THE CONGLOMERATE NETWORK

KENNETH R. AHERN[a]

LEI KONG[b]

XINYAN YAN[c]

Abstract

This paper proposes a network model of the economy in which conglomerate firms transmit idiosyncratic shocks from one industry to another. The strength of inter-industry connections in the network is determined by two factors: 1) conglomerates' market shares of total industry sales, and 2) the distribution of a conglomerate's total sales across industries. From these two factors, the model generates a new measure of cross-industry concentration that extends the widely-used Herfindahl index. Consistent with the model's predictions, the empirical results show that industry growth rates comove more strongly within industry pairs that are more closely connected in the conglomerate network.

[a] University of Southern California, Marshall School of Business, 701 Exposition Blvd., Ste. 231, Los Angeles, CA 90089, and National Bureau of Economic Research. E-mail: kenneth.ahern@marshall.usc.edu

[b] University of Alabama, Culverhouse College of Business, Box 870224, Tuscaloosa, AL 35487. Email: lkong@cba.ua.edu

[c] University of Dayton, School of Business Administration, Dayton, OH 45469. E-mail: xyan02@udayton.edu.

Conflict-of-interest Disclosure Statement

Kenneth R. Ahern
I have nothing to disclose.

Lei Kong
I have nothing to disclose.

Xinyan Yan
I have nothing to disclose.

I. Introduction

This paper studies how industrial concentration influences the transmission of idiosyncratic shocks across the economy. In a seminal paper, Gabaix (2011) shows that aggregate economic growth is driven by a small number of very large firms. Because economic activity is sufficiently concentrated in this small set of firms, their idiosyncratic shocks are not averaged out by the large number of little firms, leading to aggregate fluctuations. Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012) present a related explanation for aggregate fluctuations based on the inter-sectoral transmission of shocks through input-output (IO) links. As in Gabaix, it is the concentration of IO links within a small number of very influential sectors that causes idiosyncratic shocks to spread throughout the economy. These papers show that to understand aggregate economic fluctuations, it is imperative to understand how idiosyncratic shocks transmit from one sector to another.

We start from a simple observation: the largest firms in the economy tend to be conglomerates that span multiple industries. This means that the large, firm-level idiosyncratic shocks of Gabaix are actually transmitted across multiple industries through conglomerates. At the same time, a large division of a conglomerate might be small relative to the industry, while a small division could be relatively large. For example, one of the largest firms in the world, Amazon, receives the bulk of its revenues from its retail operations. However, Amazon commands a relatively small share of the total retail sector. In contrast, Amazon receives only 12% of its revenues from its computing services division, but maintains a dominant 33% market share within that industry. Thus, industrial concentration is likely to affect how firm-level idiosyncratic shocks transmit across the economy.

To better understand the transmission of idiosyncratic shocks, we model the economy as a bipartite, affiliation network of firms and industries. In the bipartite network, firms are only affiliated with industries, and industries are only affiliated with firms. We define the strength of connections from firms to industries, and vice versa, in proportion to their total outputs. Specifically, we assume that a growth shock transmits from a firm to an industry with a strength proportional to the firm's share of total industry sales. Similar to Gabaix (2011), an idiosyncratic shock to a firm that commands a large share of industry output has a greater influence on the industry's total fluctuations. In the opposite direction, we assume that a shock transmits from an industry to a firm with a strength proportional to the industry's share of the firm's total sales. For instance, a shock to the computing services industry will only

affect 12% of Amazon's total sales. Thus, in our network, each firm-industry pair is characterized by two directed links, each with its own strength.

Using the dual perspectives of a bipartite network, we create two separate unipartite networks, one for industry-to-industry links and another for firm-to-firm links. In the firm-to-firm network, firms are connected to other firms through common industry affiliations. In this perspective, intra-industry market forces serve as the conduit between firms. In the dual perspective, industries are connected to other industries through conglomerate firms that operate in both industries. Thus, in this perspective, conglomerate firms serve as the conduit of economic shocks between industries. The duality of firms and industries in our model is in the spirit of Alchian and Demsetz (1972), who argue that, in a frictionless setting, contracts within a firm are identical to contracts across a market.

For both unipartite networks, we derive three forms of connections. For brevity, we only describe the industry-to-industry connections at this point, but the firm-level connections are analogous. First, we calculate the strength of the inter-industry *transmission* from an industry to its affiliated firms and then from these firms to another industry. Because we allow for weighted and directed links in the network, the transmission strength from industry A to industry B is not necessarily equal to the strength from B to A.

Second, we calculate two projections from firms onto industries. In one projection, we calculate the *shared in-links* of an industry-pair. This measures the strength of industry connections based on the commonality of market shares of the firms that operate in both industries. If the same firms have more similar market shares in two industries, these industries are more connected. In the second projection, we calculate the *shared out-links* of an industry-pair. This represents the commonality in firms' exposures to industry shocks. These three measures of industry connections are based solely on the concentration of economic activity and abstract from other forms of connections between industries, such as customer-supplier links, common geographic locations, or shared institutional investors.

Though this network model is simple, it provides new insights on a number of important topics. First, the model generates new predictions on the variance and covariance of growth rates across industries. In particular, the model predicts that the variance of an industry's growth rate equals the sum of the industry-specific growth rate variance plus the variance of firm-specific growth scaled by the Herfindahl-Hirschman Index (HHI) of the industry. Second, the model predicts that the covariance of two

industries' growth rates equals to firm-specific variance scaled by the shared in-links of the two industries in the network. Intuitively, industries that share the same conglomerate firms will face common firm-level shocks. In the dual representation, we show that diversified firms, with lower concentrations of in-links from industries, have lower volatility in their growth rates.

Second, like Acemoglu et al. (2012), our model describes how idiosyncratic shocks transmit throughout an economy. In particular, our model predicts the paths and timing of the transmission of shocks from one industry to another. We show that industry centrality in the conglomerate network is directly proportional to the size of an industry's sales as a fraction of the economy's total sales. This finding provides a microfoundation for Gabaix's result on the importance of large firms.

Third, our model provides a microfoundation for standard measures of industry concentration. In particular, just as variance is a special case of covariance, our model shows that the widely-used HHI is a special case of a more general measure of cross-industry concentration that we call CoHHI. We define CoHHI as the shared in-links of two industries, as described above. Thus, CoHHI reflects the degree to which the market shares of firms in two industries overlap. If the same firms command the same market shares in each industry, then the two industries will have the same level of exposure to the same firm-specific idiosyncratic shocks. Using the covariance analogy, the CoHHI of an industry with itself is identical to HHI. Thus, our network provides a new interpretation of HHI as the concentration of in-links into an industry.

To bring our theoretical framework to the data, we use segment data from Compustat for all public firms in the US for the years 1997 to 2018. Though our theoretical framework provides both firm-level and industry-level representations of the economy, we focus on the industry-level analysis to reduce selection bias concerns. To create a balanced panel of industries, we form industry families following the procedure in Pierce and Schott (2016). In a series of robustness checks, we show that the main results are not driven by truncation bias caused by using Compustat data or particular industry definitions.

First, we show that the conglomerate network exhibits features of scale-free networks. In particular, the network is sparse with relatively few central industries and many peripheral industries. Consistent with small world networks, the maximum path length between industries in an average year is 7.5 links, out of over 500 different industries. We also show that during the sample period, the network evolved into

a more concentrated network as industries dropped links to more peripheral industries while retaining links to more central industries. These statistics provide a new perspective on the changing importance of conglomerate firms in the economy.

Next, we show that the conglomerate network is empirically distinct from other forms of industry connections. First, the correlation between conglomerate network links and input-output links is low. This is consistent with the notion that conglomerate firms diversify for reasons beyond vertical integration, such as economies of scope (Teece, 1980), co-insurance (Lewellen, 1971), or agency conflicts (Jensen, 1986). Second, the pattern of conglomerate ownership is not driven by product similarity, as measured by Hoberg and Phillips (2016).

To estimate the relationship between the industry connections in the conglomerate network and the comovement of industry growth, we first estimate cross-sectional regressions. We find that CoHHI is positively correlated with the covariance of sales growth and asset growth in the cross-section of industries, consistent with our theoretical framework. Also consistent with our framework, we find that the volatilities of growth rates are positively correlated with HHI, reflecting that concentrated industries are less diversified across individual firm-level shocks.

Next, to further isolate the correlation between the conglomerate network and the covariance of industry volatilities, we estimate panel regressions where the dependent variable is the squared difference of growth rates between industries. Industry-pair fixed effects capture both time-invariant industry and industry-pair characteristics that could influence the comovement of industry growth rates, such as average volatility, the labor share in production, and the average firm size. Year fixed effects control for general macroeconomic trends that could influence the comovement of industry growth rates. We also control for input-output linkages and product market similarity of Hoberg and Phillips (2016). Thus, our empirical model isolates the correlation between abnormal time-series variation in the conglomerate network and the comovement of industry growth.

The panel regressions show that when two industries have stronger connections in the conglomerate network, their growth rates comove more closely, as predicted by the model. The results hold for both in-links, out-links, and the transmission network for sales growth and asset growth. The economic magnitude of the results is meaningful. Compared to an industry-pair with no conglomerate link, the formation of a link is associated with a squared difference in growth rates that is lower by about 30%, relative to the median squared difference.

Because conglomerate firms do not randomly choose the industries in which they operate, it is possible that our results reflect the endogenous choice of firms to operate in industries that tend to comove with each other. We address this concern in a variety of ways. First, we note that the conglomerate network is based on market shares, which are not endogenously chosen by firms, but are determined by market forces. Second, the industry-pair fixed effects in our tests control for all time-invariant factors that cause industries to comove, while the time-varying variables control for the most common economic determinants of comovement. Though we cannot rule out all alternative explanations, our research design implies that for our results to be spurious, there would need to be an omitted time-varying factor that is highly correlated with time-varying market shares, but also orthogonal to major economic determinants of diversification.

To further address endogeneity concerns, we use a quasi-natural experiment to identify the transmission of economic shocks through the conglomerate network. Following Pierce and Schott (2016), we exploit cross-sectional variation in industries' exposure to tariff rate shocks from the granting of normal trade relations to China in 2000. Using the conglomerate network from 1999 to control for reverse causation, and controlling for industry fixed effects, year fixed effects, and customer-supplier links, we find that industries with stronger connections in the network to those industries most affected by the tariff shock had larger declines in employment following the shock. These results show that a specific, identifiable industry shock can be traced through the conglomerate network.

Finally, we present a series of robustness tests. First, because we construct the conglomerate network using only publicly-traded firms' segment data, we test for the generalizability of our results using employment data from the US Census County Business Patterns data, which covers nearly all establishments in the private sector. We find similar results using these data, which implies that the predictive power of the conglomerate network is not limited to public firms. Second, we find similar results to our baseline tests when we estimate the same relationships using lagged measures of the conglomerate network. These results further mitigate concerns of reverse causation. Third, we increase the minimum size threshold of firms in our sample to show that there is little evidence of truncation bias from using relatively large, public firms to construct our conglomerate network. Fourth, we show that our results persist when we construct our network using coarser industry definitions. We also provide evidence of the transmission of shocks within conglomerates, as shown

in prior research, and provide suggestive evidence on the role of financial constraints for the transmission of shocks.

