Cascades and Fluctuations in an Economy with an Endogenous Production Network[*]

Mathieu Taschereau-Dumouchel[†]

Cornell University

February 2020

Abstract

This paper studies an economy in which the network structure of production is endogenously determined by the firms' extensive margin of operation. In the model, a finite number of firms are connected through input-output linkages and must pay a fixed cost to produce. When economic conditions are poor, some firms might decide not to produce, thereby severing the links with their neighbors. Together, the operating decisions of the firms therefore determine the structure of the production network. Since producers benefit from having access to additional suppliers, the economy features complementarities between the operating decisions of nearby firms. As a result, if a firm exits following a severe shock, a cascade of shutdowns might spread from neighbor to neighbor as the network reorganizes itself. While well-connected firms are better able to withstand shocks, they trigger larger cascades upon shutdown, a prediction confirmed by U.S. data. The theory also predicts how the structure of the production network interacts with the business cycle. As in the data, downturns feature less clustering among firms, and are associated with degree distributions with thinner tails. In addition, the endogenous reorganization of the network substantially dampens aggregate fluctuations relative to a scenario in which the network is kept fixed.

JEL Classifications: E23, D85, C67

[*]First draft: February 2017. I thank Enghin Atalay, Julieta Caunedo, Jeremy Greenwood, Christopher Huckfeldt, Ben Lester, Kristoffer Nimark, Ezra Oberfield, Nicolas Petrosky-Nadeau, Carlos Ramírez, Stephen Redding, Gill Segal, Ali Shourideh, Gianluca Violante, and seminar participants at Wharton, the Philadelphia Fed, Drexel, FGV São Paulo, the Cornell/Penn State workshop, UT Austin, SED Edinburgh, SAET Faro, the ITAM-PIER Conference, the Tepper-LAEF Conference, the CIRANO – Sam M. Walton College of Business Workshop on Networks, LSE, Cornell, the annual meeting of the CMSG, Rutgers, the New York Fed, the Midwest Macro Meetings, Minnesota Macro, USC Marshall, Queen's, the NYU Search Conference, Princeton, the 2019 AEA, the Atlanta Fed, Binghamton and HKUST for helpful comments. I gratefully acknowledge support from Cornell's Institute for the Social Sciences. Replication files are available at the author's website `https://www.mathtd.info`.

[†]Email: mt763@cornell.edu; Address: 480 Uris Hall, Cornell University, Ithaca NY 14853

1 Introduction

Production in modern economies involves a complex network of specialized firms, each using inputs from suppliers and providing their own output to downstream producers. In such an environment, the structure of the production network determines how idiosyncratic shocks propagate from firm to firm and, as a result, how they aggregate into macroeconomic fluctuations (Acemoglu et al., 2012). But the production network is also constantly changing as ailing firms go out of business—severing links with their previous neighbors—and as new firms are born—creating links with new customers and suppliers.[1] To better understand the origin of aggregate fluctuations, this paper proposes a simple theory of production in which the structure of the network is endogenously determined by the firms' extensive margin of operation.

In the model, a finite number of firms produce differentiated goods using labor and a set of inputs from other producers. Production requires the payment of a fixed cost, so that firms operate or not as a function of economic conditions. When a firm does operate, it makes an additional input available to all of its customers, and it purchases intermediate goods from its suppliers, thereby creating new input-output relationships. Together, the operating decisions of the firms therefore determine the structure of the production network. I show that every stable equilibrium is efficient in this environment and therefore focus on the problem of a social planner.

In the model, firms combine intermediate inputs from their suppliers using a standard CES production technology. As a result, having access to an additional input lowers the marginal cost of production and makes the firm effectively more productive. Because of these gains from input variety, firms with multiple suppliers are more likely to operate in the efficient allocation—their high productivity more than compensate for the fixed cost of operation. Similarly, firms with many customers provide a valuable input to multiple producers and are also more likely to operate. These forces also create complementarities between the operating decisions of nearby firms: neighbors tend to operate, or not, together.

These complementarities have important implications for the structure of the production network. First, they lead to the creation of clusters of firms that are tightly connected with one another. By organizing production in this way, firms increase their number of customers and suppliers so as to take full advantage of the gains from input variety. Second, cascades of firm failures can arise in the efficient allocation. If a firm faces a severe shock and stops production, its customers, having lost a valuable input, and its suppliers, now producing a less useful product, are also more likely to shut down. The same logic applies to the firm's second neighbors which are more likely to shut down as well, and so on. As a result, the initial shock can trigger a cascade of firm shutdowns that propagates upstream and downstream through the production network.

[1] In U.S. data, the extensive margin of production is a key source of changes in the production network. According to Factset Revere, a large dataset that covers firm-level input-output linkages in the U.S., about 40% of all link destructions occur when either the supplier or the customer (or both) stops producing. See Section 5.1 for details.

One other, and perhaps unusual, consequence of the operating complementarities between neighbors is that they can lead to large reorganizations of the network after a small change in the environment. For instance, a small decline in the productivity of a centrally located firm can lead to the shutdown of its whole neighborhood, as economic activity moves to a more promising part of the network. Through that mechanism, large changes in the distribution of firm-level outcomes can occur in response to arbitrarily small shocks.

The model is also able to shed light on the relationship between the structure of the production network and the business cycle. Since the number of firms in the economy is finite, fluctuations in aggregate output emerge from the idiosyncratic shocks faced by the firms. As these shocks also affect the structure of the production network, the model features comovements between the macroeconomy and the structure of the network.

Two features of the environment make the problem of the social planner particularly challenging to solve. First, since the decision to operate a firm is binary, the planner's optimization problem has a non-convex feasible set. Second, the increasing returns to scale generated by the fixed costs break the concavity of the objective function. As a result, the objective function generally features multiple local maxima and standard algorithms are unable to identify the global maximum.[2] To overcome these difficulties, I propose a novel solution approach that involves reshaping the original optimization problem such that 1) this reshaped problem can be solved easily, and 2) its solution coincides with that of the original problem. I establish sufficient conditions under which this approach is guaranteed to find the efficient allocation. But even when those conditions are not met, numerical simulations show that it provides a rapid and robust way of tackling a class of challenging network formation problems (Carvalho and Tahbaz-Salehi, 2018).

I provide a basic calibration of the model to the United States economy. To better understand how economic forces shape the production network, I compare the efficient network, designed optimally by the planner, to a neutral benchmark whose structure is randomly determined. I find that the efficient network features in-degree (number of suppliers) and out-degree (number of customers) distributions with thicker right tails, as well as a higher amount of clustering between firms. These differences show that the planner takes advantage of the operating complementarities by creating tightly connected clusters of economic activity centered around well-connected firms.

I use the calibrated economy to investigate how cascades of firm shutdowns arise and propagate through the network. I find that, as in the data, highly connected firms are more resilient to shocks but that, upon shutting down, they create larger cascades that lead to the exit of several of their neighbors. I also investigate how cascades affect macroeconomic aggregates. While the average cascade has a negligible effect, a cascade that originates from a highly-connected firm can have a substantial negative impact on output. Finally, I describe through sensitivity analysis how the parameters of the model influence how cascades propagate and their aggregate impact. In

[2] These nonconvex mixed integer nonlinear problems are notoriously difficult to solve (Garey and Johnson, 1990).

particular, cascades become more damaging when intermediate inputs are less substitutable.

One contribution of this paper is to highlight novel business cycle correlations between aggregate output and the structure of the production network. I find that, in the data and in the model, recessions are periods in which, 1) the tails of the degree distributions are thinner, and 2) there is less clustering between firms. These correlations are naturally explained through the lens of the model. Expansions are periods in which it is easy to leverage the complementarities at work in the economy by creating productive clusters of firms around highly-connected producers. In contrast, recessions are periods in which creating these clusters would be too costly, perhaps because a few influential firms are facing bad shocks, and in which production therefore involves a more diffused, and less productive, network.

Finally, I consider how the endogenous formation of the network interacts with firm-level shocks to influence aggregate fluctuations. To do so, I compare the benchmark economy, in which the production network reorganizes itself in response to shocks, to an alternative economy in which the structure of the network is kept fixed. I find that aggregate output is 11% lower and 20% more volatile when the network is fixed. This last finding highlights the importance of considering how the production network adapts to shocks to better understand the microeconomic origin of aggregate fluctuations.

Relation to the Literature

The theory is motivated by an empirical literature documenting that losing a supplier is disruptive to a firm's operations.[3] Carvalho et al. (2014) document that firms that stopped production because of the Great East Japan Earthquake of 2011 had a significant negative impact on their customers and suppliers that were outside of the affected zone. Hendricks and Singhal (2005) find that firms facing supply chains disturbances face large and long-lasting negative abnormal stock returns. Wagner and Bode (2008) survey business executives in Germany who report that issues with supply chains, including the loss of a supplier, were responsible for significant disturbances in production. The Zurich Insurance Group (2015) also conducted a global survey of executives in small and medium enterprises. Of all the respondent, 39% report that losing their main supplier would adversely affect their operation and 14% report that they would need to significantly downsize their business, require emergency support or that they would shut down.

This paper also relates to a literature that studies how shocks to interconnected sectors contribute to aggregate fluctuations in exogenous networks (Long and Plosser, 1983; Horvath, 1998; Dupor, 1999). In an influential paper, Acemoglu et al. (2012) find that sectoral shocks can lead to large aggregate fluctuations if there is enough asymmetry in the way sectors supply to each

[3]One piece of anecdotal evidence comes from the 2008 bailout of U.S. car manufacturers. At that time, the CEO of Ford advocated for the bailout of its competitors, General Motors and Chrysler, so that their shared suppliers could remain in business.

other.[4] Acemoglu et al. (2015a) further show that inter-sectoral linkages can generate tail-risks in aggregate output. This literature emphasizes the importance of the (fixed) structure of the network in transmitting idiosyncratic shocks. In contrast, the current paper studies how endogenizing the network leads to smaller aggregate fluctuations.[5]

This paper also contributes to a recent literature in which production networks are built endogenously by the decisions of economic agents.[6] Perhaps first in that literature, Oberfield (2018) builds a model in which producers optimally choose one input from a randomly evolving set of suppliers, thereby creating a production network. He finds that star suppliers can emerge endogenously in equilibrium. Lim (2018) studies sourcing decisions in a model with sticky relationships. These papers feature a continuum of firms so that aggregate fluctuations do not arise from individual idiosyncratic shocks, a margin whose importance has been emphasized by the granularity literature (Gabaix, 2011) but that has proven challenging to incorporate in network formation models (Carvalho and Tahbaz-Salehi, 2018).

In independent contemporaneous work, Acemoglu and Azar (2018) consider a network of competitive industries in which firms select a production technique that involves different sets of suppliers. They show that the endogenous evolution of the network can generate long-run growth. Tintelnot et al. (2018) build a model of endogenous network formation and international trade. In contrast to the current paper, they only consider acyclic networks. Boehm and Oberfield (2018) estimate a model of network formation using Indian micro data to study misallocation in the inputs market.

One distinguishing feature of the present work is that it proposes a macroeconomic model in which the input-output network is built endogenously through the extensive production decisions of the firms—a margin that accounts for a large fraction of link destructions in the U.S. data and that allows for cascades of firm shutdowns to arise.

Baqaee (2018) also studies the impact of cascades of firm shutdowns on the macroeconomy. To do so, he considers an exogenous sectoral network in which the mass of firms in each sector can vary. This adjustment margin can lead to further amplification of sectoral shocks in the presence of external economies of scale. In contrast, the current paper studies cascades and fluctuations when the production network evolves endogenously.

One methodological contribution of this paper, that may be of independent interest, is a new solution technique for some nonconvex optimization problems with binary variables. Several heuris-

[4]In a related paper, Gabaix (2011) shows that when the tail of the firm size distribution is sufficiently thick, firm-level shocks can have large effect on aggregates.

[5]See Carvalho (2014) and Carvalho and Tahbaz-Salehi (2018) for an overview of the literature on production networks. Recent contributions to the literature on macroeconomics and networks include di Giovanni et al. (2014), Barrot and Sauvagnat (2016), Atalay (2017), Baqaee and Farhi (2017a), Baqaee and Farhi (2017b), Bigio and La'O (2016), Caliendo et al. (2017b), Caliendo et al. (2017a), Grassi (2017) and Ozdagli and Weber (2017).

[6]A related literature studies networks that are mostly formed through an exogenous random process. See for instance Atalay et al. (2011) and Carvalho and Voigtländer (2015). König et al. (2018) proposes an hybrid model in which a firm's extensive production decision depends on its number of customers and an exogenous shock.

tics have been developed to handle these problems (Li and Sun, 2006). Closest to the present work are smoothing algorithms that attempt to get rid of the local maxima that emerge in the relaxed problem (Murray and Ng, 2010). In practice, finding an appropriate smoother is usually done through trial and error and there is no guarantee that the algorithm converges to a global maximum. In contrast, the current work explicitly describes how to reshape the objective function of the planner and proposes a rapid and robust solution method.

The next section introduces the model. Section 3 describes the solution approach. Section 4 discusses the forces at work in the economy. Section 5 illustrates how the methodology can be used to explain some patterns in U.S. data. The last section concludes. The proofs are in Appendix C and the algorithms are in Appendix D.

2 Model

The model is static, and there are three types of agents: firms, a final good producer and a representative household. There is a set $\mathcal{N} = \{1, \ldots, n\}$ of firms, each of which produces a differentiated good that can be used as intermediate input by the final good producer and the other firms. The final good producer uses the CES production technology

$$C = \left(\sum_{j \in \mathcal{N}} \beta_j^{\frac{1}{\sigma}} c_j^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

with elasticity of substitution $\sigma > 1$ and factor intensities $\{\beta_j\}_{j \in \mathcal{N}}$, to convert intermediate inputs $\{c_j\}_{j \in \mathcal{N}}$ into aggregate output C. The representative household consumes C and supplies $L > 0$ units of labor inelastically.

To produce, a firm $j \in \mathcal{N}$ must employ $f_j L \geq 0$ units of labor as a fixed cost, in which case we say that j is *operating*. This fixed cost captures overhead labor, such as managers and other non-production workers, that is necessary for production.[7] The vector $\theta \in \{0, 1\}^n$ keeps track of the operating decisions of the firms, such that $\theta_j = 1$ if j operates and $\theta_j = 0$ otherwise.

When it operates, firm j has access to a technology that converts l_j units of labor and a vector of intermediate inputs $\{x_{ij}\}_{i \in \mathcal{N}}$ into y_j units of good j according to the production function

$$y_j = \frac{A}{\alpha_j^{\alpha_j} (1 - \alpha_j)^{1-\alpha_j}} z_j \theta_j \left(\sum_{i \in \mathcal{N}} \Omega_{ij}^{\frac{1}{\varepsilon_j}} x_{ij}^{\frac{\varepsilon_j - 1}{\varepsilon_j}} \right)^{\frac{\varepsilon_j}{\varepsilon_j - 1} \alpha_j} l_j^{1-\alpha_j}, \tag{2}$$

where $\Omega_{ij} \geq 0$ denotes the factor intensity of input i, $\varepsilon_j > 1$ is the elasticity of substitution between

[7] Empirical studies have found these fixed costs to be important to explain firm-level production. For instance, Bresnahan and Ramey (1994) show that the extensive margin is responsible for about 80% of plant-level output fluctuations in the automobile industry.

inputs, $0 < 1 - \alpha_j < 1$ is the labor intensity, and $A > 0$ and $z_j > 0$ are aggregate and firm-specific total factor productivities. Without loss of generality, assume that each firm has access to at least one input, so that $\sum_{i \in \mathcal{N}} \Omega_{ij} > 0$ for all $j \in \mathcal{N}$, otherwise firm j cannot produce and we can redefine the economy without it.[8]

We see from (2) that a firm j can only use inputs from a supplier i if $\Omega_{ij} > 0$. As such, the matrix Ω describes a network of *potential connections* between firms. A potential connection (i, j) is *active*—with goods being traded—if firms i and j both operate, otherwise it is *inactive*.[9] The production network is therefore jointly determined by Ω and θ, and economic conditions, through their impact on the firms' operating decisions, endogenously determine the shape of the network.

Panel (a) in Figure 1 provides an example of the potential connections $\Omega_{ij} > 0$ in an economy with six firms. Each arrow represents a connection $\Omega_{ij} > 0$, with the direction of the arrow showing the potential movement of goods between the firms. The set of active connections, in blue in panel (b), is determined endogenously by the set of operating firms, also shown in blue.

While the paper focuses on the role of the firms' extensive margin of operation for the formation of the network, the model is general enough to accommodate the formation of individual links. Specifically, we can interpret a link between any two firms i and k, as a *pseudo* firm j that 1) only has potential connections in Ω with i as a supplier and k as a customer, and 2) produces a good that is not included in the production of the final good ($\beta_j = 0$). In this context, j can be interpreted as a shipping technology whose cost f_j is the fixed cost of operating the link, its productivity z_j affects the variable cost of shipping goods through that link and α_j controls how labor intensive the technology is. The variable $\theta_j \in \{0, 1\}$ then indicates whether the link between i and k is active or not.

Equilibrium

We now turn to the definition of an equilibrium in this environment. As is common in network economies, I focus on stable equilibria. Namely, firms are facing contractual obligations to purchase and deliver goods, and there are no group of firms that wants to change the terms of the contracts.[10]

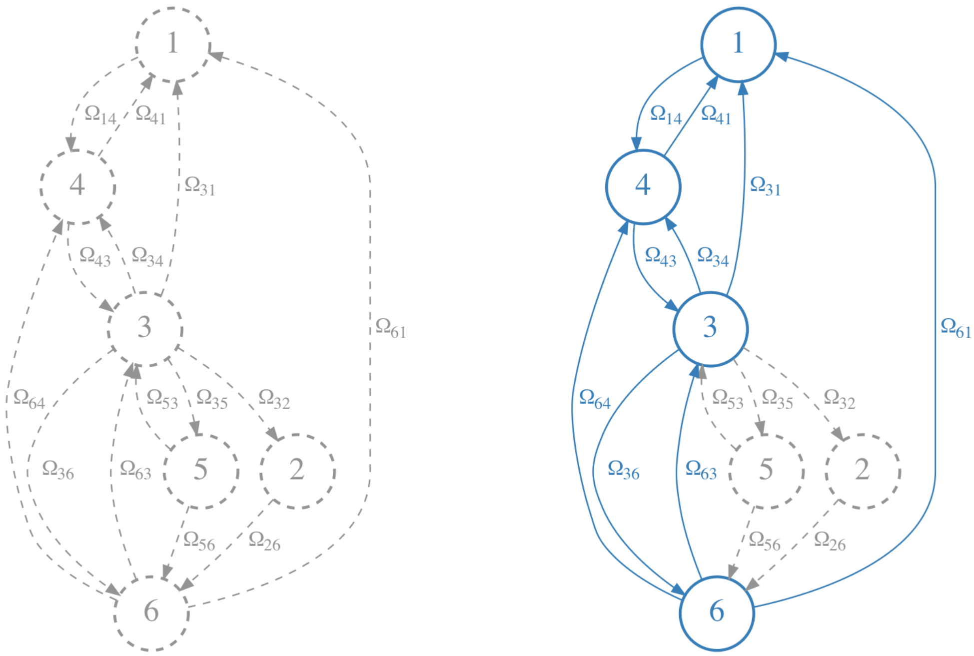
We first describe the contractual environment. Define a *contract* between two firms i and j as a pair $\{x_{ij}, T_{ij}\}$ where x_{ij} is a quantity shipped from i to j, and T_{ij} is a payment from j to i. An *arrangement* is a collection of contracts between all possible pairs of firms $\{x_{ij}, T_{ij}\}_{i,j \in \mathcal{N}^2}$.

Under a given arrangement, a firm j must supply and purchase the prescribed quantities, but it can decide on a price p_j to charge the household, an amount c_j to sell to the final good producer,

[8] The restrictions $\sigma > 1$ and $\varepsilon > 1$ are necessary to avoid a complete shutdown of the economy, or of a customer, if a single supplier does not operate.

[9] The production function (2) implies that $\partial y_j / \partial x_{ij} \to \infty$ as $x_{ij} \to 0$ if $\Omega_{ij} > 0$ and $\theta_j = 1$. As such, any potential connection between two operating firms will be active in the efficient allocation.

[10] This equilibrium concept has proven particularly convenient in network economies (Jackson and Wolinsky, 1996; Hatfield et al., 2013). The approach followed here is most closely related to Oberfield (2018).



(a) Potential connections

(b) Active firms and connections

Figure 1: The firms' operating decisions determine the production network

how much labor l_j to employ, and its operating status θ_j. It makes these decisions to maximize profits

$$\pi_j = p_j c_j - w l_j + \sum_{i \in \mathcal{N}} T_{ji} - \sum_{i \in \mathcal{N}} T_{ij} - w \theta_j f_j L, \tag{3}$$

where w is the wage, and subject to a technology constraint

$$c_j + \sum_{k \in \mathcal{N}} x_{jk} \leq y_j, \tag{4}$$

where y_j satisfies (2), and to the usual demand curve

$$c_j = \beta_j C \left(p_j / P\right)^{-\sigma} \tag{5}$$

where $P = \left(\sum_j \beta_j P_j^{1-\sigma}\right)^{1/(1-\sigma)}$ is the price index.

We say that an allocation is *feasible* if all the technology constraints (4) and the labor resource constraint $\sum_j l_j + \sum_j \theta_j f_j L \leq L$ are satisfied.

A *coalition* is a set of firms J. A *deviation* for a given coalition J consists of (i) dropping any contracts that involve at least one firm in J and (ii) altering the terms of any contract involving a buyer and a supplier that are both members of the coalition. Finally, a *dominating deviation* for a given coalition is a deviation that delivers at least the same amount of profits to all members of the coalition and strictly greater profits to at least one member.

We can now define a stable equilibrium in this environment.

Definition 1. A stable equilibrium is an arrangement $\{x_{ij}, T_{ij}\}_{i,j\in\mathcal{N}^2}$, firms' choices $\{p_j, c_j, l_j, \theta_j\}_{j\in\mathcal{N}}$ and a wage w such that (i) given the wage, total profits, and prices, the consumption choices $\{c_j\}_{j\in\mathcal{N}}$ maximize the utility of the representative household; (ii) for each $j \in \mathcal{N}$, $\{p_j, c_j, l_j, \theta_j\}$ maximizes the profits of j given the arrangement, the wage, the household's demand and the technology constraint; (iii) labor and final goods markets clear; (iv) there are no dominating deviations available to any coalition; and (v) the equilibrium allocation is feasible.

