}$ for all i, so that all firms experience the same volatility $\bar{\sigma}$. In the extreme case that all firms are homogeneous along all dimensions, $v_i = \frac{1}{n}$ for all i, and so $\sigma_y = \frac{\bar{\sigma}}{\sqrt{n}}$. Then, as $n \to \infty$, σ_y rapidly decays to zero. This is the classical diversification argument as in Lucas (1977): in an economy with potentially millions of firms, idiosyncratic shocks to individual firms have a negligible impact on the aggregate and aggregate output y rapidly converges to its mean, μ. With asymmetric firm sizes however, micro shocks can show up in the aggregate if the distribution of the influence vector is sufficiently skewed, the crux of the argument in Gabaix (2011). With the additional existence of heterogeneous input-output linkages, v_i represents the influence of firm i through the whole network structure of production, in which large firms exist from a combination of heterogeneity in firm productivities and input requirements. Also note that we can write the Euclidean norm of the influence vector as $\|\mathbf{v}\|_2 = \sqrt{\sum_{i=1}^{n} v_i^2}$. These two sources of heterogeneity then drive the Euclidean norm: σ_y clearly increases in $\|\mathbf{v}\|_2$, and with a power law distributed $\mathbf{v}$ with infinite variance, idiosyncratic productivity shocks can contribute to aggregate volatility.[6]

[6] Gabaix (2011) and Acemoglu et al. (2012) prove that economies with fat-tailed distributions of v_i, consistent with a power law distribution with tail exponents of $\beta \in [1, 2)$, can generate sizable fluctuations in the aggregate. Volatility is then proportional to $1/n^{1-1/\beta}$, much larger than the prediction from the diversification argument. We empirically confirm this condition in Section V.

E Extensive margins of the network

Idiosyncratic shocks could have substantial effects on the extensive margins of the economy, i.e. the adding or dropping of firm-to-firm relations and the entry and exit of firms as a response to productivity shocks. For instance, firms could switch suppliers instead of taking the price changes as given. In addition, when shocks are sufficiently large, productivity shocks could lead to firms entering and exiting.

In the light of our model however, there is no meaningful entry or exit of firms. Entry is governed by the *ex ante* draws of productivity and input requirements, which are fixed thereafter. Moreover, given fixed costs of production and $E(\varepsilon_i) = 0$, firms do not exit in the presence of idiosyncratic shocks: firms stay in the market as long as they cover variable costs of production.

There is also tentative empirical evidence supporting our view of a fixed production network. First, in our data, the share of entering and exiting firms is small. Out of all firms active at year t, on average 3% exit at $t+1$ while 3% entered at $t-1$. Entering and exiting firms are also smaller than incumbents: the median turnover and input consumption of entering/exiting firms is one third that of incumbents. Moreover, entering/exiting firms have fewer buyers/suppliers: the median number of buyers of entering/exiting firms is 2, with a median of 6 suppliers, compared to 11 and 28 respectively for all firms.

Second, at the level of supplier-buyer relationships, entering/exiting relationships are also smaller: the median value of new relationships is half that of continuing relationships. This is even lower for new relationships between incumbents, compared to relationships that involve an entering or exiting firm. Supplier switching between incumbents as a response to shocks would suggest that relationships of similar sizes are added and dropped, which we do not observe in the data. Finally, the extensive margin of relationships contributes little to firm growth: new relationships account for around 3% in total turnover for sellers and around 2% in total inputs of buyers.

These findings are mirrored by a large literature suggesting that the volatility of aggregate growth rates is generally dominated by the intensive margin (e.g. Bernard et al. (2009) for the US, Bricongne et al. (2012); Osotimehin (2013); di Giovanni et al. (2014) for France and Behrens et al. (2013) for Belgium).

F Identification strategy

The model is a static representation of an economy's steady state output with disturbances. While the model exhibits a stationary equilibrium, GDP is clearly not stationary in reality (which would imply that the economy is not growing over time). In most of the empirical analysis that follows, we exploit the unique panel dimensions of the Belgian data and focus on the stationary series of aggregate value-added *growth* to predict aggregate volatility.

From (8), we generate a very simple dynamic interpretation of the model. The growth rate of the economy can be written as

$$\Delta y = \ln Y_t - \ln Y_{t-1} = \mathbf{v}'(\varepsilon_t - \varepsilon_{t-1}) \tag{10}$$

Assuming independence of shocks over time, this results in the estimation equation to be used in Section IV:

$$\hat{\sigma}_{\Delta y} = \sqrt{\sum_{i=1}^{n} \hat{v}_i^2 Var(d\hat{\varepsilon}_{it})} \tag{11}$$

where $\hat{\sigma}_{\Delta y}$ is the predicted standard deviation of GDP growth, $\hat{v}_i$ denotes the empirical counterpart of v_i and $d\hat{\varepsilon}_{it} = (\hat{\varepsilon}_{it} - \hat{\varepsilon}_{it-1})$ denotes estimated TFP growth of firm i at time t.

A few remarks. First, in the model, the network structure $\mathbf{v}$ is exogenously given by the firm-level input-output matrix $\mathbf{\Omega}$. Hence, v_i is constant over time. Second, idiosyncratic productivity shocks are estimated as firm-level TFP growth. Our structural estimation of TFP follows Wooldridge (2009), implying the assumption of a first-order Markov process of TFP growth. Hence, idiosyncratic shocks are independent across firms as well as over time and mean reverting. Third, as argued by Cochrane (1994a), supply shocks are mostly transitory, after which the system returns to its steady state, as in our setup. Demand shocks in contrast, are generally more permanent in nature and are related to a shift in the mean output of the economy. This additionally underlines our choice of modeling mean-reverting output shocks as productivity shocks.

In Section IV, we fix $\mathbf{v}$ for the year 2012 and calculate firm-level volatility over the years 2002-2012 for firms that are active in 2012. We use an unbalanced panel structure and calculate $Var(d\hat{\varepsilon}_{it})$ over the observed time series per firm. We then compare this to the volatility of GDP growth over the same period.[7]

[7] This procedure is similar to di Giovanni et al. (2014), who fix the weights of individual growth rates over time

III The Belgian network of production

A The Belgian VAT system and VAT filings

The Belgian value-added tax (VAT) system requires that the vast majority of firms located in Belgium and across all economic activities, charge VAT on top of the delivery of their goods and services. This also includes foreign companies with a branch in Belgium or those firms whose securities are officially listed in Belgium. Firms that only perform financial transactions or medical or socio-cultural activities are exempt. The tax is levied in successive stages of the production and distribution process: with each transaction, firms charge VAT on top of their sales and pay VAT on inputs sourced from their suppliers, in effect only paying taxes on the value-added generated at that stage. The tax is neutral to the firm and the full burden of the tax ultimately lies with the final consumer. The standard VAT rate in Belgium is 21%, but for some goods a reduced rate of 12% or 6% applies.[8]

VAT-liable firms have to file periodic VAT declarations and VAT listings to the tax administration.[9] The VAT declaration contains the total sales value, the VAT amount charged on those sales (both to other firms and to final consumers), the total amounts paid on inputs sourced and the VAT paid on those inputs. This declaration is due monthly or quarterly, depending on some size criteria, and it is the basis for the balance of VAT due to the tax authorities every period. Additionally, at the end of every calendar year, all VAT-liable firms have to file a complete listing of their Belgian VAT-liable customers over that year. An observation in this listing refers to the yearly values of a sales relationship from firm i to firm j. The reported value is the sum of the values of all invoices from i to j. Whenever this aggregate value is larger than or equal to 250 euro, the relationship has to be reported. Sanctions for incomplete and mis-reporting guarantee a very high quality of the data.

B Data sources and construction

The empirical analysis in this paper mainly draws from five firm-level data sources for the years 2002-2012, administered by the National Bank of Belgium (NBB): (i) the NBB B2B Transactions

when aggregating to country-level growth rates.

[8] See ec.europa.eu/taxation_customs for a complete list of rates. These rates did not change over our sample period.

[9] Sample forms of the VAT declarations can be found here (French) and here (Dutch). Sample forms of the VAT listings can be found here (French) and here (Dutch).

Dataset based on the VAT listings, (ii) annual accounts from the Central Balance Sheet Office at the NBB, (iii) VAT declarations of firms as reported to the NBB, (iv) the main economic activity of the firm by NACE classification from the Crossroads Bank of Belgium, and (v) firm-level import and export data from the Foreign Trade Statistics at the NBB.

Firm-to-firm relationships The confidential NBB B2B Transactions Dataset contains the values of yearly relationships between all VAT-liable Belgian firms for the years 2002 to 2012. We use the sales values excl. VAT for the analysis. A detailed description of the collection and cleaning of this dataset is given in Dhyne et al. (2015).

Firm-level characteristics Virtually all firms have to file annual accounts at the end of their fiscal year.[10] We extract firm-level information from the annual accounts at the NBB and annualize all flow variables from fiscal years to calendar years. This transformation ensures that all firm-level information in our database is consistent with observations in the VAT listings data.[11] We extract information on turnover, intermediate material and service inputs (in euro) and employment (in average full-time equivalents (FTE)). Small firms do not have to report turnover and intermediate inputs in the annual accounts; for those firms we use the VAT declarations to extract values on turnover and inputs. We also extract information on labor costs, debts and assets (in euro) for firm-level TFP estimation and the calculation of labor shares. Finally, for some of the sensitivity analyses in Section IV, we extract export and import values of firm i in year t from the Foreign Trade Statistics at the NBB. Firms are identified by their VAT number, which is unique and common across these databases. This implies that observations are at the level of the firm, rather than at the plant or establishment level.[12]

We calculate value-added of the firm as turnover minus intermediate inputs. A detailed description of the TFP estimation procedure, following Wooldridge (2009), is provided in Appendix D. For the construction of the firm-level input-output matrix, we drop firms that generate negative value-added to ensure well-defined influence vectors and weighted growth rates. We

[10] See here for filing requirements and exceptions. See here for the size criteria and filing requirements for either full-format or abridged annual accounts.

[11] 78% of firms have annual accounts that coincide with calendar years, while 98% of firms have fiscal years of 12 months.

[12] If a VAT identity is a conglomerate of plants or establishments, this actually reinforces our results. Gabaix (2011): *"If Walmart doubles its number of supermarkets and thus its size, its variance is not divided by 2—as would be the case if Walmart were the amalgamation of many independent supermarkets. Instead, the newly acquired supermarkets inherit the Walmart shocks, and the total percentage variance of Walmart does not change."*

also drop firms that have less than one FTE or that do not report employment to account for very small firms (including management firms). The final dataset contains information on firms that are VAT liable and also report annual accounts.

Sector membership We obtain information on the main economic activity of the firm at the NACE 4-digit level from the Crossroads Bank of Belgium. The NACE Rev. 2 classification is the current official EU classification system of economic activities, active since Jan 1, 2008. The 2-digit levels coincide with the international ISIC Rev 4. classification, while the 4-digit levels are particular to the EU. We concord NACE codes over time to the NACE Rev. 2 version to cope with changes in the NACE classification over our panel from Rev. 1.1 to Rev. 2. We drop firms in NACE sectors 64-66, which have activities related to financial services and insurance. We group sectors into industries across the following NACE 2-digit codes: Primary (NACE 01-09), Manufacturing (NACE 10-33), Utilities (NACE 35-43) and Services (NACE 45-63, 68-82 and 94-96). See Table I for an overview of the sectors and industries covered in our dataset.