This paper makes two central contributions. First, we present a model and empirical evidence of inter-industry transmission of economic shocks through conglomerate firms. As discussed above, these findings build on Gabaix (2011) and Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012). More recently, Herskovic, Kelly, Lustig, and Van Nieuwerburgh (2020) find that firm size is related to aggregate fluctuations in a firm-level customer-supplier network. Additional empirical evidence of the spread of idiosyncratic shocks through production networks is found in Ahern and Harford (2014) and Barrot and Sauvagnat (2016). In contrast to these papers, we show that industry-specific shocks transmit across the economy through the internal redistribution of conglomerate firms. Our approach is also related to the notion that local shocks spread to wider geographic regions through multi-regional firms (di Giovanni, Levchenko, and Mejean, 2014; Kleinert, Martin, and Toubal, 2015; Giroud and Mueller, 2019). In contrast to shocks that are spread across geographic space, we show multi-segment firms facilitate the spread of economic shocks across industry space. Our results also relate to recent literature on the importance of common ownership (Azar, Schmalz, and Tecu, 2018; Antón, Ederer, Giné, and Schmalz, 2018). While this line of research focuses on partial ownership by institutional investors, we study controlling ownership by conglomerate firms.

This paper also relates to the large literature on the causes of within-firm reallocation of resources (Stein, 1997; Lamont, 1997; Shin and Stulz, 1998; Matsusaka and Nanda, 2002) and the motivations for diversification (Lewellen, 1971; Aggarwal and Samwick, 2003; Villalonga, 2004). Our model does not depend on any specific mechanisms and our goal is not to identify specific internal mechanisms that cause within-firm or within-industry reallocations. In contrast, we provide a macro-level, industry-to-industry perspective of reallocations, taking as given the reasons for diversification and the various micro-level mechanisms of internal reallocations documented in the prior literature.

The second contribution of this paper is to provide a novel, economic interpretation of HHI. Even though HHI was not originally developed from economic theory,[1] it has become the standard metric of industry concentration among academics, practitioners, and policy-makers. For example, during 2011 to 2020, HHI is referenced

---

[1]HHI was originally developed to meet two ad hoc criteria of concentration measures (Hirschman, 1945; Herfindahl, 1950). Hirschman acknowledges that HHI is not a unique solution.

in 4.5% of all articles published in the top five economics journals, which represents an increase of 87% relative to the 2000s. Second, HHI is a primary filter to identify non-competitive markets among a wide range of regulatory agencies, including the DOJ, FTC, FCC, FDIC, and the Federal Reserve. Our framework provides a new microfoundation for the calculation of HHI based on a network of firms and industries. To our knowledge, this provides one of the only economic interpretations of HHI.[2] Moreover, we use our framework to derive a new measure of cross-industry concentration, CoHHI.

II. The Theoretical Conglomerate Network

To construct the conglomerate network of industries, we start with a bipartite graph, also known as a two-mode network or an affiliation network, in which there are two types of nodes that are disjoint, independent sets and each type of node is connected only to nodes of the other type. Typical examples of affiliation networks include football players and football clubs, co-authors and publications, and corporate directors and corporate boards. Much of the research on networks in economics studies one-node networks with an implicit assumption of an underlying two-mode network. For instance, corporate boards typically do not have direct connections with other corporate boards, but instead, have indirect connections through shared directors in an affiliation network.

In our network, the two types of nodes are firms and industries. Firms are affiliated with industries, and industries are affiliated with firms. Because two-mode networks have two distinct types of nodes, they allow for dual perspectives on the network's structure. In our setting, one representation of the network is from the perspective of firms: firms are connected to each other through shared industry affiliations. This perspective is the commonplace view of the relationship between firms and industries. The dual representation of the network from the perspective of industries is less commonplace: industries are connected to other industries by firms that operate in multiple industries. In this perspective, conglomerate firms that span multiple industries are the conduits for economic shocks to transmit between industries. Though the firm perspective is the basis for the common assumption that firms in the same industry face the same economic shocks, the industry perspective is an

---

[2]Stigler (1964) and subsequent research shows that HHI is directly related to profitability in Cournot competition models. Our microfoundation applies to all industries, not just those with Cournot competition, which, according to Berry, Gaynor, and Morton (2019), are rare.

equally valid representation of the same underlying two-mode network. Because of the importance of conglomerates in the dual representation, throughout the paper, we call this affiliation network the conglomerate network.

The transmission of economic shocks in the conglomerate network is distinct from an input-output network where firms buy and sell directly from other firms. Instead, economic shocks in the affiliation network transmit through firms' shared exposure to industry conditions, or equally, through industries' shared exposures to firm conditions. Thus, our model assumes that conglomerates internally reallocate shocks from one sector to another and that industries reallocate shocks from one firm to another. We do not model the specific sources of these shocks, nor any specific mechanisms that cause firms to internally redistribute industry shocks among their segments. We only assume that such mechanisms exist, and then study how the shocks spread through the conglomerate network.

A large literature supports this assumption. For example, prior theoretical research argues that within-firm redistribution could be caused by corporate socialism (Scharfstein and Stein, 2000), optimal reallocation to equate marginal revenue products of capital or labor (Williamson, 1975), or the trade-off between the benefit of flexible investments versus the costs of agency-driven over-investment (Matsusaka and Nanda, 2002). A large body of empirical evidence also supports these assumptions (Maksimovic and Phillips, 2002; Seru, 2014; Tate and Yang, 2015; Giroud and Mueller, 2019). Similarly, the assumption that market forces within an industry redistribute firm-level idiosyncratic shocks from one firm to another is also supported by a large literature. As summarized in Shea (2002), these forces could be consumption complementarities, external economies of scale, or aggregate demand spillovers, among other spillover mechanisms. The generality of these assumptions reflects our focus not on the transmission of shocks within firms or within industries but on the firm-to-firm and industry-to-industry transmission of shocks.

We also do not model the formation of the conglomerate network. Instead, we take the network as given, and identify the transmission of shocks through conglomerate links. This approach mirrors the approach of research on production networks. Just as we do not attempt to identify why one firm chooses to operate in a particular set of industries, research on production networks does not typically model why a customer chooses one set of suppliers over another set (Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi, 2012; Herskovic, Kelly, Lustig, and Van Nieuwerburgh, 2020).

II.A. Formal Definitions

To formalize these assumptions, we assume the economy has $i = 1, \ldots, n$ firms and $j = 1, \ldots, m$ industries. Let S be the $n \times m$ bi-adjacency matrix in which entry $s_{i,j}$ denotes firm i's sales in industry j. Thus, the total sales for firm i is $\Sigma_j^m s_{i,j}$. The total sales for industry j is $\Sigma_i^n s_{i,j}$. Below, we use capital letters to denote matrices, lower case letters to denote matrix elements, and $\vec{x}$ to represent vectors.

We normalize S in two ways. First, we generate the matrix of market shares, H, by normalizing S by its column sums. Thus, the market share of firm i in industry j is $h_{i,j} = \frac{s_{i,j}}{\Sigma_i^n s_{i,j}}$. Similarly, we generate the matrix of industries' firm shares, F, by normalizing S by its row sums: $f_{i,j} = \frac{s_{i,j}}{\Sigma_j^m s_{i,j}}$. Thus, each entry of F represents the fraction of firm i's total sales that are attributed to industry j. By normalizing industries and firms by their total sales, we focus on the relative importance of the connections between industries and firms, rather than the size of each node.

We allow the connections between firms and industries in the conglomerate network to be directional and weighted. In particular, we assume that growth shocks that transmit from a firm node to an industry node are weighted by the firm's market share in the industry, as recorded in H. Intuitively, a firm-level shock will affect an industry's growth in proportion to the firm's fraction of the industry's total sales. Analogously, we assume that shocks that transmit from an industry to a firm are weighted by the size of the industry segment in the firm's overall operation, as recorded in F. We denote this fraction as an industry's firm share, analogous to a firm's market share. Intuitively, an industry-wide growth shock will affect a firm's growth in proportion to the industry's importance in the firm's total sales.

We combine F and H into an $(m+n) \times (m+n)$ adjacency matrix A that represents the complete, weighted and directed bipartite graph, as follows,

$$A = \begin{bmatrix} 0 & F' \\ H & 0 \end{bmatrix}. \tag{1}$$

The first m rows and columns of A refer to industries, and the last n rows and columns refer to firms. Matrix A represents the effect of a shock in the row entry on the column entry. F' represents the effect of a shock transmitting from an industry to a firm. H represents the effect of a shock transmitting from a firm to an industry. The zero matrices on the diagonals reflect that in the bipartite graph, firms and industries do not have direct connections. Also note that A is not symmetric, which reflects the directional nature of the bipartite network.

To illustrate our network setting, consider a simple example with three firms $(x,$ $y,$ and $z)$ that operate in two industries $(p$ and $q)$. Their segment sales are given in matrix S, and we normalize S by row sums and column sums to generate F and H, as follows:

$$S = \begin{array}{c} \\ x \\ y \\ z \end{array}\begin{array}{cc} p & q \\ \left[\begin{array}{cc} 3 & 3 \\ 0 & 3 \\ 1 & 4 \end{array}\right] \end{array}, \qquad F = \begin{array}{c} \\ x \\ y \\ z \end{array}\begin{array}{cc} p & q \\ \left[\begin{array}{cc} 0.50 & 0.50 \\ 0.00 & 1.00 \\ 0.20 & 0.80 \end{array}\right] \end{array}, \qquad H = \begin{array}{c} \\ x \\ y \\ z \end{array}\begin{array}{cc} p & q \\ \left[\begin{array}{cc} 0.75 & 0.30 \\ 0.00 & 0.30 \\ 0.25 & 0.40 \end{array}\right] \end{array}.$$

$$(2)$$

Figure I provides a graphical representation of this network, where blue arrows refer to the effect of firms on industries (H) and red arrows refer to effects of industries on firms (F), where the weights of the connections are determined by a firm's market share (blue arrows) or an industry's firm share (red arrows).
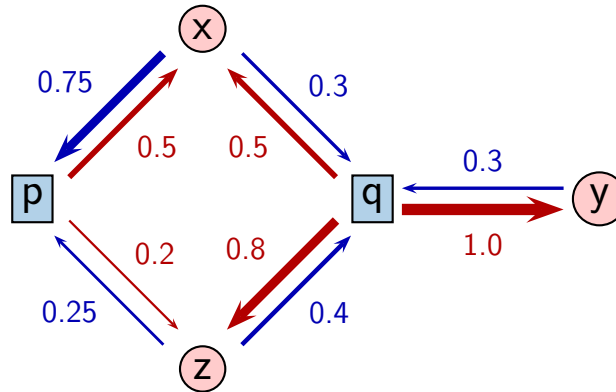
FIGURE I
A Network with Two Industries and Three Firms
This figure presents a graphical representation of an example network. The firms are x, y, and z, and the industries are p and q. Blue arrows reflect matrix H, the effect of firms on industries. Red arrows reflect matrix F, the effect of industries on firms.

