UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for New Economic Thinking

Cambridge-INET Institute

Cambridge-INET Working Paper Series No: 2018/13

Cambridge Working Papers in Economics: 1856

PRODUCTION NETWORKS: A PRIMER

Vasco M. Carvalho (University of Cambridge)

Alireza Tahbaz-Salehi (Northwestern University)

This article reviews the literature on production networks in macroeconomics. It presents the theoretical foundations for the roles of input-output linkages as a shock propagation channel and a mechanism for transforming microeconomic shocks into macroeconomic fluctuations. The article provides a brief guide to the growing literature that explores these themes empirically and quantitatively.

Production Networks: A Primer*

Vasco M. Carvalho[†] Alireza Tahbaz-Salehi[‡]

September 2018

Abstract

This article reviews the literature on production networks in macroeconomics. It presents the theoretical foundations for the roles of input-output linkages as a shock propagation channel and a mechanism for transforming microeconomic shocks into macroeconomic fluctuations. The article provides a brief guide to the growing literature that explores these themes empirically and quantitatively.

Keywords: networks, shock propagation, input-output linkages.

*Prepared for the *Annual Review of Economics*.
[†]Faculty of Economics, University of Cambridge, vmpmdc2@cam.ac.uk..
[‡]Kellogg School of Management, Northwestern University, alirezat@kellogg.northwestern.edu.

1 Introduction

The idea that the production of goods and services in any economy relies on a complex web of transactions between a wide range of suppliers and customers has a long tradition in economics. As far back as the 1950s, in his study of the structure of the American economy, Leontief (1951) observed that almost everyone is "[...] aware of the existence of some kind of interconnection between even the remotest parts of a national economy" and that "the presence of these invisible but nevertheless very real ties can be observed whenever expanded automobile sales in New York City increase the demand for groceries in Detroit, [...] when the sudden shutdown of the Pennsylvania coal mines paralyzes the textile mills in New England," and in fact "with relentless regularity in alternative ups and downs of business cycles."

This article reviews the recent theoretical and empirical literature that, under the broad heading of production networks, revisits the roles of input-output linkages in propagating shocks and transforming microeconomic disturbances into macroeconomic fluctuations. Though its motivation is classical, the modern literature on production networks relies on two fairly recent developments.

First, by drawing on tools from a diverse body of knowledge, the burgeoning field of network analysis has developed a conceptual framework and an extensive set of tools to effectively encode and measure interconnections between the units of analysis comprising a network. As we shall see below, when coupled with the language of general equilibrium theory, these serve as useful tools for assessing how shocks propagate throughout the economy, how different sectors comove over the business cycle, or how aggregate fluctuations can be traced out to localized, micro-disturbances.

Second, and keeping in tandem with developments elsewhere in economics, the availability of novel large datasets on the granular nature of production across the economy has paved the way for a wide range of empirical and quantitative analyses to answer classical questions such as the origins of aggregate fluctuations. Almost eighty years after Leontief's pioneering study of the structure of the American economy, modern input-output tables detail the complex patterns of input linkages between hundreds of industries. Going even deeper into the micro, it is now possible to identify supplier-customer relationships between millions of firms in various industrialized economies.

In this article, we provide a broad overview of the growing literature that leverages the above developments, with a particular focus on macroeconomic implications. While not a comprehensive survey, the article aims to offer a user guide to some of the recent theoretical and empirical works in the area.

We begin, in Section 2, by presenting a benchmark model of production networks that will serve as the basis for the main theoretical results in this article. We use this framework to demonstrate the role of input-output linkages as a shock propagation channel throughout the economy. In Section 3, we focus our attention on the role of input-output linkages as a mechanism for translating microeconomic shocks into aggregate fluctuations. These results provide sharp conditions for whether and when macroeconomic fluctuations can have their origins in idiosyncratic shocks to individual firms or disaggregated industries. We review the related empirical and quantitative literature in Section 4, both at the firm- and industry-level across various countries. We conclude in Section 5 by discussing a number of open questions and promising avenues for future research.

2 A Model of Production Networks

We start by presenting a baseline model of production networks, which will serve as a useful starting point for analysis. The model is a static variant of the multi-sector general equilibrium model of Long and Plosser (1983), which is also analyzed by Acemoglu et al. (2012). We discuss various modifications and generalizations of this model in the subsequent sections.

2.1 Baseline Model

Consider a static economy consisting of n competitive industries denoted by $\{1, 2, \ldots n\}$, each producing a distinct product. Each product can be either consumed by the households or used as an intermediate input for production of other goods. Firms in each industry employ Cobb-Douglas production technologies with constant returns to scale to transform intermediate inputs and labor into final products. In particular, the output of industry i is given by

$$y_i = z_i \zeta_i l_i^{\alpha_i} \prod_{j=1}^{n} x_{ij}^{a_{ij}}, \tag{1}$$

where l_i is the amount of labor hired by firms in industry i, x_{ij} is the quantity of good j used for production of good i, $\alpha_i > 0$ denotes the share of labor in industry i's production technology, z_i is a Hicks-neutral productivity shock, and ζ_i is some normalization constant whose value only depends on model parameters.[1]

The exponents $a_{ij} \geq 0$ in Equation (1) formalize the idea that firms in an industry may need to rely on the goods produced by other industries as intermediate inputs for production. In particular, a larger a_{ij} means that good j is a more important input for the production of good i, whereas $a_{ij} = 0$ means that good j is not a necessary input for i's production. Note that, in general, $a_{ij} \neq a_{ji}$, as industry i's reliance on industry j as an input-supplier may be different from j's dependence on i. Furthermore, it may also be the case that $a_{ii} > 0$, as good i may itself be used as an intermediate input for production by firms in industry i. Finally, note that the assumption that all technologies exhibit constant returns to scale implies that $\alpha_i + \sum_{j=1}^{n} a_{ij} = 1$ for all i.

In addition to the firms described above, the economy is populated by a representative household, who supplies one unit of labor inelastically and has logarithmic preferences over the n goods given by

$$u(c_1, \ldots, c_n) = \sum_{i=1}^{n} \beta_i \log(c_i/\beta_i), \tag{2}$$

where c_i is the amount of good i consumed. The constants $\beta_i \geq 0$ measure various goods' shares in the household's utility function, normalized such that $\sum_{i=1}^{n} \beta_i = 1$.

Equations (1) and (2) thus fully specify the environment. The competitive equilibrium of this economy is defined in the usual way: it consists of a collection of prices and quantities such that (i) the representative household maximizes her utility; (ii) the representative firm in each sector maximizes its profits while taking the prices and the wage as given; and (iii) all markets clear.

[1] In what follows, we set the value of this constant to $\zeta_i = \alpha_i^{-\alpha_i} \prod_{j=1}^{n} a_{ij}^{-a_{ij}}$. The sole purpose of this constant is to simplify the analytical expressions without any bearing on the results.

Before characterizing the equilibrium, it is useful to define a few key concepts that will play a central role in our subsequent analysis. First, note that we can summarize the input-output linkages between various industries with a matrix $A = [a_{ij}]$, which with some abuse of terminology, we refer to as the economy's *input-output matrix*.[2] Thus, coupled with the vector of productivity shocks $(z_1, \ldots, z_n)$, the input-output matrix A serves as a sufficient statistic for the production side of the economy. Note that the assumption that $\alpha_i > 0$ for all i implies that A is an element-wise non-negative matrix with row sums that are strictly less than 1. This in turn guarantees that the spectral radius of A — defined as the largest absolute value of its eigenvalues — is also strictly less than 1 (Berman and Plemmons, 1979, p. 37).

The input-output linkages between various industries can alternatively be represented by a weighted and directed graph on n vertices. Each vertex in this graph — which we refer to as the economy's *production network* — corresponds to an industry, with a directed edge with weight $a_{ij} > 0$ present from vertex j to vertex i if industry j is an input-supplier of industry i. While the production network representation of the economy is equivalent to the representation using the input-output matrix — in fact, in graph theory terminology, the input-output matrix A is nothing but the adjacency matrix of the economy's production network — it can provide a conceptually simpler framework for summarizing (and visualizing) input-output linkages.

Finally, we define an industry's *Domar weight* as that industry's sales as a fraction of GDP. More specifically, the Domar weight of industry i is defined as

$$\lambda_i = \frac{p_i y_i}{\text{GDP}}, \tag{3}$$

where p_i is the price of good i and y_i is industry i's output. These weights will play a key role in the analysis in Section 3.

We now proceed to determining the equilibrium prices and quantities. First, note that firms in industry i choose their demand for labor and intermediate goods in order to maximize profits, $\pi_i = p_i y_i - w l_i - \sum_{j=1}^{n} p_j x_{ij}$, while taking all prices $(p_1, \ldots, p_n)$ and the wage w as given. Thus, the first-order conditions corresponding to firms in industry i are given by $x_{ij} = a_{ij} p_i y_i / p_j$ and $l_i = \alpha_i p_i y_i / w$. Plugging these expressions into firm i's production function in Equation (1) and taking logarithms imply that

$$\log(p_i/w) = \sum_{j=1}^{n} a_{ij} \log(p_j/w) - \epsilon_i,$$

where $\epsilon_i = \log z_i$ is the (log) productivity shock to firms in industry i. Since the above relationship has to hold for all industries i, it provides a system of equations to solve for all relative prices in terms of productivity shocks. More specifically, rewriting this system of equations in matrix form implies that $\hat{p} = A\hat{p} - \epsilon$, where A is the economy's input-output matrix and $\hat{p} = (\log(p_1/w), \ldots, \log(p_n/w))'$ and $\epsilon = (\epsilon_1, \ldots, \epsilon_n)'$ denote the vectors of log relative prices and productivity shocks, respectively.

[2]More generally, the input-output matrix $\Omega = [\omega_{ij}]$ of an economy is defined in terms of input expenditures as a fraction of sales, that is, $\omega_{ij} = p_j x_{ij} / p_i y_i$. However, in the special case that all technologies and preferences are Cobb-Douglas, ω_{ij} coincides with the exponent a_{ij} in Equation (1).

Consequently, the equilibrium vector of (log) relative prices is given by

$$\hat{p} = -(I - A)^{-1}\epsilon \tag{4}$$

in terms of industry-level shocks and the economy's production network.

Before proceeding any further, it is useful to comment on some of the key properties of matrix $L = (I - A)^{-1}$ in Equation (4), commonly known as the economy's *Leontief inverse*. First, the fact that the input-output matrix A is nonnegative with a spectral radius that is strictly less than 1 means that $I - A$ is a non-singular M-matrix, which in turn guarantees that the Leontief inverse $L = (I - A)^{-1}$ always exists and is element-wise nonnegative.[3] Second, the observation that the spectral radius of A is strictly less than 1 also implies that the Leontief inverse can be expressed as the infinite sum of the powers of the input-output matrix A (Stewart, 1998, p. 55), i.e.,

$$L = (I - A)^{-1} = \sum_{k=0}^{\infty} A^k. \tag{5}$$

This decomposition illustrates that the (i, j) element of the Leontief inverse measures the importance of industry j as a *direct* and *indirect* input-supplier to industry i in the economy. To see this, note that for any $i \neq j$, Equation (5) implies that $\ell_{ij} = a_{ij} + \sum_{r=1}^{n} a_{ir}a_{rj} + \dots$, with the first term in this expression accounting for j's role as a direct supplier to i, the second term accounting for j's role as a supplier to i's suppliers, and so on. Interpreted in terms of the production network representation of the economy, ℓ_{ij} accounts for all possible directed walks (of various lengths) that connect industry j to industry i over the network.

Returning to equilibrium characterization, recall that the firms' first-order conditions imply that the quantity demanded by industry i from industry j is given by $x_{ij} = a_{ij}p_iy_i/p_j$, while the representative household's logarithmic utility implies that she demands $c_j = \beta_j w/p_j$ units of good j. Plugging these expressions into the market-clearing condition for good j, which is given by $y_j = c_j + \sum_{i=1}^{n} x_{ij}$, implies that $p_jy_j = \beta_j w + \sum_{i=1}^{n} a_{ij}p_iy_i$. Dividing both sides of this equation by w and noting that that the value added in this economy is equal to the household's labor income, we obtain

$$\lambda_j = \beta_j + \sum_{i=1}^{n} a_{ij}\lambda_i,$$

where λ_i is the Domar weight of industry i defined in Equation (3). Rewriting the above equation in matrix form and solving for the vector of Domar weights implies that $\lambda = (I - A')^{-1}\beta$, or equivalently, $\lambda_i = p_iy_i/\text{GDP} = \sum_{j=1}^{n} \beta_j \ell_{ji}$. Furthermore, recall from Equation (4) that $\log(p_i/\text{GDP}) = -\sum_{j=1}^{n} \ell_{ij}\epsilon_j$, thus leading to the following result:

Theorem 1. *The log output of industry i is given by*

$$\log(y_i) = \sum_{j=1}^{n} \ell_{ij}\epsilon_j + \delta_i, \tag{6}$$

where δ_i is some constant that is independent of the shocks.