C Descriptive statistics

Table II reports the share of the aggregate industries covered in terms of total gross output and GDP. These industries cover 81% of gross output of the Belgian economy and 71% of GDP in 2012 as reported by the National Accounts; these shares are very stable over our sample period. The financial sector, which is not part of the VAT system, accounts for the majority of the remaining output and value-added.

Table III presents summary statistics for the main variables in 2012. By far the most novel information is that on business relationships r_{ij}. In our final sample, there are 3,505,207 domestic relationships between 79,788 firms in 2012. The average relationship value between any two Belgian firms is 46,403 euro, with a standard deviation of 1.2 million euro. The distribution is clearly skewed: the median relationship value is 2,000 euro, with extreme outliers over 300 million euro. There are 67,466 firms that have at least one business customer in 2012. The median firm has 11 business customers. On the input side, there are 79,689 firms with at least one supplier in the network and the median firm sources from 28 suppliers.

Turnover of the median firm is 0.8 million euro, with a standard deviation of 150 million euro. Input consumption follows very similar patterns with a median input of 0.5 million euro and a standard deviation of 141 million euro. The median firm employs 4 FTE, with a standard deviation of 244 FTE. Each of these firm size distributions is very skewed, as is well-documented in the literature (e.g. see Axtell (2001) and Gabaix (2011) for US firms). Results for other years in our sample are very similar (not reported).

D Construction of the influence vector

In order to estimate $\mathbf{v} = \frac{\alpha}{Y}\left[\mathbf{I} - (1-\alpha)\mathbf{\Omega}\right]^{-1}\mathbf{b}$, we need information on input shares $\omega_{ij} \in \mathbf{\Omega}$ and final demand $b_i \in \mathbf{b}$. We obtain intermediate input shares to j, $\hat{\omega}_{ij} \in \hat{\mathbf{\Omega}}$, as the value of relationships r_{ij} in total input consumption of j in 2012. Hat notation denotes the empirical counterpart to theoretical variables.[13] We construct sales to final demand of firm i, $\hat{b}_i$, as the residual of turnover minus the total value of its business sales observed in the dataset: $\hat{b}_i = r_i - \sum_{j=1}^{n} r_{ij}$. Final demand thus contains final consumption, government expenditure, exports and domestic sales to other firms not in the observed network.[14]

The last two rows of Table III report the summary statistics of $\hat{\mathbf{\Omega}}$ and $\hat{\mathbf{b}}$. The median input share across all suppliers in the economy is 0.3% with a standard deviation of 7.5%. The distribution is clearly skewed: most suppliers represent only a small fraction of input requirements, while some linkages represent key input supplier-buyer relationships. This distribution is very similar when we compare the list of input shares firm-by-firm (not reported). The median firm sells for 0.5 million euro to final demand with a standard deviation of 133 million euro. It is clear that the distributions of both input shares and final demand are heavily skewed, providing a first empirical support for our choice of the two sources of heterogeneity in this paper.

To get a simple estimate for the labor share in production, α, we consider three measurements and take their simple average as our baseline estimate. First, from a macro perspective, we calculate the labor share year-by-year from the National Accounts as the labor cost aggregated over all 2-digit sectors in our analysis, divided by total gross output of these sectors. These shares are very stable over our sample period, ranging between 0.18 and 0.20. We take the simple average across years, resulting in $\hat{\alpha} = 0.19$. Second, from a micro perspective, we regress firm turnover on labor costs and inputs (all in logs) for the years 2002-2012, using OLS and controlling for year fixed effects, resulting in $\hat{\alpha} = 0.17$.[15] Third, we calculate the labor share in

[13] Since $\sum_{i=1}^{n} \omega_{ij} = 1$ in the model, we renormalize each entry in the relationship data so that input shares sum to one per buyer, or $\hat{\omega}_{ij} = \frac{r_{ij}}{\sum_{i=1}^{n} r_{ij}}$. This ensures that our dataset lines up with the model: all value-added is generated from transactions in the network and sales to final demand.

[14] In Appendix C, we develop a simple extension of the model that includes capital, so that final demand additionally contains investments in line with the typical National Accounting Identity. Note also that while changes in inventories are part of aggregate final demand in the National Accounts, we do not observe these inventories in the VAT data. However, these changes contribute typically to less than 1% to GDP.

[15] The estimated coefficients are 0.17 and 0.81 respectively, good support for our constant returns to scale as-

production in 2012 from a combination of micro and macro data as

$$\hat{\alpha} = \frac{\text{value added}}{\text{total output}} \times \frac{\text{total labor cost}}{\text{value added}} = \left[\frac{1}{n}\sum_{i=1}^{n}\left(1 - \frac{\text{inputs}_i}{\text{turnover}_i}\right)\right] \times \frac{\text{total labor cost}}{\text{GDP}} = 0.31 \times 0.69 = 0.21$$

For the first factor, we use annual accounts information; for the second, we use the average labor share in GDP for Belgium over the period 2002-2010 of 0.69, as reported by the OECD, resulting in $\hat{\alpha} = 0.21$. The arithmetic average of these estimates then results in our baseline $\hat{\alpha} = 0.19$.

Finally, as we only observe a subset of firms in the economy contributing to observed GDP (we do not observe the financial sector and some small firms in other sectors), we calibrate Y by normalizing $\mathbf{v}$ so that $\sum_i v_i = 1$; i.e. we construct a synthetic GDP over the set of firms we use for estimation in our sample. Alternatively summing over all value-added of firms in our sample generates identical results.

E Estimation of firm-level volatilities

We use firm-level TFP growth, g_{it}, as our measure of productivity shocks, so that[16]

$$g_{it} = \Delta \ln TFP_{it} = \ln TFP_{it} - \ln TFP_{it-1}$$

Following Gabaix (2011) and di Giovanni et al. (2014), we winsorize growth rates at $g_{it} = \pm 1$ (i.e. a doubling or halving of productivity from $t-1$ to t) to account for extreme outliers such as mergers and acquisitions and possible measurement error in the micro data.[17] Our measure of firm-level volatility is then obtained as the standard deviation of individual growth rates over all years T: $\hat{\sigma}_i = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(g_{it} - \frac{1}{T}\sum_{t=1}^{T} g_{it}\right)^2}$.

Figure I shows the histograms of firm-level growth rates (g_{it}) in panel (a) (pooled over the period 2002-2012) and firm-level volatilities (st. dev.(g_{it})) in panel (b). In panel (a), the distribution of growth rates is symmetric with mean 0.5%, close to zero. Hence by the law of large numbers, micro shocks could cancel out in the aggregate with homogeneous firms as proposed by Lucas (1977). This distribution is very similar when compared year-by-year, across

sumption in the model.

[16] Remark that we use yearly growth rates as in similar studies (e.g. Gabaix (2011); Acemoglu et al. (2012); di Giovanni et al. (2014)), as our data dimensions are mainly driven by availability of the different firm-level variables. This is slightly different from the bulk of the real business cycles literature, which mostly uses quarterly or monthly aggregate output data.

[17] The cutoff of 1 we use is extremely liberal. Gabaix (2011) winsorizes growth rates for the top 1,000 firms in his Compustat sample at 20%, while di Giovanni et al. (2014) trim (and thus drop) observations with double and half growth rates. In our data, this procedure leads to 2% of observations to be winsorized for TFP growth. Alternative cutoff values for this procedure have no material impact on our results.

industries or using alternative measures of growth rates (not reported). In panel (b), we see that most firms experience a relatively low volatility. The mean volatility is 25.7% with a standard deviation of 17.7%; the median is 20.7%. Again, results are presented across the economy, but the moments are very similar across aggregated industries and using alternative growth rate measures.

As a simple way to account for aggregate shocks, we follow Stockman (1988); Gabaix (2011); Carvalho and Gabaix (2013) and extract common comovement at aggregate levels by demeaning individual growth rates. In particular, we extract economy-wide comovement as $d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$, where $\bar{g}_t$ is the growth rate of the economy at time t. To construct $\bar{g}_t$, we use the weighted arithmetic mean of individual growth rates of all firms in our panel, $\bar{g}_t = \sum_i \theta_{it} g_{it}$, where θ_{it} is the share of value-added of firm i in total value-added in the sample at time t. Similarly, we account for sector-level comovement by generating sector-demeaned growth rates by subtracting the 2-digit sector growth rate from the individual growth rates: $d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$, where $\bar{g}_{I_2 t}$ is the weighted mean growth rate of the 2-digit sector I at time t, to which firm i belongs. Similarly for comovement at the 4-digit sector level, as $d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$, where $\bar{g}_{I_4 t}$ is the weighted mean growth rate of 4-digit sector I at time t. We demean individual growth rates with sector aggregates only if there are at least ten firms in the same sector.[18]

IV Empirical validation

This Section describes the empirical implementation of the model. We then turn to the underlying micro origins of aggregate fluctuations. Finally, we present several sensitivity analyses.

A *Aggregate volatility*

With estimates for v_i and $d\varepsilon_i$, we obtain the model's prediction of aggregate volatility from

$$\hat{\sigma}_{\Delta y} = \sqrt{\sum_{i=1}^{n} \hat{v}_i^2 Var(d\hat{\varepsilon}_{it})}$$

We estimate the model and three benchmark specifications of $\mathbf{v}$, shutting down one or more sources of heterogeneity in the data. In particular:

1. **Model**: both $\mathbf{\Omega}$ and $\mathbf{b}$ are heterogeneous, so that $\mathbf{v} = \frac{\alpha}{Y} \left[I - (1-\alpha)\mathbf{\Omega}\right]^{-1} \mathbf{b}$.

[18] From our definition of g_{it}, and using the same methodology as in di Giovanni et al. (2014), growth rates are conditional on firms surviving from $t-1$ to t, but firms are allowed to drop in and/or out the sample over time. We then calculate the volatility over the observed growth rates in the sample.

2. **Network benchmark** (Acemoglu et al. (2012)): $\mathbf{\Omega}$ is heterogeneous and $\mathbf{b}$ is homogeneous, so that $\mathbf{v} = \frac{\alpha}{n}\left[I - (1-\alpha)\mathbf{\Omega}\right]^{-1}\mathbf{1}$.

3. **Final Demand benchmark**: $\mathbf{\Omega}$ is homogeneous and $\mathbf{b}$ is heterogeneous, so that $\mathbf{v} = \frac{\alpha}{Y}\left[I - \frac{(1-\alpha)}{n}\right]^{-1}\mathbf{b}$.

4. **Diversification benchmark** (Lucas (1977)): both $\mathbf{\Omega}$ and $\mathbf{b}$ are homogeneous, so that $\mathbf{v} = \frac{1}{n}$.

Table IV presents the main results of this paper. The top left cell shows the model's predicted volatility of 1.11% over the period 2002-2012, while the observed volatility of GDP ($\sigma_{\Delta y}$) over this period was 1.99%. The second to fourth rows show aggregate volatility predictions, accounting for comovement at different levels of aggregation, up to the 4-digit sector level. Predicted volatility decreases monotonically as we demean growth rates at more disaggregated levels, but most of aggregate volatility remains after accounting for aggregate comovement. The second column reports predictions of the Network benchmark (Acemoglu et al. (2012)). Predicted volatility is then 0.69%. Again predictions naturally decrease when accounting for aggregate comovement. The results from the Final Demand benchmark are given in column three, with a predicted volatility of 0.32%. The last column reports predicted volatility under the assumption of the Diversification benchmark (Lucas (1977)). Predicted volatility is then 0.09%, an order of magnitude smaller than the model prediction.