Planner's problem

We can also define the problem $\mathcal{P}$ of a social planner that maximizes the utility of the representative household:

$$\mathcal{P} : \max_{\substack{c\geq 0,\, x\geq 0,\, l\geq 0\\ \theta\in\{0,1\}^n}} \left(\sum_{j\in\mathcal{N}} \beta_j^{\frac{1}{\sigma}} c_j^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{6}$$

subject to a resource constraint for each intermediate good $j \in \mathcal{N}$,

$$c_j + \sum_{k\in\mathcal{N}} x_{jk} \leq y_j, \tag{7}$$

where y_j is given by (2), and a resource constraint for labor,

$$\sum_{j\in\mathcal{N}} l_j + \sum_{j\in\mathcal{N}} \theta_j f_j L \leq L. \tag{8}$$

We say that an allocation is *efficient* if it solves $\mathcal{P}$.

The following proposition shows how equilibria and the efficient allocation are related.

Proposition 1. *Every stable equilibrium is efficient.*

Proof. All proofs are in Appendix C. □

This proposition shows that every equilibrium allocation is a solution to the planner's problem $\mathcal{P}$.[11] We therefore focus from now on on solving $\mathcal{P}$ directly, with the understanding that we are also implicitly characterizing equilibrium outcomes in this economy.[12]

[11] Note that since the household is not part of any coalition, the grand coalition does not seek to maximize social welfare.

[12] We can also define a monopolistic competition equilibrium as in Dixit and Stiglitz (1977). In this case, multiple equilibria generally exist and an equilibrium is not efficient without taxes and subsidies. Some pathological allocations, such as a no-production economy ($\theta_j = 0$ for all j), are also equilibria under that definition, which makes it unwieldy as an equilibrium concept in the current environment.

3 Solving $\mathcal{P}$

To solve the planner's problem $\mathcal{P}$, it is useful to first find the best allocation when the production network, or equivalently θ, is fixed. We can then take a step back and find the efficient network θ.

3.1 Planner's problem with exogenous firm status θ

Before solving this problem, it is useful to characterize the set of firms that will produce for a given matrix of potential connection Ω and a vector of operating firm θ. For that purpose, define an *operating cycle* as a sequence of operating firms $\{s_1, \ldots, s_k\}$, for some $k \geq 1$, such that 1) $\Omega_{s_i,s_{i+1}} > 0$ for $i \in \{1, \ldots, k-1\}$, and 2) $\Omega_{s_k,s_1} > 0$. We say that a firm j *has access to an operating cycle* if there exists a sequence of operating firms $\{s_1, \ldots, s_j\}$ such that 1) s_1 is part of an operating cycle, and 2) $\Omega_{s_i,s_{i+1}} > 0$ for $i \in \{1, \ldots, j-1\}$.

These definitions are useful to determine which firms produce in the efficient allocation and which ones do not. For a firm to produce, it must receive some intermediate input from a supplier, and this supplier must also receive some inputs from a supplier, and so on. Since n is finite, this sequence of suppliers must contain a cycle for production to take place. As a result, a firm without access to an operating cycle cannot produce. The converse is also true. Since, from (1), $\lim_{c_j \to 0} \partial C / \partial c_j = \infty$ any operating firm that has access to an operating cycle produces in the efficient allocation.

We can now turn to the solution of the planner's problem. For a fixed θ, $\mathcal{P}$ is a convex maximization problem and the usual Karush-Kuhn-Tucker conditions are necessary and sufficient to characterize its solution. Denote by λ_j the Lagrange multiplier on the resource constraint (7) for good j and by w the multiplier on the labor resource constraint (8). The first-order conditions imply that $(1 - \alpha_j)\, y_j \lambda_j = w l_j$ so that, as in Oberfield (2018), we can define $q_j = w / \lambda_j$ as a measure of firm j's productivity.

The planner's first-order conditions, together with the production function, yield the following result.

Proposition 2. *In the efficient allocation, the productivity vector q satisfies, for all $j \in \mathcal{N}$,*

$$q_j = z_j \theta_j A \left(\sum_{i \in \mathcal{N}} \Omega_{ij} q_i^{\varepsilon_j - 1} \right)^{\frac{\alpha_j}{\varepsilon_j - 1}}. \tag{9}$$

Furthermore, there is a unique q that solves (9) and such that $q_j > 0$ if j operates and has access to an operating cycle, and $q_j = 0$ otherwise.

Equation (9) fully characterizes the vector of productivities q_j in this environment, and that vector can be found rapidly by iterating on the recursive mapping (Kennan, 2001).

Several features of (9) are worth emphasizing. First, its recursive structure implies that a change in the productivity q of a firm propagates downstream through production chains. For instance, if a firm j faces a negative TFP shock, the amount of labor needed to produce one unit of good j increases, which leads to a higher unit labor cost for j's customers, and for its customers' customers and so on.[13] Second, (9) implies that a firm that has access to a greater set of active suppliers (more terms in the summation), is more productive (higher q_j). Intuitively, with a more diverse set of inputs a firm might be able to use better production techniques that would otherwise be unavailable. The elasticity ε_j governs how substitutable these inputs are with each other and is the key parameter determining the strength of this mechanism. When ε_j is small, intermediate inputs are poor substitutes and the benefit of having an additional supplier is large. In contrast, when ε_j is large, firm j's productivity is almost entirely driven by its most productive supplier. As we will see in Section 4, these mechanisms have important implications for the structure of the network and for the propagation of shocks in this economy.

With q in hand, it is straightforward to derive all other quantities in the efficient allocation (see Appendix A for the equations). In particular, the following lemma shows that aggregate output C can be computed as the product of aggregate productivity

$$Q = \left(\sum_{j \in \mathcal{N}} \beta_j q_j^{\sigma-1} \right)^{\frac{1}{\sigma-1}}. \tag{10}$$

and the amount of labor available after fixed costs have been paid.

Lemma 1. *In the efficient allocation, aggregate output is*

$$C = Q \left(1 - \sum_{j \in \mathcal{N}} \theta_j f_j \right) L. \tag{11}$$

As we can see from (10), aggregate productivity Q is a CES aggregator of the underlying firm-level productivities q_j, with an elasticity of substitution controlled by σ. When σ is small, the differentiated goods are poor substitutes and each additional good is highly valued by the planner. When, instead, σ is large the household derives utility primarily from the most productive firm.[14] As a result, σ affects the planner's incentives to operate more firms and will also play an important role in shaping the structure of the network.

Lemma 1 completes the solution of the planner's problem with fixed θ and we can now take a step back to consider the full problem $\mathcal{P}$ in which the network itself is a choice variable.

[13] From the first-order condition we can write the unit labor cost of production as $l_j/y_j = (1-\alpha_j)\, q_j^{-1}$.

[14] For $\sigma \to 1$, Q is Cobb-Douglas with intensities $\{\beta_j\}$ and aggregate consumption C is zero unless all firms in the economy produce. For $\sigma \to \infty$, Q becomes a max function and only the largest q_j matters for C.

3.2 Planner's problem with θ as a choice variable

By combining Lemma 1 and Proposition 2, we can rewrite $\mathcal{P}$ as the problem of finding a vector θ^* that maximizes consumption (11), and where q solves (9). The left-hand side of Figure 2 lays out that version of $\mathcal{P}$.

The two equations (9) and (11) highlight the trade-offs faced by the planner when deciding whether to operate a firm j. Because of the recursive structure of (9), operating j improves the productivity q not only of j itself, but also of all its downstream customers, which benefits aggregate productivity Q. On the other hand, operating j also takes $f_j L$ units of labor away from other uses.

$\mathcal{P}$ is hard to solve for two reasons.[15] First, θ is limited to the *corners* $\{0,1\}^n$ of the n-dimensional unit hypercube—a non-convex set. But even if θ could move freely over $[0,1]^n$, the fixed costs create firm-level increasing returns to scale that break the (quasi) concavity of the objective function. As a result, there are usually multiple local maxima, and the standard Karush-Kuhn-Tucker conditions are not helpful to find the global maximum.[16]

There is however a brute-force way of solving $\mathcal{P}$. Since there are only a finite number of vectors θ in the feasible set $\{0,1\}^n$, one can try them all. For each θ, q can be found by iterating on (9) and the objective function can then be computed using (11). While this *exhaustive search* approach is guaranteed to find the correct solution, it is limited to economies with only a few firms. Since there are 2^n possible θ in $\{0,1\}^n$, the number of vectors to try explodes as n grows.[17]

Reshaping the planner's problem

To handle economies with n large, I propose a novel solution method that is less computationally intensive. The key idea is to find an alternative optimization problem that is easy to solve and whose solution coincides with that of $\mathcal{P}$. This alternative problem, denoted by $\mathcal{R}$, is obtained by *relaxing* and *reshaping* $\mathcal{P}$.[18] $\mathcal{R}$ is defined on the right-hand side of Figure 2, alongside $\mathcal{P}$ to highlight their differences.

$\mathcal{R}$ differs from $\mathcal{P}$ in two important ways—emphasized in bold and blue in Figure 2. First, the binarity constraint $\theta \in \{0,1\}^n$ is relaxed, and θ can now take values *inside* the unit hypercube, $[0,1]^n$. While this relaxation has the advantage of convexifying $\mathcal{P}$'s feasible set, it also augments the planner's problem with points that have no real economic meaning. For instance, $\theta_j = 0.5$

[15] It belongs to the class of Mixed Integer Nonlinear Problems (MINLP). Their combinatorial nature makes them notoriously challenging to solve and they are, in general, NP-Hard (Garey and Johnson, 1990).

[16] The key difference with the fixed-network problem considered previously is that the production function (2) is no longer concave when θ is a choice variable.

[17] In some models, it is possible to order firms in some way and to progressively shut down the "worst" ones until the desired allocation is found. Here, however, firms differ along several dimensions of heterogeneity so that this approach does not apply directly.

[18] I have also experimented with other techniques, such genetic algorithms and branch-and-bound algorithms, to solve $\mathcal{P}$ but they are overall slow and unreliable.

$\mathcal{P}$: Original planner's problem	$\mathcal{R}$: Relaxed and reshaped problem
$$\max_{\theta \in \{0,1\}^n} Q \left(1 - \sum_{j \in \mathcal{N}} \theta_j f_j\right) L$$ where q solves, for each $j \in \mathcal{N}$, $$q_j = z_j \theta_j A \left(\sum_{i \in \mathcal{N}} \Omega_{ij} q_i^{\varepsilon_j - 1}\right)^{\frac{\alpha_j}{\varepsilon_j - 1}} \quad (9)$$	$$\max_{\theta \in \boldsymbol{[0,1]^n}} Q \left(1 - \sum_{j \in \mathcal{N}} \theta_j f_j\right) L$$ where q solves, for each $j \in \mathcal{N}$, $$q_j = z_j \theta_j^{\boldsymbol{a_j}} A \left(\sum_{i \in \mathcal{N}} \boldsymbol{\theta_i^{b_{ij}}} \Omega_{ij} q_i^{\varepsilon_j - 1}\right)^{\frac{\alpha_j}{\varepsilon_j - 1}} \quad (12)$$

Figure 2: Differences between the original and the reshaped problems

for some j does not correspond to any physical reality in the economic environment. Since they have no interest on their own, we can change the value of the objective function over these new points—and only over these new points—to help us solve $\mathcal{P}$.

This is done in (12), which is a transformed version of (9) that includes the *reshaping constants* $a_j > 0$ and b_{ij}. These constants modify the shape of the optimization problem everywhere except over $\mathcal{P}$'s original feasible set: $\{0,1\}^n$. Indeed, for $\theta \in \{0,1\}^n$, $\theta_j^{a_j} = \theta_j$ for all j. Similarly, for b_{ij}, if $\theta_i = 0$ then $q_i = 0$ anyway, and if $\theta_i = 1$ then $\theta_i^{b_{ij}} = 1$.[19] In both cases, the term in the summation is unchanged. This reshaping procedure therefore preserves the ranking, in terms of utility, of the corners $\{0,1\}^n$—the only points with actual economic meaning—while elsewhere changing the shape of the optimization problem.

Crucially, we are free to pick the constants a_j and b_{ij} to help us solve the planner's problem. In particular, we can pick these constants to increase the concavity of $\mathcal{R}$ with the goal of removing the undesirable local maxima that prevent an easy resolution of the relaxed problem. On the other hand, too much concavity can create a new global maximum somewhere in the middle of $[0,1]^n$, in which case the solutions of $\mathcal{P}$ and $\mathcal{R}$ would clearly differ. Specific values for a_j and b_{ij} provide the right balance and are needed for the results of this section to hold. From now on, I therefore set

$$a_j = \frac{1}{\sigma - 1} \quad \text{and} \quad b_{ij} = 1 - \frac{\varepsilon_j - 1}{\sigma - 1}. \tag{$\star$}$$

Proposition 6 below will explain where these particular values come from.

Sufficiency of the first-order conditions

For this new approach to be useful, it must be that the reshaped problem is easy to solve and that it's solution coincide with that of the original problem of the planner. The rest of this subsection addresses these two points in turn.

[19]For 12 to be well-defined as $\theta \to 0$, I impose $b_{ij} \geq -a_i(\varepsilon_j - 1)$ for all i, j (note that $\theta_i^{a_i}$ is part of q_i). This inequality is satisfied by the condition ($\star$) below.

The following propositions establish conditions under which $\mathcal{R}$ can be solved easily using standard algorithms.

Proposition 3. *Let $\varepsilon_j = \varepsilon$ and $\alpha_j = \alpha$. If $\Omega_{ij} = d_i e_j$ for some vectors d and e then the Karush-Kuhn-Tucker conditions are necessary and sufficient to characterize a solution to $\mathcal{R}$.*

A similar result holds for a different set of matrices Ω. Define $\bar{\Omega} = \omega (\mathbb{1}_n - I_n)$ where $\mathbb{1}_n$ is the $n \times n$ all-one matrix, I_n is the $n \times n$ identity matrix and $\omega > 0$. The matrix $\bar{\Omega}$ describes a network of potential connections in which firms are connected to each other, but not with themselves, with the same intensity ω. The following proposition shows that $\mathcal{R}$ is easy to solve when Ω is close to $\bar{\Omega}$.

Proposition 4. *Let $\sigma = \varepsilon_j$ for all j. Suppose that the $\{\beta_j\}_{j \in \mathcal{N}}$ are not too far from each other and that the fixed costs $f_j > 0$ are not too large. If the matrix Ω is close enough to $\bar{\Omega}$, then the Karush-Kuhn-Tucker conditions are necessary and sufficient to characterize a solution to $\mathcal{R}$.*[20]

Propositions 3 and 4 establish conditions under which a feasible point θ^* that satisfies the first-order conditions and the complementary slackness conditions solves $\mathcal{R}$.[21] As a result, standard numerical algorithms, such as gradient ascent, can rapidly solve $\mathcal{R}$ even for economies with thousands of firms.

Equivalence of the solutions

With a solution to $\mathcal{R}$ in hand, the next proposition establishes conditions under which that solution also solves $\mathcal{P}$.

Proposition 5. *If a solution θ^* to $\mathcal{R}$ is such that $\theta^* \in \{0,1\}^n$ then θ^* also solves $\mathcal{P}$.*

This result follows directly from the fact that the feasible set of $\mathcal{R}$ contains the feasible set of $\mathcal{P}$ and that both of their objective functions coincide over $\{0,1\}^n$ by construction.

Together, Propositions 3 to 5 offer a convenient way to solve $\mathcal{P}$. First, find a solution θ^* to $\mathcal{R}$ using a standard algorithm for convex problems. If θ^* belongs to $\{0,1\}^n$ then it also solves $\mathcal{P}$. This last condition can of course be tested in practice, but the whole solution approach would not be very useful if θ^* rarely belonged to $\{0,1\}^n$. Fortunately, condition ($\star$) is such that solutions to $\mathcal{R}$ are naturally pushed toward $\{0,1\}^n$.

To understand why, consider a stylized version of the first-order condition associated with the operating status θ_j of some firm j. That condition can be written as

$$\text{Marginal Benefit}_j \left(\theta_j, F_j (\theta)\right) - \text{Marginal Cost}_j \left(\theta_j, G_j (\theta)\right) = \overline{\mu}_j - \underline{\mu}_j \tag{13}$$

[20] To be precise, let $\bar{\beta}$ be a $n \times 1$ vector with identical elements. Then there exists a ball $\mathcal{B} = \left\{(\Omega, \beta) : \left\|(\Omega, \beta) - \left(\bar{\Omega}, \bar{\beta}\right)\right\| < \delta\right\}$ for $\delta > 0$ such that the statement holds for $(\Omega, \beta) \in \mathcal{B}$. An earlier version of this paper included a variant of Proposition 4 for binary matrices Ω.

[21] The assumptions about Ω in Propositions 3 and 4 impose that the firms' input sets are not too different from each other. For instance, $\Omega_{ij} = d_i e_j$ implies that the factor intensity vectors of the firms (the columns of Ω) are proportional to each other. Appendix D provides a fast algorithm to solve $\mathcal{R}$.

where $\overline{\mu}_j$ and $\underline{\mu}_j$ are the Lagrange multipliers associated with the constraints $\theta_j \leq 1$ and $\theta_j \geq 0$, respectively, and where both the "Marginal Benefit" and the "Marginal Cost" of increasing θ_j are functions that depend not only on θ_j itself, but also on the operating status of the other firms through some functions F_j and G_j.[22] These functions depend on the whole vector θ, instead of the specific θ_j, and I therefore refer to them as *aggregates*.

The following proposition shows how setting the reshaping constants a_j and b_{ij} to their ($\star$) values affects the "Marginal Benefit" and the "Marginal Cost" functions in (13).

Proposition 6. *Under the ($\star$) condition, the marginal benefit and the marginal cost of increasing θ_j only depend on θ_j through the aggregates F_j and G_j.*

This proposition shows that it is possible to find values for a_j and b_{ij} such that the "Marginal Benefit" and the "Marginal Cost" functions no longer depend on θ_j directly. In this case, θ_j affects the first-order condition (13) only through the aggregates F_j and G_j. These aggregates are summations over many firms and, as the number of firms increases, they become more and more independent of θ_j itself, and so does the left-hand side of (13).[23] This pushes solutions to $\mathcal{R}$ toward $\{0,1\}^n$. To understand why, consider for instance a gradient ascent algorithm that begins at $\theta_j = 1/2$. Suppose that the marginal benefit of increasing θ_j is larger than the marginal cost, the planner increases θ_j slightly. But since the marginal benefit and marginal cost themselves are independent of θ_j, the marginal benefit remains larger and the planner keeps increasing θ_j until it reaches 1, at which point the Lagrange multiplier $\overline{\mu}_j$ becomes positive to enforce the constraint $\theta_j \leq 1$. The opposite happens if the marginal benefit is initially lower than the marginal cost. As a result, the solution θ^* to $\mathcal{R}$ is pushed toward $\{0,1\}^n$, and Proposition 5 guarantees that θ^* also solves $\mathcal{P}$.

3.3 Example with two firms

To better understand how the solution approach works, it is helpful to consider a simple economy with two firms $j \in \{1,2\}$ and a complete set of potential connections between them ($\Omega_{ij} = 1$ for all i,j). The objective function $V(\theta)$ of the relaxed planner's problem without any reshaping ($a_j = 1, b_{ij} = 0$) is shown in Figure 3a, where warmer colors represent higher utility levels for the planner. The horizontal and vertical axes refer to the operating decisions θ_j of firms 1 and 2, respectively. We see that V is shaped like a saddle with local maxima at $(\theta_1, \theta_2) = (1,0)$ and $(0,1)$, and local minima at $(0,0)$ and $(1,1)$. The global maximum is at $(1,0)$.

[22] The expressions for the "Marginal Benefit" and "Marginal Cost" functions can be found in the proof of Proposition 6 in Appendix C.

[23] The proof of Proposition 6 shows that F_j and G_j are functions of Q and $B_j = \left(\sum_{i\in\mathcal{N}} \theta_i^{b_{ij}} \Omega_{ij} q_i^{\varepsilon_j - 1}\right)^{1/(\varepsilon_j - 1)}$. When Ω is very sparse, it is possible that q_i depends strongly on θ_j itself, in which case B_j would remain affected by θ_j even when n is large. These cases are easily identified in practice: the numerical solution to $\mathcal{R}$ will be outside of $\{0,1\}^n$. Appendix B.3 shows that the solution approach performs well even in these more pathological cases. Appendix B.4 shows the performance of the algorithm when Ω is sparse.

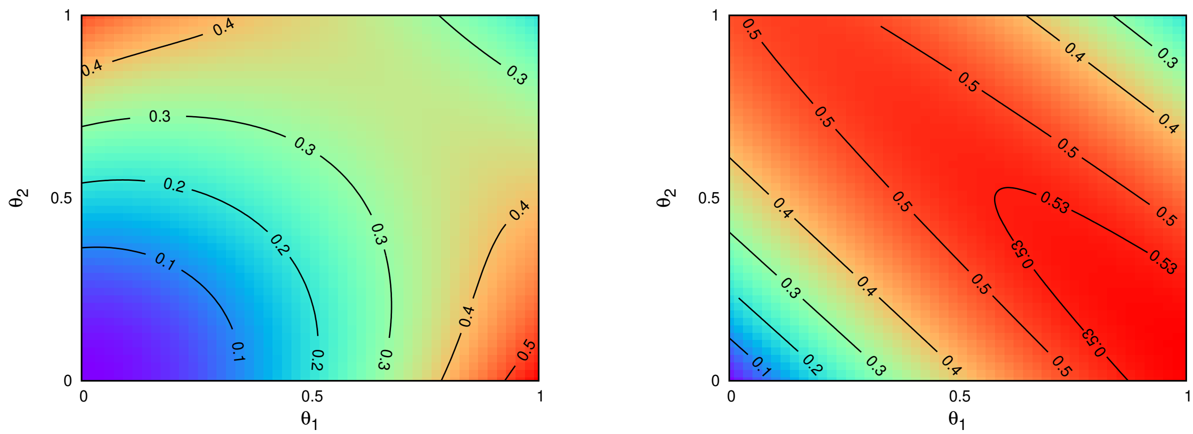
Since V is not concave, the first-order conditions are not sufficient to characterize the global maximum—they are indeed satisfied at both $(0,1)$ and $(1,0)$. As a result, this problem cannot be solved reliably with standard numerical algorithms. Starting from an initial point, these algorithms generally move locally by following the steepest slope, so that they can easily converge to the local (but not global) maximum at $(0,1)$.