[3] A square matrix Q is called an M-matrix if there exist a nonnegative square matrix B and a constant $r \geq \rho(B)$ such that $Q = rI - B$, where $\rho(B)$ is the spectral radius of B. If $r > \rho(B)$, then Q is a non-singular M-matrix. Plemmons (1977, Theorem 2) shows that the inverse of any non-singular M-matrix is element-wise nonnegative.

The above theorem has a few important implications. First, the mere fact that the output of industry i may depend on the shocks to industries $j \neq i$ indicates that the input-output linkages in the economy can function as a mechanism for the propagation of shocks from one industry to another. Second, it shows that the resulting propagation patterns are captured by the economy's Leontief inverse L (and not its input-output matrix A). This means that the input-output linkages can result in both direct and indirect propagation of the shocks over the production network. Third, the fact that the impact of a shock to industry j on i's output is captured by ℓ_{ij} means that productivity shocks in this model propagate "downstream" from one industry to its customers, its customers' customers, and so on.[4] To see this, recall from the expansion in Equation (5) that ℓ_{ij} is a measure of the importance of industry j as (direct and indirect) input-supplier to industry i.[5]

The intuition underlying Theorem 1 is fairly straightforward. Suppose that industry j is hit by a negative shock that reduces its production and hence increases the price of good j. Such a price increase adversely impacts all the industries that rely on good j as an intermediate input for production, thus creating a direct impact on j's customer industries. But this initial impact will in turn result in further propagation over the production network: the prices of goods produced by industries affected in the first round of propagation will rise, creating an indirect negative effect on their own customer industries, and so on. The overall effect of these direct and indirect downstream propagation of the initial shock is summarized by the corresponding element of the economy's Leontief inverse.

But why is it that shocks in this model only propagate from an industry to its (direct and indirect) customers but not its suppliers? The absence of such "upstream" propagation is a consequence of three specific features of the model: (i) Cobb-Douglas preferences and technologies, (ii) a single factor of production (in this case labor), and (iii) constant returns to scale. The latter two features together guarantee that productivity shocks do not impact upstream prices (relative to the wage): the price of good i is equal to industry i's marginal cost, which only depends on the productivities of i and its upstream industries. On the other hand, as we already showed, in a Cobb-Douglas economy, the Domar weight of all industries are invariant to the shocks ($\lambda_i = p_i y_i / \text{GDP} = \sum_{j=1}^n \beta_j \ell_{ji}$). Hence, no upstream effect on (relative) prices has to translate into no upstream effect on quantities.

We now turn to determining the production networks' macroeconomic implications by studying how they shape aggregate economic variables. Recall from Equation (4) that the price of good i satisfies $\log(p_i/w) = -\sum_{j=1}^n \ell_{ij}\epsilon_j$. Multiplying both sides by β_i and summing over all industries i leads to $\log(\text{GDP}) = \sum_{i,j=1}^n \beta_i \ell_{ij}\epsilon_j + \sum_{i=1}^n \beta_i \log p_i$. On the other hand, choosing the consumption good bundle, whose price is given by $P_c = \prod_{i=1}^n p_i^{\beta_i}$, as the numeraire implies that $\sum_{i=1}^n \beta_i \log p_i = 0$. We therefore have the following result:

Theorem 2. *The economy's (log) real value added is given by*

$$\log(\text{GDP}) = \sum_{i=1}^n \lambda_i \epsilon_i, \tag{7}$$

[4]As we show in the subsequent sections, demand-side shocks exhibit significantly different propagation patterns.

[5]Note that when the production network exhibits cycles (say, in an economy with roundabout production), an industry can be simultaneously upstream and downstream to another industry. What we mean by downstream propagation is that shocks transmit from one industry to another in the direction of the flow of goods and services.

where

$$\lambda_i = \frac{p_i y_i}{\text{GDP}} = \sum_{j=1}^{n} \beta_j \ell_{ji} \tag{8}$$

and ℓ_{ji} *is the* (j,i) *element of the economy's Leontief inverse* $L = (I - A)^{-1}$.

The significance of the above result is twofold. First, Equation (7) illustrates that (log) aggregate output is a linear combination of industry-level productivity shocks, with coefficients given by the industries' Domar weights. Thus, the Domar weight of industry i is a sufficient statistic for how shocks to that industry impact aggregate output. As we will discuss in Section 2.3.2, some variant of this relationship, which is commonly known as "Hulten's Theorem," holds much more generally (Hulten, 1978; Gabaix, 2011).

Second, Theorem 2 establishes that with Cobb-Douglas preferences and technologies, the Domar weights take a particularly simple form: the Domar weight of each industry depends only on the preference shares and the corresponding column of the economy's Leontief inverse. This means that, while λ_i is a sufficient statistic for how shocks to industry i impact $\log(\text{GDP})$, the value of λ_i itself depends the economy's production network. In particular, as Equation (8) illustrates, all else equal, an increase in ℓ_{ji} will increase industry i's Domar weight and hence intensify the impact of shocks to i on aggregate output. The intuition underlying this result parallels that of Theorem 1: the downstream propagation of shocks from an industry to its direct and indirect customers means that, all else equal, shocks to industries that are more important input-suppliers to the rest of the economy have a more pronounced effect on macroeconomic aggregates.

2.2 Demand-Side Shocks

We next show that demand-side shocks lead to propagation patterns that are substantially different from those of supply-side productivity shocks studied so far.

To incorporate demand-side shocks into the model, we follow Acemoglu, Akcigit, and Kerr (2016) and modify our benchmark model by assuming that the government purchases an exogenously given quantity g_i of good i. This modification implies that good i's market-clearing condition is given by $y_i = c_i + g_i + \sum_{j=1}^{n} x_{ji}$. Thus, changes in government spending on various goods correspond to demand-side shocks that affect industries differentially. To simplify the derivations, we abstract away from supply-side shocks by assuming $z_i = 1$ for all i.

Solving for the economy's competitive equilibrium is straightforward. Plug industry i's first-order conditions — given by $x_{ij} = a_{ij} p_i y_i / p_j$ and $l_i = \alpha_i p_i y_i / p_j$ — into Equation (1) and solve the resulting system of equations, which implies that $p_i = w$ for all i. This means that, unlike productivity shocks, demand-side shocks do no impact relative prices. On the other hand, the representative household's budget constraint is given by $\sum_{i=1}^{n} p_i c_i = w - T$, where $T = \sum_{i=1}^{n} p_i g_i$ is the total amount of government spending, financed by lump sum taxes on the household. Therefore, the market-clearing condition for good i reduces to $y_i = \beta_i (1 - \sum_{j=1}^{n} g_j) + g_i + \sum_{j=1}^{n} a_{ji} y_j$. Rewriting the resulting system of equations in matrix form, we obtain $y = (1 - g'\mathbf{1})\beta + g + A'y$, where $g = (g_1, \ldots, g_n)'$ is the vector of quantities demanded by the government and $\mathbf{1}$ is a vector with all entries equal to 1. Solving this system of equations leads to the following result:

Theorem 3. *The output of industry i is given by*

$$y_i = \sum_{j=1}^{n} \ell_{ji} g_j + \left(1 - \sum_{k=1}^{n} g_k\right) \left(\sum_{j=1}^{n} \ell_{ji} \beta_j\right), \tag{9}$$

where $L = (I - A)^{-1}$ is the economy's Leontief inverse matrix.

Contrasting Theorems 1 and 3 illustrates the stark difference in how supply- and demand-side shocks propagate: whereas the impact of a productivity shock to industry j on the output of industry i is captured by ℓ_{ij}, the impact of a demand shock to j on i is captured via ℓ_{ji}. This means that, unlike supply-side shocks that propagate downstream, demand-side shocks propagate upstream from one industry to its direct and indirect suppliers.[6] The intuition underlying this propagation pattern is as follows: a positive demand shock to industry j increases j's demands for inputs, which is in effect a positive demand shock to j's suppliers. A similar logic implies that the original demand shock would propagate further upstream.

We conclude this discussion by noting that the baseline model we focused on thus far is special along multiple dimensions: it is a perfectly competitive economy with a single factor of production and Cobb-Douglas technologies and preferences. In the remainder of this section, we briefly discuss the implications of relaxing some of these assumptions.

2.3 More General Production Technologies

We start by illustrating how relaxing the assumption that all production technologies are Cobb-Douglas alters shocks' propagation patterns as well as their aggregate implications.

2.3.1 Propagation Patterns

One of the consequences of assuming Cobb-Douglas production technologies is that an industry's expenditure on various inputs as a fraction of its sales is invariant to the realization of the shocks. In particular, for any pair of industries i and j, the ratio $\omega_{ij} = p_j x_{ij}/p_i y_i$ is equal to the exponent a_{ij} in Equation (1), which is an exogenously given parameter of the model. Such an invariance, however, no longer holds for more general production technologies. This in turn can lead to richer patterns of shock propagation over input-output linkages.

These effects are explored by Carvalho et al. (2016), who focus on a generalization of the baseline model by replacing the production functions in Equation (1) by a nested CES structure. Since a closed-form characterization is in general not possible, they use a first-order approximation to show that when the elasticities of substitution between various intermediate inputs or between the intermediates and primary factors of production are different from 1, a negative productivity shock to industry i impacts the output of other industries via two distinct channels. First, the resulting increase in good i's price adversely impacts all industries that rely on good i as an intermediate input for production, thus leading to a downstream propagation of the shock to i's direct and indirect

[6] Note that in addition to the upstream propagation channel highlighted above, the expression in Equation (9) also includes a term $(1 - \sum_{k=1}^{n} g_k)(\sum_{k=1}^{n} \ell_{ji}\beta_j)$ that corresponds to a resource-constraint effect: an increase in government spending requires higher taxes on the households, and hence, fewer resources for private consumption.

customers. This "output effect" thus leads to propagation patterns that mirror those in a Cobb-Douglas economy. Second, the negative productivity shock to industry i may also result in reallocation of resources across different industries depending on the elasticities of substitution across various inputs. For instance, the increase in the price of good i in response to such a shock results in an increase (respectively, decrease) in demand by i's customers for input j if goods i and j are gross substitutes (respectively, complements) in these customers' production technologies. Hence, in contrast to the Cobb-Douglas economy, the impact of a shock to industry i may not remain confined to i's downstream industries.

These results are further extended by Baqaee and Farhi (2018a) to a general class of economies with heterogenous agents, arbitrary nested CES production structures, and multiple (and potentially industry-specific) factors of production. For the purposes of this article, we find it instructive to focus on a special case with a single factor of production (labor) and a single CES nest to clarify the two propagation channels highlighted in the previous paragraph. In particular, suppose that the production technology of firms in industry i is given by

$$y_i = z_i \zeta_i l_i^{\alpha_i} \left(\sum_{j=1}^{n} a_{ij}^{1/\sigma_i} x_{ij}^{1-1/\sigma_i} \right)^{(1-\alpha_i)\sigma_i/(\sigma_i-1)}, \tag{10}$$

where $\alpha_i + \sum_{j=1}^{n} a_{ij} = 1$, σ_i denotes the elasticity of substitution between the various inputs, and the normalization constant $\zeta_i = \alpha_i^{-\alpha_i}(1-\alpha_i)^{-(1-\alpha_i)\sigma_i/(\sigma_i-1)}$. This economy reduces to the baseline model with Cobb-Douglas technologies in Equation (1) when $\sigma_i \to 1$ for all i. As we show in Supplemental Appendix A, log-linearization of equilibrium conditions implies that the effect of a shock to industry j on the output of industry i is given by

$$\left.\frac{d\log(y_i)}{d\epsilon_j}\right|_{\epsilon=0} = \ell_{ij} + \frac{1}{\lambda_i}\sum_{k=1}^{n}(\sigma_k - 1)\lambda_k \left(\sum_{r=1}^{n} a_{kr}\ell_{ri}\ell_{rj} - \frac{1}{1-\alpha_k}\Big(\sum_{r=1}^{n} a_{kr}\ell_{ri}\Big)\Big(\sum_{r=1}^{n} a_{kr}\ell_{rj}\Big) \right) \tag{11}$$

up to a first-order approximation. As before, λ_i denotes industry i's Domar weight of and $L = (I-A)^{-1}$ is the economy's Leontief inverse, where $A = [a_{ij}]$.

The first term on the right-hand side of Equation (11) coincides with the expression in Equation (6) and captures the downstream output effect that is also present in a Cobb-Douglas economy. The second term captures the reallocation effect: in response to a negative shock to industry j, all industries k that are downstream to j may readjust their demand for all other inputs. Crucially, the impact of such readjustments by any given k on the output of industry i depends on (i) the elasticity of substitution σ_k in k's production function and (ii) the extent to which the supply chains that connect i and j to k coincide with one another.