This example illustrates our definition of the strengths of connections in the conglomerate network. Because firm x receives half of its sales from industry p, a growth shock in industry p will affect half of firm x's sales. In contrast, the same industry shock in p only affects firm z by 0.2, because firm z only receives 20% of its sales from industry p. An identical interpretation exists for the dual of the network. A growth shock in firm x has a larger effect in industry p than q because firm x has a market share of 75% in industry p, but only 30% in industry q.

II.B. Network Transformations

To study the inter-industry and inter-firm connections, we transform the bipartite graph in matrix A into a unipartite graph in three ways. The first transformation represents the strength of transmissions between nodes of the same type based on the compound effect of shocks from one industry to another through affiliated firms, or from one firm to another through industry affiliations. The second and third transformations are projections from one set of nodes onto the other. The projections reflect the strength of shared in-links or shared out-links between nodes of the same type.

B.1. Network Paths

The first transformation of the network generates the strength of the paths that lead from one node to another of the same type. In particular, we denote

$$\text{Transmission Matrix} = A^2 = \begin{bmatrix} F'H & 0 \\ 0 & HF' \end{bmatrix}. \tag{3}$$

In a bipartite network, it takes two links to connect nodes of the same type (e.g., one link from an industry node to firm nodes, and a second link from firm nodes to industry nodes). If A was an unweighted adjacency matrix consisting of zeros and ones, A^2 would count the number of unique paths with a length of two that connect two industry nodes. If more paths connect two industries, they would have a stronger connection. In our case, using weighted links, the entry in the j'th row and k'th column of $F'H$ reflects the compound effect of a transition from industry j to industry k through conglomerate firms that operate in both industries. Likewise, in the bottom-right quadrant of A^2, HF' represents the compound effect of inter-firm transitions through industries. Note also that A^2 is a left stochastic matrix, where each column sums to one.

In our numerical example, the effect of moving from p to q is the effect of moving from p to x, then x to q ($0.50 \cdot 0.30$) plus the effect of moving from p to z then z to q ($0.20 \cdot 0.40$), which equals 0.230. Panel A of Figure II presents a visual representation of the transition matrix. Notice that this matrix is not symmetric. The effect of moving from p to q is 0.230 compared to the effect of moving from q to p, which is 0.575. The asymmetry is caused by asymmetry in the strength of the nodes' in-links relative to their out-links. In Figure I, note that the strength of links that lead out of industry p are weaker than the links that lead into industry p. In contrast, the

strength of the links that lead out of industry q are stronger than the links that lead into it.

Second, the diagonal entries in A^2 represent the transmission from a node back to itself after two links in the network. For example, the first entry in $F'H$ represents the effect of a shock transitioning from industry p back to industry p. In particular, this is the effect of moving from p to x, then x to p $(0.50 \cdot 0.75)$ plus the effect of moving from p to z then z to p $(0.20 \cdot 0.25)$, which equals 0.425. In a network setting, the diagonals of A^2 represent a feedback loop, or an 'echo,' as denoted in Sharifkhani and Simutin (2021). In the sense of the redistribution of a shock, we can also think of the diagonal of A^2 as the residual fraction of the shock that is not transmitted to other nodes.

Panel B of Figure II presents the firm-to-firm transition of our numerical example. A shock in firm x has the greatest effect back on firm x and the least effect on firm y. As in industries, the transition matrix for firms is asymmetric.

B.2. Network Projections

The second type of transformation is a projection from firms onto industries and vice versa. The first projection reflects the strength of shared in-links:

$$\text{Shared in-links} = A'A = \begin{bmatrix} H'H & 0 \\ 0 & FF' \end{bmatrix}. \tag{4}$$

$A'A$ reflects the strength of the in-links that two nodes share and the diagonal of the matrix is the sum of the squared weights of the in-links for each node. This reflects how similar two nodes are to each other based on the strength of their common exposures. If two industries receive shocks from the same firms, in the same proportions, then they will be more closely related in this projection. Because the projection is based on shared in-links, it is a symmetric matrix.

Panels C and D of Figure II present a visual representation of the strength of shared in-links for our numerical example. At the firm-level, firm z has a smaller connection to x than it does to y. This is because firms z and y share strong common in-links from industry q, whereas firms x and z share weak common in-links from industry p. Thus firm z has a more similar exposure to firm y from industry shocks than it does to firm x.

(A) Industry Transition Matrix $F'H$

(B) Firm Transition Matrix HF'

(C) Industry Shared In-links $H'H$

(D) Firm Shared In-links FF'

(E) Industry Shared Out-links $F'F$

(F) Firm Shared Out-links HH'
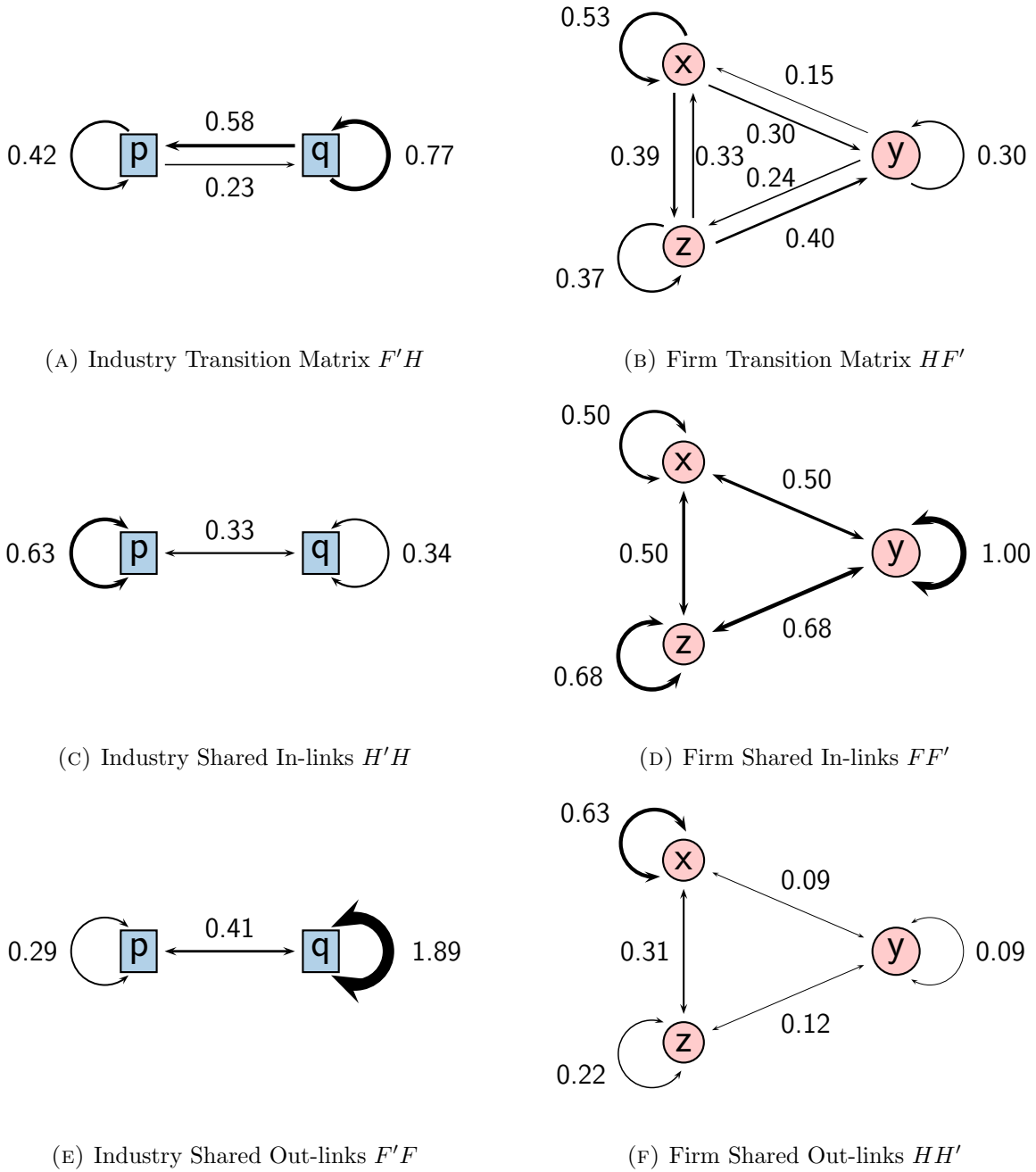
FIGURE II

Three Transformations of the Conglomerate Network

Panels A and B present transition matrices for industries (A) and firms (B). Panels
C and D present projection matrices of shared in-links for industries (C) and firms
(D). Panels E and F present projection matrices of shared out-links for industries
(E) and firms (F).

The second projection reflects the strength of shared out-links:

$$\text{Shared out-links} = AA' = \begin{bmatrix} F'F & 0 \\ 0 & HH' \end{bmatrix}. \tag{5}$$

If A was an unweighted, binary matrix, AA' would reflect the number of out-links with the same destination that two nodes share in common. Using weighted connections, as in our case, AA' reflects the strength of the out-links that two nodes share and the diagonal of the matrix is the sum of the squared weights of the out-links for each node. This reflects how similar two nodes are to each other based on the strength of the commonality of destinations for shocks. If two industries tend to have similar effects on the same firms, then the industries have higher connections in this projection.

II.C. Concentration Measures in the Conglomerate Network

The conglomerate network provides a new interpretation of standard measures of industry concentration. First, note that the columns of H are n-dimensional vectors representing the market shares of the n firms in each industry. Denote an arbitrary column j in H as $\vec{h}_j$. As defined above, the entries of the matrix of shared in-links, $H'H$, are equivalent to the dot products of the columns of H. Therefore, for two industries, j and k, the row j and column k entry of $H'H$ is $\vec{h}_j \cdot \vec{h}_k = h_{1,j}h_{1,k} + h_{2,j}h_{2,k} + \cdots + h_{n,j}h_{n,k}$. The diagonal entries of $H'H$ are the dot products of an industry's market share vector with itself. For industry j, this is $h_{1,j}^2 + h_{2,j}^2 + \cdots h_{n,j}^2$. Thus, the diagonal entries of $H'H$ are the Herfindahl-Hirschman Indices (HHI) of industry concentration.

Using the conglomerate network, we can extend this derivation of HHI to generate a measure of cross-industry concentration. For two industries, j and k, we define,

$$\text{Industry CoHHI}_{j,k} = (H'H)_{j,k} = \vec{h}_j \cdot \vec{h}_k = h_{1,j}h_{1,k} + h_{2,j}h_{2,k} + \cdots + h_{n,j}h_{n,k}. \tag{6}$$

Industry CoHHI is the commonality in the pattern of firms' market shares across two industries. If the same firms have similar market shares in both industries, then Industry CoHHI will be larger. Thus, Industry CoHHI reflects the similarity of concentration of sales across firms in different industries. In turn, this means that the standard definition of Industry HHI is a special case of Industry CoHHI with itself.