Figure 3b shows the objective function $V_R(\theta)$ of the same optimization problem but, this time, reshaped according to condition $(\star)$. Three things are worth noticing. First, V and V_R coincide, by construction, at the corners $\{0,1\}^2$. As a result, the ranking of these corners, in terms of utility, is the same in both problems. Second, the reshaping stretches the objective function so that V_R is concave. The first-order conditions are therefore sufficient to characterize the global maximum. Third, the reshaping did not create another maximum somewhere inside $[0,1]^2$, and V_R's maximum is also the maximum of V. As a result, starting from any initial θ_0 in $[0,1]^2$, a simple gradient ascent algorithm will converge to the global maximum at $(1,0)$. This point also solves $\mathcal{P}$ by Proposition 5.

Building on this example, we can get some intuition about why the specific values given by $(\star)$ are useful. Consider for instance the role played by the constant a_j. Suppose that we begin the optimization at the local maximum $\theta=(0,1)$ in Figure 3a, and are looking in its neighborhood to improve the utility of the planner. Equation (11) implies that increasing θ_1 slightly, to move toward the global maximum, would incur a marginal cost proportional to fQL. But, without reshaping the problem, the marginal benefit of this move is of the order of $\partial Q/\partial\theta_1 \propto \theta_1^{\sigma-2}$, a negligible amount for $\theta_1 \approx 0$ particularly when goods are highly substitutable (high σ).[24] Intuitively, when σ is large, goods 1 and 2 are essentially the same to the household and its utility comes mostly from the operation of the highest-q firm (see equation (10)). Since at $\theta=(0,1)$, $q_1=0$ and $q_2 \gg 0$, increasing θ_1 has a negligible benefit but a non-negligible cost. It is therefore not *locally* advantageous for the algorithm to deviate and it remains stuck at the local maximum.

If instead the optimization problem is reshaped, the marginal benefit of increasing θ_1 becomes $\partial Q/\partial\theta_1 \propto \theta_1^{a_j(\sigma-1)-1}$, so that the reshaping constant a_j is effectively changing how substitutable goods are. When a_j satisfies $(\star)$, the marginal benefit is of the same order of magnitude as the marginal cost fQL and the algorithm is free to move away from $(0,1)$ and towards the global maximum. A similar intuition applies for the constant b_{ij}, except that in this case the elasticity of substitution between intermediate inputs ε is the key parameter guiding the reshaping procedure.

[24] When $\sigma<2$, we have the opposite problem, there is too much concavity in the objective function and, without reshaping, the maximum is generally in $[0,1]^n$. In this case, the reshaping flattens the objective function to push the solution to a corner.



(a) The objective function $V(\theta)$ of the relaxed (but not reshaped) problem is not concave

(b) The objective function $V_R(\theta)$ of the relaxed and reshaped problem is concave

Figure 3: Reshaping the planner's problem in a simple economy

3.4 Numerical tests

The theoretical results of the last section provide *sufficient* conditions under which reshaping the planner's problem provides the solution to $\mathcal{P}$, but these conditions are far from being necessary. In this section, I show through numerical simulations that the solution approach also works well when these conditions are not satisfied. As they involve different testing procedures, I first consider economies with a few firms and then present results for economies with a large number of firms.

Few firms

When the economy contains a small number of firms, it is possible to find the true solution to the planner's problem by comparing the utility provided by the 2^n possible vectors $\theta \in \{0,1\}^n$, as in the exhaustive search algorithm mentioned above. We can then compare this allocation to the solution of the relaxed problem with and without reshaping.

Appendix B provides the details of the simulations. They involve a broad range of economies with firms that differ along all the dimensions of heterogeneity allowed by the model. They also cover matrices Ω with different shapes and various degrees of sparsity. The results for economies with up to $n = 14$ firms are presented in Table 1. We see that reshaping the planner's problem (first two columns) attributes the correct status θ to more than 99.9% of the firms. It also finds output values that are within 0.001% of their correct values. In contrast, without reshaping the problem (last two columns), about 15% of the firms are assigned the wrong status θ and the average error in output can reach above 0.9%, a large number when studying aggregate fluctuations. The table also shows that the performance of the reshaping algorithm stays relatively constant as n increases, in contrast to the non-reshaped solution which performs worse as the number of firms increases.[25]

Table 1: Testing the solution approach for n small

	With reshaping		Without reshaping	
n	Correct θ	Error in C	Correct θ	Error in C
8	99.9%	0.001%	86.5%	0.791%
10	99.9%	0.001%	85.2%	0.855%
12	99.9%	0.001%	84.5%	0.903%
14	99.9%	0.001%	84.0%	0.926%

Notes: See Appendix B for the details of the simulations.

These simulations cover a broad range of environments but we can also focus the tests on more empirically relevant economies. To do so, I set the parameters of the model to their calibrated values, which will be described in Section 5, except for the number of firms n which must remain small so that the true solution of the planner's problem can be found. Table 2 presents the results. We see that the solution approach still performs well, although the errors in C and θ are larger than in the general simulations of Table 1.[26] More importantly, the error in C are slowly declining as n increases which is reassuring about the algorithm's performance for realistic n. In contrast, without reshaping the performance of the algorithm degrades quickly as n increases.

Table 2: Testing the solution approach with calibrated parameters

	With reshaping		Without reshaping	
n	Correct θ	Error in C	Correct θ	Error in C
8	98.2%	0.009%	89.6%	0.229%
10	98.9%	0.008%	87.7%	0.274%
12	98.8%	0.008%	86.8%	0.289%
14	98.8%	0.008%	85.3%	0.322%
16	98.8%	0.008%	84.5%	0.339%
18	98.9%	0.007%	84.2%	0.348%
20	98.8%	0.007%	83.3%	0.367%

Notes: Parameters as in the calibrated economy of Section 5 except for n. For each n, 15,000 random economies are simulated. See also Appendix B for more details about the procedure.

[25]Table 1 shows the fraction of firms that are assigned the correct status θ_j. An alternative measure of success would be to consider the fraction of economies for which the whole vector θ is correct. Table 11 in Appendix B.2 presents these results. On average, the algorithm finds the correct vector θ in 99.7% of economies with reshaping. Without reshaping that number is 19.9%.

[26]The calibrated matrices Ω are very sparse, which makes these economies more challenging for the algorithm.

Many firms

When n is large, finding the true solution to $\mathcal{P}$ through an exhaustive search would take an infeasibly long time. We can, however, verify whether there exist welfare-improving deviations from the solutions to the relaxed problems. To do so, I change the operational status θ_j of each firm to see if it improves the utility of the planner. I keep repeating this procedure as long as there are deviations to be found. I then compare this deviation-free solution to the original one. The precise algorithm is described in Appendix D.3.

Since this procedure is computationally costly, I only consider economies that follow the calibration of Section 5. The results are presented in Table 3. Again, the reshaping approach performs very well. After all the possible deviations are accounted for, more than 99.9% of the firms have kept the same operating status θ_j and aggregate output has changed by a negligible amount.[27] In contrast, without reshaping more than 30% of the firms are assigned the wrong operating status, and the error in aggregate output amounts to 0.58%. While this test does not guarantee that the solution approach finds the correct efficient allocation, it provides a good indication that there are no obvious mistakes in its solution.

Table 3: Testing the reshaping approach for n large

	With reshaping		Without reshaping	
n	Correct θ	Error in C	Correct θ	Error in C
1000	$> 99.9\%$	$< 0.001\%$	68.9%	0.58%

Notes: Parameters as in the calibrated economy (see Section 5.2). I simulate 100 different matrices Ω and, for each Ω, draw 100 productivity vectors z. I run the procedure described in Appendix D.3 on each of them and report average results. $x < 0.001\%$ indicates that $x > 0$ but that proper rounding would yield 0, and similarly for $x > 99.9\%$.

Appendix B provides several additional exercises to further test the robustness of the solution approach. It considers economies 1) with very sparse matrices Ω, 2) in which the production network is created through individual link formation, and 3) for which the solution to $\mathcal{R}$ is not in $\{0,1\}^n$. These additional tests show that the solution approach performs well in a broad set of economic environments.

4 Economic forces

We now explore how the economic forces at work in the environment influence the structure of the production network, the propagation of shocks and the distribution of firm-level outcomes.

[27] When the reshaping approach fails it is in general because it gets the wrong operating status for a firm that is fairly isolated from the rest of the network. Since, these firms are in general small, they only have little influence on aggregate production, which explains why the error in output is very small in Table 3.

Equation (9), which pins down the productivity q of every firm as a function of the production network θ, is central for that purpose.

4.1 Highly-connected firms are more likely to operate

As discussed earlier, equation (9) embeds two key mechanisms: 1) having access to an additional supplier increases a firm's productivity q_j, and 2) because of (9)'s recursive structure, any gain in productivity q_j propagates downstream to benefit all of j's (perhaps removed) customers. These two mechanisms incentivize the planner to operate firms that are well-connected in terms of both direct and indirect neighbors (second neighbors, etc). Indeed, a firm j with many upstream suppliers enjoys a larger q_j, which makes it more attractive for the planner to pay its operating fixed cost $f_j L$. Similarly, a firm with numerous downstream producers increases the productivity of its many (perhaps indirect) customers. As a result, this firm is valuable to the planner, which is then more likely to pay the fixed cost to operate it.

The following proposition formalizes this discussion.

Proposition 7. *In a large economy, operating a firm (weakly) increases the incentives to operate its (perhaps indirect) potential suppliers and customers.*[28]

4.2 Cascades of firm shutdowns can arise

One immediate consequence of this proposition is that the efficient allocation exhibits complementarities between the operating decisions of nearby firms: they tend to operate (or not) together. These complementarities, in turn, can generate cascades of firm shutdowns. Consider for instance a firm j that stops production after suffering from a severe z shock. In response, its first neighbors, having lost a useful supplier or a valuable customer, are also likely to shut down. But then j's second neighbors are also losing a neighbor and are at a greater risk of shutting down themselves. Since the same logic applies to further neighbors of j, the initial z shock can trigger a wave of shutdowns that spreads through the economy.

The magnitude of these cascades depends on the strategic importance of firm j in the production network. If j is the sole supplier to many firms, or their only customer, its exit is likely to trigger multiple shutdowns in its neighborhood. On the other hand, the planner is likely to keep strategically important firms in operation even after large adverse shocks. The model therefore predicts a negative correlation between the likelihood of a firm shutting down and the magnitude of the cascade it triggers upon exit. We will see in the next section that this correlation is also visible in U.S. data.

[28] An economy is *large* if the operating decision of any individual firm has no impact on the Lagrange multiplier w associated with the labor resource constraint (8). A potential customer (supplier) is a firm that can be reached by moving downstream (upstream) in the network described by Ω.

4.3 Clustering of economic activity

Since having multiple suppliers and customers is valuable, the planner prefers to create networks in which many firms have many neighbors. Organizing production in this way involves building tightly connected clusters of operating firms. In this type of configuration, each firm j benefits from the presence of multiple suppliers and contributes to the productivity of its many customers. As a result, these clusters reinforce the productivity q of all its members and makes organizing production in this way particularly efficient.

Example. A simple example is useful to understand what features of the environment make clustering more or less desirable. Suppose that the planner wants to operate n identical firms and can decide to do so by operating clusters of m firms that are fully-connected (so that each producer supplies to the $m-1$ other firms in the cluster).[29] Which amount m of clustering is preferred? We can easily compute the productivity q of each firm in an m-cluster. Assuming that firms share the same productivity $z=1$ for simplicity, we find $q=(m-1)^{\frac{\alpha}{1-\alpha}\frac{1}{\varepsilon-1}}$ from (9), and from (11), we can compute total output as

$$C = n^{\frac{1}{\sigma-1}} (m-1)^{\frac{\alpha}{1-\alpha}\frac{1}{\varepsilon-1}} L(1-fn). \tag{14}$$

This expression is increasing in m, so the planner prefers, all else equal, to operate a unique tightly connected group of firms instead of a large number of connected pairs of firms. Equation (14) also shows that this preference for clustering is more important when the share α of intermediate inputs is large and when the elasticity of substitution ε between these inputs is low. Under these conditions, a firm values additional suppliers more, which leads to more clustering in the efficient allocation.

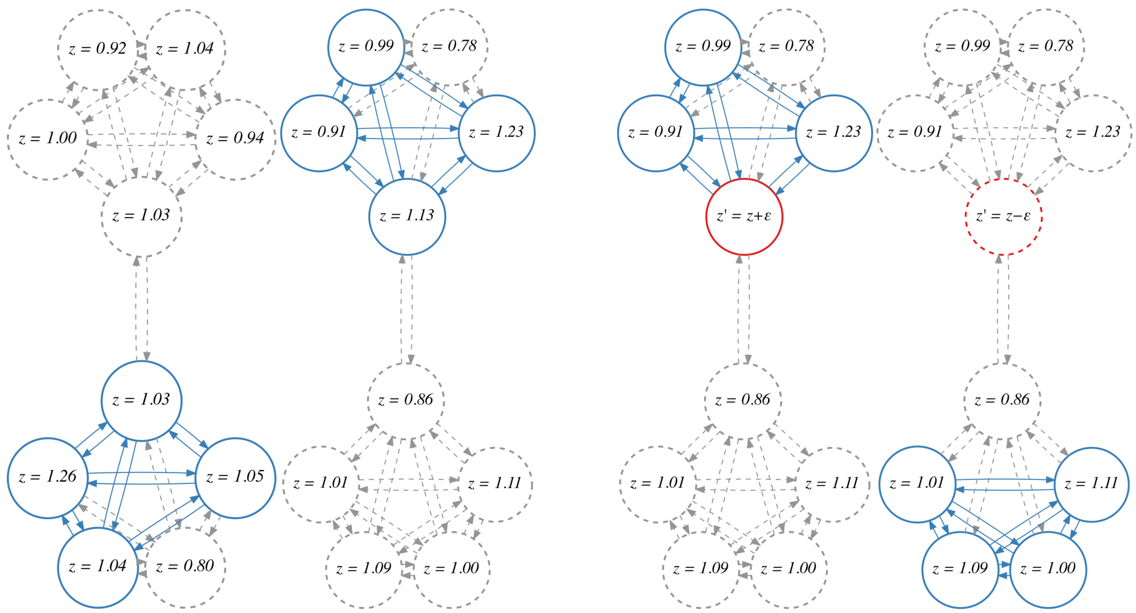
Proposition 8 formalizes the intuition from that example and shows that the planner prefers to operate groups of firms that are highly connected.

Proposition 8. *The incentives of the planner to operate a group of firms (weakly) increase with additional potential connections between them.*

The previous example featured firms with identical total factor productivity z. When z varies across firms, the efficient network tends to cluster economic activity around the most productive firms. Figure 4a presents an example of this process in an economy in which the network Ω consists of two groups of five fully-connected firms. The two groups are linked together by a single pair of connections. The figure shows the same economy under two randomly drawn vectors z. We can see that the planner clusters economic activity in either the top or the bottom group of firms.

[29] For simplicity, assume that n is large enough and that it is divisible by m. While the current example might seem abstract it fits in the framework outlined in the last section. The whole economy would be formed of groups of n firms which various degree of clustering. The exercise would then be about which of these groups the planner prefers to operate.

Operating a few firms in each group would be a poor way to organize production. In addition, the active cluster tends to be the one with the highest-z firm. By organizing production around these high-performing firms, the planner magnifies their impact on their neighbors and on aggregate output.



(a) Clustering of activity

(b) Small shock, large reorganization

Figure 4: Economic forces shape the network

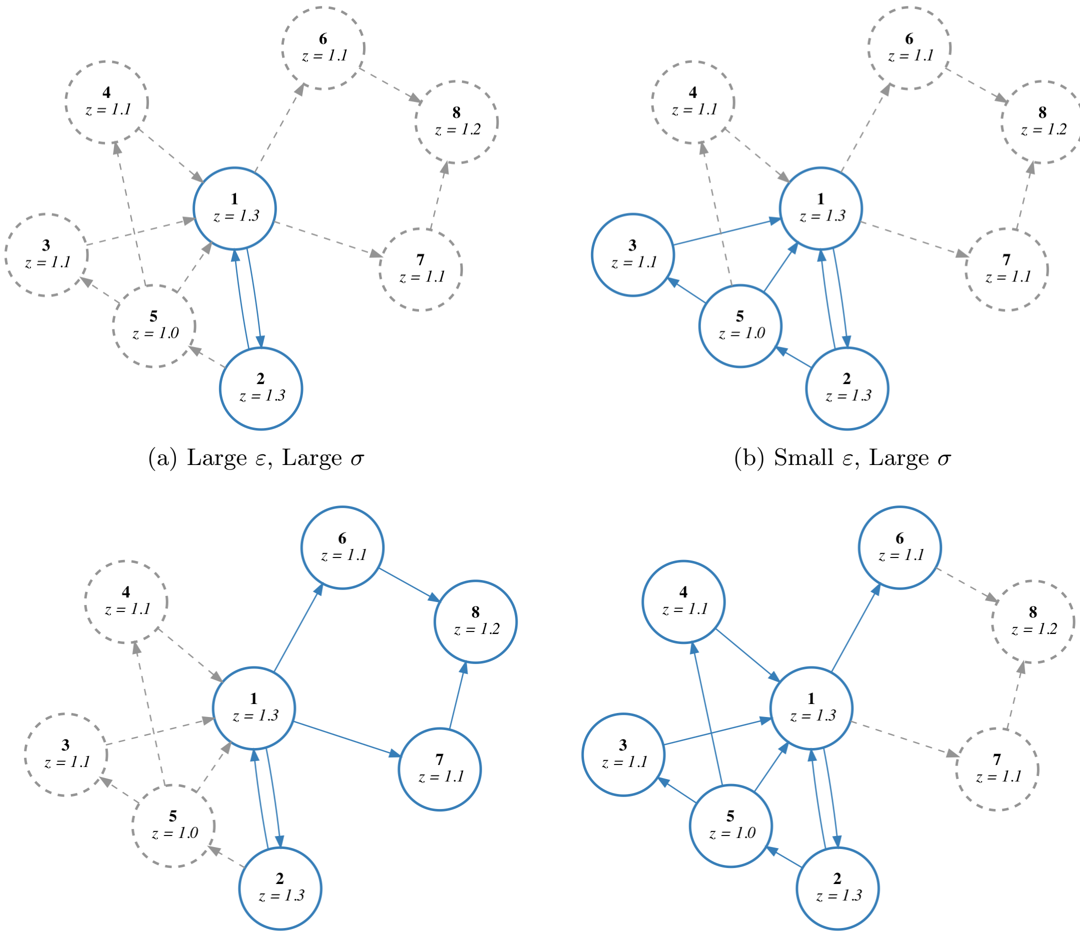
4.4 Small shocks can lead to large reorganizations

One perhaps unusual feature of the model is that a small change in the environment can trigger a large reorganization of the network. When designing the network, the planner compares the 2^n alternatives $\theta \in \{0,1\}^n$ and selects the one providing the highest utility to the household. As, say, a firm's TFP z declines there is a point at which the planner shuts the firm down. But because of the complementarities between neighbors, it might be better to shut down the whole cluster around the firm and to move production elsewhere.

Figure 4b provides an example. Both economies are identical except for the productivity z of the red firm which is slightly larger in the economy on the left. While the drop in z from left to right is negligible, it triggers a large reorganization of the network. Aggregate output, however, is barely affected. Indeed, the planner reorganizes the network precisely to limit the negative impact of the shock on output. But while aggregate output barely changes, firm-level distributions can change substantially. In this example, for instance, the dispersion in labor productivity, output and employment across firms collapses after the shock. A negligible shock can therefore have a large impact on cross-producer distributions.

4.5 Elasticities and the structure of the network

The elasticities of substitution in the aggregators for final consumption (σ) and intermediate inputs (ε) play an important role in shaping the production network. Figure 5 shows the efficient network in four economies that differ only in terms of σ and ε. In Panel (a) both elasticities are large. Since firms are essentially producing the same good, the planner prefers to focus production in the hands of a small group of very productive suppliers (firms 1 and 2). The situation is different in Panel (b) where the elasticity ε is small. In this case, goods are poor substitutes when they serve as intermediate inputs and additional suppliers are more valuable. The planner therefore provides additional inputs to firm 1 to increase its productivity. If, instead, ε remains large but σ is small, as in Panel (c), goods are poor substitutes in the consumption aggregator. The household prefers to enjoy a wide variety of products and, as a result, the planner operates producers that are downstream from firm 1. These firms can then take advantage of the high productivity of firm 1 to provide the household with cheap additional goods. When both elasticities are large, as in Panel (d), the planner moves on both margins to operate some additional downstream and upstream producers.



(a) Large ε, Large σ (b) Small ε, Large σ

(c) Large ε, Small σ (d) Small ε, Small σ

Figure 5: Role of σ and ε in shaping the production network

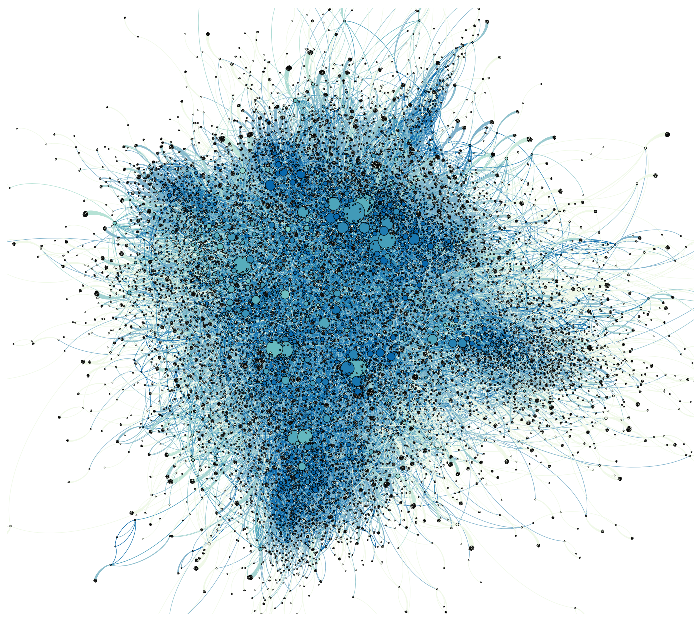
5 Quantitative exploration

This section provides a basic calibration of the model and shows that it captures salient features of the data such as cascades of firm shutdowns and movements in the structure of the production network over the business cycle.