To clarify the workings of the reallocation channel and its relationship to Equation (11), it is instructive to focus on a simple environment. Consider an industry k with $\sigma_k > 1$ that is downstream to both i and j, each of which supply k via a single production chain. This means that the constellation of these industries in the production network can take one of the following two forms: either i and j supply industry k via production chains that pass through the same supplier of k, as depicted in Figure 1(a); or the production chains supplied by i and j reach industry k via two distinct suppliers of k, as depicted in Figure 1(b).

Figure 1. Industry k is downstream to industries i and j.

It is not hard to verify that, for the economy depicted in Figure 1(a), the term in braces on the right-hand side of Equation (11) is equal to

$$\sum_{r=1}^{n} a_{kr}\ell_{ri}\ell_{rj} - \frac{1}{1-\alpha_k}\Big(\sum_{r=1}^{n} a_{kr}\ell_{ri}\Big)\Big(\sum_{r=1}^{n} a_{kr}\ell_{rj}\Big) = a_{ks}\ell_{si}\ell_{sj}\left(1 - \frac{a_{ks}}{1-\alpha_k}\right),$$

where s is the supplier of k that is downstream to both i and j. This expression is strictly positive as long as industry s is not the sole supplier of k (i.e., $a_{ks} < 1 - \alpha_k$). Hence, a negative productivity shock to industry j results in a decrease in i's output. This is, of course, fairly intuitive: the fact that $\sigma_k > 1$ implies that, in response to a negative shock to j, industry k substitutes away from the production chain supplied by j, in the process impacting industry i negatively.

The substitution channel results in a different propagation pattern in the economy depicted in Figure 1(b). Once again, a negative shock to j would force industry k to substitute away from the production chain that is supplied by j whenever $\sigma_k > 1$. But, unlike the previous case, such a substitution results in an increase in i's output precisely because the production chains supplied by i and j do not overlap with one another. Indeed, the term in braces on the right-hand side of Equation (11) is given by

$$\sum_{r=1}^{n} a_{kr}\ell_{ri}\ell_{rj} - \frac{1}{1-\alpha_k}\Big(\sum_{r=1}^{n} a_{kr}\ell_{ri}\Big)\Big(\sum_{r=1}^{n} a_{kr}\ell_{rj}\Big) = \frac{-1}{1-\alpha_k}(a_{ks}\ell_{si})(a_{k\tilde{s}}\ell_{\tilde{s}j}),$$

which is strictly negative, as expected.

Taken together, this example illustrates that when inputs are gross substitutes, more overlap in the production chains that originate from i and j translates a negative productivity shock to j into a reduction in i's output, while such overlaps have the opposite effect when inputs are gross complements.

2.3.2 Aggregate Effects and Hulten's Theorem

Recall from Theorem 2 that, with Cobb-Douglas preferences and technologies, an industry's Domar weight is a sufficient statistic for how TFP shocks to that industry impact GDP. We now show that a variant of this relationship holds much more generally: in any efficient economy, the impact on output of a TFP shock to industry i is equal to i's Domar weight up to a first-order approximation. More specifically, if z_i denotes the TFP shock to industry i, then irrespective of household's preferences and firms' production technologies,

$$\frac{d\log(\text{GDP})}{d\log(z_i)} = \lambda_i, \tag{12}$$

where $\lambda_i = p_i y_i / \text{GDP}$ is the Domar weight of industry i.

The simplicity of the relationship in Equation (12), which has come to be known as *Hulten's theorem*, makes it a useful tool in empirical studies of microeconomic origins of aggregate fluctuations. For instance, as we will discuss in subsequent sections, Gabaix (2011) uses the empirical distribution of firm-level Domar weights to measure the extent to which firm-level shocks can explain GDP volatility, while Carvalho and Gabaix (2013) rely on Hulten's theorem to investigate whether changes in the economy's microeconomic composition can account for the "great moderation" and its unraveling in major world economies.

Despite its simplicity, Hulten's theorem may appear surprising at first sight: how is it that in the presence of input-output linkages an industry's role in shaping aggregate outcomes is entirely reflected by its size, irrespective of its position in the production network?[7]

To derive and illustrate the intuition behind Equation (12), we follow papers such as Gabaix (2011) and Baqaee and Farhi (2018c) and extend the baseline model in Section 2 by allowing for general production functions, preferences, and factor markets. More specifically, consider a static economy consisting of n competitive industries, each producing a distinct product using intermediate inputs and m different primary factors of production. Firms in industry i employ constant returns to scale production technologies given by

$$y_i = z_i f_i(x_{i1}, \dots, x_{in}, l_{i1}, \dots l_{im}),$$

where z_i is the Hicks-neutral productivity shock to industry i and x_{ij} and l_{ik} are the quantities of good j and the k-th primary factor used by firms in industry i, respectively. The economy is also populated by a representative household with preferences $u(c_1, \dots, c_n)$, which we assume to be homogenous of degree 1. This representative household is endowed with h_k units of the k-th primary factor, which she supplies inelastically to the market. As before, we focus on the economy's competitive equilibrium, in which (i) all firms maximize their profits, taking the factor and intermediate good prices as given; (ii) the representative household maximizes her utility; and (iii) all good and factor markets clear.

By the first welfare theorem, the competitive equilibrium of this economy is efficient. This means that one can determine the equilibrium allocation by solving the social planner's problem:

$$\begin{aligned} W = \max_{c_i, l_{ik}, x_{ij}} \quad & u(c_1, \dots, c_n) \\ \text{s.t.} \quad & c_i + \sum_{j=1}^{n} x_{ji} = z_i f_i(x_{i1}, \dots, x_{in}, l_{i1}, \dots, l_{im}) \qquad i = 1, \dots, n \\ & \sum_{i=1} l_{ik} = h_k \qquad k = 1, \dots, m. \end{aligned}$$

[7] The apparent discrepancy between one's intuition and Hulten's theorem is probably best captured in a speech by Summers (2013): "[...] electricity was only 4% of the economy, and so if you lost 80% of electricity, you couldn't possibly have lost more than 3% of the economy [...] we would understand that somehow, even if we didn't exactly understand it in the model, that when there wasn't any electricity, there wasn't really going to be much economy."

The constraints in the above problem correspond to the resource constraints for good i and the k-th primary factor of production, respectively. The first-order condition of optimality requires that $du/dc_i = \eta_i$, where η_i is the Lagrange multiplier corresponding to good i's resource constraint. Furthermore, applying the envelope theorem to the planner's problem implies that $dW/dz_i = \eta_i f_i(x_{i1}, \ldots, x_{in}, l_{i1}, \ldots, l_{im}) = \eta_i y_i / z_i$. Consequently,

$$\frac{d \log(W)}{d \log(z_i)} = \frac{\eta_i y_i}{W}. \tag{13}$$

On the other hand, the household's optimization problem in the decentralized representation of the equilibrium is given by

$$\begin{aligned} W = \max_{c_i} \quad & u(c_1, \ldots, c_n) \\ \text{s.t.} \quad & \sum_{i=1}^{n} p_i c_i = \sum_{k=1}^{m} w_k h_k, \end{aligned}$$

where w_k is the price of the k-th primary factor of production. First-order conditions imply that $du/dc_i = \phi p_i$, where ϕ is the Lagrange multiplier corresponding to the household's budget constraint. Contrasting this with the corresponding first-order condition from the planner's problem implies that $\eta_i = \phi p_i$. Furthermore, multiplying both sides of the household's first-order condition by c_i, summing over all i, and using the fact that u is homogenous of degree 1 implies that $W = \phi \sum_{i=1}^{n} p_i c_i$. Now replacing for η_i and W in Equation (13) and normalizing the ideal price index to 1 establishes Hulten's theorem in Equation (12).

The above derivations illustrate that equilibrium efficiency and the envelope theorem lie at the heart of Hulten's theorem. In general, a positive productivity shock to industry i impacts aggregate output via two channels. First, it results in an outward shift in the economy's production possibility frontier. Second, it may result in the reallocation of resources across the various firms in the economy. However, when the original allocation is efficient, any aggregate effect due to the resource reallocation channel is second order (by the envelope theorem) and hence can be ignored in a first-order approximation.

This observation also implies that Hulten's theorem may not hold in inefficient economies. Indeed, production network models such as Jones (2013), Bigio and La'O (2017), and Liu (2018), which exhibit some form of distortions or wedges, all violate Equation (12). Extending these results, Baqaee and Farhi (2018b) provide a comprehensive analysis of a shock's first-order impact on aggregate output in a wide class of inefficient economies. In particular, they illustrate that, the shock's first-order impact can be decomposed into two separate terms: (i) a term that accounts for the shock's "pure" technology effect and (ii) an additional term that accounts for changes in the economy's allocative efficiency.

We also remark that while Hulten's theorem establishes that Domar weights are sufficient statistics for how industry-level shocks impact aggregate output, these weights are endogenous objects that are determined in equilibrium. In fact, as Theorem 2 illustrates, even in the very simple economy of Section 2.1 with Cobb-Douglas preferences and technologies, Domar weights depend on the economy's production network (via its Leontief inverse) and the household's preferences.

Finally, it is important to bear in mind that Hulten's theorem maintains that Domar weights are sufficient statistics for microeconomic shocks' aggregate impact only up to a first order. This means that while Equation (12) can be a reasonable approximation when either the shocks are small or the economy does not exhibit significant nonlinearities, it may be a fairly poor approximation more generally.[8] Baqaee and Farhi (2018c) explore the role of such nonlinearities by extending Hulten's theorem to include the second-order effects of microeconomic shocks on aggregate output. Focusing on a general economy with a nested CES structure, they illustrate that these second-order terms depend on the economy's production network, the elasticities of substitution at various CES nests, and the degree to which factors can be reallocated across industries. The presence of these nonlinearities (which include, but are not restricted to, the second-order effects) are at the core of the apparent disparity between Hulten's theorem and one's intuition regarding network linkages mentioned earlier: while Hulten's theorem is a statement about the shocks' first-order effects, the economy's production network can manifest itself via significant nonlinear effects captured by the higher-order terms.

2.4 Frictions and Market Imperfections

As already emphasized, propagation of productivity shocks in the perfectly competitive models of Sections 2.1 and 2.3.1 occurs via two channels. First, a negative shock to industry i results in an increase in the price of good i, thus increasing the production cost of industries that use i as an input for production. Second, the increase in i's price may also induce the customer industries to readjust their demand for other intermediate inputs. These observations imply that departures from the assumption of perfect competition that either (i) distort the input usage of customer industries or (ii) modify prices' responsiveness to the shocks can reshape propagation patterns over the network.

The simplest departure from the assumption of perfect competition is the introduction of exogenous wedges — say, in the form of markups — between firms' marginal revenue and marginal costs that distort their input and output choices away from efficient levels. This is the approach adopted by Jones (2013), Bigio and La'O (2017), Liu (2018), and Fadinger et al. (2018), who investigate how the production network interacts with productivities and wedges in the determination of aggregate outcomes in the Cobb-Douglas economy of Section 2.1. They find that such exogenous wedges result in misallocation of resources, which in turn manifests itself as reductions in the economy's allocative efficiency and aggregate TFP. One consequence of focusing on a Cobb-Douglas economy, however, is that productivity shocks do not impact the allocation of resources across the economy. This means that propagation patterns in the distorted and undistorted economies are identical. In particular, if μ_k denotes the markup charged by industry k to all its customers, then

$$\frac{d}{d\mu_k}\left(\frac{d\log(y_i)}{d\epsilon_j}\right) = 0$$

for all i, j, and k.

[8] Put differently, Equation (12) is obtained under the assumption that the shocks' impact on the Domar weights themselves is negligible. In general, however, the sales share of an industry may respond significantly to shocks. This observation also illustrates why Hulten's theorem holds globally (i.e., regardless of the size of the shocks) in the baseline model of Section 2.1 (Equation (7)): in the special case that all preferences and technologies are Cobb-Douglas, Domar weights are independent of the realization of productivity shocks, which implies that Hulten's first-order approximation is in fact exact.

The interaction between distortions and productivities are investigated by Caliendo et al. (2018), who consider a model of the world economy with CES technologies, and Baqaee and Farhi (2018a,b), who provide first-order approximations to the impact of productivity shocks in a fairly general class of economies while maintaining the wedges as exogenously given model primitives. Baqaee and Farhi (2018a) characterize how, away from the Cobb-Douglas benchmark, the presence of distortions can change the economy's allocative efficiency and hence the productivity shocks' propagation patterns over the network.

While simple, reduced-form exogenous wedges or markups are not adequate for capturing how specific market imperfections shape propagation dynamics. Such an analysis requires a micro-founded model for the interaction between shocks and wedges. Grassi (2017) takes a first step in this direction by considering a model of production networks with oligopolistic market structures. In such an environment, firm-level productivity shocks affect not only prices but also markups via changes in the firms' competitiveness vis-à-vis other firms in the same industry. This means that the responsiveness of prices to shocks is no longer invariant to the realization of shocks, thus impacting the extent of downstream propagation. Furthermore, changes in market concentration impact industries' demand for intermediate inputs, thus inducing an upstream propagation channel that would be absent in a model with exogenous markups. In the same spirit, Baqaee (forthcoming) endogenizes the mass of firms active in each industry in the context of an economy with imperfect competition and external economies of scale due to firm entry and exit. He shows that exits in an industry can change the profitability of firms in other industries and hence trigger endogenous adjustments in the mass of active firms. This creates an amplification channel in the form of upstream and downstream cascades of exits.