A few remarks. First, these results suggest that even in the presence of many firms, aggregate volatility from micro origins persists. Moreover, as in Gabaix (2011) and Acemoglu et al. (2012), aggregate fluctuations are of the same order of magnitude as those observed in reality. This is evidence in favor of the hypotheses in Gabaix (2011) and Acemoglu et al. (2012), now evaluated over the observed firm-level production network.

Second, the bulk of aggregate volatility remains explained by the model after accounting for aggregate comovement up to highly disaggregated sector levels. If sectoral, rather than firm-specific, idiosyncrasies would be the underlying source of aggregate volatility, this remaining variance would be zero. These results also imply that the micro origins of aggregate volatility in this paper are complementary to existing macro shocks and other amplification mechanisms (e.g. synchronization of investment as in Gourio and Kashyap (2007)).

Third, both sources of heterogeneity contribute to aggregate fluctuations: counterfactually shutting down either channel in the benchmarks generates substantially lower predictions of aggregate volatility relative to our model. Confirming earlier findings, the classical diversification argument grossly underestimates observed aggregate volatility from micro origins.

B The micro origins of aggregate fluctuations

The above results suggest that there is a significant role for individual firms in explaining aggregate volatility. We now consider these micro origins in more detail. First, from (11), it is straightforward to additively decompose the contribution of the top k firms to total volatility of the economy as

$$\hat{\sigma}^2_{\Delta y} = \sum_{i=1}^{k} \hat{v}_i^2 Var(d\hat{\varepsilon}_i) + \sum_{i=k+1}^{n} \hat{v}_i^2 Var(d\hat{\varepsilon}_i) \tag{12}$$

where $k = 1,000$ and 100 captures the variance of the model explained by shocks to the top 1,000 and top 100 firms respectively. We calculate the variance of aggregate movement ($\hat{\sigma}^2_{\Delta y}$) instead of the standard deviation ($\hat{\sigma}_{\Delta y}$) to ensure that volatility shares sum to one. In particular, the contribution of the top k firms is obtained from

$$\hat{\sigma}^2_{\Delta y|i=\{1,\ldots,k\}} = \frac{\sum_{i=1}^{k} \hat{v}_i^2 Var(d\hat{\varepsilon}_i)}{\sum_{i=1}^{k} \hat{v}_i^2 Var(d\hat{\varepsilon}_i) + \sum_{i=k+1}^{n} \hat{v}_i^2 Var(d\hat{\varepsilon}_i)}$$

The share of variance explained by the k most influential firms is then expressed as $\frac{\hat{\sigma}^2_{\Delta y|i=\{1,\ldots,k\}}}{\hat{\sigma}^2_{\Delta y}}$. Table V presents these results. In the baseline setting, 98.9% of volatility predicted by the model is generated by the top 1,000 firms out of almost 80,000 firms used in estimation. The top 100 most influential firms still contribute to 91.3% of explained variance. Again, demeaning has a slightly decreasing impact on prediction, but the vast majority of variance remains explained by the top k firms.

These results indicate that there is a key role for a very small subset of firms. For instance, the top 100 firms represent just 0.15% of observations. The classical diversification argument would then allocate a role of 0.09% $\times$ 0.15% = 0.0135% to these top 100 firms, compared to around 90% in the data. Again, this underlines the rationale for a framework in which heterogeneous firms contribute to aggregate volatility.

Turning to the sector membership of these top firms, Figure II shows the distribution of the top 100 most influential firms across 4-digit sectors predicted by (12).[19] Firms in Solid,

[19] Due to confidentiality reasons, we only report sectors with at least 3 firms in a 4-digit sector.

Liquid and Gaseous Fuels (4671) and in Renting and Leasing of Cars and Light Motor Vehicles (7711) are among the most influential firms in the economy. Other inputs such as Wholesale of Computers, Computer Peripheral Equipment and Software (4651), Sale of Cars and Light Motor Vehicles (4511) and Temporary Employment Agencies (7820) are also prominent. Interestingly, there are only two Manufacturing sectors present in this ranking: Manufacture of Other Organic Basic Chemicals (2014) and Manufacture of Motor Vehicles (2910).

Finally, we evaluate the distribution of aggregate volatility across industries. We allocate each i to industry $I = \{\text{Primary, Manufacturing, Utilities, Services}\}$ and obtain volatility shares similar to (12), but now across I's:

$$\hat{\sigma}^2_{\Delta y|i\in I} = \frac{\sum_{i\in I} \hat{v}_i^2 Var(d\hat{\varepsilon}_i)}{\sum_{i\in I} \hat{v}_i^2 Var(d\hat{\varepsilon}_i) + \sum_{i\notin I} \hat{v}_i^2 Var(d\hat{\varepsilon}_i)}$$

Table VI reports the shares of variance explained by firms in these industries. Clearly the Primary industry has a negligible impact on aggregate volatility. Conversely, Manufacturing accounts for 23.8% of aggregate fluctuations in the baseline setting, while Utilities and Services account for 35.6% and 40.5% respectively. However, considering that Manufacturing and Services contribute respectively 15% and 47% to Belgian GDP over our sample period (Table II), these results suggest that Manufacturing is relatively more volatile than the Services industry. This underlines the need for a comprehensive view on the economy, spanning all economic activities: studies with access to Manufacturing data only might incorrectly extrapolate findings for Manufacturing to the rest of the economy.

C Sensitivity analysis

We present a series of sensitivity analyses to the baseline empirical procedure. First, final demand is obtained as the residual of turnover minus sales to other firms in the Belgian economy. This residual contains final consumption, government spending, exports and sales to other firms not observed in the network. In reality however, exports contain both exports of intermediate goods and sales to foreign final demand. This generates an upward bias towards the estimation of the final demand channel in the model. To evaluate the impact of this bias, we perform two additional estimations. First, we estimate the model with only domestic turnover (including that of exporters), ignoring exports and assuming that all final demand is domestic. Second, we alternatively split up exports into business sales and final demand sales, applying the ratio of both channels for firms their domestic sales. This implies assuming that the exports of any firm contain the same ratio of sales to final demand and sales to other businesses as their domestic sales. In particular, we obtain firm-level domestic sales as turnover minus exports, and subsequently domestic sales to final demand as domestic sales minus sales to other firms in the network. We then construct the share of domestic final demand out of domestic sales and apply this share to exports, assuming that foreign final consumption follows the same ratio as in the Belgian data. We then adjust the final demand vector **b** downwards and re-estimate the model. Table A.III presents these additional results, showing that predictions for aggregate volatility are very robust to these alternatives.

Second, in the baseline model we assume that all inputs are sourced within the domestic production network and hence normalize input shares to sum to one. There are two main reasons for this approach. First, we do not observe firm-level counterparts for foreign suppliers, only product-country-level information. Given the different units of observation in the trade data compared to the domestic data, it is impossible to capture import transactions in the same way as inputs from domestic suppliers. Second, it is also difficult to consider changes in imports as true idiosyncrasies at the firm level (see e.g. Kramarz et al. (2015) for correlated shocks across exporters and importers). Given these data limitations and the structure of the model, we cannot capture the impact of international trade shocks on domestic growth rates. See Dhyne et al. (2017) for a further analysis of imported shocks. Here, we simply allow for imports to be used as inputs by calculating input shares as $\omega_{ij} = \frac{r_{ij}}{\sum_i r_{ij} + m_j}$ where m_i denotes the import value of firm j. We then re-estimate the model using these adjusted input shares. Table A.IV presents these extra results, showing that the model now generates a slightly higher aggregate volatility of 1.20% in the baseline case. When accounting for aggregate comovement however, this difference evaporates, suggesting that imported shocks can be correlated at higher levels of aggregation.

Third, we alternatively use unweighted averages ($\bar{g}_{It} = \frac{1}{n} \sum_{i \in I} g_{it}$) in the calculation of aggregate growth rates. When using weighted averages, firms with high value-added shares drive the growth rate of their sector. This can lead to an under- or overestimation of the model predictions of aggregate volatility, guided by the growth rates of the highest value-added firms. Table A.V presents these results. The baseline results are identical to the main results of this paper, but the unweighted version actually generates slightly increasing volatility when demeaning at more disaggregated levels.

Finally, we use several other measures for firm growth, as it is difficult to measure firm-level productivity movement due to the measurement in micro data, identifying assumptions on the link between productivity and observables etc. (Gabaix (2011)). We derive firm growth using labor productivity (expressed as value-added per employee), sales per worker, sales, value-added and FTE as as alternative measures to calculate g_{it}. Table A.VI presents these additional results. The main findings are robust to these alternative measures. Note that labor productivity generates higher predictions for aggregate volatility. This is consistent with the observation in Acemoglu et al. (2012), their footnote 28: "To the extent that total factor productivity is measured correctly, it approximates the variability of idiosyncratic sectoral shocks. In contrast, the variability of sectoral value-added is determined by idiosyncratic shocks as well as the sectoral linkages, as we emphasized throughout the paper."

V Additional evidence

In this Section, we first evaluate the non-linear effect of the labor share α on the model predictions. Then, we empirically confirm that our influence vector satisfies the sufficient condition for micro shocks to surface in the aggregate as stipulated in Gabaix (2011) and Acemoglu et al. (2012).

A Counterfactual analysis of a change in the labor share

We investigate how the relative contributions of the network structure and sales to final demand change, as the labor share α varies across a sequence of counterfactual economies. In particular, when $\alpha = 1$, the intermediate input share is zero, output only requires labor, all sales are to final demand and aggregate value-added equals gross output. The propagation mechanism through the network of production is then mute and all aggregate variance in the model directly stems from large firms selling to final demand. As the labor share α decreases, relatively more intermediate inputs are used in production. This has a non-linear impact on the influence vector from the interaction between an increase in the Leontief inverse $[\mathbf{I} - (1-\alpha)\mathbf{\Omega}]^{-1}$ and a decrease in the constant term $\frac{\alpha}{Y}$. When $\alpha \to 0$, the propagation mechanism of the network dominates the contribution of micro shocks to aggregate volatility. We simulate $\mathbf{v}$ for incremental steps in α and re-estimate aggregate volatility for the four specifications from Section IV. We use the baseline specification of $d\hat{\varepsilon}_{it} = g_{it}$ for exposition below, but results are quasi identical for the demeaned growth rates.

Figure III shows the comparative static results of changes in α. The X-axis shows the counterfactual α, ranging from 0.01 to 0.99; the Y-axis shows predicted aggregate volatility under this restriction. First consider the analysis of the benchmarks. The dotted line represents the prediction of the classical diversification argument as $\hat{\sigma}_{\Delta y} \propto \frac{1}{\sqrt{n}}$, or 0.09% in our data. By construction, this is independent of the labor share in the economy.