5.1 Data

The main firm-level input-output dataset is the Factset Revere Supply Chain Relationships Data, which provides detailed annual data from 2003 to 2016. These data are gathered by analysts from a variety of sources such as 10-Q and 10-K fillings, annual reports, investor presentations, websites, press releases, etc. I restrict the sample to links in which at least one partner is located in the United States. In an average year, the sample includes almost 13,000 firms and more than 40,000 relationships.

Figure 6 shows the production network of the U.S. economy in 2016, as described by the Factset data. Each circle represents a firm with the size of the circle increasing in its number of connections. The largest circles are all household names. In 2016, the firm with the largest in-degree was Walmart with 448 suppliers and the firm with the largest out-degree was Microsoft with 332 customers. I describe these data in more details below.



Notes: The size of a circle represents the number of connections of a firm. Darker colors denote firms with higher local clustering coefficients. Image generated using Gephi with the Yifan Hu layout. Vector image; zoom in for additional details. There are 20,702 firms and 62,474 links.

Figure 6: 2016 Factset Revere U.S. firm-level production network

We can use these data to evaluate how much of the variation in the production network can be explained by the firms' extensive margin of operation. In an average year, more than 40% of all link destructions observed in Factset occur at the same time as either the supplier or the customer (or both) stops producing. This number suggests that the firms' extensive margin plays a sizable role in shaping input-output relationships.[30]

To verify the robustness of some empirical patterns, I also rely on Compustat as another source of annual data. Compustat gathers information about a firm's major customers, defined as buyers of more than 10% of total sales, from annual financial statements. This reporting is mandated by Financial Accounting Standards No. 131. Since firms are not required to report smaller customers, we rarely see a firm supplying to more than 10 clients in the data.[31] Another limitation of these data

[30] The analogous exercise for link creations finds a similar number. To remove high-frequency gaps in the data, I assume that a link is created during the first year it appears in the dataset and is destroyed during its last. Similarly, a firm is considered as shutting down during the last year that it is in the sample.

[31] As a result, the tail of the out-degree distribution in Compustat is likely to be artificially thinner. Factset also relies on these data but supplements them using a variety of other sources. This issue is therefore likely to be less important in the Factset sample.

is that the names of the customers are self-reported, so General Motors might enter the database as "General Motors", "GM", "General Mtrs", etc. To address this issue, Cohen and Frazzini (2008) (CF) and Atalay et al. (2011) (AHRS), have used a combination of automatic algorithms and manual matching to properly identify each firm and to construct the production networks. Their samples cover the periods 1980 to 2004 and 1976 to 2009, respectively. I use both samples below to show the robustness of some empirical patterns.[32]

The Compustat dataset cover fewer firms than Factset—about 1,300 firms and 1,500 relationships in an average year—but over a longer time period. Since it is available since 1976 it provides a more accurate picture of the evolution of the production network over the business cycle.

5.2 Parametrization

The model is parametrized at an annual frequency, and I normalize $A = 1$ and $L = 1$. For the share of intermediate goods in the production function, I follow Jorgenson et al. (1987) and set $\alpha = 0.5$. Jones (2011) surveys the literature on the share of intermediate goods in different countries and also suggests $\alpha = 0.5$ as a benchmark.

I assume that the log of the productivities z_{it} are drawn from independent AR(1) processes with persistence ρ_z and a standard deviation σ_z for its ergodic distribution. Bartelsman et al. (2013) measure the dispersion in firm-level physical productivity in a number of countries and find $\sigma_z = 0.39$ for the United States. Their estimation technique controls for the usage of intermediate inputs. For the persistence, I follow Foster et al. (2008) and set $\rho_z = 0.81$. Since the model itself is static, the persistence in z_{it} is the only inter-temporal linkage in the economy. The idiosyncratic shocks to z_{it} will also generate aggregate fluctuations because of the finite number of firms.

There is no consensus in the literature about the cost of overhead labor f. Since employment in management occupations accounts for about 5% of total employment, I set f so that $f \times n = 5\%$.[33] For the number of firms, I set $n = 1000$ as a good trade-off between realism and computation time.[34]

Broda and Weinstein (2006) use disaggregated trade data for the U.S. to estimate the elasticity of substitution between product varieties. I set $\sigma = 5$ as an average of their estimates. The empirical literature provides little guidance about the elasticity of substitution between intermediate inputs at the firm level. I therefore set $\sigma = \varepsilon$. I describe below how changes in these parameters affect some of the results.

I construct Ω by assuming that the number of *potential* incoming and outgoing connections, for

[32] I thank the authors for sharing their data.

[33] These fixed costs have often been estimated using models of firm turnover. For instance, Bartelsman et al. (2013) estimate a model of firm entry and exit on the manufacturing sector and estimate that overhead labor accounts for 14% of total employment in the industry. This number is likely to be lower for firms outside of manufacturing. Higher fixed costs do not affect the qualitative findings of this section.

[34] See Appendix E.1 for simulations with $n = 20,000$ firms and aggregate shocks. The results are similar.

any given firm, is drawn from a bivariate power law of the first kind. This family of distributions is entirely described by a unique shape parameter ξ.[35] I set $\xi = 1.79$ so that the distribution of *active* incoming connections generated by the model is close to its empirical counterpart in the Factset data.[36] These two distributions are well approximated by power laws, with the empirical distribution close to following Zipf's law (see Section 5.3 below). I therefore target a power law exponent of 1 for the distribution generated by the model. This indirect inference approach ensures that the calibrated economy is consistent with a key feature of the empirical production network.[37] So that the results do not hinge on one particular matrix Ω, I randomly draw 20 different Ω's and, for each of them, simulate the economy for 100 periods.[38] The results are averages over these simulations.

Table 4 shows the parameters of the calibrated economy.

Parameter	Value
Time period	1 year
Average productivity	$A = 1$
Labor supply	$L = 1$
Number of firms	$n = 1000$
Intermediate goods intensity	$\alpha = 0.5$
Elasticity of substitution for final goods	$\sigma = 5$
Elasticity of substitution for intermediate goods	$\varepsilon = 5$
Standard deviation of individual productivities	$\sigma_z = 0.39$
Persistence of individual productivities	$\rho_z = 0.81$
Fixed cost of operation	$f = 0.05/n$
Shape of Ω	$\xi = 1.79$

Table 4: Parameters of the calibrated economy

5.3 Calibrated economy

Table 5 shows how the calibrated production network compares to the U.S. data. I focus on three key moments to describe the overall structure of the network. The first two moments are the in-degree and out-degree distributions. In the model and in the data, these distributions are close to power laws so that their exponent parameters provide a good description of the full

[35] The probability that a firm has x_{in} and x_{out} potential inbound and outbound links in Ω is $p(x_{in}, x_{out}) = \xi(\xi - 1)(x_{in} + x_{out} - 1)^{-\xi - 1}$.

[36] I target moments from Factset, instead of Compustat, as it is the most comprehensive data source for linkages. I also target the in-degree distribution, instead of its out-degree counterpart, as it is less affected by the 10% reporting threshold described in Section 5.1. Online Appendix E.2 shows how the production network is affected by changes in ξ.

[37] All power law exponents are estimated using the estimator of Gabaix and Ibragimov (2011). Appendix D.5 provides the algorithm to construct Ω.

[38] I discard and redraw simulations for which iterating on the first-order conditions does not converge to a point θ in $\{0, 1\}^n$. This rarely happens and, overall, the rejected networks do not look different. Keeping all the simulations in the sample yields very similar results.

distributions. As is now well-known, these exponents have an important influence for the aggregate impact of idiosyncratic shocks (Acemoglu et al., 2012).[39] The third moment is the global clustering coefficient which measures how tightly connected firms are with one another—a key moment given the importance of clustering for productivity.[40]

We see from Table 5 that the calibrated economy, despite its simplicity, fits the Factset data relatively well but there are some discrepancies with the Compustat datasets which is not surprising given its coverage. These discrepancies are particularly important when looking at the out-degree distribution—a consequence of the 10% truncation threshold described above.

Table 5: Production network in the calibrated economy and in the data

	Model	Datasets		
		Factset	Compustat	
			AHRS	CF
Power law exponents				
In-degree distribution	1.00	0.97	1.13	1.32
Out-degree distribution	0.96	0.83	2.24	2.22
Global clustering coefficient (normalized)	3.31	3.46	0.08	0.09

Notes: Power law exponents are estimated following Gabaix and Ibragimov (2011). Global clustering coefficients are multiplied by the square roots of the number of nodes. See footnote 40 for details.

Figure 7 shows the degree distributions in the model and in the Factset data for 2016, the most recent year in the sample. To highlight the shape of these distributions, the figure uses a log-log scale and plots the complementary cumulative distributions (CCDF) on the vertical axis. The roughly linear shapes confirm that they are close to power laws. The model fits the in-degree distribution well but the fit of the out-degree distribution is less precise. Again, the data limitations discussed above are the likely culprit for the departure from the power law observed in the right tail of the out-degree distribution.[41]

[39] The out-degree distribution in the calibrated economy is closely related to the centrality distribution, which is the focus of Acemoglu et al. (2012). Conducting the exercises of this section with the centrality distribution instead yields similar results.

[40] The global clustering coefficient is computed on the undirected graph. It equals three times the number of triangles (three fully connected nodes) divided by the number of triplets (three connected nodes). In power law graphs, the global clustering coefficient declines naturally with n. Following Ostroumova Prokhorenkova and Samosvat (2014), I therefore normalize the means of the coefficients by multiplying them by the square root of the number of nodes. This normalization allows for a better comparison of networks across datasets.

[41] The model generates similar in- and out-degree distributions. In countries with better firm-level network data such as Japan, the in- and out-degree distributions look similar (Bernard et al., 2015).

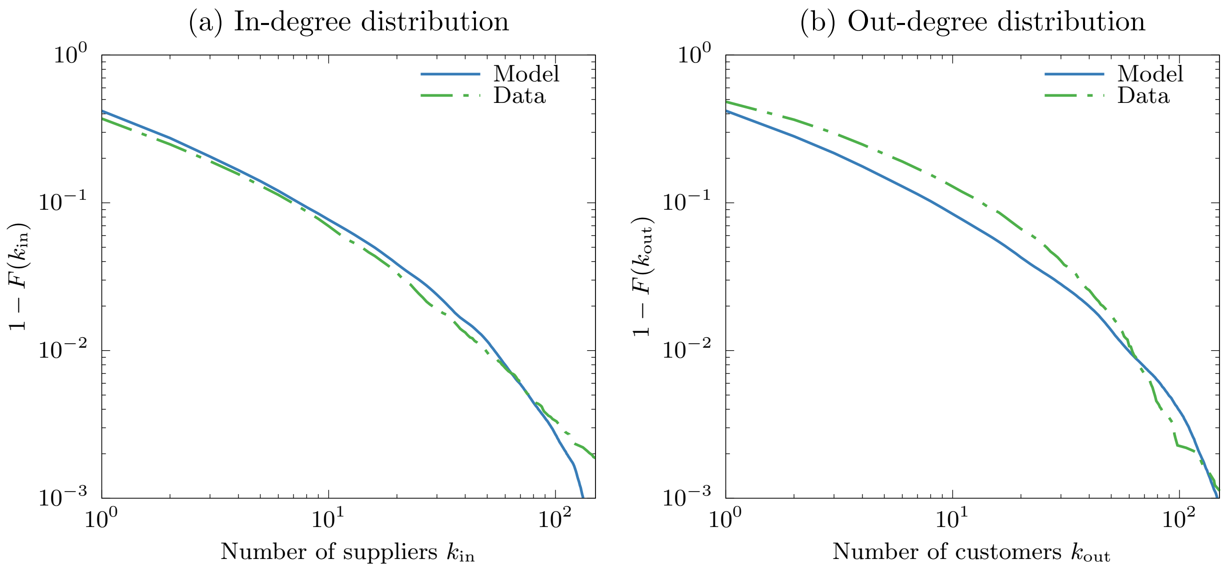


Figure 7: Distribution of the number of suppliers and customers

5.4 Comparison with a neutral network

To highlight which features of a network are desirable for efficiency, we can compare the calibrated network, which has been designed optimally by the planner, to a *neutral* benchmark built randomly by operating each firm with some probability $p > 0$.[42] All other quantities—except for the network itself—are chosen optimally by the planner. Since it is completely random, any discrepancies between the neutral benchmark and the efficient network are design decisions taken by the planner to improve efficiency.

Table 6 shows how both networks differ. The power law exponents are smaller in the efficient network, indicating thicker tails than in the neutral benchmark. The efficient network therefore features a larger share of highly connected suppliers and customers. The clustering coefficient is also larger in the efficient network.[43] These moments highlight the planner's preferred way of organizing production: tightly connected clusters of economic activity centered around firms with many connections. By building the network in this way, the planner takes full advantage of the gains from input variety present in the environment.

[42] p is set so that both networks have the same number of active firms.

[43] This exercise shows that a large amount of clustering among firms comes from the mechanisms of the model. To further disentangle the role of the endogenous forces versus the exogenous matrix Ω, I also compute the local clustering coefficient of each firm in the endogenous network and compare it to its local clustering coefficient in the exogenous network Ω. I find that there is 40% more clustering in the endogenous network, such that the forces at work in the economy generate substantial clustering on their own.

Table 6: Comparing the efficient network and the neutral benchmark

Network	Power law exponents		Clustering coefficient
	In-degree	Out-degree	
Efficient	1.00	0.96	3.31
Neutral	1.16	1.15	2.25

Notes: Power law exponents are estimated following Gabaix and Ibragimov (2011). Global clustering coefficients are multiplied by the square roots of the number of nodes (see footnote 40).

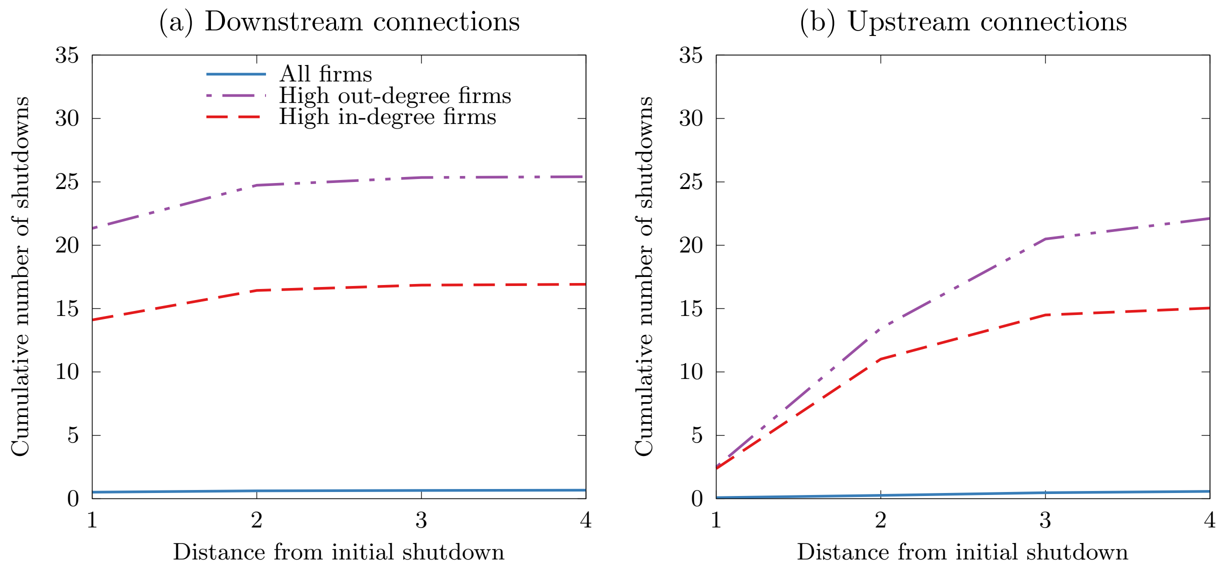
5.5 Cascades of firm shutdowns

We can use the calibrated model to evaluate how cascades of firm shutdowns arise and propagate through the production network. To do so, I select a random active firm in the calibrated economy and set its productivity z to zero so that it stops production. I then compute the new efficient allocation and count the neighbors of the firm that shut down. Figure 8 shows the outcome of this exercise. The left panel looks at the firm's downstream neighbors, and the vertical axis shows the cumulative number of shutdowns as we move away from the shuttered firm. The right panel provides the same information but for upstream neighbors. The figure also differentiates between cascades originating from the average firm, and from firms with a high number of neighbors (above 99th percentile).

We see that the shutdown of an average firm is likely to only create a small cascade: about 0.6 of its downstream neighbors, and even fewer of its upstream neighbors, shut down. As we move to high degree firms the cascades become larger: for important suppliers about 26 downstream neighbors are wiped by the cascade and the production network is extensively reorganized. Figure 8 also shows that cascades mostly propagate downstream, from customer to customer, instead of upstream. This is a consequence of the gains from input variety, embedded in equation (9), which make losing a supplier particularly costly in terms of productivity. That mechanic also implies that the exit of high out-degree producers is more damaging than that of high in-degree ones.

We can use the model to evaluate how cascades propagate in counterfactual economies with different parameters. One parameter with an important role for the cascades is the elasticity of substitution between intermediate inputs, ε. Figure 9 shows the outcome of the same exercise as that of Figure 8 but in an economy in which $\varepsilon = 3$, such that intermediate inputs are worse substitutes than in the calibrated economy. By comparing the figures, we see that the lower elasticity affects the cascades in two important ways: 1) shocks to high degree firms now trigger larger cascades (notice the different scales) and, 2) these cascades have more substantial upstream propagation compared to those in the benchmark economy.

Why cascades are larger is easy to understand. With ε small, intermediate inputs are poor substitutes and losing a supplier has a larger negative impact on a firm's productivity, which leads



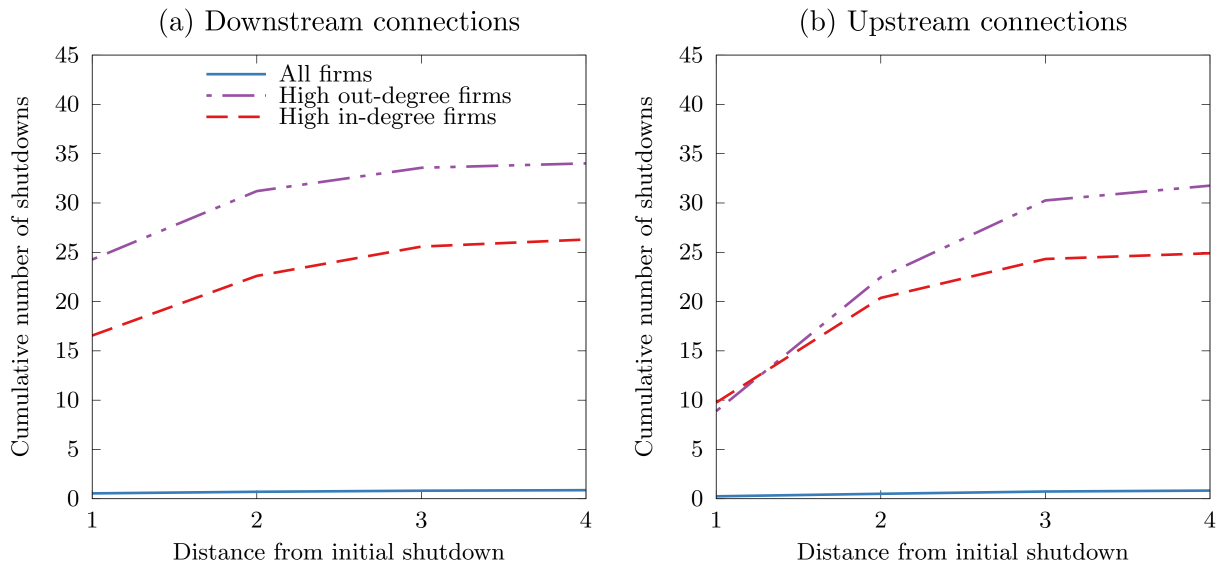
Notes: Cumulative number of exits at different distances from shuttered firm. "High degree" refers to the firms above the 99th percentile. Simulations of 100 randomly drawn matrices Ω, for each of which 1000 cascades are created.

Figure 8: Cumulative cascades by degree of originator

to more shutdowns. To understand why it also makes cascades have more upstream propagation, it is useful to think about the planner's incentives to operate a firm in this economy. Since the elasticity of substitution in the consumption aggregator, σ, is relatively large, equation (10) implies that firms with high productivity q are particularly valued by the household. But because ε is small, these high-q firms are likely to get their high productivity from a large number of suppliers. As a result, if one of the high-q firm shuts down, its many suppliers are no longer useful (they don't contribute much to Q) and the planner is likely to shut them down as well, thereby triggering an upstream cascade. In contrast, in the benchmark economy, where $\varepsilon = \sigma = 5$, the planner puts a higher value on the direct contribution to final consumption of these many suppliers, and they are therefore more likely to remain if one of their large customers shuts down.

Welfare cost of cascades. We can also use the calibrated model to evaluate the welfare cost of cascades. Table 7 shows that the model generates a positive correlation between the size of a cascade and its negative impact on output. While the impact of the exit of an average firm is negligible, a cascade that originates from a high degree firm is responsible for a 2.3 percent drop in output on average. Firm with high out-degree—the star suppliers—have a disproportionate impact on aggregate output upon shutting down. Since they help to improve the productivity of many producers, their exits lower the aggregate productivity of the network substantially.