2.5 Endogenous Production Networks

Our discussion up to this point was based on the assumption that while input-output linkages can function as a shock propagation mechanism, the structure of the production network itself is invariant to the shocks. In reality, however, firms systematically respond to changes in economic conditions by altering their trading partners. For instance, they may source new inputs to take advantage of technological innovations or may enter into relationships with new customers in response to a customer's exit. Such endogenous changes in the production network can, in turn, significantly alter the economy's response to exogenous disturbances.

To accommodate the production network's response to shocks, a small but growing literature focuses on developing a joint theory of production and endogenous network formation. Developing such a theory, however, faces a central challenge. The complexity inherent to direct and indirect network effects coupled with the combinatorial nature of graphs means that the relevant state space for firm-level decision making can become prohibitively large, even in fairly small economies consisting of a handful of firms.

A first set of papers sidestep this challenge by proposing statistical models of network formation. Atalay et al. (2011) develop a model in which links between firms are created through a variant of the preferential attachment model, while Carvalho and Voitgländer (2015) propose an industry-level network formation model based on the friendship model of Jackson and Rogers (2007), according to which existing input-output linkages are used to search for new inputs for production. Using industry-level data, and consistent with the model's central mechanism, they find that producers are more likely to adopt inputs that are already in use by their current (direct or indirect) upstream suppliers.

While statistical models like the ones mentioned above are able to match some of the key attributes of real-world production networks, by their nature, they abstract from firms' link formation incentives. These incentives are explicitly incorporated by Oberfield (2018) into a dynamic model of network formation in which producers optimally choose one input from a randomly evolving set of suppliers. He finds that such endogenous choice results in the emergence of star suppliers that sell their goods to many other firms for intermediate use. Oberfield overcomes the "curse of dimensionality" discussed earlier by (i) considering an economy consisting of a continuum of firms and (ii) restricting attention to single-input production technologies. These assumptions simplify the analysis by guaranteeing that equilibrium production networks that exhibit cycles are of measure zero.

Acemoglu and Azar (2018) consider an alternative model in which firms in each one of n industries decide which subset of the other $n-1$ industries to use as input-suppliers, with each input combination leading to a different constant returns to scale production technology. The key assumption in the model is that markets are "contestable" in the sense that a large number of firms have access to the same menu of technologies. This assumption ensures that, when choosing its input combination, each firm can take the production network and all prices as given, thus bypassing complex strategic considerations of how its choice may reverberate through the network. In such an environment, aside from its standard effect of reducing all downstream prices (relative to the wage w), a positive technology shock to an industry alters the incentives of firms in downstream industries to adopt a wider set of inputs.

A related propagation mechanism is explored by Taschereau-Dumouchel (2018), who develops a firm-level model of network formation in which firms exit if they cannot meet fixed costs of production. Such extensive margin adjustments create strong complementarities between the firms' operating decisions: a negative shock that results in a firm's exit reduces the profitability of its suppliers and customers, thus creating the potential for a cascade of shutdowns that changes the shape of the production network.[9] The resulting propagation mechanism implies that periods of low economic activity feature a less clustered production network, a prediction that is consistent with the data.[10]

3 The Network Origins of Aggregate Fluctuations

The discussion in the previous section illustrates that the economy's production network can function as a mechanism for propagating shocks from one firm or industry to the rest of the

[9]The firm-level nature of the model, alongside the binary decision faced by the firms, makes the model analytically and computationally intractable. However, Taschereau-Dumouchel (2018) illustrates that under certain conditions, a relaxed version of the social planner's problem can be solved numerically.

[10]A related set of papers, such as Antràs and Chor (2013), Chaney (2014), Antràs, Fort, and Tintelnot (2017), and Tintelnot et al. (2018) studies firms' sourcing decisions in the international trade context. See Chaney (2016), Johnson (2018), and Bernard and Moxnes (2018) for a general overview of network models in international trade.

economy. But can such a propagation mechanism translate idiosyncratic microeconomic shocks into significant fluctuations at the aggregate level? The answer to this question can shed light on whether macroeconomic fluctuations can have their origins in idiosyncratic shocks to individual firms or disaggregated industries.

Going as far back as Lucas (1977), the possibility that significant fluctuations in aggregate economic variables may originate from microeconomic shocks was downplayed by the literature. This dismissal was based on a "diversification argument," which maintained that in an economy consisting of n industries hit by independent shocks, the standard deviation of aggregate fluctuations would be roughly proportional to $1/\sqrt{n}$, a negligible effect at high levels of disaggregation (corresponding to large values of n). This argument, however, ignores the possibility that shocks may propagate from one firm or industry to another over input-output linkages: with such a propagation mechanism at work, sectoral outputs would be correlated and hence may not wash out upon aggregation, even when the shocks themselves are independent.

In this section, we use Theorems 1 and 2 to revisit Lucas's argument and characterize the conditions under which input-output linkages in the economy can indeed generate sizable aggregate fluctuations from purely idiosyncratic shocks. We do so in two steps. We first use Equation (7) to relate the economy's aggregate volatility to the distribution of sectoral Domar weights in the economy. We then rely on Equation (8) to provide a characterization of "network-originated" macro fluctuations in terms of the economy's production network structure.

3.1 Micro Shocks and Macro Fluctuations

To illustrate the key ideas in the most transparent manner, we impose the following regularity assumptions on the baseline model from Section 2.1. First, we assume that the productivity shocks $\epsilon_i = \log(z_i)$ are independent and identically distributed with mean zero and finite standard deviation σ, thus ensuring that the economy is only subject to industry-level idiosyncratic shocks. Second, we suppose that the share of labor is the same across all industries, i.e., $\alpha_i = \alpha$ for all i.

Equation (7) implies that the aggregate volatility that is due to idiosyncratic microeconomic shocks is given by

$$\sigma_{\text{agg}} = \text{stdev}(\log(\text{GDP})) = \sigma\|\lambda\|, \tag{14}$$

where $\|\lambda\| = (\sum_{i=1}^n \lambda_i^2)^{1/2}$ denotes the second (uncentered) moment of Domar weights. Since Domar weights sum up to $\sum_{i=1}^n \lambda_i = \sum_{i,j=1}^n \beta_j \ell_{ji} = 1/\alpha$, we can then rewrite the above equation as

$$\sigma_{\text{agg}} = \frac{\sigma/\alpha}{\sqrt{n}}\sqrt{1 + n^2\alpha^2 \text{var}(\lambda_1, \ldots, \lambda_n)}. \tag{15}$$

This relationship has two immediate implications. First, it implies that when all Domar weights are identical, σ_{agg} is proportional to $1/\sqrt{n}$, consistent with the diversification argument. Second, the fact that in general σ_{agg} depends on the variance of Domar weights indicates that the argument put forth by Lucas may break down if sectoral Domar weights exhibit significant heterogeneity. In particular, Equation (15) illustrates that, all else equal, more dispersion in Domar weights results in higher levels of aggregate volatility emerging from purely idiosyncratic shocks.

These observations are at the heart of what Gabaix (2011, 2016) refers to as the *granularity hypothesis*: in the presence of significant heterogeneity at the micro level, the incompressible "grains" of economic activity (comprised of firms or disaggregated industries) can matter for the behavior of macroeconomic aggregates. This is driven by the fact that such heterogeneity reduces the extent to which various shocks cancel each other out at the aggregate level. Importantly, Gabaix (2011) also shows that when the distribution of Domar weights is sufficiently heavy-tailed, aggregate volatility can be significantly larger than Lucas's $1/\sqrt{n}$ benchmark, even at high levels of disaggregation. For example, suppose that Domar weights have a Pareto distribution with exponent $\gamma \geq 1$, in the sense that the fraction of industries with Domar weights greater than any given λ is proportional to $\lambda^{-\gamma}$, with a smaller γ corresponding to more heterogeneity in Domar weights. One can show that when $\gamma \in (1, 2)$, then $\|\lambda\|$ in Equation (14) is proportional to $n^{1/\gamma-1}$ as $n \to \infty$.[11] Thus, a sufficiency skewed distribution of Domar weights can result in significantly higher levels of aggregate volatility compared to the benchmark of $1/\sqrt{n}$.

3.2 Network-Originated Macroeconomic Fluctuations

We now turn to our main question of interest, namely, whether the economy's production network can translate idiosyncratic shocks into sizable macroeconomic fluctuations.

We first note that while Equation (15) readily establishes that the extent of micro-originated GDP fluctuations is tightly linked to the heterogeneity in Domar weights, these weights are endogenous objects that are determined in equilibrium. We thus use Equation (8) to obtain a characterization in terms of the economy's structural parameters. Furthermore, to isolate the role of input-output linkages, we normalize the preference shares such that $\beta_i = 1/n$ for all i. Such a normalization ensures that any heterogeneity in Domar weights only reflects differences in the roles of different industries in the economy's production network. In particular, $\lambda_i = v_i/n$, where $v_i = \sum_{j=1}^{n} \ell_{ji}$ is the i-th column sum of the economy's Leontief inverse and measures the importance of industry i as a direct or indirect input-supplier to all sectors in the economy.

We can now use Equations (8) and (15) to relate the volatility of log output in this economy to its production network:

$$\sigma_{\text{agg}} = \frac{\sigma}{\sqrt{n}} \sqrt{\alpha^{-2} + \text{var}(v_1, \dots, v_n)}. \tag{16}$$

Equation (16) recovers the key insight of Acemoglu et al. (2012): sufficient heterogeneity in various industries' roles as input-suppliers can lead to significantly higher levels of aggregate volatility compared to the $1/\sqrt{n}$ rate predicted by the diversification argument. For example, if the v_i's have a Pareto distribution with exponent $\gamma \in (1, 2)$, then σ_{agg} will be proportional to $n^{1/\gamma-1}$ as $n \to \infty$. The intuition underlying this result is tightly linked to the nature of the propagation mechanism as captured by Theorems 1 and 2. Microeconomic shocks wash out at the aggregate level if they impact the aggregate output roughly symmetrically. But when industries are highly asymmetric in their roles

[11] If $\gamma > 2$, then $\|\lambda\|$ scales as $1/\sqrt{n}$ as $n \to \infty$, whereas in the knife-edge case of $\gamma = 2$, it scales as $\sqrt{\log(n)/n}$. See the proofs of Proposition 2 of Gabaix (2011) and Corollary 1 of Acemoglu et al. (2017) for detailed derivations.

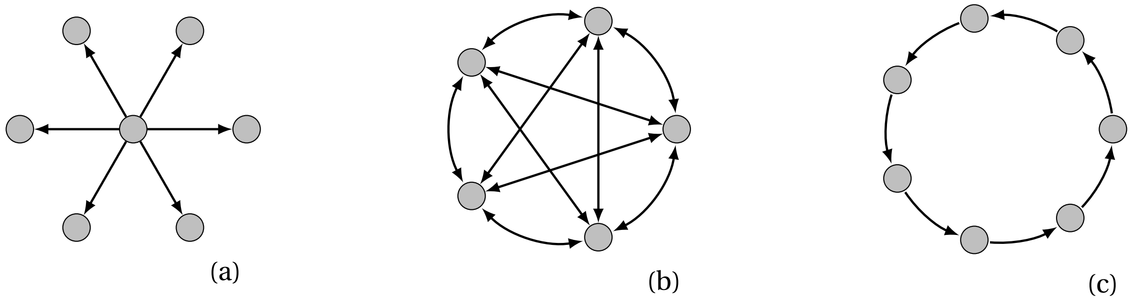

Figure 2. Production networks corresponding to three economies with non-trivial input-output linkages. Each vertex corresponds to an industry, with a direct edge present from one vertex to another if the former is an input-supplier to the latter.

as input-suppliers, shocks to industries that are more important suppliers propagate more widely and hence do not wash out with the rest of the shocks upon aggregation.