The dashed line shows predicted volatility using the Network benchmark. Volatility decreases monotonically as α increases: with a low α, the network structure is important in production, and multiplier effects generate sizable aggregate fluctuations. However, as α increases, multiplier effects die out, $\mathbf{v}$ converges to $\frac{1}{n}$ in the limit as $\alpha \rightarrow 1$, and aggregate volatility declines exponentially. This observation is consistent with Acemoglu et al. (2012), who provide conditions for aggregate volatility to remain bounded away from zero, even as the number of production units in the economy tends to infinity.[20]

The dash-dotted line depicts aggregate volatility for the Final Demand benchmark. In this setting, all firms source equally from all other firms in the economy so that $\omega_{ij} = \frac{1}{n}, \forall i, j$. The production structure thus represents a complete network, in which the homogeneous intensity of sourcing depends on α. For large n, we can approximate the influence vector as $\mathbf{v} \simeq \frac{\alpha}{Y}\mathbf{b}$. Predicted volatility as a function of α is then a straight line, with slope given by $\sqrt{\sum_{i=1}^{n}\left(\frac{b_i}{Y}\right)^2}$: the more skewed the distribution of sales to final demand, the steeper the slope.

Now consider the predictions of the model, depicted by the solid line. Predicted volatility follows a U-shaped curve as a function of α: if α is close to zero, almost all aggregate volatility is generated from the propagation of shocks in the network. Increasing α however, leads to a drop in volatility until around $\alpha = 0.2$, after which volatility increases again. The dominant mechanism in aggregate volatility from micro origins then becomes the channel of sales to final demand. As in the limit $\alpha \rightarrow 1$, all volatility is generated from sales to final demand, coinciding with the sales shares vector of the economy.

The model captures both sources of heterogeneity in the presence of a production network. The simulations show that the level of α matters in explaining which channel is dominant in generating aggregate volatility. This is not surprising in view of $\mathbf{v}$ as a centrality measure: $\mathbf{v}$ collapses to a Bonacich (1987) centrality in the case of Acemoglu et al. (2012), where all the adjustments are made inside the network. The formulation of our model however is consistent

[20] Note that Acemoglu et al. (2012) perform a counterfactual analysis for a sequence of economies as $n \rightarrow \infty$. We keep n fixed and vary the importance of the network structure in aggregate output through α.

with a generalized Bonacich centrality where differences in final demand matter, outside the network of intermediate goods. The interaction of α in the model then results in non-linearities in the prediction of aggregate volatility through the influence vector.

These mechanisms are intuitive when we think of different types of economies. A post-industrial economy heavily relies on labor or human capital as input. The majority of idiosyncratic shocks has to surface directly in the aggregate, as few transmission mechanisms are active. Conversely, a highly industrialized economy heavily depends on its production network and shocks then dominantly propagate throughout this network. Our model could serve as a framework for real economies with any share of labor and inputs.

B Empirical distribution of the influence vector

Gabaix (2011) and Acemoglu et al. (2012) derive a sufficient condition under which the skewed distribution of the influence vector generates aggregate fluctuations from micro origins. We now show that the distribution of the influence vector from our generalized model satisfies this condition.

Figure IV depicts the empirical counter-cumulative distribution function (CCDF) of the influence vector predicted by our model. The X-axis labels the values of the elements of the influence vector $v_i \in (0, 1)$, while the Y-axis represents the probability that i's influence is larger than the observed value. Both axes are in log scales. Most firms have a negligible total impact on the economy, in the order of a share of 10^{-9} to 10^{-7}. However, some firms are very central in the economy, with a total influence of more than 1%, as indicated by the outliers in the plot.

Idiosyncratic shocks to firms can thus contribute to aggregate fluctuations if some firms are very influential relative to the rest of the economy. In particular, if the tail of the distribution of the influence vector can be approximated by a power law distribution with exponent $\beta \in [1, 2)$, aggregate fluctuations are predicted to have a size of $\frac{\sigma}{n^{(1-1/\beta)}}$. This is much larger than $\frac{\sigma}{\sqrt{n}}$, the size suggested by the diversification argument as in Lucas (1977). We define a *power law distribution* as follows:

DEFINITION: The distribution of a variable x is consistent with a power law distribution if, there exists a $\beta > 1$ and a slowly varying function $L(x)$ so that for all $a > 0$, $\lim_{x\to\infty} L(ax)/L(x) = 1$, the counter-cumulative distribution function can be written as[21]

[21] Typical representations are $L(x) = \kappa$, or $L(x) = \kappa \ln x$, where κ is some non-zero constant (Gabaix (2011)). In

$$\Pr(v > x) = x^{-\beta}L(x)$$

The shape parameter β captures the scaling behavior of the tail of the distribution. As β decreases, the distribution becomes more skewed, generating more observations with extreme values. For the range $\beta \in [2, \infty)$, the first two moments of the distribution are well-defined, and Gabaix (2011) and Acemoglu et al. (2012) show that micro fluctuations average out in the aggregate at a rate consistent with the diversification argument. However for $\beta \in [1, 2)$, the second moment diverges, i.e. the distribution is fat-tailed as the variance tends to infinity (see e.g. Gabaix (2009)). In this parameter range, the law of large numbers does not hold and aggregate volatility decays more slowly than presented by the standard diversification argument. For $\beta = 1$ (i.e. Zipf's law), volatility decays at a rate proportional to $1/\ln(n)$. For values of $\beta < 1$, none of the moments are defined.

We fit a power law to the tail of the distribution of the influence vector generated by our data. Results are in Table VII; the solid line in Figure IV shows the graphical representation of the fit. We estimate β using the numerical maximum likelihood Hill estimator with endogenous cutoff for the tail (Clauset et al. (2009)). The estimated coefficient for the influence vector from the model is $\hat{\beta} = 1.12$, with a standard error of 0.03, again confirming that micro shocks can surface in the aggregate.[22]

The majority of the literature imposes exogenous cutoffs that represent some fraction of the observations (e.g. top 5% observations) or a visual cutoff to define the minimum value of x, x_{min}, above which the fit is performed. However, estimated β's can be very sensitive to changes in the cutoff, since there is much less mass in the tail. This generates biased results for the scaling behavior of the distribution and we are particularly interested in the behavior of these outliers. The endogenous cutoff method ensures the best fit given the data by minimizing the Kolmogorov-Smirnov distance between the proposed fit and the data. Additionally, many papers resort to an OLS estimation on the log-density function $\ln p(x) = \ln L(x) - \beta \ln x$, where $L(x)$ is constant for large enough values of x. Using the x_{min} endogenously dictated by the MLE method and estimating β using OLS, we find an estimate of $\hat{\beta} = 1.20$. In any case, these results confirm that the distribution of the influence vector is sufficiently skewed and consistent

its most familiar form, the log of the probability density function can then be written as $\ln p(x) = \ln L(x) - \beta \ln x$ for sufficiently large values of x.

[22] For the maximum likelihood estimation, approximate standard errors are calculated as $\frac{\hat{\beta}-1}{\sqrt{N}}$ (Clauset et al. (2009)).

with a power law with infinite variance, satisfying the condition stipulated in Gabaix (2011) and Acemoglu et al. (2012).

Furthermore, it is possible to fit a power law to any distribution to get an estimate for β and other distributions might actually provide better fits. As a natural alternative, we fit a log-normal distribution on the influence vector using MLE (see Saichev et al. (2010) for a lengthy discussion on power law versus log-normal in firm sizes and other examples). We impose the same cutoffs for these estimations as in our earlier estimations. The log-normal fit is represented by the dashed line in Figure IV. We then perform a Vuong (1989) likelihood ratio test to compare the fits of both models. In particular, the test statistic is given by $R = \ln \frac{L(\theta_1|x)}{L(\theta_2|x)}$, where L is the likelihood function and θ_1 is a vector of parameters for model 1 (power law). Similarly for model 2 (log-normal). The sign of R indicates which model is closer to the true (unobserved) model: if $R > 0$, the test statistic presents evidence in favor of model 1. Results are presented in the last two rows of Table VII. R presents evidence that the log-normal distribution is closer to the true model, with a p-value of 9%. While our model is data driven and does not crucially depend on the power law assumption, this does suggest however that in general, there is room for alternative models with log-normal distributions of firm size and influence.

VI Conclusion

This paper shows that firm-level idiosyncrasies are important drivers for aggregate fluctuations. We have developed a model in which both firm productivity and the network structure of production interact, providing channels for random growth to surface in the aggregate output of an economy. This view generalizes existing contributions on the microeconomic sources of aggregate volatility in which either of both channels matter.

More importantly, we are the first to evaluate the impact of these two sources of heterogeneity on aggregate fluctuations at this level of detail. We have calibrated and estimated the model using firm-to-firm relationship data for Belgian firms and found that a large part of aggregate volatility originates at the firm level, even after accounting for aggregate movement and highly disaggregated sector comovement.

This paper presents a framework in which heterogeneous firms depend on each other in the network structure of production. We believe this has interesting applications in several domains. Using this type of micro data, the model proposes an important policy tool for the evaluation and simulation of different types of micro shocks on the economy at large. It provides a yardstick for the impact of idiosyncrasies to the largest firms, the most connected firms, geographically clustered firms etc. Importantly, the model reveals patterns that remain hidden when using a standard firm-level framework: for instance, the amount of fragmentation of the value chain dictates the importance of additional network multipliers to proliferate.

All European countries have the same VAT system and tax authorities in every Member State have to collect similar data for VAT transactions as does Belgium. If this data is made available for research, it would be very useful to test the model across different economies with different levels of interaction between the network structure and sales to final demand.

This paper embeds the Melitz (2003) model, where inputs now contain the whole network of (intermediate) production rather than only labor. Future work might explore how the network structure changes when firms open up to international trade and reshuffling appears as firms drop below the productivity thresholds for production from international competition. A first step in this direction is di Giovanni and Levchenko (2012), who show that opening up to trade leads to higher volatility from the reallocation of factors to larger and more productive firms. The effect on the network structure of production however, is not explored yet.

Glenn Magerman: Université libre de Bruxelles (ECARES), National Bank of Belgium

Karolien De Bruyne: KU Leuven

Emmanuel Dhyne: National Bank of Belgium, Université de Mons

Jan Van Hove: KU Leuven, KBC Group

A Model derivation

Demand Each household is endowed with one unit of labor, supplied inelastically and the size of the economy is normalized to 1 so that labor market clearing implies $\sum_i l_i = 1$, where l_i is the amount of labor needed to produce good i. Each household then maximizes utility over CES preferences:

$$U = \left(\sum_{i=1}^{n} q_i^{\rho} \right)^{1/\rho} \tag{13}$$

subject to its budget constraint $\sum_{i=1}^{n} p_i q_i = Y$, where q_i is the quantity consumed of good i and $\eta = \frac{1}{1-\rho} > 1$ is the elasticity of substitution, common across goods. Labor is paid in wages w and it is the only source of value-added in this economy, so that $Y = w$, where Y represents total spending on final goods. Residual demand then follows $q_i = \frac{p_i^{-\eta}}{P^{1-\eta}} Y$.