Table 7 also shows that the elasticity of substitution ε between intermediate inputs matters for the impact of the cascades on output. When these inputs are less substitutable, the exit of a high-degree firm creates bigger cascades (as seen in Figure 9) and these cascades have a larger



Notes: Cumulative number of exits at different distances from shuttered firm. "High degree" refers to the firms above the 99th percentile. Simulations of 100 randomly drawn matrices Ω, for each of which 1000 cascades are created.

Figure 9: Cumulative cascades by degree of originator, $\varepsilon = 3$.

negative effect on aggregate output.[44]

	Impact on output	
	Benchmark	$\varepsilon = 3$
Average firm	−0.1%	−0.1%
High in-degree	−1.8%	−3.2%
High out-degree	−2.4%	−3.7%
High degree	−2.3%	−3.7%

Notes: "High degree" refers to firms above the 99th percentile. Simulations of 100 randomly drawn matrices Ω, for each of which 1000 cascades are created.

Table 7: Correlation between output drop and firm degree

Cascades in the model and the data. In the data, many firms are simultaneously hit by (unobserved) shocks and multiple cascades might overlap, so that there is no straightforward way to use the exercises above to evaluate the empirical performance of the model. We can, however, use simple regressions to capture the impact of an exiting firm on its neighbors. Since these regressions can be run in the data and in the calibrated economy, they provide a good test of the model's ability to generate realistic cascades.

Specifically, I compute the fraction of each firm j's neighbors that exit in a given period and

[44]In empirical work, Barrot and Sauvagnat (2016) have found that, similarly, a shock to a producer of a highly specific input is more damaging to its customers.

regress that number on whether j itself shuts down. I run separate regressions for upstream and downstream neighbors at various distances from j. To be precise, denote by DX_{jdt} and UX_{jdt} the fraction of firm j's downstream (DX) and upstream (UX) neighbors located at a distance d that exit between t and $t+1$. I regress

$$\mathrm{DX}_{jdt} = \alpha^D + \beta_d^D \mathrm{Exit}_{jt} + \mathrm{Controls}_{jt} + \varepsilon_{jdt} \tag{15}$$

and

$$\mathrm{UX}_{jdt} = \alpha^U + \beta_d^U \mathrm{Exit}_{jt} + \mathrm{Controls}_{jt} + \varepsilon_{jdt} \tag{16}$$

where Exit_{jt} is an indicator variable that equals 1 if j exits between t and $t+1$ and 0 otherwise.[45] The coefficients β_d^D and β_d^U provide information about how cascades propagate in this economy. They capture the increase in shutdown probability associated with the exit of a neighboring firm located at a distance d.[46]

Figure 10 shows the coefficients estimated from the Factset data (green dashed lines). We see that the shutdown of a firm is associated with about a 10% increase in the probability that one of its direct suppliers or customers also exits. This number falls to about 2% for the second neighbors and keeps declining afterwards. The model (solid blue lines) is roughly able to match these patterns, suggesting that it broadly captures the joint operating decisions of nearby firms.[47,48]

While the shutdown of any firm has the potential to push a neighbor out of business, the exit of well-connected producers generally trigger larger cascades. To see this, we can first measure the size of a cascade as the total number of shutdowns, summed up to the second neighbors, associated with the exit of a firm, and then compare this statistic across firms with different numbers of neighbors.

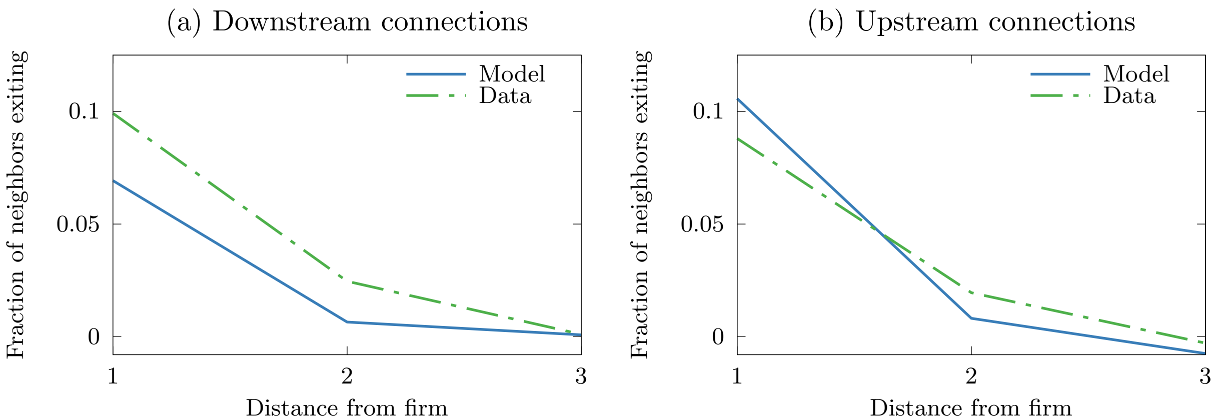
The results are presented in Table 8. The first column shows that, in the data, firms that are above the 95th percentile of the degree distribution are associated with cascades that are about three times larger than those associated with the average firm. High-degree firms are, however, less likely to actually shut down in response to shocks, as the third column shows. In the data, an average firm has a 11.8% chance of exiting in a given year, while this number drops to 2.0% for a

[45] In the data, I consider that a firm shuts down during the last year that it is in the sample. I use the SDC Platinum dataset to exclude mergers and acquisitions. The controls in (15) and (16) include time fixed effects as well as the in- and out-degree of firm j.

[46] Regressions (15) and (16) can suffer from endogeneity issues, such that β_d^D and β_d^U should not be interpreted as capturing a causal relationship. Nonetheless, they describe correlation patterns between the operating status of nearby firms and we can compare these patterns in the model and the data.

[47] One possibility is that the regressions (15) and (16) capture common shocks across firms instead of the propagation over the network. For instance, since trading partners are likely to be geographically close to each other, a local shock could directly affect both of them at the same time. To alleviate this concern, I run the same regressions on supplier/customer pairs located in different zip codes. Reassuringly, the results are essentially the same.

[48] These results, which suggest that shocks can propagate through the extensive margin of production, complement empirical work that documents that exogenous shocks such as natural catastrophes also propagate over input-output linkages through intensive margins of adjustment (Barrot and Sauvagnat, 2016; Carvalho et al., 2016).



Notes: Factset data. Estimated coefficients from regressing the fraction of exiting neighbors on whether a firm exits. Time fixed effects and in-degree and out-degree controls are included. The distance is the smallest number of connections between two firms.

Figure 10: Cascades of firm shutdowns in the model and in the data

high-degree firm.[49]

The model does well in terms of the size of the cascades and is also able to roughly replicate the exit probabilities. In the model, high-degree firms are particularly valuable to the planner and are therefore kept in operation even after severe shocks. When they do shut down, however, the planner reorganizes the whole cluster of producers that was built around them, which explains the large cascades that they trigger.

Table 8: High-degree firms are more resilient but create larger cascades

	Size of cascades		Probability of exit	
	Data	Model	Data	Model
Average firm	0.9	0.9	11.8%	16.6%
High-degree firm	3.1	3.1	2.0%	0.6%

Notes: "High degree firms" are above the 95th percentile of the degree distribution. "Size of cascades" is the sum of exiting firms up to the second neighbors downstream and upstream, computed by multiplying the regression coefficients in Figure 10 by the number of neighbors at the corresponding distance.

5.6 Aggregate fluctuations

There is a finite number of firms in the economy, so the idiosyncratic productivity shocks create aggregate fluctuations. Since these shocks also affect the production network, aggregate output is endogenously correlated with the structure of the network. We investigate that correlation in

[49] The elasticities of substitution play an important role in determining whether a firm shuts down after an adverse z shock. For instance, a smaller ε makes the network more rigid and firms are less likely to shut down. Alternatively, a higher elasticity σ, since it makes inputs in the consumption aggregator more substitutable, leads to a higher likelihood of firm exit.

this section, and we also consider how the endogenous reorganization of the network amplifies or dampens fluctuations in macroeconomic aggregates.[50]

There are no aggregate productivity shocks in the calibrated economy but one might think that these shocks can also affect the structure of the network and should therefore be included in the analysis. Fortunately, the following proposition shows that this is not the case: the network is completely unaffected by changes in aggregate productivity A. We can therefore abstract from aggregate shocks to explore the interaction between aggregate fluctuations and the shape of the network.[51,52]

Proposition 9. *If $\alpha_j = \alpha$ for all $j \in \mathcal{N}$, then the efficient network θ does not depend on aggregate productivity A.*

Comovements

Table 9 shows the correlations between aggregate output and the structure of the network in the calibrated economy and in the data. We see that in the model the exponent parameters of the degree distributions are negatively correlated with output, which indicates thicker right tails, and thus an abundance of well-connected firms, during expansions. The economy also features more clustering during booms. These correlations are similar in the data, although there are some discrepancies across datasets. The model is closest to the Factset data, which provides the most comprehensive link coverage, while the two Compustat datasets confirm that the correlations keep the same sign when computed over longer time horizons.

These patterns can be explained through the lens of the model. When well-positioned firms receive good shocks, the planner builds highly-connected clusters of suppliers and customers around them. As discussed before, these clusters are particularly productive, which generates the observed correlations between output, clustering and the degree distributions. Inversely, during recessions it might be too costly to organize these productive clusters—perhaps because a few critical firms face low z shocks. As a result, production is more dispersed and output is lower.

[50] In a richer model, firms could have to pay a cost to *change* their operating status. These costs would limit the reaction of the network in response to a temporary shock. They would also make the model dynamic, adding significant technical difficulties given the high dimensionality of the state space.

[51] The proof of Proposition 9 shows that we can write the planner's problem as

$$\max_{\theta \in \{0,1\}^n} A^{\frac{1}{1-\alpha}} Q \left(1 - \sum_{j \in \mathcal{N}} \theta_j f_j\right) L, \text{ where } q_j = z_j \theta_j \left(\sum_{i \in \mathcal{N}} \Omega_{ij} q_i^{\varepsilon_j - 1}\right)^{\frac{\alpha}{\varepsilon_j - 1}}, \forall j. \tag{17}$$

Since A only enters as a multiplicative constant in the objective function, it has no impact on the optimal θ and thus on the production network. (17) also shows that it is straightforward to compute the variance of output with aggregate shocks. If the variance of log output is x without fluctuations in A, then the overall variance of log output would simply be $x + (1-\alpha)^{-2} y$ if we added shocks with a variance of y to $\log(A)$.

[52] Appendix E.1 provides a version of the calibrated economy with a much larger number of firms and aggregate shocks. The results are similar.

Table 9: Correlation between the network and aggregate output

	Model	Datasets		
		Factset	Compustat	
			AHRS	CF
Power law exponents				
In-degree distribution	−0.59	−0.87	−0.35	−0.12
Out-degree distribution	−0.71	−0.97	−0.31	−0.11
Global clustering coefficient	0.54	0.76	0.18	0.11

Notes: All time series are in logs. In the data, output is annual real GDP. Output is detrended linearly in sample. Since there are only 13 years in the Factset data I use the CBO 10-year projection for real GDP growth at the beginning of the sample in 2003 (2.58%) to detrend the series.

Figure 11 provides a simple example of these mechanisms in an economy under two realizations of the idiosyncratic shocks z. In Panel (a), the realization of the shocks is such that the planner organizes a cluster of producers around firms 1 and 2. Firms in this cluster have many connections and take advantage of the gains from input variety to improve their productivity. As a result, the clustering coefficient is large, the degree distributions have thick tails, as seen in Panel (c), and the economy is booming. Panel (b) shows the same economy but under a realization of shocks such that operating a cluster around firms 1 and 2 is not efficient. In this case, economic activity moves to the outskirts of the network, where there are fewer connections between firms. As a result, the clustering coefficient is smaller, the degree distributions have thinner tails and the economy is in a downturn.

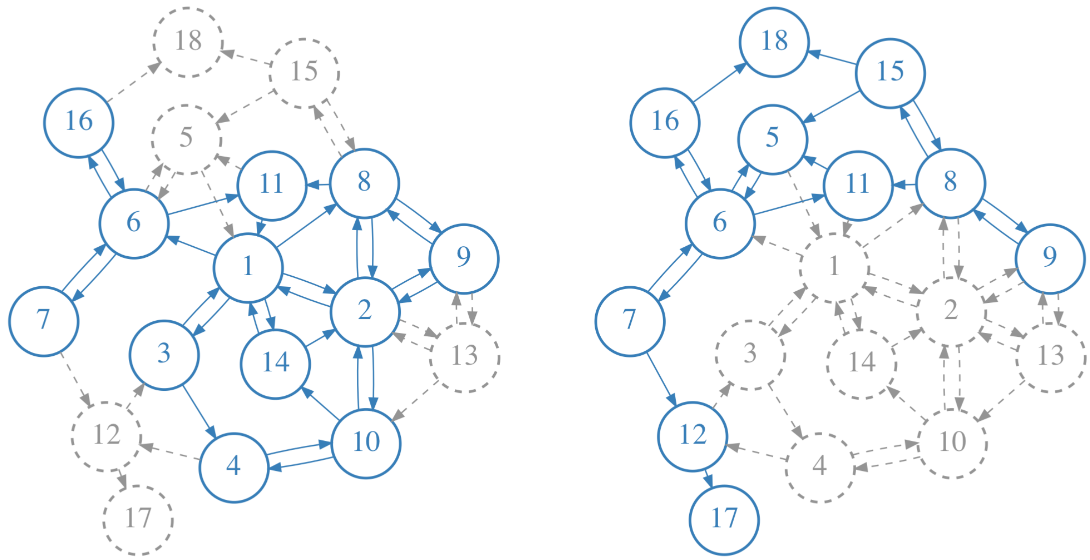
Level and volatility of output

We now consider the impact that the endogenous formation of the network has on the level and the volatility of aggregate output. To do so, it is useful to compare the efficient allocation, in which the network is constantly reorganized in response to shocks, to an alternative economy in which the network is designed efficiently in the first period but then kept completely fixed afterward. The differences between these two economies capture the role played by the endogenous response of the network to shocks.

I find large differences between these two economies. First, aggregate output is 10% lower when the network is kept fixed, which suggests that frictions that impede the reorganization of the network can have large welfare consequences.[53] Second, aggregate output is 20% more volatile when the network is fixed, which shows the importance of the endogenous evolution of the network for the aggregation of firm-level shocks into macroeconomic fluctuations.[54] To get a sense of magnitude

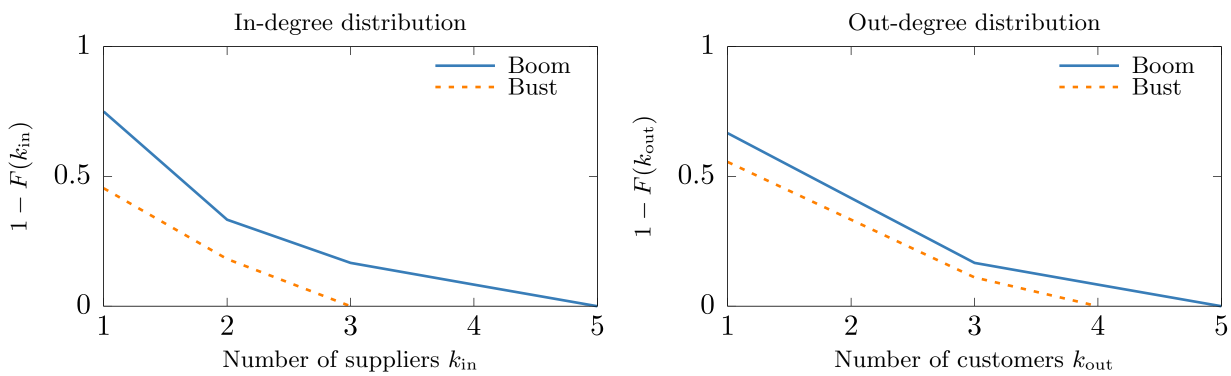
[53] Since the planner has more tools when the network is flexible it is not surprising that output is larger in this case. The magnitude of the increase is however informative of the importance of the forces involved.

[54] It is possible to find examples in which the flexible network economy is more volatile than its fixed network



(a) Boom: $C = 3.3$, clustering coefficient $= 0.37$

(b) Bust: $C = 2.9$, clustering coefficient $= 0.14$



(c) Degree distributions

Figure 11: Booms and busts in the production network

for this last number we can also consider an economy in which, in addition to the network itself, all the inputs of the firms are kept fixed. In this case output volatility increases by 40% compared to the flexible network benchmark. On its own, the endogenous formation of the network is therefore able to explain half of the reduction in volatility generated by all the adjustment margins together.

The elasticities of substitution σ and ε matter for the differences between the flexible and the fixed networks. As Table 10 shows, these differences become larger when ε and σ are large. In this case, the network is very flexible so that small shocks can create big movement. Preventing these movement can lead to a steep decline in output and a large increase in volatility. In contrast, when the elasticities are low, the network is very rigid so that the differences between the flexible and

counterpart. But for reasonable changes in the parameters of the calibration, however, the flexible network economy remains the least volatile.

fixed networks are minimal.[55]

Elasticities ε and σ	7	5	3
Change in mean C from fixed network	22%	10%	3%
Change in std C from fixed network	−30%	−20%	−1%

Table 10: Flexible vs fixed network under different elasticities

To understand how the reorganization of the network dampens fluctuations, it helps to think of the planner as choosing, for each productivity vector z, the best network θ out of 2^n possibilities. Total production $C(z)$ can therefore be written as

$$C(z) = \max_{k \in \{1,\ldots,2^n\}} C_k(z)$$

where $C_k(z)$ is the aggregate output produced by the kth network. By itself, each network k is associated with a probability distribution for C_k where the randomness comes from the underlying shocks z. The mean and the variance of these distributions vary with k but for the networks that are actually selected by the planner the differences are limited and the distributions overlap substantially with one another. Figure 12 provides an example with five potential networks $k \in \{1,\ldots,5\}$. Each line represents the PDF of aggregate output under a fixed network k. The figure also shows the output produced by each network under four different productivity vectors z. Each z is associated with a symbol whose locations on the graph represent the output $C_k(z)$ of network k under shock z. For instance, we see that network $k = 1$ performs poorly under the productivity shock ■, while network $k = 5$ performs very well. The symbols in blue indicate which network performs the best under a given z and is therefore chosen by the planner. We see that, for any fixed network, the PDFs are spread out and the variance of output is relatively large. In contrast, the output produced by the best network—the blue symbols in the right tails of the distributions—are close to each other indicating that the variance of output $C_{k^*}(z)$ under the efficient network $k^*(z) = \arg\max C_k(z)$ is relatively small.[56]

[55]The correlations in Table 9 stay roughly the same when the elasticities σ and ε change.

[56]This intuition is reminiscent of results from extreme value theory that show that the variance of the maximum of a large number m of independent normal random variables declines rapidly with m.

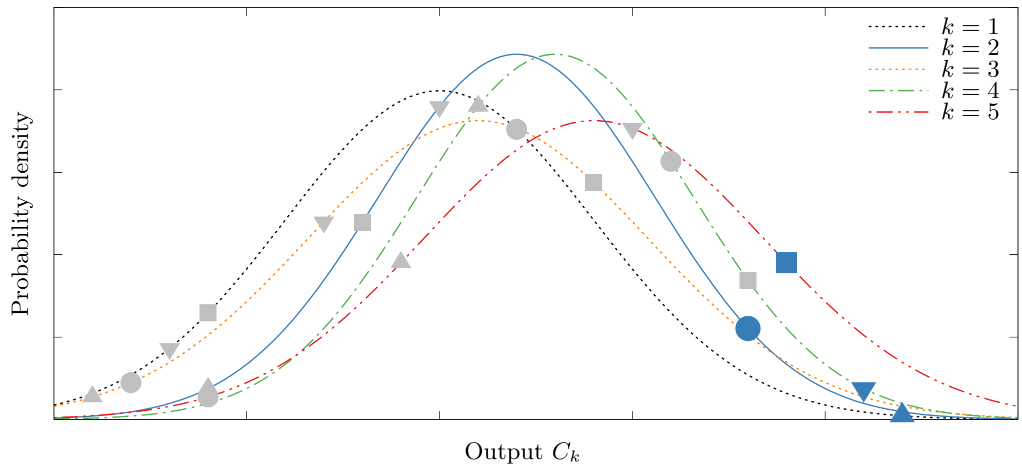


Figure 12: Distribution of output C_k under different networks k

6 Conclusion

This paper proposes a theory of network formation that operates through the firms' extensive margin of production. Because of a complementarity between firms' operating decisions, production tends to be organized in tightly connected clusters of producers centered around the most productive firms. The complementarities also give rise to cascades of firm shutdowns. As in the data, highly connected firms are more resilient to shocks but trigger larger cascades upon shutdown. The theory also predicts how the shape of the network comoves with aggregate output. In particular, expansions feature more clustering among firms as well as thicker tails for the degree distributions. These correlations are also present in U.S. data. Finally, the theory predicts that the optimal reorganization of the network is responsible for a substantial decline in the size of aggregate fluctuations. These findings highlight the importance of endogenizing the production network to better understand the origin of aggregate fluctuations.

One contribution of the paper is to provide a novel solution method to solve some non-convex optimization problems with discrete adjustment margins. Using this method, it is possible to quickly find the efficient allocation in economies that are particularly challenging to solve using standard techniques. This new tool could potentially be useful in a broad range of economic environments, including models with menu costs, or fixed costs of investment or hiring.

The model also opens the door to future research projects such as the study of industry clusters, like Silicon Valley for technology, and New York or London for finance. In particular, the theory could be used to quantify how well-connected, high-productivity firms affect other participants' incentives to join the cluster, and therefore shed light on how these clusters are created in the first place. Finally, the model could be extended to provide a quantitative framework that could guide policymakers when considering whether they should bail out distressed firms.

References

Acemoglu, D. and P. D. Azar (2018): "Endogenous Production Networks," Working Paper 24116, National Bureau of Economic Research.

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012): "The Network Origins of Aggregate Fluctuations," *Econometrica*, 80, 1977–2016.

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2015a): "Microeconomic Origins of Macroeconomic Tail Risks," Working Paper 20865, National Bureau of Economic Research.

Atalay, E. (2017): "How Important Are Sectoral Shocks?" *American Economic Journal: Macroeconomics*, 9, 254–80.