This representation provides a microfoundation for the calculation of industry concentration as the sum of squared market shares, which is not justified in Hirschman

(1945) or Herfindahl (1950).[3] In addition, using the conglomerate network to derive HHI reveals that HHI is a special case of a more general measure of CoHHI, which reflects the similarity between industries' market share distributions. Using the network perspective, these results also show that standard HHI measures can be thought of as the concentration of in-links of an industry. In other words, if two industries receive the same shocks from the same set of firms, then the two industries have high co-concentration.

We can apply a similar idea to the projection of out-links, AA'. At the industry level, $F'F$ reflects the commonality of out-links from industries to firms. If two industries tend to affect the same set of firms, then the two industries have higher co-concentration of destinations. This would happen when firms have the same fraction of sales from the same industries. Industries with focused firms will tend to have higher out-link concentration. Thus, this projection represents a new measure of industry concentration that is complementary to standard HHI.

Because of the duality of the bipartite graph, we can also provide similar measures of concentration at the firm level. In particular, FF' represents the Firm CoHHI matrix in which the diagonals are the Firm HHIs of concentration, and the off-diagonal elements are the Co-HHI between firms. Firm HHI measures the concentration of a firm's sales across industries. A firm with equal sales in two industries has a lower Firm HHI than a firm with the majority of its sales in one industry. The Firm Co-HHI reflects the commonality of two firm's distribution of sales across industries. Two firms that tend to sell the same fractions of their total sales in each industry will have a higher Co-HHI.

II.D. The Variance of Growth Shocks in the Conglomerate Network

In this section, we use our model to derive predictions on the covariance of industry and firm growth rates. We assume at time $\tau = 0$, firm i receives a shock ε_i and industry j receives a shock η_j. Both shocks are random variables with mean 0 and

---

[3]Hirschman designed HHI to meet two criteria that he argued that any measure of industry concentration should include: 1) concentration should be related to the dispersion of market shares, and 2) concentration should be declining with the number of firms in an industry. Hirschman's measure accomplished these two goals, though the square of market shares was not justified. In fact, Hirschman recognized that his concentration measure is not the only measure that could meet these criteria. Herfindahl (1950) justified using the square of market shares by considering HHI as a weighted average of market shares where the weights were the market shares themselves. However, he did not justify using the shares as weights versus logged market shares, the square root of market shares, or any other weighting scheme.

standard deviations σ_ε and σ_η, such that $cov(\varepsilon_k, \varepsilon_l) = 0$ for all (k, l) when $k \neq l$; $cov(\eta_k, \eta_l) = 0$ for all (k, l) when $k \neq l$; and $cov(\varepsilon_k, \eta_l) = 0$ for all (k, l). Thus, ε represents firm-level growth shocks after removing industry-level growth shocks, η, and vice versa. For simplicity, we assume that σ_ε is the same across firms and σ_η is the same across industries. In vector form, $\vec{\eta}$ is the $m \times 1$ vector of industry shocks and $\vec{\varepsilon}$ is the $n \times 1$ vector of firm shocks.

We assume shocks transmit from one node to another over time. In a bipartite graph, one step in the network, enacted by an application of the A matrix, represents an aggregation of a node's own shock plus the weighted average of the shocks of connected nodes. Two applications of the adjacency matrix (A^2) to a shock vector represents a complete transition of shocks back to their original node type. Therefore, to study the transition of shocks through the network, we take snapshots of the network after every complete transition of shocks, where the initial shock is a node's own shock plus the aggregation of its connected nodes' shocks. During every complete transition, we assume the shocks decay with rate δ.[4]

In particular, the snapshot of the growth rate of industry j at $\tau = 0$ can be written as the industry-specific growth shock plus the weighted average of the growth rates of the firms operating in industry j. This is $g_{j,0} = \eta_j + \sum_{i=1}^{n} h_{i,j} \varepsilon_i$. Therefore, the vector of industry growth rates is $\vec{g}_{ind,0} = \vec{\eta} + H'\vec{\varepsilon}$. The growth rate of firms follows the same pattern: the firm's specific growth rate plus the industry-specific growth rates weighted by the industry's firm share. This is, $\vec{g}_{firm,0} = \vec{\varepsilon} + F\vec{\eta}$. In matrix notation, the initial growth rates at $\tau = 0$ are

$$\vec{g}_0 = (I + A')\vec{\nu} = \begin{bmatrix} I & H' \\ F & I \end{bmatrix} \begin{bmatrix} \vec{\eta} \\ \vec{\varepsilon} \end{bmatrix} = \begin{bmatrix} \vec{\eta} + H'\vec{\varepsilon} \\ \vec{\varepsilon} + F\vec{\eta} \end{bmatrix}. \tag{7}$$

The variance-covariance matrix of growth rates at $\tau = 0$ is

$$Cov(\vec{g}_0) = \begin{bmatrix} \sigma_\eta^2 I + \sigma_\varepsilon^2 H'H & \sigma_\eta^2 F' + \sigma_\varepsilon^2 H' \\ \sigma_\eta^2 F + \sigma_\varepsilon^2 H & \sigma_\varepsilon^2 I + \sigma_\eta^2 FF' \end{bmatrix}. \tag{8}$$

The upper-left entry of this matrix represents the variance-covariance matrix of industry growth rates. The diagonal elements reflect the variances of industry growth

---

[4]The decay reflects the fraction of the initial shock that is passed from one industry to another, with the remainder absorbed either within the firm or industry. For example, a negative cash flow shock to a segment that accounts for 10% of a firm's sales may absorb more than 10% of the shock.

rates which equal the variance of industry-specific shocks plus the variance of firm-specific shocks scaled by the industry's HHI. For industry j, the variance is

$$Var(g_{j,0}) = \sigma_\eta^2 + \sigma_\varepsilon^2 HHI_j. \tag{9}$$

Thus, assuming all firm-level shocks are equally distributed, more concentrated industries have higher variance of growth rates. This is driven by concentrated industries' greater exposure to relatively few idiosyncratic firm-specific shocks.

The off-diagonal elements of the industry-level variance-covariance matrix equal the variance of firm-specific shocks scaled by the CoHHI between two industries. Thus, the covariance in growth rates at $\tau = 0$ is

$$Cov(g_{j,0}, g_{k,0}) = \sigma_\varepsilon^2 CoHHI_{j,k}. \tag{10}$$

This derivation shows that CoHHI is directly proportional to the covariance of growth rates, just as HHI is directly proportional to the variance of growth rates.

The variance-covariance matrix of firm-level growth rates is $\sigma_\varepsilon^2 I + \sigma_\eta^2 FF'$. This is the dual interpretation of the industry-level matrix. On the diagonal, firm growth rates have a variance equal to the firm-level variance plus industry-level variance scaled by the firm's HHI across segments. Diversified conglomerate firms with operations in multiple sectors face lower industry-specific variance, compared to focused, single-segment firms. The off-diagonal elements in the firm-level variance-covariance matrix are equal to industry-specific variance scaled by firm-level CoHHI. Firms that operate in the same industries have greater shared in-links and thus, have higher covariance in their growth rates.

The off-diagonal $n \times m$ matrix, $\sigma_\eta^2 F' + \sigma_\varepsilon^2 H'$, in $Cov(\vec{g}_0)$ reflects the covariance in growth rates between firms and industries. This covariance is equal to the sum of the firm-specific variance σ_ε^2 scaled by the strength of the link from the firm to the industry plus the industry-specific variance scaled by the strength of the link from the industry to the firm. Intuitively, the covariance of growth rates between firms and industries is the sum of the firm and industry specific variance scaled by the strength of the connection between the firm and industry.

Moving shocks forward one cycle in the network, the variance-covariance matrix of growth rates at $\tau = 1$ is

$$Cov(\vec{g}_1) = \delta^2 \begin{bmatrix} \sigma_\eta^2 H'FF'H + \sigma_\varepsilon^2 H'FH'HF'H & \sigma_\eta^2 H'FF'HF' + \sigma_\varepsilon^2 H'FH'HF' \\ \sigma_\eta^2 FH'FF'H + \sigma_\varepsilon^2 FH'HF'H & \sigma_\varepsilon^2 FH'HF' + \sigma_\eta^2 FH'FF'HF' \end{bmatrix}. \tag{11}$$

This variance-covariance matrix represents that as shocks pass through the network over time, they repeatedly transmit from firms to industries and back to firms.

To help interpret the covariance matrix, we again focus on the industry-to-industry portion of the covariance matrix. At $\tau = 1$, shocks have made a full cycle in the network, represented by the transmission network A^2. For brevity of notation, we denote the entry in row j and column k of $F'H$ as $t_{j,k}$, which records the strength of transmission from industry j to industry k. Using this notation, the variance of the growth rate for industry j at $\tau = 1$, after a full cycle in the network, is as follows:

$$Var(g_{j,1}) = \delta^2 \left[\sigma_\eta^2 \sum_{r=1}^{m} t_{r,j}^2 + \sigma_\varepsilon^2 \sum_{r=1}^{m} t_{k,j}^2 HHI_r + 2\sigma_\varepsilon^2 \sum_{r=2}^{m} \sum_{s=1}^{r-1} t_{s,j} t_{r,j} CoHHI_{s,r} \right]. \quad (12)$$

This equation shows that after one cycle in the network, the variance of industry j's growth rate has three components. The first component is the variance of industry-level shocks scaled by the sum of the square of transmission links from all other industries into industry j. This reflects industry j's exposure to all other industry shocks through the conglomerate network. The second component is the firm-level variance scaled by the sum of the transmission strength into industry j from all other industries weighted by the other industries' HHI measures. Thus, if industry j has stronger connections to concentrated industries, its variance is higher because of higher exposure to firm-level shocks in connected industries. Finally, the last component is the sum of all combinations of industries' CoHHI measure multiplied by their transmission strengths and the variance of firm-level shocks.

The covariance of the growth rate of industry j and industry k at $\tau = 1$, is as follows:

$$Cov(g_{j,1} g_{k,1}) = \delta^2 \left[\sigma_\eta^2 \sum_{r=1}^{m} t_{r,j} t_{r,k} + \sigma_\varepsilon^2 \sum_{r=1}^{m} t_{r,j} t_{r,k} HHI_r + \sigma_\varepsilon^2 \sum_{r=2}^{m} \sum_{s=1}^{r-1} \left(t_{s,j} t_{r,k} + t_{r,j} t_{s,k} \right) CoHHI_{s,r} \right].$$
$$(13)$$

This equation shows that the covariance in the growth rates of industries j and k is determined by the similarity of their exposure to industry shocks through the transmission network plus the similarity of their exposure to firm-level shocks through within-industry concentration and common in-links (CoHHI).