Firm environment After payment of an investment cost $f_e > 0$, each firm i observes its own productivity ϕ_i (drawn from a Pareto distribution) and also receives a contingent blueprint of production, stipulated by particular input requirements ω_{ji}. After payment of fixed costs, output x_i follows a Cobb-Douglas production technology with constant returns to scale:

$$x_i = (z_i l_i)^{\alpha} \prod_{j=1}^{n} x_{ji}^{(1-\alpha)\omega_{ji}} \tag{14}$$

with associated unit cost function:

$$c_i = B_i \left(\frac{w}{z_i} \right)^{\alpha} \prod_{j=1}^{n} p_{ji}^{(1-\alpha)\omega_{ji}} \tag{15}$$

Cost minimization can be about cheaper, better or more novel inputs, all isomorphic to the model. Marginal costs are given by $\frac{c_i}{\phi_i}$ and from monopolistic competition and CES preferences, prices are set as a constant markup over marginal cost $p_i = \frac{c_i}{\rho \phi_i}$. Hence, prices are also stochastic through c_i.

Competitive equilibrium A static competitive equilibrium is given by the following equilibrium quantities: prices $(p_1, ..., p_n)$, final demands $(q_1, ..., q_n)$, quantities (l_i, x_{ij}, x_i), profits

$(\pi_1, ..., \pi_n)$ and cutoff productivities $(\phi_1^*, ..., \phi_n^*)$. Total cost is given by $\Gamma_i = \left[f + \frac{x_i}{\phi_i}\right] c_i$.[23] The firm's problem is to optimize the amount of inputs l_i and x_{ji}, delivering its output price p_i, taking as given the prices of inputs w and p_{ji}, and its total cost function Γ_i. The firm's problem can then be written as:

$$\pi_i = p_i x_i - w l_i - \sum_{j=1}^{n} x_{ji} p_j - w f - f \sum_{j \in \mathcal{S}_i} p_j \tag{16}$$

subject to $x_i = (z_i l_i)^{\alpha} \prod_{j=1}^{n} x_{ji}^{(1-\alpha)\omega_{ji}}$, and where we have used the fact that l_i is the amount of variable labor needed to produce x_i/ϕ_i output and x_{ji} is the amount of variable input j needed to produce x_i/ϕ_i. $\mathcal{S}_i$ is the set of input suppliers j to i. Plugging in x_i and taking first-order conditions with respect to l_i : $\frac{\partial \pi_i}{\partial l_i} = \alpha p_i z_i^{\alpha} l_i^{\alpha-1} \prod_{j=1}^{n} x_{ji}^{(1-\alpha)\omega_{ji}} = w$ and with respect to x_{ji}: $\frac{\partial \pi_i}{\partial x_{ji}} = (1-\alpha)\omega_{ji} p_i (z_i l_i)^{\alpha} x_{ji}^{(1-\alpha)\omega_{ji}-1} \prod_{k \neq j}^{n-1} x_{ki}^{(1-\alpha)\omega_{ki}} = p_j$, leads to optimal factor and inputs demands:

$$\begin{cases} w l_i = \alpha p_i x_i \\ x_{ji} p_j = (1-\alpha)\omega_{ji} p_i x_i \end{cases}$$

Plugging x_{ij} back into goods market clearing $x_i = \sum_j x_{ij} + q_i$ leads to equilibrium revenues per firm:

$$x_i = \sum_{j=1}^{n} x_{ij} + q_i$$

$$\Leftrightarrow r_i \equiv \underbrace{p_i x_i}_{\text{total revenue}} = \underbrace{\sum_{j=1}^{n} p_i x_{ij}}_{\text{downstream intermediate revenue}} + \underbrace{p_i q_i}_{\text{final demand revenue}}$$

$$\Leftrightarrow r_i = \sum_{j=1}^{n} p_i \frac{(1-\alpha)\omega_{ij} p_j x_j}{p_i} + p_i q_i$$

$$\Leftrightarrow r_i = (1-\alpha) \sum_{j=1}^{n} \omega_{ij} r_j + \left(\frac{p_i}{P}\right)^{1-\eta} Y \tag{17}$$

[23] This specification is similar to Bernard et al. (2007) when there are no intermediate goods, in which case Γ_i collapses to $\left[f + \frac{q_i}{\phi_i}\right] c_i$. It is also similar to Melitz (2003) when there are no factors of production, only labor, in which case Γ_i collapses to $f + \frac{q_i}{\phi_i}$, where wages are normalized to one.

where the last equation follows from plugging in optimal demands from the utility maximization. From constant markups, this allows us to write profits as:

$$\pi_i = \frac{r_i}{\eta} - c_i f = \underbrace{\frac{1-\alpha}{\eta}\sum_{j=1}^{n}\omega_{ij}r_j}_{\text{downstream}} + \underbrace{\left(\frac{\rho\phi_i P}{c_i}\right)^{\eta-1}\frac{Y}{\eta}}_{\text{final demand}} - c_i f$$

The network structure of production We can write (17) in matrix form as

$$\mathbf{r} = (1-\alpha)\mathbf{\Omega}\mathbf{r} + \mathbf{b}$$

where directed edges from i to j are given by the non-negative elements of the adjacency matrix $\omega_{ij} \in \mathbf{\Omega}$. This leads to

$$\mathbf{r} = \left[\mathbf{I} - (1-\alpha)\mathbf{\Omega}\right]^{-1}\mathbf{b} \tag{18}$$

where $\mathbf{I}$ is the $n \times n$ identity matrix and $\mathbf{\Omega}$ is the input-output matrix, with input shares ω_{ij} as elements.

Finally, we derive the relationship between the revenue vector and the influence vector. A variant of Hulten (1978) theorem as in Acemoglu et al. (2012), states that, when individual firms are hit with Harrod-neutral productivity shocks ε_i, aggregate output Y changes as $dY = \frac{p_i x_i}{\sum_j p_j x_j} d\varepsilon_i$. Whenever production follows a Cobb-Douglas specification, Hulten (1978) holds, and accounting for heterogeneous firms does not influence the outcome. The result is thus the same as in Acemoglu et al. (2012). Let production be given by $x_i = e^{\alpha\varepsilon_i} f(x_{i1}, ..., x_{in}, l_i)$. Goods market clearing in the economy is given by $x_i = \sum_j x_{ij} + q_i$. The social optimum is given by

$$\max_{q_i, l_i, x_{ij}} \quad U(q_1, ..., q_n) \quad s.t. \begin{cases} \sum_{j=1}^{n} x_{ij} + q_i = e^{\alpha\varepsilon_i} f(\cdot) \\ \sum_{i=1}^{m} l_i = 1 \end{cases}$$

The Lagrangian of the economy can be written as

$$\mathcal{L} = U(\cdot) + \sum_{i=1}^{n} p_i \left[e^{\alpha\varepsilon_i} f(\cdot) - q_i - \sum_{j=1}^{n} x_{ij}\right] + Y\left[1 - \sum_{i=1}^{m} l_i\right]$$

Aggregate output is given by, $Y = w = \sum_{i=1}^{n} p_i q_i = \alpha \sum_{i=1}^{n} p_i x_i$. If a Harrod-neutral shock hits individual firms, welfare changes as

$$dY = \sum_{i=1}^{n} p_i \left[e^{\alpha\varepsilon_i} f(\cdot) d\alpha\varepsilon_i\right] = \sum_{i=1}^{n} p_i x_i d\alpha\varepsilon_i$$

Finally, since $Y = \alpha \sum_{i=1}^{n} p_i x_i$, $dY = \frac{p_i x_i}{\alpha \sum_{j=1}^{n} p_j x_j} d\alpha\varepsilon_i$ and so $v_i \equiv \frac{dY}{d\varepsilon_i} = \frac{p_i x_i}{\sum_{j=1}^{n} p_j x_j}$. $\square$

Aggregate output and the influence vector for the economy From the firm's production function, plug in optimal inputs l_i and x_{ji}:

$$x_i = \left(\frac{z_i \alpha p_i x_i}{w}\right)^{\alpha} \prod_{j=1}^{n} \left(\frac{(1-\alpha)\omega_{ji} p_i x_i}{p_j}\right)^{(1-\alpha)\omega_{ji}}$$

Taking logarithms and using $\sum_j \omega_{ji} = 1$:

$$\begin{aligned} \ln x_i \;\; = \;\; & \alpha \underbrace{\ln z_i}_{\equiv\varepsilon_i} + \underbrace{\alpha \ln \alpha + (1-\alpha)\ln(1-\alpha)}_{\equiv C} + \alpha \ln p_i + \alpha \ln x_i - \alpha \ln w \\ & + (1-\alpha) \sum_{j=1}^{n} \omega_{ji} \ln \omega_{ji} + (1-\alpha)\ln p_i + (1-\alpha)\ln x_i - (1-\alpha)\sum_{j=1}^{n} \omega_{ji} \ln p_j \end{aligned}$$

where $H = -\sum_j \omega_{ji} \ln \omega_{ji}$ is the Shannon (1948) entropy of the system.[24] Collecting terms and dividing by α:

$$\ln w = \varepsilon_i + C/\alpha + \ln p_i + \frac{(1-\alpha)}{\alpha} \sum_{j=1}^{n} \omega_{ji} \ln \omega_{ji} - \frac{(1-\alpha)}{\alpha} \sum_{j=1}^{n} \omega_{ji} \ln p_j$$

Premultiply with the i-th element of the influence vector $\mathbf{v} = \frac{\alpha}{Y}\left[\mathbf{I} - (1-\alpha)\boldsymbol{\Omega}\right]^{-1} \mathbf{b}$ and sum over all firms i:

$$\ln w \equiv y = \mu + \mathbf{v}'\varepsilon$$

where we have used $\sum_{i=1}^{n} v_i = 1$ and μ is a mean shifter for the output of the economy, independent of the vector of shocks ε: $\mu = C/\alpha + \sum_{i=1}^{n} v_i \ln p_i + \frac{(1-\alpha)}{\alpha} \sum_{i=1}^{n} \sum_{j=1}^{n} v_i \omega_{ji} \ln \omega_{ji} - \frac{(1-\alpha)}{\alpha} \sum_{i=1}^{n} \sum_{j=1}^{n} v_i \omega_{ji} \ln p_j$.

[24] Shannon entropy represents the average amount of information contained in an event and ranges between 0 and $\ln n$. A higher entropy indicates more potential states of the system. If $\omega_{ji} = 1$ for all i, then all firms only use 1 firm as input and the economy is represented by a circle graph. As $\omega_{ji} \to 0$ for all i and j, the number of possible states increases. Interestingly in our model, H shows that the realized production recipes determine the possible outcome space of the economy, dictated by the phase space of the system.

Cutoff productivities Firms pay a fixed investment cost $f_e > 0$, before entering the market, paid in terms of labor and all inputs. After payment, this cost is sunk and it entitles the firm to a productivity draw ϕ_i and a blueprint for production of a particular product (i.e. its set of ω_{ji}). ϕ_i is drawn from a Pareto distributed cumulative distribution function $G(\phi) \equiv 1 - \left(\frac{\phi_{min}}{\phi}\right)^{\theta}$ with support $[\phi_{min}, \infty)$, where $\phi \geq \phi_{min} > 0$ and we require $\theta > 1$ for the distribution to have a finite mean. θ is the shape parameter of the distribution, common across all firms and is inversely related to the variance of the distribution. The blueprint is contingent, in that it stipulates how to produce a product in-house if it is not available on the market as an input. After observing its productivity and blueprint for production, firms decide whether to start producing at their technologies Γ_i or exit immediately without producing. The existence of fixed costs of production dictates that there is a cutoff productivity ϕ^*, below which firms cannot make positive profits in expectation and it is endogenously determined by the zero profit condition for the marginal firm:

$$\phi_i^{*\eta-1} = \frac{\eta \left(c_i f - \frac{1-\alpha}{\eta} \sum_{j=1}^{n} \omega_{ij} r_j\right) \left(\frac{c_i}{\rho}\right)^{\eta-1}}{Y P^{\eta-1}}$$

Note that cutoff productivities are i-specific in our setup, due to the input requirements of downstream firms and the obtained blueprint for production by i. Also note that we take as given the productivities of other firms.