To further clarify how input-output linkages may shape aggregate volatility, it is instructive to consider the graph-theoretic interpretations of vector $v = (v_1, \dots, v_n)$ and Equation (16). By its definition, the Leontief inverse satisfies $L = I + LA$. Consequently, $v_i = \sum_{j=1}^{n} \ell_{ji}$ can be expressed in a recursive form as

$$v_i = 1 + \sum_{j=1}^{n} a_{ji} v_j.$$

This representation indicates that v_i coincides with the so-called *Bonacich centrality* of vertex i in the graph that represent's the economy's production network (Bonacich, 1987): an industry i is more "central" in the production network if it is a more important input-supplier to other central industries. Thus, according to this interpretation, Equation (16) establishes that microeconomic shocks can generate sizable aggregate fluctuations when the economy's production network consists of industries with widely disparate centralities. Figure 2(a) provides an example of one such economy (albeit an extreme one), in which a single industry serves as the sole input-supplier to all other industries. As such, microeconomic shocks to this central industry propagate widely throughout the economy and hence do not wash out with the rest of the shocks. In fact, it is easy to verify that, among all economies with the same labor share α, the star network in Figure 2(a) exhibits the maximal dispersion in industrial centralities, $\mathrm{var}(v_1, \dots, v_n)$. The economies depicted in Figures 2(b) and 2(c) are at the other end of the spectrum, with all industries taking symmetric roles as input-suppliers in the economy. Thus, despite the fact that micro shocks propagate over these networks, the fact that they propagate symmetrically means that they will cancel each other out, leading to minimal aggregate effects. In fact, by Equation (16), the volatility of aggregate fluctuations driven by idiosyncratic shocks in these economies is proportional to $1/\sqrt{n}$.[12]

[12] Whether input-output linkages can turn microeconomic shocks into sizable aggregate fluctuations was debated by Horvath (1998, 2000) and Dupor (1999) back in the 1990's, with the former arguing in favor, while the latter providing analytical results for a broad class of economies with fairly dense production networks — such as Figure 2(b) — in which micro shocks wash out at a fairly rapid rate upon aggregation. Our discussion above and Equation (16) illustrate that Dupor's results were driven by his focus on economies in which all industries have identical centralities (despite the presence of non-trivial input-output linkages).

3.3 Comovements

Our discussion thus far shows that sufficient heterogeneity in Domar weights can translate microeconomic shocks to highly disaggregated industries into macroeconomic fluctuations. Importantly, this is the case regardless of whether such heterogeneity is driven by asymmetry in the economy's production network (as in Section 3.2) or is due to other reasons (for example, as a result of heterogeneity in preference shares $(\beta_1, \ldots, \beta_n)$). By now, however, it should be clear that even when the source of heterogeneity in Domar weights may not matter for aggregate fluctuations, economies that exhibit higher levels of "network heterogeneity" exhibit higher levels of comovements: propagation of shocks over the economy's production network increases the likelihood that more industries move in tandem over the business cycle.

To formalize this statement, we consider the benchmark model from Section 2.1 under the restriction that the economy's input-output matrix is a symmetric circulant matrix with diagonals that are greater than $1/n$.[13] While somewhat restrictive, the focus on this subclass of economies provides us with enough symmetry to present the key ideas in the most transparent manner. We maintain the assumption that microeconomic shocks $(\epsilon_1, \ldots, \epsilon_n)$ are independent and identically distributed with mean zero and standard deviation σ.

Given two economies in this class with input-output matrices A and $\tilde{A}$ and identical Domar weights ($\lambda_i = \tilde{\lambda}_i$ for all i) and labor shares ($\alpha_i = \tilde{\alpha}_i = \alpha$ for all i), we say the latter economy is *more interconnected* than the former if $\tilde{a}_{ij} = \gamma a_{ij} + (1-\gamma)(1-\alpha)/n$ for all pairs of industries i and j and some $\gamma \in [0, 1]$. Under this definition, the intensity of input-output linkages between any two industries in the more interconnected economy is more evenly distributed, with a small value of γ corresponding to an economy that is more similar to the complete network in Figure 2(b).[14] We have the following novel result, the proof of which is provided in Supplemental Appendix A.

Theorem 4. *Consider a pair of economies with identical Domar weights and suppose that the latter is more interconnected than the former. Then,*

(a) the average pairwise correlation of (log) outputs is higher in the more interconnected economy;

(b) the industries in the less interconnected economy are more volatile.

Statement (a) of the above result thus formalizes our earlier claim regarding the importance of input-output linkages in creating comovements across different industries: given two economies with identical Domar weights, the one with higher levels of interconnectivity leads to higher average pairwise correlations, despite the fact that the two economies are indistinguishable at the aggregate level. Statement (b) then establishes that this increase in comovements is coupled with a reduction in sectoral volatilities. This is a consequence of the fact that, in a more interconnected economy with a more even distribution of input-output linkages, each industry is more diversified with respect to the upstream risk emanating from its suppliers, its suppliers' suppliers, and so on.

[13] A matrix is said to be circulant if each row is a single-element rotation of the previous row.

[14] While related, this notion of interconnectivity is distinct from the one defined in Acemoglu et al. (2017). The transformation $\tilde{a}_{ij} = \gamma a_{ij} + (1-\gamma)(1-\alpha)/n$ coincides with the concept of γ-convex combination of two networks in Acemoglu et al. (2015).

Taken together, the two parts of Theorem 4 illustrate a key distinction between the nature of economic fluctuations in (i) an economy with high levels of network heterogeneity and (ii) an economy with an identical Domar weight distribution but with low levels of network heterogeneity. Aggregate fluctuations in the latter economy arise as a consequence of fluctuations in sectors with high Domar weights. In contrast, aggregate fluctuations that arise from the interplay of microeconomic shocks and the production network exhibit significant comovements across a wide range of sectors within the economy.

3.4 Macroeconomic Tail Risks

Our focus in the previous subsections was on how input-output linkages can shape (i) the economy's aggregate volatility and (ii) the comovements between various industries as measured by their variance-covariance matrix. The economy's production network, however, has implications for the distribution of sectoral and aggregate outputs, well beyond their second moments.

These implications are the focus of Acemoglu et al. (2017), who show that the economy's production network may fundamentally reshape the distribution of output by increasing the likelihood of large economic downturns from infinitesimal to substantial. Such an analysis requires a notion for measuring "macroeconomic tail risks." A natural candidate is to measure macro tail risks in terms of systematic departures in the frequency of large contraction in aggregate output from what would prevail under a normal distribution with an identical variance. The central result is that an economy with a non-trivial production network that is subject to thin-tailed shocks may exhibit deep recessions as frequently as economies that are subject to heavy-tailed shocks. Importantly, whereas the second moment of the distribution of Domar weights (i.e., $\|\lambda\|$ in Equation (14)) is a sufficient statistic for the extent of micro-originated volatility, the extent of macroeconomic tail risks is determined by a statistic that also depends on the largest Domar weight in the economy. This disparity implies that the role of production networks in generating macroeconomic tail risks is distinct from their role in generating high levels of aggregate volatility. Hence, macroeconomic tail risks may vary significantly even across economies that exhibit otherwise identical behavior for moderate deviations.

4 Empirical and Quantitative Studies

In the previous sections, we used a simple model and a few of its variants to present the theoretical foundations for the role of production networks in macroeconomics. In this section, we provide a brief guide to the literature that explores these themes empirically and quantitatively.

4.1 Properties of Production Networks

We start with an overview of some of the well-documented stylized facts concerning various firm- and industry-level production networks.

Perhaps the most widely used industry-level data are the Input-Output Accounts Data compiled by the Bureau of Economic Analysis (BEA). This database contains the most disaggregated sectoral data available worldwide, providing a detailed breakdown of the U.S. economy into hundreds of industries.

As documented by Carvalho (2010, 2014) and Acemoglu et al. (2012), among others, the BEA data indicate that the U.S. industry-level production network exhibits a few notable properties. First, the industry-level network is highly sparsely connected, in the sense that narrowly-defined specialized industries supply inputs to only about 11 other industries on average. Second, it is dominated by a small number of hubs: general purpose industries that supply a wide range of industries in the economy. This is reflected in a highly skewed distribution of (weighted) outdegrees, which is well-approximated by a Pareto distribution. Third, the production network exhibits what has come to be known as the "small-world" property, where though most industry-pairs are not directly linked by an input-supply relation, they are indirectly linked by hub-like sectors, resulting in a network with short average path length distance and a small diameter. Finally, the production network exhibits a highly skewed distribution of sectoral (Bonacich) centralities, also well-approximated by a Patero distribution with diverging second moments. As reviewed in Section 3.2, this last property indicates significant enough heterogeneity in centralities for the breakdown of the diversification argument, implying the possibility that micro shocks generate sizable aggregate fluctuations. As we shall see below, this possibility is indeed confirmed by a host of quantitative studies.

Industry-level input-output data are also available for many other countries, albeit at considerably coarser levels. The STAN database (containing benchmarked input-output data for 47 industries across 37 OECD countries) and the Global Trade Analysis database (with better coverage of low-income countries at a slightly higher level of aggregation) allow for cross-country comparative studies of production networks. Using these data, McNerney et al. (2013), Blöchl et al. (2011), and Fadinger et al. (2018) document that, consistent with the patterns in the U.S., the distributions of sectoral outdegrees and centralities are highly heterogeneous in a wide range of countries. In addition, Blöchl et al. (2011) document that different groups of countries cluster around different central industries, while Fadinger et al. (2018) document that central industries in richer countries are relatively less productive.

The recent availability of large scale firm-level transactions data has made analyses at a more granular level possible. One of the most extensive of such datasets is from a large (private) credit reporting agency in Japan, named Tokyo Shoko Research (TSR), which in the course of issuing credit scores for firms, obtains the identity of firms' customers and suppliers. This yields information on the buyer-supplier relations of close to a million firms, virtually covering the universe of Japanese firms with more than 5 employees. Another important source of firm-level transactions data is value-added tax (VAT) records from countries where such tax is levied. The tax authorities require the reporting of transactions between any two VAT-liable entities. The best studied of such datasets are based on VAT records from Belgium, containing the universe of all domestic supplier-customer relations at the firm level. Relative to the Japanese data mentioned above, this data is richer as it also contains the transaction amounts associated with each of the firm-to-firm links. Carvalho et al. (2016) and Bernard et al. (forthcoming) report some stylized facts emerging from the Japanese data, while Bernard et al. (2018) do the same for Belgium. While based on two different countries, these studies suggest a number of salient characteristics of firm-level production networks. First, as in the case of industries, the firm-level networks exhibit extensive heterogeneity in the role of firms as input-suppliers, with outdegree distributions that are close to Pareto. Second, and this time in contrast to industry-level networks, the indegree distributions are also very skewed, indicating the presence of firms that rely on a large number of suppliers. Third, larger firms in terms of sales or employees are the firms with larger numbers of buyers and suppliers. Finally, geographical distance is an important determinant in firm-to-firm link formation, with most linkages being local.

Unfortunately, the structure of the U.S. firm-level production network has received less attention as data is more scant. The most widely used U.S. dataset on buyer-supplier relations comes from the Compustat database, which is based on the financial accounting regulations that require publicly-listed firms to disclose the identity of any customer representing more that 10% of their reported sales. Clearly, this induces a double selection bias: the data only contains linkages with publicly-traded firms at both ends and typically correspond to small firms supplying to relatively larger customers. Nonetheless, the data can still provide valuable information about the granular nature of production in the economy. For example, Atalay et al. (2011) are able to document that, as in Japan and Belgium, the indegree distribution of the network of publicly-listed firms in the U.S. is also highly skewed.

4.2 Propagation Patterns

We next review some of the empirical evidence on the propagation of shocks at the industry and firm levels.

4.2.1 Industry-Level Evidence

Acemoglu, Akcigit, and Kerr (2016) provide a first pass at testing the propagation mechanism implied by the baseline model in Section 2.1 at the industry level. Their starting point is the expression in Equation (6) for the equilibrium output of each industry as a function of the economy's production network and microeconomic productivity shocks. Taking first differences, this expression implies that

$$\Delta \log(y_i) = \Delta\epsilon_i + \sum_{j=1}^{n}(\ell_{ij} - \mathbb{I}_{\{j=i\}})\Delta\epsilon_j,$$

thus decomposing the output growth of industry i into an "own effect" (the result of i's own productivity shock, $\Delta\epsilon_i$) and a "network effect" due to the propagation of shocks to other industries. They operationalize this decomposition by constructing the Leontief inverse using the input-output tables compiled by the BEA and sourcing detailed sectoral output data from the NBER-CES manufacturing industry database. As a proxy for productivity shocks, they use lagged realizations of sector-level TFP growth from the same database so as to minimize concerns regarding contemporaneous joint determination of output and TFP. Combining input-output data with this candidate measure for shocks yields the main regressor of interest, defined as $\text{Downstream}_{i,t-1} = \sum_{j=1}^{n}(\ell_{ij} - \mathbb{I}_{\{j=i\}})\Delta\text{TFP}_{j,t-1}$. This is a weighted average of shocks hitting i's direct and indirect suppliers, using the entries of the Leontief inverse as weights (as instructed by the model). The labelling of this regressor reflects our discussion in Section 2.1 that in a Cobb-Douglas economy productivity shocks should only propagate downstream. This in turn implies that the corresponding upstream measure, $\text{Upstream}_{i,t-1} = \sum_{j=1}^{n}(\ell_{ji} - \mathbb{I}_{\{j=i\}})\Delta\text{TFP}_{j,t-1}$, should have no effect on i's output dynamics. Finally, define $\mathrm{Own}_{i,t-1} = \Delta\mathrm{TFP}_{i,t-1}$ as an industry's own direct productivity shock. The following regression can then be used to test the propagation patterns implied by our baseline setup:

$$\Delta \log(y_{it}) = \delta_t + \psi \Delta \log(y_{it-1}) + \beta_{\text{own}} \mathrm{Own}_{it-1} + \beta_d \mathrm{Downstream}_{it-1} + \beta_u \mathrm{Upstream}_{it-1} + \varepsilon_{it}.$$

This specification additionally allows for the presence of lagged dependent variables and year fixed effects to deal with possibly correlated error structures, either across time or in the cross-section.