Atalay, E., A. Hortaçsu, J. Roberts, and C. Syverson (2011): "Network structure of production," *Proceedings of the National Academy of Sciences*, 108, 5199–5202.

Baqaee, D. R. (2018): "Cascading Failures in Production Networks," *Econometrica*, 86, 1819–1838.

Baqaee, D. R. and E. Farhi (2017a): "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem," Working Paper 23145, National Bureau of Economic Research.

——— (2017b): "Productivity and Misallocation in General Equilibrium." Working Paper 24007, National Bureau of Economic Research.

Barrot, J.-N. and J. Sauvagnat (2016): "Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks," *The Quarterly Journal of Economics*, 131, 1543–1592.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2013): "Cross-Country Differences in Productivity: The Role of Allocation and Selection," *American Economic Review*, 103, 305–34.

Bernard, A. B., A. Moxnes, and Y. U. Saito (2015): "Production Networks, Geography and Firm Performance," Working Paper 21082, National Bureau of Economic Research.

Bigio, S. and J. La'O (2016): "Financial Frictions in Production Networks," Working Paper 22212, National Bureau of Economic Research.

Boehm, J. and E. Oberfield (2018): "Misallocation in the Market for Inputs: Enforcement and the Organization of Production," Working Paper 24937, National Bureau of Economic Research.

Bresnahan, T. F. and V. A. Ramey (1994): "Output Fluctuations at the Plant Level," *The Quarterly Journal of Economics*, 593–624.

Broda, C. and D. E. Weinstein (2006): "Globalization and the Gains From Variety," *The Quarterly Journal of Economics*, 121, 541–585.

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2017a): "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy," Working paper.

Caliendo, L., F. Parro, and A. Tsyvinski (2017b): "Distortions and the Structure of the World Economy," Working Paper 23332, National Bureau of Economic Research.

Carvalho, V. M., , M. Nirei, and Y. Saito (2014): "Supply Chain Disruptions: Evidence from the Great East Japan Earthquake," Working Paper 14-E-035, The Research Institute of Economy, Trade and Industry.

Carvalho, V. M. (2014): "From Micro to Macro via Production Networks," *Journal of Economic Perspectives*, 28, 23–48.

Carvalho, V. M., M. Nirei, Y. U. Saito, and A. Tahbaz-Salehi (2016): "Supply Chain Disruptions: Evidence from the Great East Japan Earthquake," Working paper.

Carvalho, V. M. and A. Tahbaz-Salehi (2018): "Production networks: A primer," .

Carvalho, V. M. and N. Voigtländer (2015): "Input Diffusion and the Evolution of Production Networks," Working Paper 20025, National Bureau of Economic Research.

Cohen, L. and A. Frazzini (2008): "Economic Links and Predictable Returns," *The Journal of Finance*, 63, 1977–2011.

di Giovanni, J., A. A. Levchenko, and I. Mejean (2014): "Firms, Destinations, and Aggregate Fluctuations," *Econometrica*, 82, 1303–1340.

Dixit, A. K. and J. E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," *The American Economic Review*, 67, pp. 297–308.

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

Foerster, A. T., P.-D. G. Sarte, and M. W. Watson (2011): "Sectoral versus Aggregate Shocks: A Structural Factor Analysis of Industrial Production," *Journal of Political Economy*, 119, 1–38.

Foster, L., J. Haltiwanger, and C. Syverson (2008): "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *The American Economic Review*, 98, 394–425.

Gabaix, X. (2011): "The Granular Origins of Aggregate Fluctuations," *Econometrica*, 79, 733–772.

Gabaix, X. and R. Ibragimov (2011): "Rank - 1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents," *Journal of Business & Economic Statistics*, 29, 24–39.

Garey, M. R. and D. S. Johnson (1990): *Computers and Intractability; A Guide to the Theory of NP-Completeness*, New York, NY, USA: W. H. Freeman & Co.

Grassi, B. (2017): "IO in I-O: Size, Industrial Organization and the Input-Output Network Make a Firm Structurally Important," Working paper.

Hatfield, J. W., S. D. Kominers, A. Nichifor, M. Ostrovsky, and A. Westkamp (2013): "Stability and Competitive Equilibrium in Trading Networks," *Journal of Political Economy*, 121, 966–1005.

Hendricks, K. B. and V. R. Singhal (2005): "An Empirical Analysis of the Effect of Supply Chain Disruptions on Long-Run Stock Price Performance and Equity Risk of the Firm," *Production and Operations Management*, 14, 35–52.

Horvath, M. (1998): "Cyclicality and Sectoral Linkages: Aggregate Fluctuations from Independent Sectoral Shocks," *Review of Economic Dynamics*, 1, 781 – 808.

Jackson, M. O. and A. Wolinsky (1996): "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71, 44 – 74.

Jones, C. I. (2011): "Intermediate Goods and Weak Links in the Theory of Economic Development," *American Economic Journal: Macroeconomics*, 3, 1–28.

Jorgenson, D. W., F. Gollop, and B. Fraumeni (1987): *Productivity and U.S. Economic Growth*, Harvard University Press, reprinted- Universe, 1999.

Kennan, J. (2001): "Uniqueness of Positive Fixed Points for Increasing Concave Functions on R^n: An Elementary Result," *Review of Economic Dynamics*, 4, 893 – 899.

König, M. D., A. Levchenko, T. Rogers, and F. Zilibotti (2018): "Aggregate Fluctuations in Adaptive Production Networks," Working paper.

Li, D. and X. Sun (2006): *Nonlinear integer programming*, vol. 84, Springer Science.

Lim, K. (2018): "Endogenous Production Networks and the Business Cycle," working paper.

Long, J. B. and C. I. Plosser (1983): "Real Business Cycles," *Journal of Political Economy*, 91, 39–69.

Murray, W. and K.-M. Ng (2010): "An algorithm for nonlinear optimization problems with binary variables," *Computational Optimization and Applications*, 47, 257–288.

Oberfield, E. (2018): "A Theory of Input-Output Architecture," *Econometrica*, 86, 559–589.

Ostroumova Prokhorenkova, L. and E. Samosvat (2014): *Global Clustering Coefficient in Scale-Free Networks*, Cham: Springer International Publishing, 47–58.

Ozdagli, A. and M. Weber (2017): "Monetary Policy through Production Networks: Evidence from the Stock Market," Working paper.

Tintelnot, F., A. K. Kikkawa, M. Mogstad, and E. Dhyne (2018): "Trade and Domestic Production Networks," Working Paper 25120, National Bureau of Economic Research.

Wagner, S. M. and C. Bode (2008): "An Empirical Examination of Supply Chain Performance Along Several Dimensions of Risk," *Journal of Business Logistics*, 29, 307–325.

Zurich Insurance Group (2015): "Potential business impact of loss of the main supplier for small and medium enterprises (SMEs) in 2015," .

Appendix

A Efficient allocation for a given θ

This appendix shows how to derive all the quantities in the efficient allocation for a given network θ. Lemma 1 already gives the expression for C. From an intermediate step in its proof, we can find the efficient allocation for labor l. Define Γ as the matrix with elements

$$\Gamma_{jk} = \frac{\alpha_k \Omega_{jk} q_j^{\varepsilon_k - 1}}{\sum_{i \in \mathcal{N}} \Omega_{ik} q_i^{\varepsilon_k - 1}}.$$

Then, (25) defines a linear system in l. Inverting this system yields

$$l = \left[(I_n - \Gamma) \operatorname{diag} \left(\frac{1}{1-\alpha} \right) \right]^{-1} \left(\beta \circ \left(\frac{q}{Q} \right)^{\circ(\sigma - 1)} \frac{C}{Q} \right) \tag{18}$$

where β denotes the vector of intensities β_j, $\operatorname{diag}\left(\frac{1}{1-\alpha}\right)$ denotes a matrix with the vector $\frac{1}{1-\alpha}$ on the diagonal and $\circ$ denotes Hadamard (element-by-element) operations such that $(A \circ B)_{ij} = (A)_{ij} \circ (B)_{ij}$ and $\left(A^{\circ B}\right)_{ij} = (A)_{ij}^{(B)_{ij}}$ for arbitrary matrices A and B. We q_j and l_j in hand, we can use the first-order condition (20) we can then write y_j as

$$y_j = (1 - \alpha_j)^{-1} q_j l_j.$$

Finally, we can rewrite the first-order condition (22) as

$$x_{ij} = \left(\frac{q_i}{q_j} \right)^{\varepsilon_j} \alpha_j \left(A z_j \theta_j \left(\frac{1}{q_j} \right)^{1-\alpha_j} \right)^{\frac{\varepsilon_j - 1}{\alpha_j}} \Omega_{ij} y_j$$

when $q_i, q_j > 0$ and $x_{ij} = 0$ otherwise.

B Additional numerical tests

This appendix provides the details of the numerical simulations of Section 3.4 as well as several additional exercises to show the robustness of the solution approach.

B.1 Details of the simulations of Table 1

The numerical simulations of Table 1 involve a large number of economies that are generated randomly from a broad set of parameters.

Aggregate parameters. The aggregate parameters are selected from: $n \in \{8, 10, 12, 14\}$ for the number of firms and $\sigma \in \{4, 6, 8\}$ for the elasticity of substitution in the consumption aggregator. The matrix Ω is such that each firm has on average 3, 4, ... up to n *potential* incoming connections (non-zero Ω_{ij}).[57] I restrict the matrices Ω to have empty diagonals, as in the data. Each non-zero element in Ω is drawn from $\Omega_{ij} \sim \text{iid } U([0, 1])$. Appendix D.1 describes the precise algorithm used to build Ω.

Firm-level parameters. The firm-level parameters are drawn from: $\log(z_k) \sim \text{iid } \mathcal{N}(0, 0.25^2)$ for the productivities, $f_j \sim \text{iid } U([0, 0.2/n])$ for the fixed costs, $\alpha_j \sim \text{iid } U([0.25, 0.75])$ for the intermediate input shares, $\varepsilon_j \sim \text{iid } U([4, 8])$ for the elasticities of substitutions between intermediate inputs, and $\beta_j \sim \text{iid } U([0, 1])$ for the factor intensities in the production of the final good.

Procedure. For every potential combination of the aggregate parameters I simulate 500 economies. In each case, the matrix Ω and the individual characteristics of the firms are drawn from the distributions described above. I then use the exhaustive search algorithm described in Appendix D.2 to compute the true solution to $\mathcal{P}$. I also use the algorithm of Appendix D.4 to compute solutions to the reshaped and non-reshaped versions of the planner's problem. These two solutions are then compared to the true solution and the results are reported in Table 1. I exclude from the simulations pathological cases in which the algorithms find an aggregate consumption of 0.[58] For the benchmark tests, an economy is kept in the sample only if the first-order conditions of the reshaped problem yield a solution in $\{0, 1\}^n$. Appendix B.3 shows that the algorithm performs well when these simulations are kept in the sample.

B.2 Additional measure of success

Table 1, in the main text, presents the fraction of firms with the correct operating status θ_j. One can also use the fraction of economies in which the full vector θ is correct as an alternative measure of success. Table 11 provides that information. We see that with reshaping the algorithm finds the correct vector θ 99.7% of the time on average, which is much better than without reshaping, where the equivalent number is 19.9%.

The table also shows that the performance of the reshaping method degrades slightly as n increases. This is not surprising since the likelihood that any one firm will be assigned the wrong status increases with n. Since the importance of any one firm for aggregate outcomes also declines with n, the error in C remains small.

[57] The corresponding average numbers of *active* incoming connections are 2.1, 3.0, 3.8, 4.5, 5.3 and 5.8, respectively. See Appendix B.4 below for tests on very sparse matrices Ω.

[58] This happens, for instance, when Ω is so sparse that a closed-loop of suppliers does not exist.

Table 11: Using the full vector θ as a measure of success

	With reshaping		Without reshaping	
n	Correct vector θ	Error in C	Correct vector θ	Error in C
8	99.8%	0.001%	32.3%	0.791%
10	99.7%	0.001%	21.6%	0.855%
12	99.7%	0.001%	15.3%	0.903%
14	99.6%	0.001%	10.6%	0.926%

Notes: Same simulations as Table 1. See Section B.1 in the Appendix for details.

B.3 When the the solution to the reshaped problem is not in $\{0,1\}^n$

The results presented in Table 1 exclude economies in which $\mathcal{R}$'s first-order conditions are such that $0 < \theta_j < 1$ for at least one firm j, which happens in less than a tenth of the simulations. This section shows that, even when these cases are not excluded from the sample, the solution approach performs well. To see this, Table 12 shows the outcome of the same simulations as Table 1 but without excluding the economies such that $\theta \notin \{0,1\}^n$. We see that on average the error in aggregate output C is less than 0.008% and about 99.7% of firms are assigned the correct status. In contrast, without reshaping the average error in output is 0.867%—108 times more.[59]

Table 12: Testing the reshaping on small networks without exclusions

	With reshaping		Without reshaping	
n	Correct θ	Error in C	Correct θ	Error in C
8	99.7%	0.011%	86.5%	0.789%
10	99.7%	0.006%	85.3%	0.852%
12	99.7%	0.007%	84.6%	0.901%
14	99.7%	0.007%	84.0%	0.926%

Notes: Same simulations as Table 1 but without excluding economies such that $\theta \notin \{0,1\}^n$. See Section B.1 in the Appendix for details.

A similar exercise can be done for economies with a large number of firms. This exercise is analogous to that of Table 3 and is presented in Table 13. We see that even when the first-order conditions of $\mathcal{R}$ yield a solution $\theta \notin \{0,1\}^n$, the error in aggregate output is negligible. This last test suggests that the solution approach works particularly well in realistic economies with a large number of firms.

[59]Note that we can see whether $\theta \in \{0,1\}^n$ or not when solving the problem. If extreme precision is needed, we can be extra careful when $\theta \notin \{0,1\}^n$ and use additional tests (for instance we can look for beneficial deviations) to check the robustness of the solution.

Table 13: Testing the reshaping on large networks without exclusions

	With reshaping		Without reshaping	
n	Correct θ	Error in C	Correct θ	Error in C
1000	$> 99.9\%$	$< 0.001\%$	68.9%	0.58%

Notes: Same as Table 3 except all simulations are kept in the sample. $x < 0.001\%$ indicates that $x > 0$ but that proper rounding would yield 0, and similarly for $x > 99.9\%$.

B.4 Performance with very sparse matrices Ω

Table 14 shows the same simulations as Table 1 but with matrices Ω that are drawn so that firms have only 1 or 2 *potential* incoming connections on average. As a result, the networks of potential connections described by the matrices Ω are extremely sparse. The algorithm still performs well, with an average error in aggregate output that is 66 times smaller than when the problem is not reshaped.

Table 14: Testing the reshaping on sparse networks Ω

	With reshaping		Without reshaping	
n	Correct θ	Error in C	Correct θ	Error in C
8	99.6%	0.008%	91.0%	0.508%
10	99.5%	0.007%	91.1%	0.496%
12	99.6%	0.007%	90.8%	0.503%
14	99.5%	0.009%	90.5%	0.531%

Notes: Same simulations as Table 1 but with matrices Ω in which firms have on average only 1 or 2 potential connections.

B.5 Formation of the network link by link

This appendix provides the results of two exercises that show that reshaping the planner's problem is also useful when the production network is constructed link by link instead of through the extensive margin of the firms. In both exercises, the economy contains m real firms that are always active ($f_j = 0$).[60] Any two of these real firms are connected to each other by a link: for any ordered pair of real firms i, j with $i \neq j$, there exists a "link firm" k such that $\Omega_{ik} > 0$ and $\Omega_{kj} > 0$). There are no other connections in Ω. These link firms will operate or not as a function of economic conditions.

[60] This assumption is made to focus on the link formation aspect of the problem. I have experimented with economies in which $f_j > 0$ for the real firms and the results are similar.

Individual link formation in small networks

When the number m of real firms is small, we can use the same approach as in Section 3.4 and find the true solution to the planner's problem by comparing the welfare provided by each possible network θ (see algorithm in Appendix D.2). There are at most $m(m-1)$ links in an economy, in which case the utility provided by $2^{m(m-1)}$ networks must be compared. Since this quantity grows rapidly with m, Table 15 shows the results of these tests when there are only $m \in \{3,4,5\}$ real firms. As before, the outcome of this exhaustive search is compared to the allocation found by reshaping the planner's problem.

We see from Table 15 that the reshaping algorithm works well. Over all the simulations, more than 99.7% of the links are assigned the proper operating status θ and the errors in aggregate output are small. Without reshaping, large fractions of the links are assigned the wrong operating status and the error in aggregate output can be sizable.

Table 15: Individual links formation with few firms

Number of firms		With reshaping		Without reshaping	
Real firms m	Link firms $n-m$	Correct θ	Error in C	Correct θ	Error in C
3	6	99.9%	0.001%	90.9%	0.25%
4	12	99.8%	0.004%	85.9%	0.39%
5	20	99.7%	0.004%	82.0%	0.52%

Notes: Real firms: $f_j = 0$, $\alpha_j = 0.5$, $\sigma = \varepsilon_j = 6$, $\sigma_z = 0.25$. Link firms: $f_{link} \sim$ iid $U([0.0, 0.1/n])$, $\alpha_{link} \sim$ iid $U([0.5, 1.0])$ and $\sigma_{z_{link}} = 0.25$. For simplicity all non-zero Ω_{ij} are set to 1. For each m, 500 economies are generated randomly and the algorithm of Section D.4 is used to solve the planner's problem. An economy is kept in the sample only if the first-order conditions converge to a point in $\{0,1\}^n$. More than 80% of the economies are kept in the sample.

Individual link formation in large networks

For economies with a large number of firms, the true solution to the planner's problem is unknown but we can check whether there exist welfare-improving deviations from the allocation found using the reshaped problem. The procedure is the same as in the Section 3.4. The parameters of the tests are the same as in Table 15 but the economies feature $m \in \{10, 25, 40\}$ real firms and $n \in \{100, 625, 1600\}$ total firms (real plus links). The results are presented in Table 16. Reshaping the planner's problem yields solutions with few welfare-improving deviations so that the vast majority of links are assigned the correct status and the errors in aggregate output are negligible. In contrast, a large fraction of the links are assigned the wrong status and the errors in aggregate output are significant when the problem is not reshaped.

One potential concern of using the reshaping method in this context is that the first-order conditions often converge on a vector θ such that $\theta_j \notin \{0,1\}$ for at least one firm.[61] There are two

[61]In the simulations of Table 16, the first-order conditions converge to a point $\theta_j \in \{0,1\}$ *for all* j in 43% of the

Table 16: Individual links formation with a large number of firms

Number of firms		With reshaping		Without reshaping	
Real firms m	Link firms $n-m$	Correct θ	Error in C	Correct θ	Error in C
10	90	99.9%	0.002%	76.8%	0.66%
25	600	$>$ 99.9%	$<$ 0.001%	74.0%	0.73%
40	1560	$>$ 99.9%	$<$ 0.001%	73.4%	0.74%

Notes: The parameters of these tests, except for m, are as in Table 15. An economy is kept in the sample only if the first-order conditions converge to a point in $\{0,1\}^n$. $x < 0.001\%$ indicates that $x > 0$ but that proper rounding would yield 0, and similarly for $x > 99.9\%$. For each m, 500 economies are generated randomly and the algorithm of Section D.4 is used to solve the planner's problem.

reasons for this. First, as the total number of firms increases (up to $n = 1600$ for the economies with $m = 40$) it's more likely that at least one firm ends up with $\theta_j \notin \{0,1\}$. Second, the matrices Ω considered here are extremely sparse. As a result, the forces pushing the first-order conditions to hit the bounds are weakened (see footnote 23). In practice, however, these issues have limited implications. Only a small fraction of the links end up away from the $\{0,1\}$ bounds, and their impact on aggregate output is minimal. Table 17 shows the outcome of the same simulations but without excluding any simulations. We see that the results are essentially unchanged and that the solution approach also performs well in these situations.

Table 17: Individual links formation with a large number of firms and without exclusions

Number of firms		With reshaping		Without reshaping	
Real firms m	Link firms $n-m$	Correct θ	Error in C	Correct θ	Error in C
10	90	99.8%	0.004%	76.9%	0.65%
25	600	$>$ 99.9%	$<$ 0.001%	74.3%	0.71%
40	1560	$>$ 99.9%	$<$ 0.001%	73.4%	0.73%

Notes: The parameters of these tests are the same as in Table 15. No economies are excluded from the sample. $x < 0.001\%$ indicates that $x > 0$ but that proper rounding would yield 0, and similarly for $x > 99.9\%$. For each m, 500 economies are generated randomly and the algorithm of Section D.4 is used to solve the planner's problem.

C Proofs

This section contains the proofs.

C.1 Equilibrium

The following lemma is useful to show that every stable equilibrium is efficient.

simulations for $m = 10$, 12% for $m = 25$ and 5% for $m = 40$.

Lemma 2. *Let $X, Y \subset \mathbb{R}^n$. Define Problem A as*

$$\sup_{x \in X, y \in Y} f(x, y) \;\; \textit{subject to} \;\; \sum_j y_j \leq 0$$

and Problem B as

$$\sup_{x \in X, y \in Y} g(f(x, y)) - \lambda \left(\sum_j y_j \right)$$

where g is a strictly increasing function and where λ is such that $\sum_j y_j = 0$ at any solution. Suppose that for any solution to Problem A the constraint binds, then Problems A and B have the same solutions.