II.E. The Centrality of the Conglomerate Network

To understand the long-run outcome of shocks transmitting through the network, we consider the sum of the growth rates from $\tau = 0, \ldots, \infty$, as follows

$$\sum_{\tau=0}^{\infty} \vec{g}_t = \left[I + \delta(A')^2 + \left(\delta(A')^2 \right)^2 + \left(\delta(A')^2 \right)^3 + \cdots \right] (I + A') \vec{\nu}$$

$$= \left[I - \delta(A')^2 \right]^{-1} (I + A') \vec{\nu}, \tag{14}$$

where the infinite sum converges because A^2 is a stochastic matrix. The term $[I - \delta(A')^2]^{-1}$ is the Leontief inverse. Thus, Equation 14 represents the transformation of initial shocks into industry and firm growth rates after passing through the conglomerate network an infinite number of times. This can be interpreted as the steady state outcome of the transition matrix A^2. In addition, as Carvalho (2014) points out, the Leontief inverse is equivalent to the Katz-Bonacich eigenvector centrality of a network. Thus, Equation 14 also implies that the long-run industry and firm growth rates equal the product of their Katz-Bonacich eigenvector centrality in the conglomerate network with their initial shock.

These properties of the conglomerate network are identical to the properties of the industry-level input-output networks studied in Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012) and the firm-level input-output network studied in Herskovic, Kelly, Lustig, and Van Nieuwerburgh (2020). In particular, both papers show that the Leontief inverse describes how network structure affects aggregate growth rates. Acemoglu et al. include the Leontief inverse in a measure they call the "influence vector" $\vec{v}$ of an industry, which is equivalent to both Katz-Bonacich eigenvector centrality and the "sales vector" of the economy, in which each element reflects sector i's sales as a fraction of the total sales in the economy. Acemoglu et al. note that the second representation is related to Gabaix's finding that firm-level productivity contributes to aggregate productivity in proportion with firm size.

Our conglomerate network has the same representations of the influence vector as the IO network, except the weights in our measure represent the strength of the bi-partite network of conglomerates. In particular, the eigenvector centrality of the industries and firms in our network matrix A^2 is also the sales vector of the economy, $\vec{v}$. Thus, like Acemoglu et al., given a vector of idiosyncratic industry shocks $\vec{\eta}$, the aggregate shock to the economy is $\vec{v}_{ind}{}'\vec{\eta}$. Equivalently, the aggregate shock to the economy from firm-level shocks is $\vec{v}_{firm}{}'\vec{\varepsilon}$, as in Herskovic et al.

In sum, our model generates the same implications for the importance of a single industry or firm in the conglomerate network as derived in the production network, but with an important distinction. In our model, firms are diversified, unlike in Gabaix (2011), which affects their centrality, and thus, their influence on the aggregate economy. Herskovic, Kelly, Lustig, and Van Nieuwerburgh (2020) present a similar intuition based on the concentration of customers in the input-output network. Our work is similar to Herskovic et al. because they allow for shocks at different levels of aggregation, including firm, industry, and economy-wide shocks. Our model is distinct from Herskovic et al. because we focus on connections through conglomerate ownership, whereas they focus on customer-supplier connections.

III. The Empirical Conglomerate Network

III.A. Data Sources

We collect segment level information of conglomerate firms from the Compustat Historical Segment data. For corporate segments that represent at least 10 percent or more of consolidated sales in a different industry, SFAS No. 14 requires that firms report accounting information on a segment-level basis for fiscal years ending after December 15, 1977. To rectify the inadequacies of SFAS No. 14, the Financial Accounting Standards Board (FASB) further issued SFAS No. 131 in June 1997, which requires that, for fiscal periods beginning after December 15, 1997, firms identify industry segments for external reporting purposes in the manner that management views operating segments for internal decision-making purposes. To ensure the time-series comparability of our conglomerate network, we use the Compustat Historical Segment data from 1997 to 2018 to construct our conglomerate network. Specifically, for each segment, we collect the following four variables: net sales, capital expenditures, identifiable total assets, and the primary NAICS code of the segment.[5] Though our framework provides both firm-level and industry-level predictions, we focus on the industry-level network to reduce selection bias in publicly-traded company data. In Section III.A.2, we discuss truncation bias concerns from using only publicly-traded firms.

One complication of the long-time horizon considered in this paper is that the scheme of industry classifications changes over time, such as the change from the SIC

---

[5]Though we would like to study productivity, the variables available at the segment level do not allow it.

to the NAICS in 1997 and subsequent versions of NAICS from 2002 to 2017. To obtain the time-consistent industry definitions, we follow Pierce and Schott (2016) and create "families" of industry codes that group related NAICS categories together across different industry classification schemes. For example, if an industry code splits into several codes from 1997 to 2002, the industry code in 1997 and its subsequent "children" would be grouped into the same family. Therefore, unless otherwise noted, industries in this paper refer to these families. This adjustment allows us to control for time-invariant industry properties using fixed effects.

A.1. Industry Definitions

One potential concern with our framework is that firm boundaries are definite, but industry boundaries are subjective. We can address this concern in a few ways. First, we note that NAICS codes were developed in the 1990s by a consortium of federal economic and statistical agencies, including the Bureau of Economic Analysis, US Department of Commerce, Bureau of the Census, and the Bureau of Labor Statistics. These agencies designed the classification system on the principle that industry definitions should be based on a single economic concept of the similarity of production processes. The boundaries of industries are limited by the degree of homogeneity of the production process among the establishments in the industry, subject to a minimum threshold of economic significance. Thus, though industry definitions are not as clearly delineated as corporate ownership, they are not arbitrary.[6]

Empirical evidence on the role of sectoral shocks for aggregate outcomes also supports the validity of industry codes. In particular, Carvalho and Gabaix (2013) shows that both the sectoral-level sales vector of the economy (as in Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012)) as well as the firm-level sales vector (as in Gabaix (2011)) help explain aggregate volatility. Because their formulation uses only the changing weights of sectors and firms in total output, their results suggest that the boundaries of industries are defined in an economically meaningful way and are not arbitrarily redefined to maintain equality in the size of sectors.

The second way we address this potential concern is to control for industry links based on the alternative industry definitions provided by Hoberg and Phillips (2016) (HP). HP use the text of the product definitions provided in firms' 10-K filings to identify the similarity of two firms' outputs. Thus, HP's definitions are likely to be

---

[6]More information on the development and principles guiding NAICS is available on the US Census Bureau's website: https://www.census.gov/naics/?008967

closer to a classification scheme based on the demand-side, compared to NAICS's scheme based on the supply-side. The third way we address this concern is to use a different level of industry aggregation in robustness tests, described below.

A.2. Truncation Bias

A potential concern with using Compustat data is that we only observe publicly-traded firms. We first note that both Atalay, Hortaçsu, Roberts, and Syverson (2011) and Herskovic, Kelly, Lustig, and Van Nieuwerburgh (2020) face the same issue when constructing empirical estimates of input-output networks. Atalay et al. show that though their sample is truncated, they do not observe bias in the network structure of the I-O links: the fraction of missing links is the same for central or peripheral nodes. Herskovic et al. note that measures based on a Herfindahl formulation (such as our CoHHI measure) give greater weight to larger firms, which reduces truncation bias. Nevertheless, the bias from using Compustat data reduces the correlations between key moments in their model. Thus, the truncation bias caused by Compustat data limitations makes it more difficult to identify statistically significant relationships in the empirical analysis. We also provide robustness tests later in the paper to better understand the potential for truncation bias. In particular, we re-estimate our main regressions using subsamples of Compustat data that are truncated by size thresholds to test whether our results differ when we exclude smaller firms.

III.B. The Structure of the Conglomerate Network

Networks exist across a continuum of types. On one extreme, random graphs contain nodes that are connected to each other with an equal probability (Erdős and Rényi, 1959). Thus, random graphs do not have central hubs. In addition, the number of connections to a node (degree) in a random graph exhibits relatively little variation around the average degree. Second, random graphs are not clustered, in which a nodes' neighbors are also connected with each other. At the other extreme of network types are ultra small world networks. These networks have very large hubs, with undefined degree variance. This means that the degree of an arbitrary network varies widely around the mean. The presence of prominent hubs in these networks reduces the average distance between all nodes and creates clusters of nodes. See Barabási (2016) for a detailed discussion of network types.

One way to measure the structure of a network is by its degree distribution. Random graphs have symmetric binomial degree distributions. Ultra small world networks have fat tailed degree distributions with long right tails indicating that a small number of nodes have many connections and a large number of nodes have few connections. A particular fat tailed distribution is the power law distribution, also known as the scale free distribution, $p(X > x) \sim Cx^{-\alpha}$, where $\alpha > 2$ is the scaling parameter. The lower is the α, the longer is the right tail. As α increases above 3, the network begins to resemble a random network.

B.1. The Structure of the Conglomerate Network in the Cross-Section

Figure III represents the complementary cumulative degree distribution of the conglomerate network in 1997 in log scale. A linear relationship indicates a fat tailed, power law distribution. The dashed line in the figure corresponds to an α of 3.44, estimated following Clauset, Shalizi, and Newman (2009). This α is comparable to 3.1 for the input-output network of industries, as estimated in Ahern and Harford (2014). Over our sample period, α is estimated to be 2.8 in an average year, though in 60% of years we reject the hypothesis that the network is power law distributed. Thus, these statistics show that the degree distribution of the conglomerate network has a substantially fat tail, even if not precisely a power law distribution. This means that the network is characterized by a relatively few hub industries with many connections to other industries and a relatively large number of industries with few inter-industry connections.

Additional network statistics confirm that the conglomerate network has a fat tail. In an average year, the average industry is connected to 6.8 other industries (degree centrality), though the median industry is connected to 3.3 other industries, consistent with a skewed degree distribution. The clustering coefficient of the average industry is 39%; for the median industry it is 32%, which is large relative to clustering in social networks. Finally, in an average year, the maximum path length between any two industries in the largest component is 7.5 links (7 at the median). Given that the largest component has 573 industries in an average year, this reflects that the conglomerate network exhibits small-world network features.[7]

---

[7]We discuss the statistics for the binary network for ease of exposition, but the interpretation of the weighted networks is similar. We report statistics for the giant component of the network, which is the largest set of interconnected nodes in a network. In an average year, 75% of industries are in the largest connected component. The remaining industries are typically in very small components of
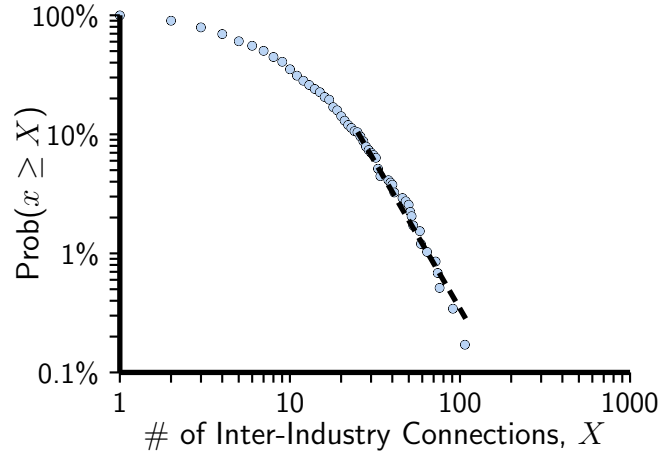
Figure III
Degree Distribution of Binary Conglomerate Network
This figure represents the distribution of degree centrality in log-log scale in the 1997 binary conglomerate network. Circles represent the degree centrality of industries, indicating how many direct connections an industry has to other industries. The dashed line is the from the estimate of α in the power distribution $P(k) = ck^{-\alpha}$.