Consistent with the theory, Acemoglu, Akcigit, and Kerr (2016) find that the downstream network effect of productivity shocks is economically and statistically significant: a one standard deviation increase in TFP growth is associated with a downstream effect of about 6% on output growth. Second, by comparison, the upstream effect of productivity shocks is much smaller economically and its statistical significance is not robust to alternative output measures. These findings are in broad accordance with the predictions of Theorem 1 for the baseline Cobb-Douglas environment.

This simple empirical framework is flexible enough to also test the propagation patterns of demand shocks. Recall from Theorem 3 that, in the baseline Cobb-Douglas economy, demand-side shocks to a given industry should only propagate upstream to its direct and indirect suppliers. Acemoglu, Akcigit, and Kerr (2016) use changes in federal government spending to construct one such shock: they interact an industry's (initial) share of sales to the federal government with aggregate growth of federal spending. As in the previous exercise, the regressors of interest are obtained by lagging the shocks and constructing weighted averages of the shocks across direct and indirect suppliers (for upstream effects) and customers (for downstream effects), with weights given by the corresponding elements of the Leontief inverse. Consistent with the models' predictions, these results indicate significant upstream — rather than downstream — network effects that dominate the industry's own effect.

We conclude by noting that while the above exercise is indicative of propagation patterns that are broadly consistent with the predictions of Theorems 1 and 3, one should be careful in interpreting these estimates as causal. Even though the use of lagged predetermined shocks seeks to minimize endogeneity concerns, these shocks — and in particular, TFP growth — may be endogenous to decisions in the recent past that affect current realizations of the left- and right-hand side variables in the regression equations.

4.2.2 Firm-Level Evidence

A different strand of literature uses more granular data across a host of different countries to document the propagation of shocks at the firm level. These studies serve two important purposes. First, while industry-level evidence for the propagation of shocks indicates that production networks can have empirically relevant implications, ultimately, any actual propagation happens at the level of firms. Therefore, firm-level studies can provide more direct evidence for the nature of the underlying propagation mechanisms. Second, the possibility of identifying arguably exogenous shocks at the firm level (such as localized natural disasters), coupled with the more extensive variation in exposure to such shocks, means that one can overcome the endogeneity concerns that may arise at more aggregated levels.

Barrot and Sauvagnat (2016) investigate the propagation of firm-specific shocks by combining data on the timing and location of major natural disasters (in the form of blizzards, earthquakes, floods, and hurricanes) in the U.S. with information on the physical headquarters' location and supplier-customer linkages of publicly-listed firms from Compustat. Given the limitations of the observable production network in the Compustat database, Barrot and Sauvagnat (2016) focus on "local" propagation patterns from a firm to its immediate suppliers and customers by regressing changes in quarterly sales of firms on a dummy variable capturing whether the firm's direct suppliers were located in a county hit by a natural disaster in a recent quarter. They document that exposures to the natural disaster results in a 2 to 3 percentage point drop in sales growth of the disrupted firm's direct customers. Importantly, this drop is particularly pronounced when the disrupted supplier is producing hard-to-substitute relation-specific inputs, in which case the shock further propagates to other (non-affected) suppliers of the customer firm. This evidence suggests that, while the Cobb-Douglas model may serve as a good approximation at the industry-level, it may break down at the more micro-level, where easily substitutable inputs coexist with relation-specific inputs that are more difficult to substitute (at least in the short run).

A similar pattern is documented by Boehm et al. (forthcoming), who use U.S. Census Bureau micro data to study firm-level cross-country transmission of supply chain disruptions caused by the Great East Japan Earthquake of 2011. Combining reduced-form evidence with structural estimates of production elasticities, they find that the U.S. affiliates of Japanese multinationals experienced a roughly one-for-one decline in output in response to declines in imports. This finding indicates that the short-run elasticity of substitution between imported and domestic inputs is close to zero.

While the above studies provide credible evidence for the propagation of shocks from a firm to its direct suppliers and customers, a shock's impact on the aggregate economy also depends on the extent to which it eventually propagates to more distant, only indirectly-connected, firms. Testing such a hypothesis, however, requires large-scale and detailed information on firm-to-firm linkages across the economy. This is the approach taken by Carvalho et al. (2016), who use the TSR data to trace the disruption caused by the 2011 earthquake and tsunami throughout the Japanese production network. Consistent with the results surveyed above, they find a significant post-earthquake impact on the sales growth rates of firms with direct suppliers in the disaster areas. In addition, they also find that the disruption (i) propagated further downstream to the disaster area firms' indirect customers and (ii) resulted in significant upstream propagation to the direct and indirect suppliers of earthquake-hit firms. The evidence on indirect propagation effects, coupled with the "small world" nature of the production network, suggests that localized disturbances like the earthquake can have non-trivial aggregate consequences: while the individual, firm-level impact of the disruption may not be very large — particularly when considering indirectly-exposed firms — its aggregate effect can be significantly higher when a large fraction of firms in the economy is only two or three input-links away from disrupted firms.

Understanding whether and how shocks propagate in production networks is currently the subject of a fast-expanding literature that combines novel production network data with a host of different shocks, going beyond the early interest in productivity disturbances. Demir, Javorcik, Michalsk, and Örs (2018) study the propagation and amplification of financial shocks by liquidity-constrained firms. Combining extensive VAT firm-to-firm transaction data from Turkey with an unexpected policy change levying a tax on trade credit financing by Turkish importers (which effectively made it more costly to finance input purchases from abroad), they find that liquidity-constrained importers exposed to the shock transmitted it to their downstream customers. Further afield, Carvalho and Draca (2018) use detailed military procurement data by the U.S. government and Compustat data on supply chain linkages for publicly-listed firms to document that an increase in demand expands innovation efforts not only by final demand producers, but also their upstream suppliers through recursive market size effects. Noting that expansionary monetary policy shocks — by acting as final demand shocks — should propagate upstream through the production network, Ozdagli and Weber (2017) investigate the role of such network effects as a possible transmission mechanism of monetary policy shocks. Finally, Auer et al. (forthcoming) show that input-linkages across country-sector pairs contribute systematically to (producer price) inflation comovements across countries.[15]

4.3 Comovements and Aggregate Fluctuations

We next briefly survey the literature aimed at quantifying the role of production networks in generating comovements and aggregate fluctuations.

A reduced-form approach to account for industrial comovements is to appeal to a small number of common factors driving (correlated) output dynamics in many industries. Such an approach entails estimating a so-called approximate factor model on a panel of sectoral output growth rates, $\Delta \log y_t$, in the form of $\Delta \log y_t = \Lambda F_t + u_t$, where F_t is a low-dimensional vector of latent factors, Λ is a matrix of factor loadings of appropriate size, and u_t is a vector of industry-specific disturbances, assumed to satisfy weak cross-sectional dependence. However, recall from our discussion in Section 3.3 and Theorem 4 that production networks can induce significant comovements from purely idiosyncratic industry-specific shocks. Thus, what could appear to the econometrician as "common shocks" may instead be the result of endogenous comovements generated by the equilibrium interactions between various industries in a production network. Properly accounting for such a possibility calls for a structural approach that takes the input-output linkages explicitly into account.

Foerster, Sarte, and Watson (2011) adopt one such structural approach to decompose the dynamics of disaggregated U.S. industrial production indices into components arising from aggregate and industry-specific shocks. They use a dynamic variant of the baseline Cobb-Douglas economy in Section 2.1, featuring capital accumulation, capital goods' linkages across industries, and more general preferences. By inverting the model-implied mapping from disturbances to observables in order to recover the underlying structural shocks, they find that idiosyncratic productivity shocks alone account for 50% of aggregate industrial production fluctuations between 1984 and 2007. Thus, while statistical models may perceive economic dynamics as being led almost solely by common macro shocks, a non-trivial fraction of aggregate fluctuations can instead be traced to idiosyncratic shocks propagating across the production network.[16]

[15] A smaller literature in finance investigates the asset pricing implications of production networks. Some recent examples include Herskovic et al. (2017), Herskovic (2018), and Gofman, Segal, and Wu (2018).

[16] For comparison, Foerster et al. (2011) also show that adopting the reduced-form approach of approximate factor models

While important, these conclusions rely on the assumption that the structural model coincides with the true data generating process and, as a result, on the particular propagation mechanics imparted by Cobb-Douglas technologies. Yet, as we discussed earlier, empirical micro studies of propagation patterns suggest important departures from the Cobb-Douglas benchmark, thus questioning the robustness of the quantitative inferences above. Atalay (2017) tackles this problem by showing how to extend the Foerster et al. (2011) methodology to an economy with CES technologies and preferences. To calibrate his model, Atalay (2017) uses annual input-output tables constructed by the BEA to estimate elasticities of substitution in industries' production functions, obtaining a value of at most 0.2 for the elasticity of substitution between intermediate inputs. The strong complementarity among inputs suggested by these estimates imply stronger propagation and hence more pronounced aggregate effects originating from microeconomic shocks compared to the Cobb-Douglas benchmark. Indeed, under his benchmark parameter estimates, Atalay (2017) concludes that 83% of the variations in aggregate output growth are attributable to idiosyncratic industry-level shocks.

The estimates by Foerster et al. (2011) and Atalay (2017) are also in broad accordance with the earlier attempts of Horvath (2000) and Carvalho (2010) at quantifying the macroeconomic importance of idiosyncratic shocks by directly calibrating large-scale multi-sector models under the assumption of uncorrelated disturbances. Both studies concluded that the interplay of idiosyncratic shocks with input-output linkages can account for about two thirds of aggregate fluctuations.[17] Reassuringly, though exploiting a different variance decomposition methodology, di Giovanni, Levchenko, and Méjean (2014) find similarly sized effects for the contribution of linkages to aggregate volatility. Taken together, this body of works suggests that the economy's production network is a major driver of comovement and GDP fluctuations.

Much like the empirical literature on propagation patterns surveyed earlier, there is an active interest in expanding the range of quantitative insights derived from taking production networks explicitly into account. Baqaee and Farhi (2018c) quantify the effects of CES-induced nonlinearities in production networks and show that such nonlinearities (i) amplify the effects of negative sectoral shocks while mitigating positive shocks; (ii) generate significant negative skewness and excess kurtosis in aggregate output dynamics even when the underlying structural shocks are symmetric and thin tailed; and (iii) can lead to significant welfare costs of business cycles, ranging from 0.2% to 1.3%, an order of magnitude larger than standard estimates in the literature. Bigio and La'O (2017) apply their model of production networks featuring Cobb-Douglas technologies and exogenous wedges to measure the impact of sectoral financial distortions during the Great Recession. They conclude that the production network amplified industry-level financial shocks from 1.7 to 2.4 times more than an equivalent economy with no linkages. Grassi (2017) instead calibrates a model of interlinked oligopolistic market structures and finds that aggregate volatility arising from independent *firm*-level shocks accounts for 34% of what is observed in the data.[18] Relatedly, Magerman et al. (2017) and Kikkawa et al. (2018) exploit extensive Belgium VAT data on firm-to-firm trade to calibrate detailed

would lead one to conclude that two factors account for 87% of variability in aggregate industrial production between 1984 and 2007.

[17]See also the related contributions of Shea (2002) and Conley and Dupor (2003) for earlier studies documenting aspects of sectoral comovement.

[18]See also the related firm-level calibration exercises of Gabaix (2011) and Carvalho and Grassi (forthcoming).

models of firm-level production networks. They conclude, respectively, that firm-level idiosyncratic shocks account for 57% of aggregate volatility and that firm-to-firm prodution networks entail a substantial amount of double marginalization, increasing by about 50% the welfare gains of reducing firm markups relative to a simpler roundabout economy featuring no network. In the same vein, Baqaee and Farhi (2018b) find that eliminating markup distortions entirely in an environment with production networks and CES production functions would raise TFP by 20%. Further afield, Caliendo et al. (forthcoming) consider the interplay between production networks and the spatial structure of production, acknowledging that sectors tend to be spatially agglomerated. Their quantitative framework offers aggregate GDP elasticities to sector-region shocks. Finally, Tintelnot et al. (2018) consider a quantitative model where domestic production networks coexist with international trade and where domestic firm-to-firm linkages can be endogenously rewired in response to international trade shocks. They find that introducing network formation attenuates the costs of large negative trade shocks while amplifying the gains from trade following large positive ones.[19]

5 Concluding Remarks

In this article, we provided a brief overview of the growing theoretical and empirical literature on the role of production networks in shaping economic outcomes. We relied on a simple benchmark model and several of its variants to illustrate how production networks can (i) function as a mechanism for the propagation of shocks throughout the economy and (ii) translate microeconomic shocks into sizable fluctuations in macroeconomic aggregates. We also surveyed the literature that tests these mechanisms empirically and quantifies their implications. We conclude by discussing several open questions and promising avenues for future research.