Proof. Take a point $\left(x^A, y^A\right)$ that solves Problem A and such that, since the constraint binds, $\sum_j y_j^A = 0$. Towards a contradiction, suppose $\left(x^A, y^A\right)$ does not solve Problem B. Then there is another point $(\tilde{x}, \tilde{y})$ such that $\sum_j \tilde{y}_j = 0$ (by the definition of λ) and such that $g(f(\tilde{x}, \tilde{y})) - \lambda\left(\sum_j \tilde{y}_j\right) > g\left(f\left(x^A, y^A\right)\right) - \lambda\left(\sum_j y_j^A\right)$. Since g is strictly increasing this implies that $f(\tilde{x}, \tilde{y}) > f\left(x^A, y^A\right)$ but, since $(\tilde{x}, \tilde{y})$ is in the feasible set of Problem A, this implies that $\left(x^A, y^A\right)$ was not a solution to Problem A, which is a contradiction. Conversely, take a point $\left(x^B, y^B\right)$ that solves Problem B. Then by the definition of λ it must be that $\sum_j y_j^B = 0$. Towards a contradiction, suppose $\left(x^B, y^B\right)$ does not solve Problem A. Then there is another point $(\tilde{x}, \tilde{y})$ such that $\sum_j \tilde{y}_j = 0$ (since the constraint in Problem A binds at the optimum) and such that $f(\tilde{x}, \tilde{y}) > f\left(x^B, y^B\right)$. Since g is strictly increasing this implies that $g(f(\tilde{x}, \tilde{y})) - \lambda\left(\sum_j \tilde{y}_j\right) > g\left(f\left(x^B, y^B\right)\right) - \lambda\left(\sum_j y_j^B\right)$ so that $\left(x^B, y^B\right)$ is not a solution to Problem B, which is a contradiction. □

Proposition 1. Every stable equilibrium is efficient.

Proof. The proof proceeds by establishing restrictions that any stable equilibrium must satisfy. It then shows that any allocation that satisfies these restrictions must be efficient.

Consider a coalition made of all the firms in the economy. For the equilibrium to be stable there cannot be an alternative arrangement that would yield larger aggregate profits. Otherwise, transfers could be designed to make one firm better off while keeping the other firms at the same profit level. The arrangement $\{x_{ij}, T_{ij}\}_{i,j}$ must therefore maximize $\sum_{j \in \mathcal{N}} \pi_j$. But, by the definition of an equilibrium, this maximization is subject to the behavior of the firms. Any equilibrium allocation therefore solves

$$\max_{\{x_{ij}, T_{ij}\}_{i,j}} \sum_{j \in \mathcal{N}} \left\{ \max_{\{p_j, c_j, l_j, \theta_j\}} \pi_j \left(p_j, c_j, l_j, \theta_j, \{x_{ij}\}_{ij} \right) \text{ s.t. } (4) \text{ and } (5) \right\}. \tag{19}$$

It is, however, equivalent to let the coalition itself directly optimize over $\{p_j, c_j, l_j, \theta_j\}_j$. To see this, notice that, conditional on the arrangement, the inner maximization problems in (19) are all independent from each other. In other words, the decisions of a firm i have no effect on the profit of a firm j as long as the contracts specified by the arrangement are fulfilled. As a result, we can write (19) as $\max_{\{x_{ij}\}_{ij},\{c_j,l_j,\theta_j\}_j} \sum_{j\in\mathcal{N}} \pi_j$ subject to the constraints (4) and (5) for all firms. By including the household's demand curves directly in the objective function, and by using the definition of π_j, the absence of dominating deviations therefore implies that the allocation must solve

$$\max_{\{x_{ij}\}_{ij},\{c_j,l_j,\theta_j\}_j} C^{\frac{1}{\sigma}} P \sum_{j\in\mathcal{N}} \beta_j^{\frac{1}{\sigma}} c_j^{\frac{\sigma-1}{\sigma}} - w \sum_{j\in\mathcal{N}} (l_j + \theta_j f_j L)$$

subject to (4) for all $j \in \mathcal{N}$, and where C and P are taken as given. Now, by Lemma 2 this problem is equivalent to an alternative problem in which the coalition maximizes $\left(\sum_{j\in\mathcal{N}} \beta_j^{\frac{1}{\sigma}} c_j^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$ subject to $\sum_j l_j + \theta_j f_j L \leq L$ in addition to the other constraints.[62] This reformulated problem is identical to the problem $\mathcal{P}$ of the social planner such that any stable equilibrium must be efficient. □

C.2 Results about q

The proof of the uniqueness of q relies on the following definitions from Kennan (2001).

Definition. A function $g : \mathbb{R}^n \to \mathbb{R}^n$ is radially quasiconcave ("R-concave") if $g(x) = 0$ and $x > 0$ and $0 \leq \lambda \leq 1$ implies $g(\lambda x) \geq 0$. If (in addition) $0 < \lambda < 1$ implies $g(\lambda x) > 0$, then g is strictly R-concave.

Definition. A function $g = (g_1, g_2, \ldots, g_n) : \mathbb{R}^n \to \mathbb{R}^n$ is quasi-increasing if $y_i = x_i$ and $y_j \geq x_j$ for all j implies $g_i(y) \geq g_i(x)$.

The following Lemma is used as an intermediate step to prove the uniqueness of the vector q.

Lemma 3. *Denote by $\tilde{\mathcal{N}}$ any subset of $\mathcal{N}$ with $\tilde{n}$ firms and such that $\sum_{i\in\tilde{\mathcal{N}}} \Omega_{ij} > 0$ for all $j \in \tilde{\mathcal{N}}$. The function $g : \mathbb{R}^{\tilde{n}} \to \mathbb{R}^{\tilde{n}}$ defined, for all $j \in \tilde{\mathcal{N}}$, as*

$$g_j(p) = (z_j A)^{\varepsilon_j} \left(\sum_{i\in\tilde{\mathcal{N}}} \Omega_{ij} p_i^{\frac{\varepsilon_j - 1}{\varepsilon_j}}\right)^{\alpha_j \frac{\varepsilon_j}{\varepsilon_j - 1}} - p_j$$

is strictly R-concave.

[62] The corresponding function g is $g(x) = x^{\frac{\sigma}{\sigma-1}}$. The constraints (4) can be included directly in the function f in Lemma 2 by setting $f = -\infty$ for points outside the constraint set.

Proof. Suppose that there exists a $p^* > 0$ such that $g(p^*) = 0$. Then, for $0 \leq \lambda \leq 1$,

$$\begin{aligned}
g_j(\lambda p^*) &= \lambda^{\alpha_j} (z_j A)^{\varepsilon_j} \left(\sum_{i \in \tilde{\mathcal{N}}} \Omega_{ij} (p_i^*)^{\frac{\varepsilon_j - 1}{\varepsilon_j}} \right)^{\alpha_j \frac{\varepsilon_j}{\varepsilon_j - 1}} - \lambda p_j^* \\
&\geq \lambda (z_j A)^{\varepsilon_j} \left(\sum_{i \in \tilde{\mathcal{N}}} \Omega_{ij} (p_i^*)^{\frac{\varepsilon_j - 1}{\varepsilon_j}} \right)^{\alpha_j \frac{\varepsilon_j}{\varepsilon_j - 1}} - \lambda p_j^* \\
&\geq \lambda g_j(p^*) \\
&\geq 0
\end{aligned}$$

where the first inequality is strict for $0 < \lambda < 1$ since $0 < \alpha_j < 1$ and $\sum_{i \in \tilde{\mathcal{N}}} \Omega_{ij} > 0$ by assumption. □

Proposition 2. In the efficient allocation, the productivity vector q satisfies, for all $j \in \mathcal{N}$,

$$q_j = z_j \theta_j A \left(\sum_{i \in \mathcal{N}} \Omega_{ij} q_i^{\varepsilon_j - 1} \right)^{\frac{\alpha_j}{\varepsilon_j - 1}}. \tag{9}$$

Furthermore, there is a unique q that solves (9) and such that $q_j > 0$ if j operates and has access to an operating cycle, and $q_j = 0$ otherwise.

Proof. We focus on the firms that have access to an operating cycle, as defined in Section 3, since the problem of the other firm is trivial. The first-order conditions of $\mathcal{P}$ with respect to l_j and x_{ij} are

$$w l_j = \lambda_j (1 - \alpha_j) y_j \tag{20}$$

$$\lambda_i x_{ij} = \lambda_j \frac{A}{\alpha_j^{\alpha_j} (1 - \alpha_j)^{1 - \alpha_j}} \alpha_j z_j \theta_j \left(\sum_{k \in \mathcal{N}} \Omega_{kj}^{\frac{1}{\varepsilon_j}} x_{kj}^{\frac{\varepsilon_j - 1}{\varepsilon_j}} \right)^{\frac{\varepsilon_j}{\varepsilon_j - 1} \alpha_j - 1} l_j^{1 - \alpha_j} \Omega_{ij}^{\frac{1}{\varepsilon_j}} x_{ij}^{\frac{\varepsilon_j - 1}{\varepsilon_j}}. \tag{21}$$

when firms i and j have access to an operating cycle.[63] Combining these conditions with the production function yields

$$x_{ij} \lambda_i^{\varepsilon_j} = \lambda_j^{\varepsilon_j} \alpha_j \left(A z_j \theta_j \left(\frac{\lambda_j}{w} \right)^{1 - \alpha_j} \right)^{\frac{\varepsilon_j - 1}{\alpha_j}} \Omega_{ij} y_j. \tag{22}$$

Plugging (20) and (22) back in the production function we find (9).

I follow Kennan (2001) to prove the uniqueness of q. Consider the change of variable $p_j = q_j^{\varepsilon_j}$,

[63] A firm j without access to an operating cycle cannot produce and the planner simply sets $l_j = 0$, $y_j = 0$, and $x_{ij} = 0$ and $x_{ji} = 0$ for all i.

and let $\tilde{\mathcal{N}}$ be the set of firms that operate and that have access to an operating cycle. Denote the number of such firms by $\tilde{n}$. Clearly, $p_j = 0$ for $j \notin \tilde{\mathcal{N}}$. We can rewrite (9) as the following mapping from $\mathbb{R}^{\tilde{n}}$ to $\mathbb{R}^{\tilde{n}}$:

$$p_j = (z_j A)^{\varepsilon_j} \left(\sum_{i \in \tilde{\mathcal{N}}} \Omega_{ij} p_i^{\frac{\varepsilon_j - 1}{\varepsilon_j}} \right)^{\alpha_j \frac{\varepsilon_j}{\varepsilon_j - 1}}, \tag{23}$$

for all $j \in \tilde{\mathcal{N}}$. Denote the right-hand side of (23) by $f_j(p)$ and define $g : \mathbb{R}^{\tilde{n}} \to \mathbb{R}^{\tilde{n}}$ as $g(p) = f(p) - p$. A solution to (23) is therefore a vector p such that $g(p) = 0$. By Lemma 3, g is strictly R-concave. Note also that g is quasi-increasing.

Consider the mapping $h : \mathbb{R} \to \mathbb{R}^{\tilde{n}}$ defined as $h(s) = f(\mathbb{1}_{\tilde{n}} s)$ where $\mathbb{1}_{\tilde{n}}$ is the all-one vector of size $\tilde{n}$. Then $h(s)$ is strictly concave, strictly increasing and differentiable with $h(0) = 0$, $\lim_{s \to 0} h'(s) = \infty$ and $\lim_{s \to \infty} h'(s) = 0$, in all dimensions.[64] As a result, there exist constants $\bar{p} > \underline{p} > 0$ such that $h(\underline{p}) > \underline{p}$ and $h(\bar{p}) < \bar{p}$. Then, Theorems 3.1 and 3.2 in Kennan (2001) apply: (23) has a unique positive fixed point p^* and there is therefore a unique positive q^* that satisfies (9). It is such that $q_j^* = \left(p_j^*\right)^{\frac{1}{\varepsilon_j}}$ if j operates and has access to an operating cycle, and $q_j^* = 0$ otherwise. Note that the proof is essentially unchanged if we use the reshaped equation (12) instead of (9). $\square$

C.3 Taking θ as fixed

Lemma 1. In the efficient allocation, aggregate output is

$$C = Q \left(1 - \sum_{j \in \mathcal{N}} \theta_j f_j \right) L \tag{11}$$

where $Q = \left(\sum_{j \in \mathcal{N}} \beta_j q_j^{\sigma - 1} \right)^{\frac{1}{\sigma - 1}}$.

Proof. The first-order condition of $\mathcal{P}$ with respect to c_j is

$$c_j = \beta_j \left(\frac{q_j}{w} \right)^{\sigma} C. \tag{24}$$

Raising both sides to the power $\frac{\sigma - 1}{\sigma}$ and summing across j's yields $w = Q$. Using the first-order conditions (20), (22) and (24) into the resource constraints (7), we find

$$0 \geq \beta_j \left(\frac{q_j}{Q} \right)^{\sigma - 1} \frac{C}{Q} + \sum_{k \in \mathcal{N}} \alpha_k \frac{\Omega_{jk} q_j^{\varepsilon_k - 1}}{\sum_{i \in \mathcal{N}} \Omega_{ik} q_i^{\varepsilon_k - 1}} \frac{l_k}{1 - \alpha_k} - \frac{l_j}{1 - \alpha_j} \tag{25}$$

for all $j \in \mathcal{N}$. Summing across j's and simplifying yields (11). Note that once q is known we can

[64] Note that $\sum_{i \in \tilde{\mathcal{N}}} \Omega_{ij}$ must be strictly positive for all j, otherwise that j is not part of an operating cycle.

find l by inverting (25). We can then find y and x using the first-order conditions (20) and (22). $\square$

C.4 Reshaping the planner's problem

Let $V_R : [0,1]^n \to \mathbb{R}$ be the objective function of $\mathcal{R}$ defined as

$$V_R(\theta) = \left(\sum_{j\in\mathcal{N}} (q_j(\theta))^{\sigma-1}\right)^{\sigma-1} \left(1 - \sum_{j\in\mathcal{N}} f_j\theta_j\right) L \tag{26}$$

where q_j is implicitly defined by (12). Similarly, let $V_P : \{0,1\}^n \to \mathbb{R}$ be the objective function of $\mathcal{P}$.

Preliminary result

The proof of Proposition 4 relies on the following lemma.

Lemma 4. *Let $F = A - fB$ where $f > 0$, A is the all-one $n \times n$ matrix and B is an $n \times n$ matrix. If B is negative definite on the subspace $S : \sum_{i=1}^n x_i = 0$ then F is positive definite for $f > 0$ small enough.*

Proof. The negative definiteness of B on S implies that $x'Bx \leq -d\|x\|^2$ for $x \in S$ and some $d > 0$. We can write any vector z as $z = x + y$ where $x \in S$ and $y \perp S$. Then,

$$\begin{aligned} z'(A - fB)z &= n\|y\|^2 - fx'Bx - fy'By - 2fy'Bx \\ &\geq (n - 1/2)\|y\|^2 + df\|x\|^2 - 2f\|B\|\|x\|\|y\| \end{aligned}$$

for f small enough. For f small enough, this last expression is strictly convex in $(\|x\|, \|y\|)$ with a minimum of 0 at $(0,0)$ or, equivalently, at $z = 0$. Since $z'(A - fB)z > 0$ for any $z \neq 0$, it follows that F is positive definite. $\square$

Proofs of concavity

Proposition 3. Let $\varepsilon_j = \varepsilon$ and $\alpha_j = \alpha$. If $\Omega_{ij} = d_i e_j$ for some vectors d and e then the Karush-Kuhn-Tucker conditions are necessary and sufficient to characterize a solution to $\mathcal{R}$.

Proof. Raise both sides of (12) to the power $\varepsilon - 1$, multiply by $d_j\theta_j^b$ and sum across j's to find

$$\sum_{j\in\mathcal{N}} d_j\theta_j^b (q_j(\theta))^{\varepsilon-1} = \left(\sum_{j\in\mathcal{N}} d_j e_j^{\alpha} (Az_j)^{\varepsilon-1} \theta_j^{a(\varepsilon-1)+b}\right)^{\frac{1}{1-\alpha}}$$

so that, once combined with (12), we find

$$q_j(\theta) = Az_j\theta_j^a d_j^{\frac{\alpha}{\varepsilon-1}}\left(\sum_{i\in\mathcal{N}} d_i e_i^\alpha (Az_i)^{\varepsilon-1}\theta_i^{a(\varepsilon-1)+b}\right)^{\frac{\alpha}{1-\alpha}\frac{1}{\varepsilon-1}}.$$

Computing the log of Q, we get

$$\begin{aligned}\log(Q) =& \frac{1}{\sigma-1}\log\left(\left(\sum_{j\in\mathcal{N}}\beta_j\left(z_j\theta_j^a A d_j^{\frac{\alpha}{\varepsilon-1}}\right)^{\sigma-1}\right)\left(\sum_{i\in\mathcal{N}} d_i e_i^\alpha (Az_i)^{\varepsilon-1}\theta_i^{a(\varepsilon-1)+b}\right)^{\frac{\alpha}{1-\alpha}\frac{\sigma-1}{\varepsilon-1}}\right)\\ =& \frac{1}{\sigma-1}\log\left(\sum_{j\in\mathcal{N}}\beta_j\left(z_j\theta_j^a A d_j^{\frac{\alpha}{\varepsilon-1}}\right)^{\sigma-1}\right)+\frac{1}{\varepsilon-1}\frac{\alpha}{1-\alpha}\log\left(\sum_{i\in\mathcal{N}} d_i e_i^\alpha (Az_i)^{\varepsilon-1}\theta_i^{a(\varepsilon-1)+b}\right)\end{aligned}$$

If $0 < a \leq (\sigma-1)^{-1}$ and $-a(\varepsilon-1) \leq b \leq 1-a(\varepsilon-1)$ (and in particular if ($\star$) holds) the exponents on θ are all between 0 and 1 so that the summations in $\log(Q)$ are concave functions of θ. The log of a concave function is concave so $\log(Q)$ is also concave. Moving towards the full objective function, the term $1-\sum_{j\in\mathcal{N}}\theta_j f_j$ is concave and so is $\log V_R$. Since, in addition, the constraint set $\theta\in[0,1]^n$ is convex and the Slater's qualification condition is obviously satisfied, the Karush-Kuhn-Tucker conditions are necessary and sufficient to characterize an optimal solution to the maximization of $\log(V_R(\theta))$ on the set $\theta\in[0,1]^n$. Since log is an increasing transformation, a solution to this problem also solves $\mathcal{R}$. □

Proposition 4. Let $\sigma=\varepsilon_j$ for all j. Suppose that the $\{\beta_j\}_{j\in\mathcal{N}}$ are not too far from each other and that the fixed costs $f_j>0$ are not too large. If the matrix Ω is close enough to $\bar{\Omega}$, then the Karush-Kuhn-Tucker conditions are necessary and sufficient to characterize a solution to $\mathcal{R}$.

Proof. To simplify the notation, define $p_j = q_j^{\sigma-1}$ and let

$$g^j = \frac{p_j}{z_j^{\sigma-1}\left(\sum_i \Omega_{ij}p_i\right)^{\alpha_j}}.$$

$\mathcal{R}$ can then be written as

$$\min_{p\in P} -\frac{1}{\sigma-1}\log\left(\sum_{j\in\mathcal{N}}\beta_j p_j\right)-\log\left(1-\sum_{j\in\mathcal{N}} f_j g^j(p)\right)$$

where $P=\left\{p\in\mathbb{R}^n_{\geq 0}: p_j \leq z_j^{\sigma-1}\left(\sum_{i\in\mathcal{N}}\Omega_{ij}p_i\right)^{\alpha_j}, \forall j\right\}$.

Denote the objective function by Λ. Its Hessian matrix has typical element

$$\frac{\partial^2 \Lambda}{\partial p_k \partial p_l} = \underbrace{\frac{1}{\sigma - 1}\beta_k \beta_l \left(\sum_j \beta_j p_j\right)^{-2}}_{A_{kl}} + \underbrace{\frac{\sum_j f_j g^j_{kl}(p)}{1 - \sum_j f_j g^j(p)}}_{B_{kl}} + \underbrace{\frac{\left(\sum_j f_j g^j_k(p)\right)\left(\sum_j f_j g^j_l(p)\right)}{\left(1 - \sum_j f_j g^j(p)\right)^2}}_{C_{kl}}, \tag{27}$$

and define A, B and C as the matrices with typical elements A_{kl}, B_{kl} and C_{kl}.

We will show that in the limit as $\Omega \to \bar{\Omega}$ and $\beta_j \to \bar{\beta}$ for all j the Hessian is positive definite when the largest fixed cost $\max_j f_j$ is small enough. To do so, we will rely on Lemma 4 above. For that purpose, notice that, in the limit, A is a positive multiple of the all-one matrix.

Pick $\bar{f} > 0$ and $\tilde{f}_j \in [0,1]^n$ so that $f_j = \bar{f}\tilde{f}_j$ for all j. Taking the derivatives of g, we find

$$B_{kl} = \frac{1}{L - \sum_j f_j g^j(p)} \left(\underbrace{\sum_{j \in \mathcal{N}} f_j \frac{\alpha_j(\alpha_j + 1) p_j \Omega_{kj}\Omega_{lj}}{z_j^{\sigma-1}\left(\sum_i \Omega_{ij} p_i\right)^{\alpha+2}}}_{B^1_{kl}} \right. \tag{28}$$

$$\left. - \bar{f} \underbrace{\left(\tilde{f}_k \frac{\alpha_k \Omega_{lk}}{z_k^{\sigma-1}\left(\sum_i \Omega_{ik} p_i\right)^{\alpha_k+1}} + \tilde{f}_l \frac{\alpha_l \Omega_{kl}}{z_l^{\sigma-1}\left(\sum_i \Omega_{il} p_i\right)^{\alpha_l+1}} \right)}_{B^2_{kl}} \right). \tag{29}$$

B^1 is a Gramian matrix where B^1_{kl} is the scalar product of a pair of vectors v_k and v_l defined as

$$v_m = \left[\sqrt{\frac{f_1\alpha_1(\alpha_1+1)p_1}{z_1^{\sigma-1}\left(\sum_i \Omega_{i1}p_i\right)^{\alpha_1+2}}}\Omega_{m1} \quad \cdots \quad \sqrt{\frac{f_j\alpha_j(\alpha_j+1)p_j}{z_j^{\sigma-1}\left(\sum_i \Omega_{ij}p_i\right)^{\alpha_j+2}}}\Omega_{mj} \quad \cdots \quad \sqrt{\frac{f_n\alpha_n(\alpha_n+1)p_n}{z_n^{\sigma-1}\left(\sum_i \Omega_{in}p_i\right)^{\alpha_n+2}}}\Omega_{mn} \right]'.$$

Since Gramian matrices are positive semi-definite, so is B^1. For B^2, we will show that, in the limit, it is negative definite on the subspace $S : \sum_{i=1}^n x_i = 0$. Define b as a vector with typical element $b_j = \tilde{f}_j \alpha_j \left[z_j^{\sigma-1}\left(\sum_i \Omega_{ij}p_i\right)^{\alpha_j+1}\right]^{-1}$. We can write $B^2 = \Omega \text{diag}(b) + (\Omega \text{diag}(b))'$. Take any vector $x \in S$, then in the limit,

$$\begin{aligned} x'B^2x &= x'\left[\omega\left(\mathbb{1}_n - I_n\right)\text{diag}(b) + \left(\omega\left(\mathbb{1}_n - I_n\right)\text{diag}(b)\right)'\right]x \\ &= x'\left[-2\omega I_n \text{diag}(b)\right]x < 0 \end{aligned}$$

for any $x \neq 0$. The matrix B^2 is therefore negative definite on S. Using Lemma 4, $A - \bar{f}B^2$ is therefore positive definite for $\bar{f} > 0$ small enough. Finally, the matrix C in (27), is also a Gramian matrix and its contribution to the Hessian is thus positive semi-definite.