B.2. Time-Series Evolution of the Conglomerate Network

Figure IV plots the time series of network statistics. First, the power law scaling parameter α has decreased over this period, while the variation in degree across nodes and clustering increased. This indicates that the network has evolved towards an ultra small world network with more prominent hubs. Second, from 2000 to 2018, average degree centrality decreased by more than 30% while eigenvector centrality decreased less noticeably. These results indicate that the average industry reduced its number of connections by removing connections to more peripheral industries.

These results highlight the changes in conglomerate ownership over the last two decades. While fewer industries are connected through conglomerate firms, the connections that remain are stronger and more centralized. In addition, industries are now more clustered together through conglomerates than in the past. This is consistent with Hoberg and Phillips (2021) which uses text of product descriptions to show that conglomerate firms have become more focused in related industries over our sample period.

---

one or two industries. We also exclude self-loops from the statistics, where an industry is connected to itself.
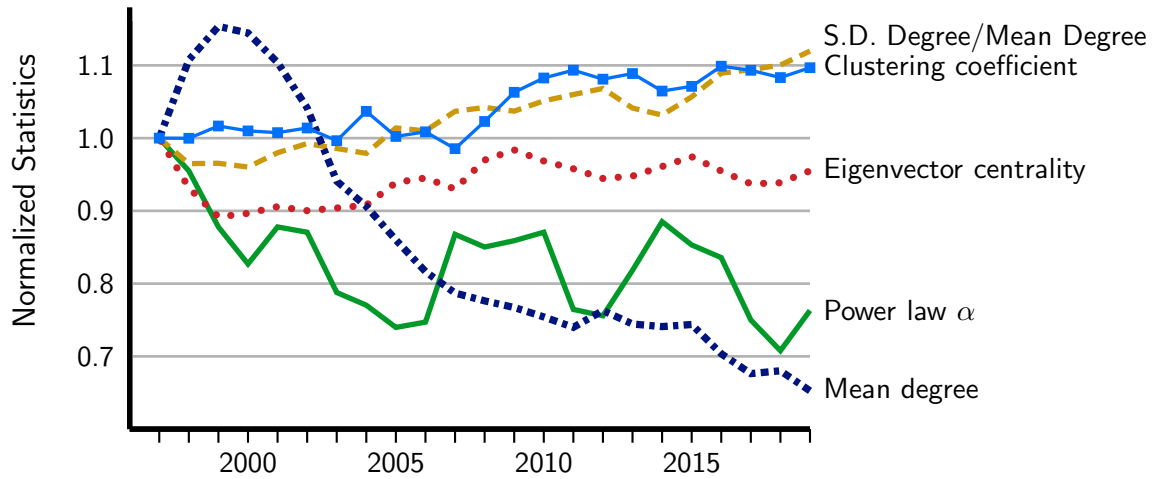
FIGURE IV
Time Series of the Binary Conglomerate Network
Five summary statistics are calculated yearly on the binary conglomerate network.
Mean degree is the number of inter-industry links into an average industry. Eigen-
vector centrality is the eigenvector value for the largest eigenvalue of the network
for an average industry. Clustering coefficient is the fraction of industries that are
connected to nodes that are also connected to each other for the average industry.
Power law α is the estimate of the scaling parameter of the power law distribution
$P(x) = Cx^{-\alpha}$. S.D. Degree/Mean Degree is the standard deviation of industry de-
gree divided by the average degree. All statistics are normalized by dividing by the
values in 1997.

Table I
Most Central Industries in the Conglomerate Network
This table lists the most central industries in the conglomerate network, using four different networks. The binary network is an undirected, unweighted network that records any connection between industries through conglomerate firms. The transmission network represents directed, weighted connections through the transmission from industries to firms and back to industries. The shared in-link network represents an undirected, weighted network of industry connections based on common conglomerate firms that span industries. The shared out-link network represents an undirected, weighted network of industry connections based on firms' common exposures to industry shocks. Panel A presents industries based on degree centrality. For the binary network, this is a count of links. For the weighted networks, the degree centrality is the sum of the weights per industry. Eigenvector centrality is calculated as Katz-Bonacich centrality with an attenuation factor equal to 90% of the max attentuation. This measure accounts for the centrality of the industries to which an industry is connected, but also allows all industries a minimum level of centrality. The top five industries listed are the industries that appeared most often in the yearly top five industry list across 1997 to 2018.

Binary Network	Transmission Network	Shared In-Link Network	Shared Out-Link Network
Panel A: Degree Centrality			
General industrial machinery	Motor vehicle parts	Cheese; natural/processed	General industrial machinery
Personal/business credit	Electric services	Frozen specialties	Electric services
Patent owners and lessors	Petroleum refining	Roasted coffee	Gas production/distribution
Mortgage banks/dep. functions	Motor vehicles/bodies	House refrig./freezers	Crude petroleum/nat. gas
Other plastics/mechanical rubber	Apparel and accessories	Cookies and crackers	Computer systems design
Panel B: Eigenvector Centrality			
Personal/business credit	Petroleum refining	Roasted coffee	Electric services
Mortgage banks/dep. functions	Crude petroleum/nat. gas	Cheese; natural/proc.	Gas production/distribution
Computer integrated systems	Motor vehicle parts	Frozen specialties	Crude petroleum/nat. gas
General industrial machinery	Electric services	Cookies and crackers	Natural gas transmission
Patent owners and lessors	General industrial machinery	Groceries & rel. prods.	Petroleum refining

B.3. Central Industries in the Conglomerate Network

Table I presents the most central industries in four versions of the conglomerate network: the binary network, transmission network, shared in-links, and shared out-links. The binary network is an unweighted network where links are defined by the existence of any conglomerate link between industries, without regard to their strength. The top five industries in each network are similar whether we define centrality using the sum of the connections (degree centrality) in Panel A or eigenvector centrality in Panel B.

Within the transmission network, the most central industries include motor vehicle parts, electric utility services, and petroleum related industries. High centrality in this network indicates that these industries are at the center of economic activity that is transmitting through conglomerate firms. The most central industries in the Co-HHI network of shared in-links are concentrated in the food industry: coffee, cheese, frozen foods, etc. High centrality in this industry reflects that these industries have high CoHHI scores with many other industries. In particular, the food industry is dominated by firms that operate in multiple industries with similar market shares in each. Finally, the most central industries in the shared out-link network are general industrial machinery, electric services, natural gas and petroleum, and computer systems design. These are industries that are highly connected to other industries through conglomerates with similar levels of exposure to common industry shocks.

IV. Comovement of Industry Growth Rates: Empirical Evidence

IV.A. Cross-Sectional Tests

First, we estimate the cross-sectional industry-to-industry variance-covariance matrix in Equation 8. In particular, we regress the time-series covariance of industry growth rates on the time-series average strength of industry connections in the CoHHI network, plus a dummy variable that indicates the diagonal entries in the matrix (i.e., an industry paired with itself).[8] We study two industry growth rates in our baseline tests: sales growth and asset growth.

The regression results are as follows,

---

[8]Equation 8 relates instantaneous growth rates with industry network at $t = 0$. Because we do not observe instantaneous growth rates at each period, we proxy for these relationships using growth rates observed over the time series.

$$Cov(g_{sales,i}, g_{sales,j}) = \underset{(0.029)}{0.406} + \underset{(1.097)}{5.171} \times \text{CoHHI} + \underset{(0.571)}{8.864} \times \text{Same Industry Dummy}$$

$$Cov(g_{assets,i}, g_{assets,j}) = \underset{(0.029)}{0.293} + \underset{(1.085)}{3.028} \times \text{CoHHI} + \underset{(0.660)}{10.354} \times \text{Same Industry Dummy},$$

where coefficients and industry-paired clustered standard errors are reported in percentages with a sample size of roughly 250,000 industry-pair observations.

These results show that there is a positive cross-sectional correlation between Co-HHI and the covariance of sales growth and asset growth rates. This implies that industries with higher CoHHI connections also have higher comovement of fundamental economic growth rates. The comovement of industry growth rates is driven by conglomerate firms with large market shares in multiple industries.[9]

Next, the the positive coefficient on the dummy variable reflects that more concentrated industries have more volatile growth rates. This result is consistent with the model's intuition that concentrated industries have greater exposure to firm-specific shocks. These results show that industry volatility is driven, in part, by large firms as in Gabaix (2011), but only if the firms have large market shares within the industry.

IV.B. Panel Tests

Second, we estimate panel regressions with fixed effects to isolate the effect of changes in the conglomerate network on within-industry-pair changes in comovement. To estimate a panel model with yearly observations, we cannot use the time-series covariance as our dependent variable. Instead, to motivate our empirical model, note that the squared difference of two industries' growth rates at $\tau = 0$, is as follows:

$$(g_j - g_k)^2 = (\eta_j - \eta_k)^2 + 2(\eta_j - \eta_k)\left(\sum_{i=1}^{n} h_{i,j}\varepsilon_i - \sum_{i=1}^{n} h_{i,k}\varepsilon_i\right) + \left(\sum_{i=1}^{n} h_{i,j}\varepsilon_i - \sum_{i=1}^{n} h_{i,k}\varepsilon_i\right)^2. \tag{15}$$

In expectation, the squared difference of growth rates is:

$$E\left[(g_{j,1} - g_{k,1})^2\right] = E(g_{j,1}^2) + E(g_{k,1}^2) - 2E(g_{j,1}g_{k,1})$$
$$= 2\sigma_\eta^2 + \sigma_\varepsilon^2(HHI_j + HHI_k) - 2\sigma_\varepsilon^2 CoHHI_{j,k}. \tag{16}$$

---

[9]In unreported tests, we verify that our results hold when we estimate the network correlations using exponential random graph models, following Ahern and Harford (2014).

The signs of the coefficients imply that, all else equal, two concentrated industries are likely to have less similar time series of growth rates, while two co-concentrated industries are likely to have more similar growth rates. This reflects that concentrated industries have greater idiosyncratic firm-specific risk, but co-concentrated industries share common firm-level risk.