While supplier-customer relationships that give rise to a production network are formed at the level of firms, most of the literature focuses on models that are better approximations to the nature of these interactions at the industry level. In particular, aside from a few exception discussed in the previous sections, the literature abstracts from important issues such as firm-specific relationships, market power, endogenous formation of supplier-customer linkages, and the possibility of firm failures. This is despite the fact that firm-level forces can have non-trivial implications for the micro and macro dynamics of production networks. Developing models that take such firm-level forces seriously can help capture the theoretical and empirical richness that are currently missing from the literature.

Relatedly, while the literature has mostly focused on how production networks can alter our understanding of the nature of business cycles, the implications for long-term growth have been left largely unexplored. A few studies, such as Ciccone (2002), Jones (2011), and Acemoglu and Azar (2018), have argued for the importance of input-linkages for industrialization and long-run growth. Development of richer firm-level models of production networks coupled with the availability of ever more detailed data can provide fruitful synergies with the resurgent literature on endogenous growth that incorporates extensive heterogeneity at the micro level.

[19] Relatedly, di Giovanni, Levchenko, and Méjean (2018) conclude that the international trade linkages of French firms account for one-third of the comovement between France and the rest of the world.

Finally, another promising avenue for future research is to investigate the implications of input-output linkages in models with nominal rigidities. The contributions of Christiano (2016) and Pasten, Schoenle, and Weber (2018) already suggest that accounting for the network structure of production may result in quantitatively large welfare costs of inflation, affect the slope of the Phillips curve, and alter the real effects of monetary policy. This, in turn, may have implications for the design of optimal monetary policy. Ascertaining the theoretical and quantitative relevance of production networks for the conduct of monetary policy can be of first-order importance for policymakers.

Disclosure Statement

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

Acknowledgements

We are grateful to Daron Acemoglu for various discussions and conversations.

References

Acemoglu, Daron, Ufuk Akcigit, and William Kerr (2016), "Networks and the macroeconomy: An empirical exploration." In *National Bureau of Economic Research Macroeconomics Annual* (Martin Eichenbaum and Jonathan Parker, eds.), volume 30, 276–335, University of Chicago Press.

Acemoglu, Daron and Pablo D. Azar (2018), "Endogenous production networks." Working paper.

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2012), "The network origins of aggregate fluctuations." *Econometrica*, 80(5), 1977–2016.

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2015), "Systemic risk and stability in financial networks." *American Economic Review*, 105(2), 564–608.

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2017), "Microeconomic origins of macroeconomic tail risks." *American Economic Review*, 107(1), 54–108.

Antràs, Pol and Davin Chor (2013), "Organizing the global value chain." *Econometrica*, 81(6), 2127–2204.

Antràs, Pol, Teresa C. Fort, and Felix Tintelnot (2017), "The margins of global sourcing: Theory and evidence from U.S. firms." *American Economic Review*, 107(9), 2514–2564.

Atalay, Enghin (2017), "How important are sectoral shocks?" *American Economic Journal: Macroeconomics*, 9(4), 254–280.

Atalay, Enghin, Ali Hortaçsu, James Roberts, and Chad Syverson (2011), "Network Structure of Production." *Proceedings of the National Academy of Sciences*, 108(13), 5199–5202.

Auer, Raphael, Andrei A. Levchenko, and Philip Saure (forthcoming), "International inflation spillovers through input linkages." *Review of Economics and Statistics*.

Baqaee, David R. (forthcoming), "Cascading failures in production networks." *Econometrica*.

Baqaee, David R. and Emmanuel Farhi (2018a), "Macroeconomics with heterogeneous agents and input-output networks." Working paper.

Baqaee, David R. and Emmanuel Farhi (2018b), "Productivity and misallocation in general equilibrium." Working paper.

Baqaee, David R. and Emmanuel Farhi (2018c), "The macroeconomic impact of microeconomic shocks: Beyond Hulten's theorem." Working paper.

Barrot, Jean-Noël and Julien Sauvagnat (2016), "Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks." *Quarterly Journal of Economics*, 131(3), 1543–1592.

Berman, Abraham and Robert J. Plemmons (1979), *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, New York.

Bernard, Andrew B., Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes (2018), "The origins of firm heterogeneity: A production network approach." Working paper.

Bernard, Andrew B. and Andreas Moxnes (2018), "Networks and trade." *Annual Review of Economics*, 10(1), 65–85.

Bernard, Andrew B., Andreas Moxnes, and Yukiko U. Saito (forthcoming), "Production networks, geography, and firm performance." *Journal of Political Economy*.

Bigio, Saki and Jennifer La'O (2017), "Distortions in production networks." Working paper.

Blöchl, Florian, Fabian J. Theis, Fernando Vega-Redondo, and Eric O'N. Fisher (2011), "Vertex centralities in input-output networks reveal the structure of modern economies." *Physical Review E*, 83, 046127.

Boehm, Christoph E., Aaron Flaaen, and Nitya Pandalai-Nayar (forthcoming), "Input linkages and the transmission of shocks: Firm-level evidence from the 2011 Tohoku Earthquake." *Review of Economics and Statistics*.

Bonacich, Phil (1987), "Power and centrality: A family of measures." *American Journal of Sociology*, 92, 1170–1182.

Caliendo, Lorenzo, Fernando Parro, Esteban Rossi-Hansberg, and Pierre-Daniel Sarte (forthcoming), "The impact of regional and sectoral productivity changes on the U.S. economy." *Review of Economic Studies*.

Caliendo, Lorenzo, Fernando Parro, and Aleh Tsyvinski (2018), "Distortions and the structure of the world economy." National Bureau of Economic Research Working Paper No. 23332.

Carvalho, Vasco M. (2010), "Aggregate fluctuations and the network structure of intersectoral trade." Department of Economics and Business, Universitat Pompeu Fabra Economics Working Paper 1206.

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

Carvalho, Vasco M. and Mirko Draca (2018), "Cascading innovation." Working paper.

Carvalho, Vasco M. and Xavier Gabaix (2013), "The great diversification and its undoing." *American Economic Review*, 103(5), 1697–1727.

Carvalho, Vasco M. and Basile Grassi (forthcoming), "Large firm dynamics and the business cycle." *American Economic Review*.

Carvalho, Vasco M., Makoto Nirei, Yukiko Saito, and Alireza Tahbaz-Salehi (2016), "Supply chain disruptions: Evidence from the Great East Japan Earthquake." Becker-Friedman Institute working paper 2017-01.

Carvalho, Vasco M. and Nico Voitgländer (2015), "Input diffusion and the evolution of production networks." NBER Working Paper 20025.

Chaney, Thomas (2014), "The network structure of international trade." *American Economic Review*, 104(11), 3600–3634.

Chaney, Thomas (2016), "Networks in international trade." In *Oxford Handbook of the Economics of Networks* (Yann Bramoulle, Andrea Galeotti, and Brian Rogers, eds.), Oxford University Press, Oxford.

Christiano, Lawrence J. (2016), "Comment on 'Networks and the macroeconomy: An empirical exploration'." *NBER Macroeconomics Annual*, 30(1), 346–373.

Ciccone, Antonio (2002), "Input chains and industrialization." *Review of Economic Studies*, 63(3), 565–587.

Conley, Timothy G. and Bill Dupor (2003), "A spatial analysis of sectoral complementarity." *Journal of Political Economy*, 111(2), 311–352.

Demir, Banu, Beata Javorcik, Tomasz K. Michalsk, and Evren Örs (2018), "Financial constraints and propagation of shocks in production networks." Working paper.

di Giovanni, Julian, Andrei A. Levchenko, and Isabelle Méjean (2014), "Firms, destinations, and aggregate fluctuations." *Econometrica*, 82(4), 1303–1340.

di Giovanni, Julian, Andrei A. Levchenko, and Isabelle Méjean (2018), "The micro origins of international business cycle comovement." *American Economic Review*, 108(1), 82–108.

Dupor, Bill (1999), "Aggregation and irrelevance in multi-sector models." *Journal of Monetary Economics*, 43(2), 391–409.

Fadinger, Harald, Christian Ghiglino, and Mariya Teteryatnikova (2018), "Income differences and input-output structure." Working paper.

Foerster, Andrew T., Pierre-Daniel G. Sarte, and Mark W. Watson (2011), "Sectoral versus aggregate shocks: A structural factor analysis of industrial production." *Journal of Political Economy*, 119(1), 1–38.

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

Gabaix, Xavier (2016), "Comment on 'Networks and the macroeconomy: An empirical exploration'." *NBER Macroeconomics Annual*, 30(1), 336–345.

Gofman, Michael, Gill Segal, and Youchang Wu (2018), "Production networks and stock returns: The role of vertical creative destruction?" Working paper.

Grassi, Basile (2017), "IO in I-O: Size, industrial organization, and the input-output network make a firm structurally important." Working paper.

Herskovic, Bernard (2018), "Networks in production: Asset pricing implications." *Journal of Finance*, 73(3), 1785–1818.

Herskovic, Bernard, Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh (2017), "Firm volatility in granular networks." Working paper.

Horvath, Michael (1998), "Cyclicality and sectoral linkages: Aggregate fluctuations from sectoral shocks." *Review of Economic Dynamics*, 1(4), 781–808.

Horvath, Michael (2000), "Sectoral shocks and aggregate fluctuations." *Journal of Monetary Economics*, 45(1), 69–106.

Hulten, Charles R. (1978), "Growth accounting with intermediate inputs." *The Review of Economic Studies*, 45(3), 511–518.

Jackson, Matthew O. and Brian W. Rogers (2007), "Meeting strangers and friends of friends: How random are social networks?" *American Economic Review*, 97(3), 890–915.

Johnson, Robert C. (2018), "Measuring global value chains." *Annual Review of Economics*, 10(1), 207–236.

Jones, Charles I. (2011), "Intermediate goods and weak links: A theory of economic development." *American Economic Journal: Macroeconomics*, 3(2), 1–28.

Jones, Charles I. (2013), "Misallocation, economic growth, and input-output economics." In *Proceedings of Econometric Society World Congress* (Daron Acemoglu, Manuel Arellano, and Eddie Dekel, eds.), 419–455, Cambridge University Press.

Kikkawa, Ayumu Ken, Glenn Magerman, and Emmanuel Dhyne (2018), "Imperfect competition in firm-to-firm trade." Working paper.

Leontief, Wassily W. (1951), *The Structure of American Economy, 1919-1939: An Empirical Application of Equilibrium Analysis*. Oxford University Press, New York, NY.

Liu, Ernest (2018), "Industrial policies in production networks." Working paper.

Long, John B. and Charles I. Plosser (1983), "Real business cycles." *Journal of Political Economy*, 91(1), 39–69.

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

Magerman, Glenn, Karolien De Bruyne, Emmanuel Dhyne, and Jan Van Hove (2017), "Heterogeneous firms and the micro origins of aggregate fluctuations." National Bank of Belgium Working Paper No. 312.

McNerney, James, Brian D. Fath, and Gerald Silverberg (2013), "Network structure of interindustry flows." *Physica A: Statistical Mechanics and its Applications*, 392(24), 6427–6441.

Oberfield, Ezra (2018), "A theory of input-output architecture." *Econometrica*, 86(2), 559–589.

Ozdagli, Ali and Michael Weber (2017), "Monetary policy through production networks: Evidence from the stock market." Federal Reserve Bank of Boston working paper.

Pasten, Ernesto, Raphael Schoenle, and Michael Weber (2018), "Price rigidity and the origins of aggregate fluctuations." Working paper.

Plemmons, Robert J. (1977), "M-matrix characterizations. I: Nonsingular M-matrices." *Linear Algebra and its Applications*, 18(2), 175–188.

Shea, John (2002), "Complementarities and comovements." *Journal of Money, Credit and Banking*, 34(2), 412–433.

Stewart, Gilbert W. (1998), *Matrix Algorithms: Basic Decompositions*. Society for Industrial and Applied Mathematics.

Summers, Lawrence (2013), "Crises yesterday and today." Speech at 14th Jacques Pollack Annual IMF Research Conference in Honor of Stanley Fischer. http://larrysummers.com/imf-fourteenth-annual-research-conference-in-honor-of-stanley-fischer.

Taschereau-Dumouchel, Mathieu (2018), "Cascades and fluctuations in an economy with an endogenous production network." Working paper.

Tintelnot, Felix, Ken Kikkawa, Magne Mogstad, and Emmanuel Dhyne (2018), "Trade and domestic production networks." Working paper.