Putting the pieces together, we have shown that the Hessian of the objective function Λ is positive definite for $\max_j f_j$ small enough when $\Omega = \bar{\Omega}$ and $\beta = \bar{\beta}$. Now, each element of the Hessian is also a continuous function of (Ω, β) in a neighborhood of $\left(\bar{\Omega}, \bar{\beta}\right)$.[65] Since the eigenvalues are continuous functions of the elements of a matrix, they are also continuous functions of Ω and β. There is therefore a ball $\mathcal{B} = \left\{(\Omega, \beta) : \left\|(\Omega, \beta) - \left(\bar{\Omega}, \bar{\beta}\right)\right\| < \delta\right\}$ for $\delta > 0$ such that the Hessian is also positive definite for $(\Omega, \beta) \in \mathcal{B}$.[66] □

Proof of the equivalence of the solutions

Proposition 5. If a solution θ^* to $\mathcal{R}$ is such that $\theta^* \in \{0,1\}^n$ then θ^* also solves $\mathcal{P}$.

Proof. By construction, the objective function V_{RR} of $\mathcal{R}$ and the objective function V_{SP} of $\mathcal{P}$ coincide over $\{0,1\}^n$. Therefore $V_R\left(\theta^*\right) = V_P\left(\theta^*\right)$. Since the feasible set of $\mathcal{R}$, $[0,1]^n$, contains the feasible set of $\mathcal{P}$, $\{0,1\}^n$, it must be that $V_P\left(\theta^*\right) \geq V_P\left(\theta\right)$ for $\theta \in \{0,1\}^n$, otherwise θ^* would not be a solution to $\mathcal{R}$. θ^* therefore solves $\mathcal{P}$. □

Proposition 6. Under the ($\star$) condition, the marginal benefit and the marginal cost of increasing θ_j only depend on θ_j through the aggregates F_j and G_j.

Proof. Rewrite $\mathcal{R}$ as

$$\max_{\theta \in [0,1]^n} \left(\sum_{j \in \mathcal{N}} \beta_j q_j^{\sigma-1}\right)^{\frac{1}{\sigma-1}} \left(1 - \sum_{j \in \mathcal{N}} f_j \theta_j\right) L$$

subject to $q_j \leq A z_j \theta_j^{a_j} B_j^{\alpha_j}$ for all $j \in \mathcal{N}$ and where $B_j = \left(\sum_{i \in \mathcal{N}} \theta_i^{b_{ij}} \Omega_{ij} q_i^{\varepsilon_j - 1}\right)^{\frac{1}{\varepsilon_j - 1}}$. This problem is equivalent to $\mathcal{R}$ since the inequality constraints always bind at the optimum. The first-order conditions with respect to θ_k and q_k are

$$f_k Q + \bar{\mu}_k - \underline{\mu}_k = \zeta_k \frac{\partial q_k}{\partial \theta_k} + \sum_{j \in \mathcal{N}} \zeta_j \frac{\partial q_j}{\partial B_j} \frac{\partial B_j}{\partial \theta_k}$$

$$\frac{\partial Q}{\partial q_k} \left(1 - \sum_{j \in \mathcal{N}} f_j \theta_j\right) L = \zeta_k - \sum_{j \in \mathcal{N}} \zeta_j \frac{\partial q_j}{\partial B_j} \frac{\partial B_j}{\partial q_k}$$

where ζ_j is the Lagrange multiplier on the j-th inequality constraints, and $\bar{\mu}_j$ and $\underline{\mu}_j$ are the Lagrange multipliers on the constraint $\theta_j \leq 1$ and $\theta_j \geq 0$.[67] Combining these first-order conditions

[65] Elements of the Hessian become infinite at the boundary of P where $\sum_i \Omega_{ij} p_i = 0$ for some j. However, these points are not relevant to the planner under our assumptions. When the fixed costs are small enough and when Ω is close enough to $\bar{\Omega}$, each firm is connected to at least one producing firm at the optimum. Therefore, we can exclude these points easily adding the constraints $\sum_i \Omega_{ij} p_i \geq D$ for some $D > 0$ small. These constraints will never bind and the solution to the planner's problem is therefore unchanged.

[66] Since all norms are equivalent in a finite dimensional space, there is no need to specify one here.

[67] The partial derivatives of q_j are to be understood for the binding inequality constraint, i.e. $\frac{\partial q_j}{\partial \theta_j} = A z_j a_j \theta_j^{a_j - 1} A B_j^{\alpha_j}$ and $\frac{\partial q_j}{\partial B_j} = A z_j \theta_j^{a_j} A \alpha_j B_j^{\alpha_j - 1}$.

yields

$$f_k Q + \bar{\mu}_k - \underline{\mu}_k = \left(1 - \sum_{j \in \mathcal{N}} f_j \theta_j\right) L \frac{\partial Q}{\partial q_k} \frac{\partial q_k}{\partial \theta_k} + \sum_{j \in \mathcal{N}} \zeta_j \frac{\partial q_j}{\partial B_j} \left(\frac{\partial B_j}{\partial \theta_k} + \frac{\partial B_j}{\partial q_k} \frac{\partial q_k}{\partial \theta_k} \right).$$

For the result to hold, we need $\frac{\partial Q}{\partial q_k} \frac{\partial q_k}{\partial \theta_k}$ and $\frac{\partial B_j}{\partial \theta_k} + \frac{\partial B_j}{\partial q_k} \frac{\partial q_k}{\partial \theta_k}$ to depend on θ_k only through aggregates. We can write

$$\frac{\partial Q}{\partial q_k} \frac{\partial q_k}{\partial \theta_k} = \left(\sum_{j \in \mathcal{N}} \beta_j q_j^{\sigma - 1} \right)^{\frac{1}{\sigma-1} - 1} \beta_k \left(A z_k \theta_k^{a_k} B_k^{\alpha_k} \right)^{\sigma - 2} \times a_k A z_k \theta_k^{a_k - 1} B_k^{\alpha_k}$$

where we see that θ_k drops out of the equation if $a_k = \frac{1}{\sigma - 1}$. Similarly, we can write

$$\begin{aligned} \frac{\partial B_j}{\partial \theta_k} + \frac{\partial B_j}{\partial q_k} \frac{\partial q_k}{\partial \theta_k} = \frac{1}{\varepsilon_j - 1} \left(\sum_{i \in \mathcal{N}} \theta_i^{b_{ij}} \Omega_{ij} q_i^{\varepsilon_j - 1} \right)^{\frac{1}{\varepsilon_j - 1} - 1} \Big(& b_{kj} \theta_k^{b_{kj} - 1} \Omega_{kj} q_k^{\varepsilon_j - 1} \\ & + \theta_k^{b_{kj}} \Omega_{kj} (\varepsilon_j - 1) q_k^{\varepsilon_j - 2} a_k A z_k \theta_k^{a_k - 1} B_k^{\alpha_k} \Big) \end{aligned}$$

and, by taking into account that q_k depends directly on θ_k, we see that θ_k drops out of the equation if we also impose $b_{kj} = 1 - \frac{\varepsilon_j - 1}{\sigma - 1}$. The first-order condition therefore only depends on θ_k through the aggregates B_k and Q. □

C.5 Forces shaping the network

Proposition 7. In a large economy, operating a firm (weakly) increases the incentives to operate its (perhaps indirect) potential suppliers and customers.

Proof. Let j be a newly operating firm. In a large economy, operating j does not affect the Lagrange multiplier on the labor resource constraint w. Consider a firm k downstream from j in Ω. From the recursivity of (9), q_k weakly increases as a result of operating j (there is a strict increase if k is a direct neighbor). As a result, the benefit to the planner of operating k also weakly increases. On the other hand, the cost of operating k is still $w f_k L$ so that the net benefit of operating k weakly increases. Now consider a firm i that is upstream from j in Ω. From the recursivity of (9) and since j now operates, operating i would weakly increase q_j (with a strict increase if i and j are direct neighbor) which the planner values. The cost of operating i is still $w f_i L$ so the net benefit of operating i weakly increases. □

Proposition 8. The incentives of the planner to operate a group of firms (weakly) increase with additional potential connections between them.

Proof. Take a set of firms $\mathcal{J}$ and suppose that there exists $i, j \in \mathcal{J}$ such that $\Omega_{ij} = 0$. Operating these firms provides a certain benefit to the planner, and this benefit is increasing in the firms' productivities $\{q_k\}_{k\in\mathcal{J}}$. Suppose instead that $\Omega_{ij} > 0$. By equation (12), this additional potential connection weakly increases each productivity q_j for $j \in \mathcal{J}$, with a strict increase for q_j if i and j produce a positive amount, i.e. $q_i > 0$ and $q_j > 0$. The cost to the planner of operating the firms in $\mathcal{J}$ is unchanged by the inclusion of the additional potential connection Ω_{ij}. As a result, the incentives of the planner to operate the firms in $\mathcal{J}$ are larger with $\Omega_{ij} > 0$. $\square$

C.6 Aggregate fluctuations

Proposition 9. If $\alpha_j = \alpha$ for all $j \in \mathcal{N}$, then the efficient network θ does not depend on aggregate productivity A.

Proof. Using Proposition 2 and Lemma 1, we can rewrite $\mathcal{P}$ as maximizing (11) over the set $\theta \in \{0,1\}^n$ where the vector q solves, for each $j \in \mathcal{N}$, (9). Let's denote this problem as $\mathcal{P}_A$, where A refers to the aggregate productivity level. We will show that the optimal vector θ_A that solves $\mathcal{P}_A$ also solves an alternative problem $\mathcal{P}_{\tilde{A}}$ in which aggregate productivity is $\tilde{A}$ instead. Define $p_j = \left(\frac{\tilde{A}}{A}\right)^{\frac{1}{1-\alpha}} q_j$, then the objective function of $\mathcal{P}_{\tilde{A}}$ can be written as $\left(\tilde{A}/A\right)^{\frac{1}{1-\alpha}} \left(\sum_{j\in\mathcal{N}} \beta_j p_j^{\sigma-1}\right)^{\frac{1}{\sigma-1}} \left(1 - f_j \sum_{j\in\mathcal{N}} \theta_j\right) L$ and its recursive equation (9) can be written as $p_j = z_j \theta_j A \left(\sum_{i\in\mathcal{N}} \Omega_{ij} \left(p_i\right)^{\varepsilon_j - 1}\right)^{\frac{\alpha}{\varepsilon_j - 1}}$. Since the constant $\left(\tilde{A}/A\right)^{\frac{1}{1-\alpha}}$ does not affect the maximization, $\mathcal{P}_{\tilde{A}}$ is the same problem as $\mathcal{P}_A$, and, at the optimum, $p_{\tilde{A}} = q_A$ and $\theta_{\tilde{A}} = \theta_A$. The production network is therefore invariant to changes in aggregate productivity A. $\square$

D Algorithms

This appendix describes the various algorithms used in the paper.

D.1 Construction of the matrix Ω for the numerical tests.

This algorithm constructs the matrices Ω used in the numerical tests of Section 3.4 in the main text and of Sections B.1 and B.3 in the Appendix. Consider an economy with n firms, each with m incoming potential connections on average. Set $p = m/(n-1)$ and $\Omega_{ij} = 0$ for all $i, j \in \mathcal{N}^2$.[68]

1. Draw $\Omega_{ij} \sim$ iid Bernoulli(p), for all $i, j \in \mathcal{N}^2$.

2. For each $i, j \in \mathcal{N}^2$ such that $\Omega_{ij} = 1$, draw $\Omega_{ij} \sim$ iid $U[0,1]$.

3. Set $\Omega_{ii} = 0$ for all $i \in \mathcal{N}$.

[68] The division by $n-1$ is needed because the diagonal is forced to be empty.

D.2 Exhaustive search

This algorithm performs an exhaustive search of the 2^n vectors $\theta \in \{0,1\}^n$. It is used in Section 3.4 in the main text as well as in Sections B.1, B.3 and B.5 in the Appendix.

1. Order in an arbitrary way all the possible $\theta \in \{0,1\}^n$, from θ^1 to θ^{2^n}.

2. For each $p \in \{1, \ldots, 2^n\}$, use equations (9) and (11) to compute the aggregate consumption associated with θ^p.

3. The vector θ that provides the highest aggregate consumption corresponds to the efficient allocation.

This algorithm is guaranteed to find the global maximum of $\mathcal{P}$ but it is infeasible for large n given the speed at which the number of vectors in $\{0,1\}^n$ grows with n.

D.3 Deviation-free allocation

This algorithm starts from an allocation $\theta^0 \in \{0,1\}^n$ and looks for welfare-improving deviations. It is used in Sections 3.4, B.3 and B.5.

1. Initialize the 0-th iteration with θ^0.

2. For the p-th iteration, define $\tilde{\theta} = \theta^p$ and set $j = 1$.

   (a) If $\theta_j^p = 0$, set $\tilde{\theta}_j = 1$. If, instead, $\theta_j^p = 1$, set $\tilde{\theta}_j = 0$.

   (b) Using equations (9) and (11) compute the welfare associated with $\tilde{\theta}$.

   (c) If the welfare under $\tilde{\theta}$ is larger than the welfare under θ^p set $\theta^p = \tilde{\theta}$.

   (d) Set $j = j + 1$, set $\tilde{\theta} = \theta^p$ and repeat steps (a) through (d) until $j = n$.

3. Repeat step 2 above until no welfare-improving deviations are found for some θ^p.

D.4 Iterating on the first-order conditions

A convenient way to solve the reshaped planner's problem is to iterate on the first-order conditions of the log of the objective function of $\mathcal{R}$ while treating (12) as an inequality constraint. In what follows ζ_k is the Lagrange multiplier on the k-th inequality constraint (12), and $\underline{\mu}_j$ and $\overline{\mu}_j$ are the Lagrange multipliers on the constraint $\theta_j \geq 0$ and $\theta_j \leq 1$. The algorithm is as follows:

1. Initialize the 0-th iteration with $\Delta\mu_k^0 = \underline{\mu}_j^0 - \bar{\mu}_j^0 = -1$ for all $k \in \mathcal{N}$.

2. For the p-th iteration:

(a) Using the complementary slackness condition set $\theta_k^p = 1$ if $\Delta\mu_k^p \leq 0$ and $\theta_k^p = 0$ if $\Delta\mu_k^p > 0$.

(b) With θ^p, iterate on (12) until convergence to find the vector q^p.

(c) For each j, compute $B_j = \left(\sum_{i=1}^n \Omega_{ij} q_i^{\varepsilon_j - 1}\right)^{\frac{1}{\varepsilon_j - 1}}$ and $\Lambda_j = \frac{\theta_j}{B_j^{\varepsilon_j - 1}}$ if $B_j > 0$ and $\Lambda_j = 0$ otherwise.

(d) Find $\frac{\zeta_k^p q_k^p}{\theta_k^p}$ by solving the following system of linear equations derived from the first-order conditions:

$$\frac{\beta_k \left(A z_k B_k^{\alpha_k}\right)^{\sigma-1}}{\sum_{j\in\mathcal{N}} \beta_j q_j^{\sigma-1}} + \sum_{j\in\mathcal{N}} \left(A z_k B_k^{\alpha_k}\right)^{\varepsilon_j - 1} \Omega_{kj} \alpha_j \Lambda_j \frac{\zeta_j q_j}{\theta_j} = \frac{\zeta_k q_k}{\theta_k}$$

for each k, and where $\frac{\beta_k \left(A z_k B_k^{\alpha_k}\right)^{\sigma-1}}{\sum_{j\in\mathcal{N}} \beta_j q_j^{\sigma-1}}$ should be set to 0 if $\sum_{j\in\mathcal{N}} \beta_j q_j^{\sigma-1} = 0$.

(e) Compute $\Delta\mu_k$ using the following equation derived from the first-order conditions

$$\frac{f_k}{L - \sum_{j\in\mathcal{N}} f_j \theta_j} = \frac{1}{\sigma - 1} \frac{\zeta_k q_k}{\theta_k} + \sum_{j\in\mathcal{N}} \left(1 - \frac{\varepsilon_j - 1}{\sigma - 1}\right) \frac{\alpha_j}{\varepsilon_j - 1} \Omega_{kj} \Lambda_j \left(A z_k B_k^{\alpha_k}\right)^{\varepsilon_j - 1} \frac{\zeta_j q_j}{\theta_j} + \Delta\mu_k$$

for each $k \in \mathcal{N}$ and update to $\Delta\mu^{p+1} = \psi\Delta\mu + (1-\psi)\Delta\mu^p$ where $0 < \psi \leq 1$ is some parameter to control the speed of convergence.

3. Repeat step 2 above until convergence on $\Delta\mu$.

In practice, it is useful to slow down the updating rule by setting $\psi = 0.9$.

Notice that this algorithm imposes that $\theta \in \{0,1\}^n$ at every iteration. When the solution to $\mathcal{R}$ is not in $\{0,1\}^n$, the algorithm does not converge and the status θ of some firms keeps alternating between 0 and 1. In practice, I stop the algorithm when the distance between $\Delta\mu_k^{p+1}$ and $\Delta\mu_k^p$ starts to increase, which usually indicates that there will be no convergence. I then look at the set of firms for which θ keeps alternating (different sign for $\Delta\mu_k^{p+1}$ and $\Delta\mu_k^p$), and then pick the best $\theta \in \{0,1\}$ to maximize the planner's objective function.

D.5 Construction of the matrix Ω in the calibrated economy

The matrix Ω is constructed by assuming that the number of potential incoming and outgoing connections (x_{in}, x_{out}), for any given firm, is drawn from a bivariate power law of the first kind $\mathcal{K}$ for which the joint density over (x_{in}, x_{out}) is $g(x_{in}, x_{out}) = \xi(\xi - 1)(x_{in} + x_{out} - 1)^{-(\xi+1)}$. The full algorithm to construct the matrix is as follows:

1. Begin with $\Omega_{ij} = 0$ for all $i, j \in \mathcal{N}^2$.

2. For each firm $j \in \mathcal{N}$, draw from $\mathcal{K}$ a pair $\left(x_{in}^j, x_{out}^j\right)$ for the number of incoming and outgoing connections for j. Redraw until $\sum_j x_{in}^j = \sum_j x_{out}^j$ so that the total number of incoming connections is equal to the total number of outgoing connections in the economy.

3. For each $j \in \mathcal{N}$, create x_{in}^j incoming stubs and x_{out}^j outgoing stubs.

4. Randomly match each incoming stub to an outgoing stub. An incoming stub has the same probability of being matched with any outgoing stub. Set $\Omega_{ij} = 1$ where i is the firm associated with the outgoing stub and j is the firm associated with the incoming stub.

5. Since there are no self-link in the data, set $\Omega_{ii} = 0$ for all i.

6. Verify that each firm has at least one potential input, otherwise go back to step 1.

E Additional quantitative exercises

E.1 Simulations with a large number of firms and aggregate shocks

To investigate how the model behaves under a more realistic parametrization, I simulate the calibrated economy with $n = 20,000$ firms and with aggregate shocks to total factor productivity A. The number of firms was chosen to roughly match the number of operating firms in the Factset data. I assume that $\log(A_t)$ follows an AR(1) process with an autocorrelation of 0.9 and and a standard deviation parameter set to match empirical estimates about the impact of aggregate shocks on volatility.[69] Table 18 shows the correlations between aggregate output and the shape of the network. We see that the numbers are broadly similar to those of the benchmark calibration. I also compute the difference in output volatility between the flexible and fixed networks in this setting. I find that the flexible network economy is about 13% less volatile. Finally, aggregate output is also 9% larger under the flexible network, roughly the same number as in the benchmark economy.

E.2 Sensitivity to different Ω

To see how the matrix Ω affects the efficient allocation, I simulate the model under different parametrization for Ω. I still assume that Ω is drawn from a bivariate power law of the first kind but I vary its exponent to $\xi = 1.7$ and $\xi = 1.9$. The results are presented in Table 19. We see from the table that changing ξ has a direct impact on the degree distributions and the global clustering coefficient in the efficient network. Under $\xi = 1.7$, the distribution of the number of potential connections features thicker tails such that Ω offers a lot of options for the planner to

[69] Atalay (2017) generalizes an empirical strategy introduced by Foerster et al. (2011) to evaluate the impact of aggregate shocks on aggregate fluctuations. He finds that they account for 17% of volatility. I parametrize the stochastic process followed by $\log(A_t)$ to match that estimate.

Table 18: Correlation between with aggregate output with $n = 20,000$ firms and aggregate shocks

	Power law exponents		Clustering coefficient
Network	In-degree	Out-degree	
Model with $n = 20,000$ firms and aggregate shocks	-0.73	-0.73	0.68
Benchmark model	-0.59	-0.70	0.54

Notes: All time series are in logs. The parameters of the economy are as in the benchmark calibration except as mentioned in the text. Since these simulations are computationaly intensive, I simulate four economies instead of twenty in the benchmark exercises.

create highly-connected firms and dense clusters of producers. The planner takes advantage of these possibilities to increase welfare: the mean of aggregate output is 14.7 under $\xi = 1.7$, but only 13.4 under $\xi = 1.9$.

Table 19: Impact of Ω on the production network

	Power law exponent		Clustering coefficient
Network	In-degree	Out-degree	
$\xi = 1.7$	0.91	0.88	4.07
$\xi = 1.79$ (benchmark)	1.00	0.96	3.31
$\xi = 1.9$	1.12	1.07	2.25

Notes: The parameters are the same as in the benchmark calibration except for the distribution from which Ω is drawn (see text).