B Model with heterogeneous labor shares

The baseline model assumes a common labor share across all firms in the economy, in line with previous input-output models at the sector level as in Long and Plosser (1983); Acemoglu et al. (2012). We derive a simple extension with heterogeneous labor shares at the firm level and re-estimate the model with this extra source of heterogeneity. Output of firm i is then given by

$$x_i = (z_i l_i)^{\alpha_i} \prod_{j=1}^{n} x_{ji}^{(1-\alpha_i)\omega_{ji}} \tag{19}$$

where α_i denotes the labor share for firm i. Revenues can be written as

$$r_i = \sum_{j=1}^{n} (1 - \alpha_j)\omega_{ij} r_j + p_i q_i \tag{20}$$

where downstream revenues naturally depend on the intermediate input share of downstream buyers j. The impact of this heterogeneity on aggregate fluctuations is *a priori* not clear, as the diffusion of shocks depends on the net impact of all downstream buyers j.

We then estimate the model under these relaxed constraints. We first obtain firm-level labor shares as $\alpha_i = \frac{\text{labor cost}_i}{\text{inputs}_i + \text{labor cost}_i}$, using information from the annual accounts in 2012. Figure V shows the distribution of labor shares for firms used in the estimation procedure, confirming our earlier estimate of a common labor share around $\alpha = 0.2$.

We re-estimate the model using (20). Results for aggregate volatility are given in Table A.I. Accounting for heterogeneous labor shares turns out to have a relatively small impact on the model's prediction of aggregate volatility, underlining our choice for the baseline specification of the model in the main text.

C Model with capital goods

One drawback of the NBB B2B dataset is that transactions can be either intermediate inputs or capital goods (e.g. machinery, construction). We cannot directly disentangle these goods, as we only observe sales values between firms across all economic activities. This generates a bias in the estimation of the Leontief inverse in our model as capital goods potentially skew the input shares matrix of the economy towards this type of goods.

Here, we develop a simple extension to the model, allowing for the factor capital in production. Output of firm i is now given by

$$x_i = (z_i l_i)^{\alpha} \, k_i^{\beta} \prod_{j=1}^{n} x_{ji}^{(1-\alpha-\beta)\omega_{ji}} \tag{21}$$

where k_i are capital inputs for firm i with capital share β. Note that capital goods can be sold as output and used as input, but it is not part of the input-output matrix $\mathbf{\Omega}$. Capital goods become part of the final demand residual as investment, consistent with the typical National Accounting Identity. The influence vector then becomes $\mathbf{v} = \frac{\alpha+\beta}{\tilde{Y}} \left[\mathbf{I} - (1-\alpha-\beta)\tilde{\mathbf{\Omega}}\right]^{-1} \tilde{\mathbf{b}}$, where $\tilde{Y}$, $\tilde{\mathbf{\Omega}}$ and $\tilde{\mathbf{b}}$ denote model elements, corrected for capital goods. This specification implies that the baseline model over-estimates the network channel while it underestimates the final demand channel. The impact in the aggregate is *a priori* unclear.

We calibrate $\beta = 0.3$ and calculate the intermediate inputs ratio for sector I in total business sales from the sectoral Input-Output matrix for Belgium at the NACE 2-digit level in 2012 as $\frac{\text{total inputs}_I}{\text{total inputs}_I + \text{investment}_I}$. Figure VI shows the distribution of sectors and the corrected intermediate inputs ratio: 29 out of 70 sectors are affected by this correction factor. The distribution at the firm level is very similar (not reported). After correction, 14% of the total value of previous business relationships are redirected towards final demand.

Table A.II reports results for these corrected flows. The model prediction is now 1.29%, slightly higher than the baseline setting in Table IV.

D TFP estimation

TFP estimation follows the CompNet procedure (Lopez-Garcia et al. (2014)). In particular, real value-added output Y_{it} of firm i in year t is assumed to follow a Cobb-Douglas production function, so that:

$$Y_{it} = A_{it} K_{it}^{\alpha_k} L_{it}^{\alpha_L} M_{it}^{\alpha_M}$$

where A_{it} is productivity of i at time t, K_{it}, L_{it} and M_{it} are capital, labor and intermediate inputs respectively. In logs:

$$y_{it} = \alpha_0 + \alpha_{kt} k_{it} + \alpha_{lt} l_{it} + \alpha_{mt} m_{it} + \varpi_{it} + u_{it} \tag{22}$$

where $\ln A_{it} = \alpha_0 + \varpi_{it} + u_{it}$, and α_0 represents mean efficiency across firms and time, while $\varpi_{it} + u_{it}$ denotes the firm-specific deviation from the mean; ϖ_{it} is the unobserved productivity component (known by the firm) and u_{it} is the i.i.d. error term.

As the firm optimizes its inputs given its own observed productivity, variable inputs are arguably correlated with ϖ_{it}. Estimating (22) using OLS then leads to biased estimated parameters of interest. Following Wooldridge (2009), we structurally estimate (22) using a control function approach. In particular, consider the following assumptions:

ASSUMPTION 1 (Monotonicity): $m_{it} = f(k_{it}, \varpi_{it})$, where $f(\cdot)$ is strictly monotonically increasing in ϖ_{it}. Then, unobserved productivity can be inverted out, so that: $\varpi_{it} = f^{-1}(k_{it}, m_{it})$.

ASSUMPTION 2 (Capital): $K_{it} = I_{it-1} + (1-\delta)K_{it-1}$ where I_{it} is the investment of firm i at t. K_{it-1} is independent of current shocks.

ASSUMPTION 3 (Markov Process): $\varpi_{it} = E\left(\varpi_{it} | \varpi_{it-1}\right) + \nu_{it}$, so that productivity follows a first-order Markov process.

Then, we can write:

$$y_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \psi\left(m_{it}, k_{it}\right) + u_{it} \tag{23}$$

where $\psi\left(m_{it}, k_{it}\right) = \alpha_0 + \alpha_K k_{it} + f^{-1}(k_{it}, m_{it})$ is a non-parametric function, proxied by a third-order polynomial in capital and intermediate inputs. We can then estimate (23) using GMM or an instrumental variables (IV) approach, leading to efficient estimation of the parameters of interest, with robust standard errors (and without relying on bootstrap methods). We follow Petrin and Levinsohn (2012), and use a pooled IV approach where we instrument current values of l_{it} with lagged values l_{it-1}. (23) is estimated within the 2-digit NACE industry level, i.e. firms within that industry share the same technologies and deviations from mean output given the same inputs are firm-specific. For investment/inputs, we use value-added deflators at the NACE 2 digit level. Capital stock is deflated using gross fixed capital formation deflators. We include year fixed effects to purge general yearly trends and cluster standard errors at the i level.

E Identification of shocks

First, we derive the economy and individual growth rates from Section III. The dynamic interpretation of the model is given by

$$\Delta \ln Y = \ln Y_t - \ln Y_{t-1} = \mu_t + \mathbf{v}_t' \varepsilon_t - (\mu_{t-1} + \mathbf{v}_{t-1}' \varepsilon_{t-1})$$

We assume $\mu_t = \mu$ (steady state) and $\mathbf{v}_t' = \mathbf{v}'$ (fixed network from *ex ante* draws). Then $\Delta \ln Y = \mathbf{v}'(\varepsilon_t - \varepsilon_{t-1})$, and so the variance of GDP growth is

$$Var(\Delta \ln Y) = Var(\mathbf{v}'(\varepsilon_t - \varepsilon_{t-1}))$$

$$\iff Var(\Delta \ln Y) = \mathbf{v}'^2[Var(\varepsilon_t) + Var(\varepsilon_{t-1}) - 2Cov(\varepsilon_t, \varepsilon_{t-1})]$$

We assume that ε_t is independent over time, so that $\sigma_{\Delta y} = \sqrt{2\sum_{i=1}^n v_i^2 \sigma_i^2}$ (this independence assumption is also reflected in the structural estimation of TFP). Then, $g_{it} = \varepsilon_{it} - \varepsilon_{it-1}$ results in $Var(g_{it}) = 2\sigma_i^2$, plugging back in leads to $\sigma_{\Delta y} = \sqrt{\sum_{i=1}^n v_i^2 Var(g_{it})}$.

References

Acemoglu, D., Akcigit, U., and Kerr, W. (2015). Networks and the macroeconomy: An empirical exploration. *NBER Macroeconomics Annual*.

Acemoglu, D., Carvalho, V., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016.

Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2017). Microeconomic origins of macroeconomic tail risks. *American Economic Review*, 107(1):54–108.

Amiti, M. and Konings, J. (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia. *American Economic Review*, 97(5):1611–1638.

Atalay, E. (2015). How important are sector shocks? Mimeo.

Axtell, R. (2001). Zipf distribution of firm sizes. *Science*, 293(5536):1818–1820.

Baqaee, D. R. (2015). Cascading failures in production networks. Mimeo.

Behrens, K., Corcos, G., and Mion, G. (2013). Trade crisis? what trade crisis? *Review of Economics and Statistics*, 95(2):702–709.

Bernard, A., Jensen, B., Redding, S., and Schott, P. (2009). The margins of us trade. *The American Economic Review Papers and Proceedings*, 99:487–493.

Bernard, A., Moxnes, A., and Saito, Y. (2015). Production networks, geography and firm performance. *NBER Working Paper*, 21082.

Bernard, A., Redding, S., and Schott, P. (2007). Comparative advantage and heterogeneous firms. *Review of Economic Studies*, 74:31–66.

Boehm, C., Flaaen, A., and Pandalai-Nayar, N. (2015). Input linkages and the transmission of shocks: Firm-level evidence from the 2011 tohoku earthquake. Mimeo.

Bonacich, P. (1987). Power and centrality: A family of measures power and centrality. *The American Journal of Sociology, 92 (5): 1170-1182.*

Bricongne, J.-C., Fontagné, L., Gaulier, G., Taglioni, D., and Vicard, V. (2012). Firms and the global crisis: French exports in the turmoil. *Journal of International Economics*, 82:134–146.

Carvalho, V. (2014). From micro to macro via production networks. *Journal of Economic Perspectives*, 28(4):23–48.

Carvalho, V. and Gabaix, X. (2013). The great diversification and its undoing. *American Economic Review*, 103(5):1697–1727.

Carvalho, V. and Grassi, B. (2016). Large firm dynamics and the business cycle. Mimeo.

Carvalho, V., Nirei, M., Saito, Y., and Tahbaz-Salehi, A. (2015). Supply chain disruptions: Evidence from the great east japan earthquake. Mimeo.

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

Christiano, L., Eichenbaum, M., and Evans, C. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45.