A Supplemental Appendix

A.1 CES Production Technologies

Suppose that the production technology of firms in industry i is given by Equation (10). The first-order conditions of firms in industry i are therefore given by

$$l_i = \alpha_i p_i y_i / w \tag{A.1}$$

$$x_{ij} = (1-\alpha_i) a_{ij} p_i y_i p_j{}^{-\sigma_i} \left(\sum_{k=1}^{n} a_{ik} p_k^{1-\sigma_i} \right)^{-1}, \tag{A.2}$$

where we are using the fact that $\alpha_i + \sum_{j=1}^{n} a_{ij} = 1$ for all i. Plugging the above expressions back into the production function of firms in industry i implies that

$$p_i z_i = w^{\alpha_i} \left(\frac{1}{1-\alpha_i} \sum_{k=1}^{n} a_{ik} p_k^{1-\sigma_i} \right)^{(1-\alpha_i)/(1-\sigma_i)}.$$

Taking logarithms from both sides of the above equation leads to the following system of equations

$$\log(p_i/w) = -\epsilon_i + \frac{1-\alpha_i}{\sigma_i - 1} \log \left(\frac{1}{1-\alpha_i} \sum_{k=1}^{n} a_{ik} (p_k/w)^{1-\sigma_i} \right).$$

We make two observations. First, the above system of equations immediately implies that when $\epsilon_i = 0$ for all industries i, then all relative prices coincide with another, that is, $p_i = w$ for all i. Second, differentiating both sides of the above equation with respect to ϵ_j and evaluating it at $\epsilon = 0$ leads to $d\hat{p}_i/d\epsilon_j = -\mathbb{I}_{\{i=j\}} + \sum_{k=1}^{n} a_{ik} d\hat{p}_k/\epsilon_j$, where recall that $\hat{p}_i = \log(p_i/w)$ is the log relative price of good i and $\mathbb{I}$ denotes the indicator function. Rewriting the previous equation in matrix form, we obtain $d\hat{p}/d\epsilon_j = -e_j + A d\hat{p}/d\epsilon_j$, where e_j is the j-th unit vector. Consequently, $d\hat{p}/d\epsilon_j = (I-A)^{-1} e_j$, which in turn can be rewritten as

$$\left. \frac{d\hat{p}_i}{d\epsilon_j} \right|_{\epsilon=0} = -\ell_{ij}. \tag{A.3}$$

The above equation therefore illustrates how shocks to industry j change the relative prices of all other industries up to a first-order approximation.

Next, recall that the market-clearing condition for good i is given by $y_i = c_i + \sum_{j=1}^{n} x_{ji}$. Multiplying both sides by p_i and dividing by GDP implies that

$$\lambda_i = \beta_i + \sum_{k=1}^{n} \omega_{ki} \lambda_k,$$

where $\lambda_i = p_i y_i / \text{GDP}$ is the Domar weight of industry i and $\omega_{ki} = p_i x_{ki} / p_k y_k$. Note that in deriving the above equation, we are using the fact that the household's first-order condition requires that $p_i c_i = \beta_i \text{GDP}$. Differentiating both sides of the above equation with respect to $d\epsilon_j$ implies that

$$\frac{d\lambda_i}{d\epsilon_j} = \sum_{k=1}^{n} \omega_{ki} \frac{d\lambda_k}{d\epsilon_j} + \sum_{k=1}^{n} \lambda_k \frac{d\omega_{ki}}{d\epsilon_j}. \tag{A.4}$$

On the other hand, Equation (A.2) implies that $\omega_{ki} = (1-\alpha_k)a_{ki}p_i^{1-\sigma_k}/\left(\sum_{r=1}^n a_{kr}p_r^{1-\sigma_k}\right)$. Hence, differentiating both sides of this expression, evaluating them at $\epsilon = 0$, and plugging the resulting expression back into the Equation (A.4) implies that

$$\frac{d\lambda_i}{d\epsilon_j} = \sum_{k=1}^n a_{ki}\frac{d\lambda_k}{d\epsilon_j} + \sum_{k=1}^n (1-\sigma_k)a_{ki}\lambda_k\left(\frac{d\hat{p}_i}{d\epsilon_j} - \frac{1}{1-\alpha_k}\sum_{r=1}^n a_{kr}\frac{d\hat{p}_r}{d\epsilon_j}\right).$$

Hence, using Equation (A.3), we obtain

$$\frac{d\lambda_i}{d\epsilon_j} - \sum_{k=1}^n a_{ki}\frac{d\lambda_k}{d\epsilon_j} = \sum_{k=1}^n (\sigma_k - 1)a_{ki}\lambda_k\left(\ell_{ij} - \frac{1}{1-\alpha_k}\sum_{r=1}^n a_{kr}\ell_{rj}\right).$$

Multiplying both sides of the above equation by ℓ_{si}, summing over all s, and noting that $L = (I-A)^{-1}$ leads to

$$\frac{d\lambda_i}{d\epsilon_j} = \sum_{k=1}^n (\sigma_k - 1)\lambda_k\left(\sum_{s=1}^n a_{ks}\ell_{si}\ell_{sj} - \frac{1}{1-\alpha_k}\sum_{r=1}^n a_{kr}\ell_{rj}\sum_{s=1}^n a_{ks}\ell_{ss}\right). \tag{A.5}$$

On the other hand, the fact that $\lambda_i = p_i y_i/\,\text{GDP}$ implies that

$$\frac{d\log y_i}{d\epsilon_j} = -\frac{d\hat{p}_i}{d\epsilon_j} + \frac{1}{\lambda_i}\frac{d\lambda_i}{d\epsilon_j} = \ell_{ij} + \frac{1}{\lambda_i}\frac{d\lambda_i}{d\epsilon_j},$$

where the second equality is a consequence of Equation (A.3). Plugging for $d\lambda_i/d\epsilon_j$ from Equation (A.5) into the above equation leads to Equation (11). □

A.2 Proof of Theorem 4

Consider two economies with symmetric circulant input-output matrices A and $\tilde{A}$ and suppose the former is more interconnected than the latter, that is, there exists a $\gamma \in [0,1]$ such that

$$\tilde{A} = \gamma A + (1-\gamma)(1-\alpha)J,$$

where $J = (1/n)\mathbf{1}\mathbf{1}'$ is a matrix with all entries equal to $1/n$. We first prove statement (b) of the theorem by showing that the above transformation can only decrease the volatility of each industry, i.e., $\text{var}(\tilde{y}_i) \leq \text{var}(y_i)$ for all i. We then use this result to establish statement (a).

Proof of part (b). Recall from Theorem 1 that the output of industry i satisfies $\log y_i = \sum_{i=1}^n \ell_{ij}\epsilon_j$. Under our assumption that all microeconomic shocks are i.i.d. with a common variance $\sigma^2 < \infty$, it is immediate that $\text{var}(\log y_i) = \sigma^2\sum_{j=1}^n \ell_{ij}^2$. Therefore, sectoral log outputs are more volatile in the less interconnected economy (that is, $\text{var}(\log\tilde{y}_i) \leq \text{var}(\log y_i)$ for all i) if and only if $\sum_{j=1}^n \tilde{\ell}_{ij}^2 \leq \sum_{j=1}^n \ell_{ij}^2$ for all i. On the other hand, the assumption that input-output matrices A and $\tilde{A}$ are symmetric and circulant implies that $\sum_{j=1}^n \tilde{\ell}_{ij}^2 = (1/n)\sum_{i,j=1}^n \tilde{\ell}_{ij}^2 = (1/n)\,\text{trace}(\tilde{L}^2)$. Hence, it is sufficient to show that

$$\frac{d}{d\gamma}\text{trace}(\tilde{L}^2)\Big|_{\gamma=1} \geq 0. \tag{A.6}$$

To this end, first note that, by definition, $\tilde{L} = (I - \tilde{A})^{-1}$. Therefore, differentiating $\tilde{L}^2$ with respect to γ leads to

$$\begin{aligned} d\tilde{L}^2/d\gamma &= \tilde{L}^2\big(2d\tilde{A}/d\gamma - \tilde{A}(d\tilde{A}/d\gamma) - (d\tilde{A}/d\gamma)\tilde{A}\big)\tilde{L}^2 \\ &= \tilde{L}^2(d\tilde{A}/d\gamma)\tilde{L} + \tilde{L}(d\tilde{A}/d\gamma)\tilde{L}^2. \end{aligned}$$

On the other hand, $d\tilde{A}/d\gamma = A - (1-\alpha)J$. Consequently,

$$\begin{aligned} \frac{d\tilde{L}^2}{d\gamma}\Big|_{\gamma=1} &= L^2AL + LAL^2 - (1-\alpha)(L^2JL + LJL^2) \\ &= 2(L^3 - L^2) - 2(1-\alpha)\alpha^{-3}J, \end{aligned}$$

where the second equality uses $LA = AL = L - I$ and the fact that the row and column sums of L are equal to $1/\alpha$, i.e., $L\mathbf{1} = L'\mathbf{1} = (1/\alpha)\mathbf{1}$. Hence,

$$\frac{d}{d\gamma}\,\mathrm{trace}(\tilde{L}^2)\Big|_{\gamma=1} = 2\,\mathrm{trace}(L^3) - 2\,\mathrm{trace}(L^2) - 2(1-\alpha)/\alpha^3.$$

Note that the trace of a matrix is equal to the sum of its eigenvalues. Furthermore, the fact that $L = (I-A)^{-1}$ implies that $\lambda_k(L) = (1-\lambda_k(A))^{-1}$, where $\lambda_k(L)$ and $\lambda_k(A)$ are the k-th largest eigenvalues of L and A, respectively. Consequently,

$$\frac{d}{d\gamma}\,\mathrm{trace}(\tilde{L}^2)\Big|_{\gamma=1} = 2\sum_{k=1}^{n}\frac{1}{(1-\lambda_k(A))^3} - 2\sum_{k=1}^{n}\frac{1}{(1-\lambda_k(A))^2} - 2(1-\alpha)/\alpha^3 = 2\sum_{k=2}^{n}\frac{\lambda_k(A)}{(1-\lambda_k(A))^3}.$$

The second equality above is a consequence of the fact that the row sums of matrix A are all equal to $1-\alpha$, and hence, by the Perron-Frobenius theorem, its largest eigenvalue is given by $\lambda_1(A) = 1-\alpha$. Multiplying and dividing the right-hand side of the above equation by $n-1$ and using the fact that the function $g(z) = z/(1-z)^3$ is convex over the interval $(-1, 1)$ implies that

$$\frac{d}{d\gamma}\,\mathrm{trace}(\tilde{L}^2)\Big|_{\gamma=1} \geq \frac{2\sum_{k=2}^{n}\lambda_k(A)}{(1-\frac{1}{n-1}\sum_{k=2}^{n}\lambda_k(A))^3}. \tag{A.7}$$

Next, note that $\sum_{k=2}^{n}\lambda_k(A) = \mathrm{trace}(A) - \lambda_1(A) = na_{ii} - (1-\alpha) \geq 0$, where we are using the assumption that $a_{ii} \geq 1/n$ for all i. This implies that the numerator of the fraction on the right-hand side of (A.7) is nonnegative. Furthermore, the fact that $\lambda_k(A) \leq \lambda_1(A) = 1-\alpha$ guarantees that the denominator of the fraction on the right-hand side of (A.7) is strictly positive. Taken together, these two observations establish inequality (A.6). □

Proof of part (a). We now use part (b) to establish part (a) of the theorem. Recall from the previous part that the variance-covariance matrix of sectoral log outputs is given by $\tilde{L}'\tilde{L}$. On the other hand, the assumption that the input-output matrix A is symmetric and circulant guarantees that all row and column sums of $\tilde{L}$ are equal to $1/\alpha$. Therefore,

$$\sum_{i,j=1}^{n}\mathrm{cov}(\log\tilde{y}_i, \log\tilde{y}_j) = \mathbf{1}'\tilde{L}'\tilde{L}\mathbf{1} = n/\alpha^2.$$

The assumption that the economy's circulant implies that all industries are equally volatile, that is, $\text{var}(\log \tilde{y}_i) = \text{var}(\log \tilde{y}_1)$ for all i. Hence,

$$\sum_{i \neq j} \text{cov}(\log \tilde{y}_i, \log \tilde{y}_j) = n(1/\alpha^2 - \text{var}(\log \tilde{y}_1)).$$

Hence, the average pairwise correlation between sectoral log outputs is given by

$$\tilde{\rho} = \frac{1}{n(n-1)} \sum_{i \neq j} \text{corr}(\log \tilde{y}_i, \log \tilde{y}_j) = \frac{1}{(n-1)\,\text{var}(\log \tilde{y}_1)}(1/\alpha^2 - \text{var}(\log \tilde{y}_1)).$$

Similar derivations for the less interconnected economy with input-output matrix A imply that

$$\rho = \frac{1}{(n-1)\,\text{var}(\log y_1)}(1/\alpha^2 - \text{var}(\log y_1)).$$

Now comparing the right-hand sides of the above two equations completes the proof. In particular, by statement (b) of the theorem, $\text{var}(\log y_1) \geq \text{var}(\log \tilde{y}_1)$, which in turn implies that $\rho \leq \tilde{\rho}$. $\square$