To test this relationship empirically, we estimate a more generalized and flexible version of the model's prediction in the following regression:

$$(g_{j,\tau} - g_{k,\tau})^2 = \gamma_1 \text{Shared In-Links (Co-HHI)}_{jk,\tau} \tag{17}$$
$$+ \gamma_2 \text{Shared Out-Links}_{jk,\tau}$$
$$+ \phi(HHI_{j,\tau} + HHI_{k,\tau})$$
$$+ \rho_\tau + \delta_{jk} + \psi \, \text{Controls}_{jk,\tau} + \varepsilon_{jk,\tau},$$

where $g_{i,\tau}$ represents industry i's growth rate at time τ, Shared In-Links (Co-HHI)$_{ij,\tau}$, and Shared Out-Links$_{ij,\tau}$ represent network connections at time τ, ρ_τ is a time fixed effect, δ_{ij} is an industry pair fixed effect, and Controls include time-varying industry-pair control variables, discussed below. This regression is estimated using undirected industry-pairs because the explanatory variables represent undirected links. We include Shared Out-Links in our regression tests, even though they do not appear in the theoretical formulation, because shocks might transfer in the opposite direction than we have assumed. Our goal is not to calibrate our model, but to use it as a guide for understanding the patterns in the data.

The regression above does not include the transmission network because it is derived from the covariance of growth rates at the initial period. If we allow for higher order connections in the network, we need to include the transmission network. As shown in Equation 13, the covariance of growth rates after one cycle in the network is a function of HHI, CoHHI, and the transmission network. We estimate a more general empirical model, as follows:

$$(g_{j,\tau} - g_{k,\tau})^2 = \beta_1 \text{Transmission}_{jk,\tau} + \beta_2 \text{CoHHI}_{jk,\tau} \tag{18}$$
$$+ \phi(HHI_{j,\tau} + HHI_{k,\tau}) + \rho_\tau + \delta_{jk} + \psi \, \text{Controls}_{jk,\tau} + \varepsilon_{jk,\tau},$$

where the Transmission variable is defined as a dummy variable if industry pair (j, k) is positive in matrix $H'FF'H$ which reflects industries that are two links away in the transmission network.

In both regressions, the industry pair fixed effects, δ_{ij}, account for time-invariant cross-sectional variation in industry pairs (directed or undirected). This controls for

any cross-industry trait that remains stable over time, such as the nature of the product (e.g., goods vs. services), the level of government regulation, access to capital, and the importance of intangible assets. We also include time fixed effects, ρ_τ, to control for economy-wide fluctuations and to isolate within-industry pair fluctuations. Finally, we also run specifications that use a dummy variable that represents the presence of a connection in the network, without regard to the strength of the connection. For the shared links networks, this dummy variable is identical in the in-links and out-links networks.[10]

The regressions also include variables to control for customer-supplier relationships and product-market similarities. We measure the customer-supplier connections between industry pairs using data from the industry-by-industry total requirement table from the Benchmark Input-Output (IO) Accounts released by the Bureau of Economic Analysis (BEA). These data measure the dollar amount of intermediary industry output required per dollar of final demand. We use the most recent data for years 1997, 2002, 2007, 2012, and 2017. For example, from year 1997 to 2001, we use the 1997 total requirement table. In a few cases, our industry pairs cannot be matched to the IO industries; however our results are qualitative unchanged if we drop these industry pairs from the sample instead.

Second, we control for time-varying asset similarities between industry pairs based on the text-based product similarity measure of Hoberg and Phillips (2016) (HP). To convert their similarity measures to our industry pairs, we identify stand-alone firm-pairs with positive HP similarity in each industry pair. We then calculate the average similarities between these firm pairs in our industry-pairs to proxy for asset similarity. Specifically, for industry pair (i,j), with m stand-alone firm-pairs (k,l) with positive HP similarity, where k denotes firms in industry i and l denotes firms in industry j, the asset similarity of industry pair (i,j) is $\frac{\sum_m HP_{k,l}}{m}$, where $HP_{k,l}$ is the text-based product similarity of the firm pair (k,l). We assign a zero to industry pairs with missing similarity scores because this implies the HP measure is below a minimum threshold.

In sum, these regressions are designed to isolate how time-series variation in the strength of the connection between the industries in the conglomerate network explain time-series variation in the comovement of two industries' growth rates. If two industries' connection through conglomerates becomes stronger, we expect to see

---

[10]We cluster standard errors at the industry-pair level, but our results persist if we double cluster standard errors by each industry of the pair.

stronger comovement in their growth rates. After presenting these baseline results, we consider sampling bias and endogeneity in robustness tests.

IV.C. Summary Statistics

Table II provides summary statistics for all of variables used in the panel regressions. In the median industry-year, sales growth is 2.67%, asset growth is 1.1%, and employment growth is -0.90%. Squared differences in these rates vary considerably. In an average year, the squared difference of sales growth is 20.88%, compared to 5.56% at the median. Asset growth is similar, but squared differences in employee growth rates are much lower. This indicates that there is less variation across industries in employee growth than sales and asset growth. All of the network measures are skewed because the network is sparse, with most industry pairs having no connection. The dummy variables indicates that 1.4% of industry-pairs are connected in the conglomerate network, which is similar in magnitude to the Hoberg-Phillips measure and the IO links.

IV.D. Correlations of Alternative Industry Connections

Figure V presents the yearly cross-sectional correlation between the measures of the conglomerate network, the input-output network, and Hoberg-Phillips industry similarity measures. Observations are based on directed industry pairs, excluding own industry loops. Panel A shows that the strength of the transmission link between industry pairs is more correlated with the strength of shared in-links (about 50%) than with the strength of shared out-links (about 28%). This represents, in part, that there is more variation across in-links than across out-links because market shares vary more widely across industries than do industry shares across firms. Shared in-links and out-links are less correlated with each other (about 11%).

Panel B shows that shared in-links and the transmission links have persistently low correlations (less than 6%) with the Hoberg-Phillips measure of industry connections. In contrast, shared out-links in the conglomerate network display higher correlations (about 14%) and exhibit larger time-series variation. These results reflect that the similarity of product descriptions across industries from HP is tied more closely to a firm's industry share than to the firm's market share. This follows because HP similarity measures are based on the firm's perspective of its products, rather than the industry's perspective of its leading firms.

TABLE II
Summary Statistics: Panel Data
This table presents summary statistics for the variables used in the regression analysis. All variables
are in percentages. Summary statistics are presented for non-directed industry-pair observations.

| | Mean | S.D. | Percentile | | | Observations |
			25th	50th	75th	
Sales growth	-0.274	32.654	-12.292	2.668	14.639	$3,827,922$
Asset growth	-0.954	32.751	-12.480	1.102	13.506	$3,781,760$
Employment growth	-1.309	9.785	-6.085	-0.897	3.446	$3,316,726$
$(\text{Sales growth}_i - \text{Sales growth}_j)^2$	20.884	35.911	0.947	5.557	24.139	$3,570,850$
$(\text{Asset growth}_i - \text{Asset growth}_j)^2$	21.693	36.750	0.997	5.844	25.814	$3,482,085$
$(\text{Employment growth}_i - \text{Employment growth}_j)^2$	1.687	3.278	0.092	0.458	1.674	$3,269,717$
Shared in-links (Co-HHI)	0.039	1.112	0.000	0.000	0.000	$4,113,989$
Shared out-links	0.162	3.269	0.000	0.000	0.000	$4,113,989$
Shared links dummy	1.396	11.731	0.000	0.000	0.000	$4,113,989$
Transmission	0.064	1.403	0.000	0.000	0.000	$4,113,989$
Transmission dummy	1.396	11.731	0.000	0.000	0.000	$4,113,989$
HHI	52.977	31.230	26.546	47.127	83.672	$4,113,989$
Sum of HHI	102.521	44.546	66.740	103.022	132.622	$4,113,989$
Input-Output link	1.132	9.542	0.000	0.004	0.063	$4,113,989$
Hoberg-Phillips similarity	0.037	0.437	0.000	0.000	0.000	$4,113,989$

(A) Conglomerate Network

(B) Hoberg-Phillips (HP) and Conglomerate Network

(C) Input-Output (IO) and Conglomerate Network
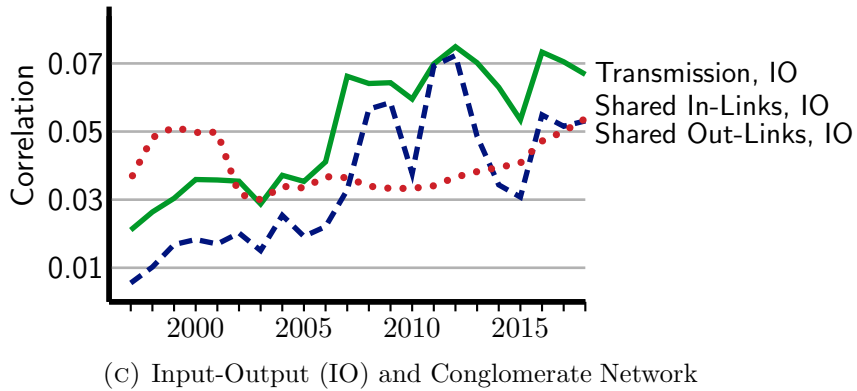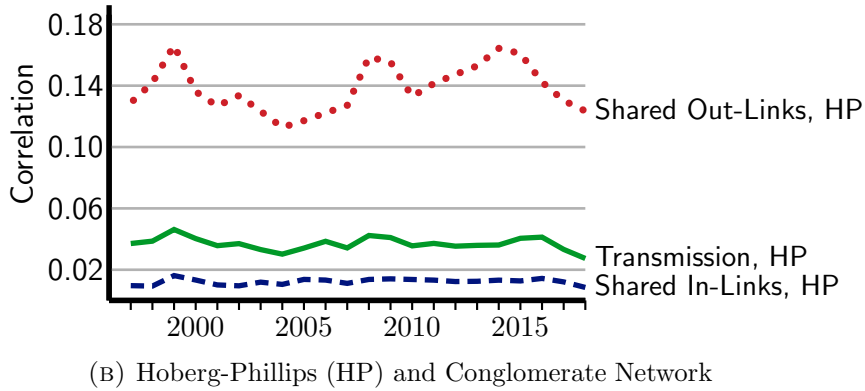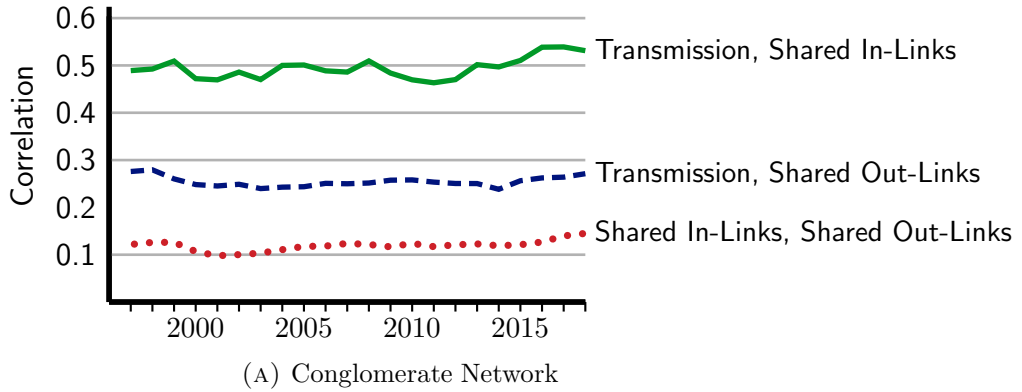
FIGURE V

Yearly Cross-Sectional Correlation of Industry Relations

Each panel presents the yearly cross-sectional correlation of an industry-pair measure, excluding own-industry pairs. Panel A presents the correlations of the conglomerate network measures, Panel B presents the correlations of the conglomerate network measures with Hoberg-Phillips (HP) measures of industry connections, and Panel C presents the correlations of the conglomerate network with the Input-Output network (IO).