Clauset, A., Shalizi, C., and Newman, M. (2009). Power-law distributions in empirical data. *SIAM Review*, 51:661–703.

Cochrane, J. (1994a). Permanent and transitory components of gnp and stock prices. *Quarterly Journal of Economics*, 109(1):241–265.

Cochrane, J. (1994b). Shocks. *Carnegie-Rochester Conference Series on Public Policy*, 41:295–364.

Commission, E. (2015). Autumn 2015 economic forecast: Moderate recovery despite challenges.

Dhyne, E., Kikkawa, A., and Magerman, G. (2017). Transmission of foreign trade shocks through the economy. Mimeo.

Dhyne, E., Magerman, G., and Rubinova, S. (2015). The belgian production network 2002-2012. *National Bank of Belgium Working Paper*, 288.

di Giovanni, J. and Levchenko, A. (2012). Country size, international trade, and aggregate fluctuations in granular economies. *Journal of Political Economy*, 120(6):1083–1132.

di Giovanni, J., Levchenko, A., and Mejean, I. (2014). Firms, destinations, and aggregate fluctuations. *Econometrica*, 82(4):1303–1340.

Dixit, A. K. and Stiglitz, J. (1993). Monopolistic competition and optimum product diversity: Reply. *American Economic Review*, 83(1):302–304.

Dupor, W. (1999). Aggregation and irrelevance in multi-sector models. *Journal of Monetary Economics*, 43:391–409.

Durlauf, S. (1993). Non ergodic economic growth. *Review of Economic Studies*, 60:349–366.

Elliot, M., Golub, B., and Jackson, M. (2014). Financial networks and contagion. *American Economic Review*, 104(10):3115–3153.

Foerster, A., Sarte, P., and Watson, M. (2011). Sector versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy*, 119:1–38.

Gabaix, X. (2009). Power laws in economics and finance. *Annual Review of Economics*, 1:255–293.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772.

Gourio, F. and Kashyap, A. (2007). Investment spikes: New facts and a general equilibrium exploration. *Journal of Monetary Economics*, 54:1–22.

Horvath, M. (1998). Cyclicality and sector linkages: Aggregate fluctuations from sector shocks. *Review of Economic Dynamics*, 1:781–808.

Horvath, M. (2000). Sector shocks and aggregate fluctuations. *Journal of Monetary Economics*, 45:69–106.

Hulten, C. R. (1978). Growth accounting with intermediate inputs. *Review of Economic Studies*, 45:511–518.

Jalan, J. and Ravallion, M. (1999). Are the poor less well insured? evidence on vulnerability to income risk in rural china. *Journal of Development Economics*, 58:61–81.

Kramarz, F., Martin, J., and Mejean, I. (2015). Volatility in the small and in the large: Diversification in trade networks. Mimeo.

Kydland, F. and Prescott, E. (1982). Time to build and aggregate fluctuations. *Econometrica*, 50(6):1345–1370.

Lim, K. (2016). Firm-to-firm trade in sticky production networks. Mimeo.

Long, J. and Plosser, C. (1983). Real business cycles. *Journal of Political Economy*, 91:39–69.

Lopez-Garcia, P., di Mauro, F., Benatti, N., Angeloni, C., Altomonte, C., Bugamelli, M., D'Aurizio, L., Navaretti, G., Forlani, E., and Rossetti, S. (2014). Micro-based evidence of eu competitiveness: the compnet database. *Working Paper Series European Central Bank*, 1634.

Lucas, R. (1977). Understanding business cycles. *Carnegie-Rochester Conference Series on Public Policy*, 5:7–29.

Melitz, M. and Redding, S. (2014). Heterogeneous firms and trade. In *Handbook of International Economics, Vol. 4*, pages 1–54. Elsevier.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Mobarak, A. M. (2005). Democracy, volatility, and economic development. *Review of Economics and Statistics*, 87(2):348–361.

Newman, M. (2010). *Networks: An Introduction*. Oxford Press.

Osotimehin, S. (2013). Aggregate productivity and the allocation of resources over the business cycle. Mimeo.

Petrin, A. and Levinsohn, J. (2012). Measuring aggregate productivity growth using plant-level data. *RAND Journal of Economics*, 43(4):705–725.

Prescott, E. (1986). Theory ahead of business cycle measurement. *Carnegie-Rochester Conference Series on Public Policy*, 25:11–44.

Ramey, G. and Ramey, V. (1995). Cross-country evidence on the link between volatility and growth. *American Economic Review*, 85(5):1138–1151.

Saichev, A., Malevergne, Y., and Sornette, D. (2010). *Theory of Zipf's Law and Beyond*. Springer.

Seppala, T. and Ali-Yrkko, J. (2011). Nokia and the evolution of mobile telephony. *Industry Studies Conference Paper*.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423.

Shea, J. (2002). Complementarities and comovements. *Journal of Money, Credit and Banking*, 42:412–433.

Stockman, A. C. (1988). Sectoral and national aggregate disturbances to industrial output in seven european countries. *Journal of Monetary Economics*, 21:387–409.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2):326–365.

Vuong, Q. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2):307–333.

Wooldridge, J. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3):112–114.

Table I: Sector Classifications.

NACE Code	Description	Type of Input
01-03	Agriculture, Forestry and Fishing	Primary
05-09	Mining and Quarrying	Primary
10-33	Manufacturing	Manufacturing
35	Electricity, Gas, Steam and Air Conditioning Supply	Utilities
36-39	Water Supply; Sewerage, Waste Management and Remediation Activities	Utilities
41-43	Construction	Utilities
45-47	Wholesale and Retail Trade; Repair of Motor Vehicles and Motorcycles	Services
49-53	Transportation and Storage	Services
55-56	Accommodation and Food Service Activities	Services
58-63	Information and Communication	Services
64-66	Financial and Insurance Activities	–
68	Real Estate Activities	Services
69-75	Professional, Scientific and Technical Activities	Services
77-82	Administrative and Support Service Activities	Services
83	Public Administration and Defense; Compulsory Social Security	–
85	Education	–
86-88	Human Health and Social Work Activities	–
90-93	Arts, Entertainment and Recreation	–
94-96	Other Service Activities	Services
97-98	Activities of Households as Employers	–
99	Activities of Extraterritorial Organizations and Bodies	–

Table II: Industry Coverage (2012).

Coverage (NACE)	Gross Output (%)	GDP (%)
Primary (01-09)	1	1
Manufacturing (10-33)	28	14
Utilities (35-43)	11	9
Services (45-63, 68-82 and 94-96)	41	47
Total Economy	81	71

Notes: Industry contribution to gross output and GDP for the Belgian economy (2012). Industries are aggregated over 2-digit NACE classes. Source: National Accounts (National Bank of Belgium).

Table IV: Aggregate Volatility (2002-2012).

$\hat{\sigma}_{\Delta y}$ (%)	Model	Network	Final Demand	Diversification
$d\hat{\varepsilon}_{it} = g_{it}$	1.11	0.69	0.32	0.09
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	1.11	0.69	0.32	0.09
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	0.96	0.66	0.24	0.09
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	0.89	0.60	0.22	0.09
$\sigma_{\Delta y}$			1.99	

Notes: Volatility predictions $\hat{\sigma}_{\Delta y}$ estimated over the period 2002-2012. $\bar{g}_t$, $\bar{g}_{I_2 t}$ and $\bar{g}_{I_4 t}$ represent the mean weighted growth rates of the economy, 2-digit and 4-digit industries respectively. $\sigma_{\Delta y}$ is the observed standard deviation of GDP growth over the same period obtained from the National Accounts.

Table III: Summary Statistics (2012).

				Percentiles						
Variable	n	Mean	St. Dev.	10th	25th	50th	75th	90th	Top 100	Top 10
Relationship Value r_{ij} (euro)	3,505,207	46,403	1.2 million	382	700	2,000	7,983	34,437	81 million	327 million
Number of Customers	67,466	52	236	1	3	11	39	109	2,260	8,020
Number of Suppliers	79,689	44	66	8	16	28	48	89	720	1,574
Turnover r_i (million euros)	79,788	8.1	150	0.2	0.3	0.8	2.3	8.6	671	4,433
Inputs (million euros)	79,788	6.6	141	0.1	0.2	0.5	1.6	6.5	573	3,848
Employment (FTE)	79,788	20	244	1	2	4	9	27	1,224	8,815
Input Shares $\hat{\omega}_{ij}$ (%)	3,505,207	2.3	7.5	0.01	0.05	0.3	1.2	4.6	100	100
Final Demand $\hat{b}_i$ (million euros)	79,788	6.1	133	0.1	0.2	0.5	1.5	5.6	529	3,730

Table V: Share Of Variance Explained, Top k Influential Firms.

Share of Variance (%)	Top 1,000	Top 100
$d\hat{\varepsilon}_{it} = g_{it}$	98.9	91.3
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	98.9	91.2
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	98.5	88.4
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	98.3	86.8

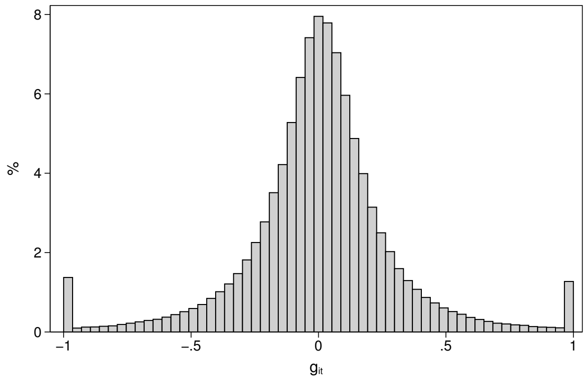
Table VI: Share Of Variance Explained, By Industries (2002-2012).

Share of Variance (%)	Primary	Manufacturing	Utilities	Services
$d\hat{\varepsilon}_{it} = g_{it}$	0.0	23.8	35.6	40.5
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	0.0	23.9	35.9	40.3
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	0.0	11.9	34.1	54.0
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	0.0	12.3	36.3	51.4

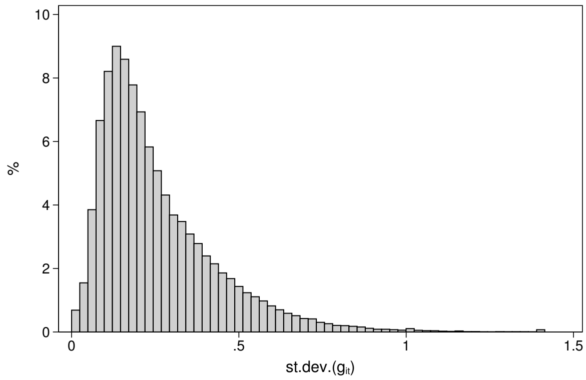
Table VII: Power Law Fit Influence Vector.

Power Law Estimation			
$\hat{\beta}_{MLE}$	1.12	$\hat{\beta}_{OLS}$	1.20
	(.003)		(.002)
x_{min}	7.99×10^{-5}		7.99×10^{-5}
N_{tail}	$1,861$		$1,861$
R value	-1.70		
p-value	0.09		

Notes: Approximate standard errors are calculated as $\frac{\hat{\beta}-1}{\sqrt{N}}$ for MLE. x_{min} is the endogenous cutoff for the estimated fit, N_{tail} denotes the number of observations in the tail fit. R value denotes the likelihood ratio test statistic and p-value denotes the significance of the R value being statistically different from zero.



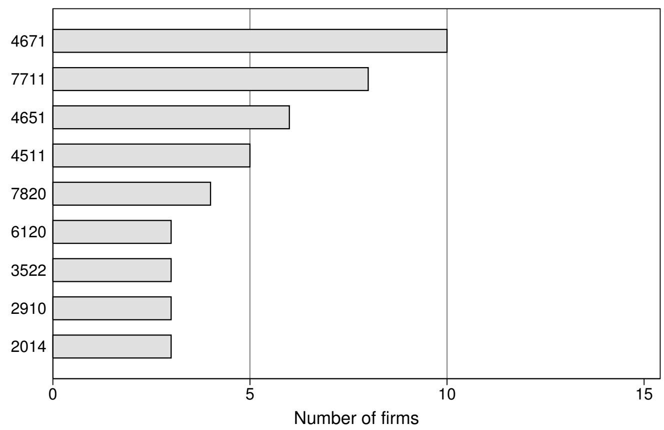
(a) TFP Growth Rates.



(b) Standard Deviation of TFP Growth Rates.

Notes: Growth rates expressed as log-differences ($\Delta \ln TFP_{it}$), winsorized at $+/-$ 1.

Figure I: Productivity Growth Rates And Volatility (2002-2012).



NACE Codes: Manufacture of Other Organic Basic Chemicals (2014); Manufacture of Motor Vehicles (2910); Distribution of Gaseous Fuels Through Mains (3522); Sale of Cars and Light Motor Vehicles (4511); Wholesale of Computers, Computer Peripheral Equipment and Software (4651); Solid, Liquid and Gaseous Fuels (4671); Wireless Telecommunication Activities (6120); Renting and Leasing of Cars and Light Motor Vehicles (7711) and Temporary employment agency activities (7820).

Figure II: Sector Membership Top 100 Firms.

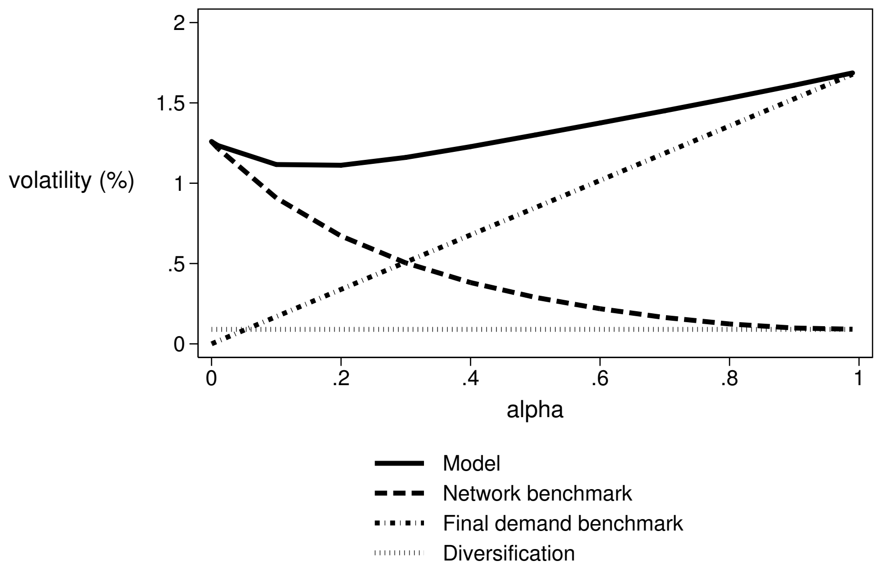
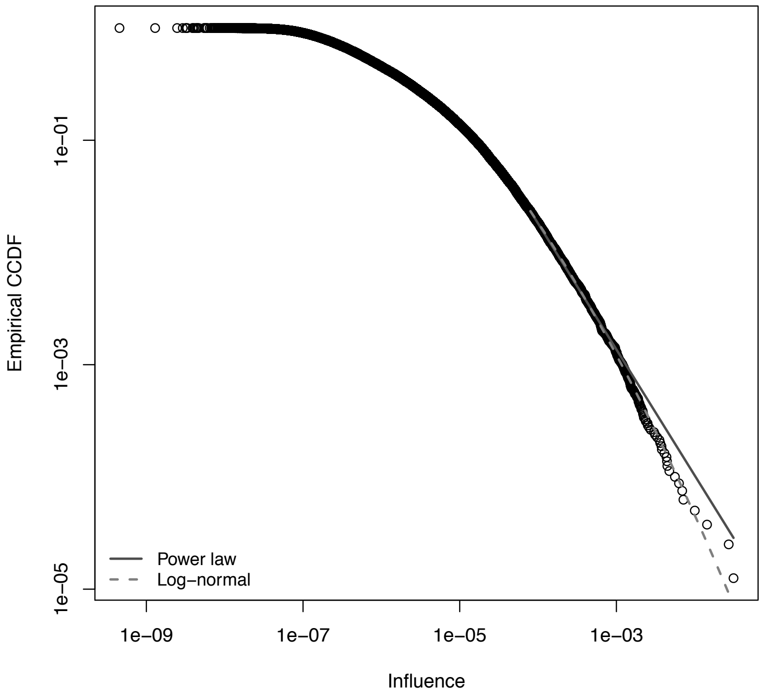


Figure III: Counterfactual Changes In Labor Share α.



Notes: Empirical counter-cumulative distribution function (CCDF), plotted against the values of the influence vector, both in log scales.

Figure IV: Influence Vector Distribution.

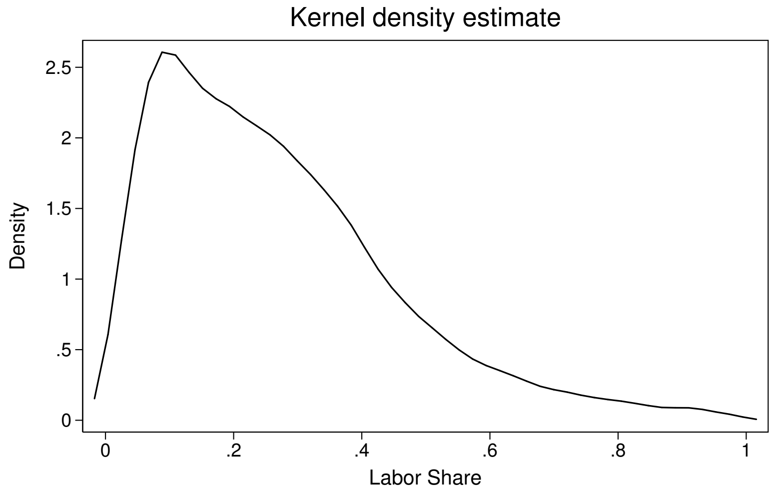


Figure V: Labor Share Distribution Across Firms (2012).

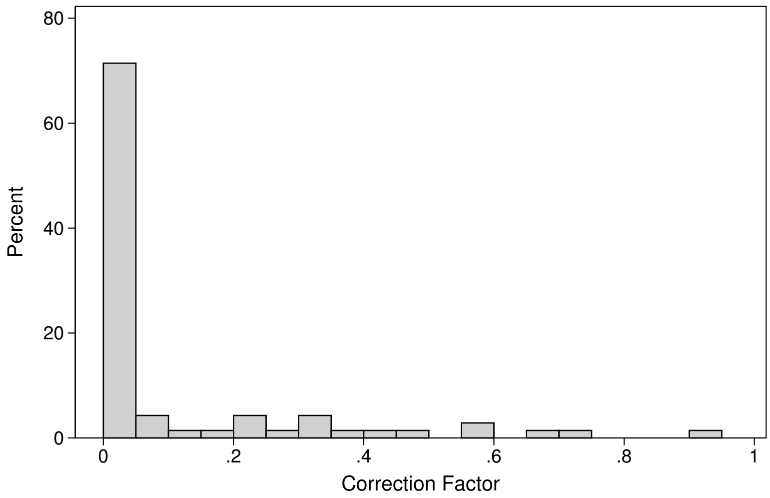


Figure VI: Corrected Intermediate Inputs Ratio (2012).

Table A.I: Aggregate Volatility, Firm-level Labor Shares.

$\hat{\sigma}_{\Delta y}$ (%)	Heterogeneous Labor
$d\hat{\varepsilon}_{it} = g_{it}$	1.12
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	1.12
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	1.11
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	1.10

Table A.II: Aggregate Volatility, Corrected For Capital Goods.

$\hat{\sigma}_{\Delta y}$ (%)	Capital Goods Correction
$d\hat{\varepsilon}_{it} = g_{it}$	1.29
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	1.29
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	1.02
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	0.96

Table A.III: Final Demand Corrections (2012).

$\hat{\sigma}_{\Delta y}$ (%)	Domestic Sales	Export Split
$d\hat{\varepsilon}_{it} = g_{it}$	1.10	1.13
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	1.10	1.13
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	0.94	0.96
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	0.87	0.88

Table A.IV: Accounting For Imports.

$\hat{\sigma}_{\Delta y}$ (%)	Imports
$d\hat{\varepsilon}_{it} = g_{it}$	1.20
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	1.20
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	0.95
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	0.89

Table A.V: Unweighted Demeaning Procedure.

$\hat{\sigma}_{\Delta y}$ (%)	Unweighted
$d\hat{\varepsilon}_{it} = g_{it}$	1.11
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	1.11
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	1.16
$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	1.20

Table A.VI: Alternative Firm Growth Measures.

	$\hat{\sigma}_{\Delta y}$ (%)	Model	Network	Final Demand	Diversification
$g_{it} \cong \Delta\ln(\text{value added}_{it}/\text{FTE}_{it})$	$d\hat{\varepsilon}_{it} = g_{it}$	1.43	0.86	0.47	0.17
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	1.40	0.84	0.46	0.17
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	1.18	0.78	0.35	0.17
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	1.10	0.74	0.32	0.17
$g_{it} \cong \Delta\ln(\text{value added}_{it})$	$d\hat{\varepsilon}_{it} = g_{it}$	1.46	0.86	0.48	0.16
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	1.44	0.85	0.47	0.16
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	1.19	0.78	0.36	0.16
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	1.11	0.74	0.32	0.16
$g_{it} \cong \Delta\ln(\text{turnover}_{it}/\text{FTE}_{it})$	$d\hat{\varepsilon}_{it} = g_{it}$	1.08	0.69	0.33	0.09
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	0.97	0.62	0.30	0.09
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	0.71	0.52	0.17	0.09
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	0.58	0.36	0.17	0.09
$g_{it} \cong \Delta\ln(\text{turnover}_{it})$	$d\hat{\varepsilon}_{it} = g_{it}$	1.05	0.67	0.34	0.07
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	0.85	0.53	0.28	0.07
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	0.64	0.46	0.17	0.07
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	0.53	0.34	0.17	0.07
$g_{it} \cong \Delta\ln(\text{FTE}_{it})$	$d\hat{\varepsilon}_{it} = g_{it}$	0.42	0.32	0.10	0.08
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_t$	0.43	0.32	0.10	0.08
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_2 t}$	0.42	0.31	0.09	0.08
	$d\hat{\varepsilon}_{it} = g_{it} - \bar{g}_{I_4 t}$	0.42	0.32	0.09	0.08