Panel C of Figure V shows that the conglomerate network has relatively little correlation with the input-output network (less than 7%), though the correlation is increasing during the sample period. This suggests that vertical integration is increasing from 1997 to 2018, relative to unrelated diversification. This is consistent with Figure IV, which shows that conglomerate firms have become more focused in central industries, rather than peripheral industries.

In sum, the results in Figure V show that the conglomerate network is not just a proxy for existing measures of industry connectedness. Instead, the conglomerate network represents a unique form of inter-industry connections. However, we include these variables in our tests as alternative explanations of industry comovement.

IV.E. Baseline Results

Table III presents estimates of the relationship between shared in-links and out-links on the comovement of industry growth rates. For sales growth and asset growth, an increase in the strength of shared links is negatively correlated with the squared differences in growth rates. Thus, consistent with our prediction, as industries become more closely connected in the conglomerate network, their growth rates comove more closely. Because we control for year fixed effects and industry-pair fixed effects, these results are not driven by cross-sectional differences in the nature of industries, nor are they explained by economy-wide fluctuations in the time-series of growth rate levels or correlations. In addition, the results are not driven by customer-supplier relationships or asset similarities.

Also consistent with the model, the sum of HHI is positively correlated with squared differences in growth rates with a high degree of statistical significance. This reflects that industries with greater internal concentration have weaker connections to other industries. Thus, an increase in an industry's HHI reduces the comovement of its growth rate with other industries.

It is easiest to interpret the magnitude of the dummy variable for the squared difference of growth rates. For both outcome variables, the presence of a connection reduces the squared difference in growth rates by about 8.5% of the average squared difference and about 31% of the median squared difference. For comparison, a one standard deviation increase in the sum of HHI is associated with an increase of squared differences in growth rates equal to 45% of the median for sales growth and 29% of the median for asset growth.

TABLE III
Comovement of Industry Growth and Shared Network Links
This table presents coefficient estimates from panel regressions where the dependent variable is $(g_k - g_j)^2$, where g_i is the growth rate of industry i for sales (Panel A) and assets (Panel B). Variable definitions are provided in the text. All regressions include industry-pair and year fixed effects. Coefficients and industry-pair clustered standard errors (in parentheses) are in percentages. Statistical significance indicated by ***, **, and * for significance at 0.01, 0.05, and 0.10.

Panel A: Sales growth				
Shared in-links (Co-HHI)	−22.494***		−21.710***	
	(2.583)		(2.594)	
Shared out-links		−4.897***	−3.552***	
		(1.214)	(1.147)	
Shared links dummy				−1.722***
				(0.215)
Sum of HHI	2.513***	2.508***	2.511***	2.499***
	(0.112)	(0.112)	(0.112)	(0.112)
Input-Output link	1.027***	1.025***	1.029***	1.036***
	(0.361)	(0.361)	(0.361)	(0.361)
Hoberg-Phillips similarity	−20.894***	−20.580***	−20.808***	−20.598***
	(4.362)	(4.355)	(4.360)	(4.355)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.068	0.068	0.068	0.068
Observations	3,563,872	3,563,872	3,563,872	3,563,872
Panel B: Asset growth				
Shared in-links (Co-HHI)	−23.258***		−22.459***	
	(2.635)		(2.643)	
Shared out-links		−5.042***	−3.634***	
		(1.455)	(1.363)	
Shared links dummy				−1.881***
				(0.226)
Sum of HHI	1.688***	1.682***	1.685***	1.672***
	(0.116)	(0.116)	(0.116)	(0.116)
Input-Output link	0.193	0.189	0.194	0.201
	(0.367)	(0.367)	(0.367)	(0.367)
Hoberg-Phillips similarity	−29.959***	−29.653***	−29.878***	−29.656***
	(5.658)	(5.641)	(5.654)	(5.639)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.054	0.054	0.054	0.054
Observations	3,474,363	3,474,363	3,474,363	3,474,363

Next, Table IV presents the estimates of the transmission network on the comovement of industry growth rates. The estimates show that the transmission network is negatively and significantly related to difference in industry growth rates for sales and asset growth. This means that when two industries become connected through two links of the transmission network, their growth rates comove more closely. As before, the sum of industry HHIs has the opposite relationship to the transmission network, consistent with the model's prediction.

We also estimate industry-level tests of the effect of distance in the conglomerate network on the comovement of industry growth rates. In particular, we estimate each industry's exposure to the growth rates of other industries through the conglomerate network at different distances. To calculate industry j's exposure to industries that are at a distance of one, we calculate the weighted average of the growth rates of all other industries, where the weights are the strength of the connections between industry j and the other industries in the transmission network. For a distance of two, we take the average growth rate of all industries that are exactly two links away from industry j in the transmission network. Similarly, we calculate the average growth rate of industries that are at industry j's maximum distance, excluding industries that are not connected at any distance to industry j.

Table V shows that industries growth rates are higher when the growth rates of connected industries' growth rates are higher. The effect is largest for industries that are at a distance of one, while the results are still positive for industries that are at a distance of two in the transmission network for sales growth. This shows that industries comove with other industries in the network, even if they are not directly connected. Growth rates of industries at the maximum distance are uncorrelated. As expected, the longer is the distance between two industries in the network, the weaker is their comovement. Because these results control for IO-weighted growth rates of other industries, year fixed effects, and industry fixed effects, the results are not driven by customer-supplier links, macroeconomic fluctuations, or cross-sectional heterogeneity in industry traits.

To provide additional evidence that the results are driven by the conglomerate network, we calculate a placebo variable in which industries are randomly assigned a transmission strength using the empirical distribution of linkages in the data. If the results were driven by macroeconomic factors, rather than connections in the conglomerate network, the results would persist in the placebo test. However, Table V shows that the placebo variable is unrelated to an industry's growth rate.

TABLE IV
Comovement of Industry Growth and the Transmission Network

This table presents coefficient estimates from panel regressions where the dependent variable is $(g_k - g_j)^2$, where g_i is the growth rate of industry i for sales (Panel A) and assets (Panel B). Transmission dummy represents industry pairs that are two links away in the transmission matrix. Shared links dummy represents industries that are one link away in the shared links matrix. All regressions include industry-pair and year fixed effects. Coefficients and industry-pair clustered standard errors (in parentheses) are in percentages. Statistical significance indicated by ***, **, and * for significance at 0.01, 0.05, and 0.10.

Panel A: Sales growth			
Transmission dummy	-1.231***	-1.201***	-1.185***
	(0.073)	(0.073)	(0.073)
Shared in-links (Co-HHI)		-20.453***	
		(2.588)	
Shared links dummy			-1.282***
			(0.216)
Sum of HHI	2.403***	2.408***	2.398***
	(0.113)	(0.113)	(0.113)
Input-Output link	1.029***	1.034***	1.040***
	(0.361)	(0.361)	(0.361)
Hoberg-Phillips similarity	-20.873***	-21.055***	-20.798***
	(4.362)	(4.366)	(4.360)
Industry-pair and year fixed effects	Yes	Yes	Yes
Adjusted R^2	0.068	0.068	0.068
Observations	3,563,872	3,563,872	3,563,872
Panel B: Asset growth			
Transmission dummy	-1.244***	-1.212***	-1.192***
	(0.077)	(0.077)	(0.078)
Shared in-links (Co-HHI)		-21.224***	
		(2.636)	
Shared links dummy			-1.437***
			(0.227)
Sum of HHI	1.577***	1.582***	1.572***
	(0.116)	(0.116)	(0.116)
Input-Output link	0.194	0.199	0.204
	(0.367)	(0.367)	(0.367)
Hoberg-Phillips similarity	-29.939***	-30.119***	-29.854***
	(5.657)	(5.667)	(5.650)
Industry-pair and year fixed effects	Yes	Yes	Yes
Adjusted R^2	0.054	0.054	0.054
Observations	3,474,363	3,474,363	3,474,363

<div align="center">

TABLE V

Industry-Level Growth and Transmission Network Links

</div>

This table presents coefficient estimates from panel regressions where the dependent variable is the industry growth rate of sales (Panel A) and assets (Panel B). Transmission-weighted growth is the growth rate of all other industries, weighted by their transmission link in the conglomerate network. Transmission-weighted placebo randomizes the transmission strength across industries. Average growth at a distance of two (longest distance) is the average growth rate of industries that are exactly two links (maximum distance) away from the focal industry. IO-weighted growth is the growth rate of other industries weighted by their IO strength to the focal industry. Coefficients and industry-clustered standard errors (in parentheses) are in percentages. Statistical significance is indicated by ***, **, and * for significance at 0.01, 0.05, and 0.10.

Panel A: Sales growth					
Transmission-weighted growth	19.157***	18.233***		18.260***	18.026***
	(2.111)	(2.262)		(2.289)	(2.297)
Transmission-weighted placebo			2.677		
			(1.815)		
Average growth at distance of two				12.928*	
				(6.720)	
Average growth at longest distance					−0.827
					(1.995)
IO-weighted growth		9.460***	13.826***	9.214***	8.988***
		(3.322)	(3.309)	(3.350)	(3.386)
Industry & year fixed effects	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.053	0.061	0.052	0.061	0.063
Observations	11,517	10,275	10,282	10,157	9,920
Panel B: Asset growth					
Transmission-weighted growth	16.082***	15.006***		14.993***	14.694***
	(1.915)	(2.066)		(2.097)	(2.099)
Transmission-weighted placebo			2.878		
			(1.821)		
Average growth at distance of two				−3.368	
				(6.816)	
Average growth at longest distance					−0.214
					(1.688)
IO-weighted growth		3.810	6.463*	3.584	3.637
		(3.541)	(3.533)	(3.559)	(3.576)
Industry & year fixed effects	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.036	0.041	0.034	0.040	0.042
Observations	11,377	10,161	10,168	10,045	9,834

V. Endogeneity Concerns: A Quasi-Natural Experiment

Conglomerate firms do not randomly choose the industries in which they operate. Instead, it is reasonable to believe that firms choose to diversify into industries that tend to comove with each other, which could produce correlations similar to the results we have shown. Thus, in contrast to our hypothesis that conglomerates transmit shocks from one industry to another, it is possible that the entirety of the shocks would have been transmitted even without the conglomerate network. In this section, we address this concern.

First, it is important to recognize that the conglomerate network is not formed completely through endogenous choice. In the model, the strength of the ties between industries is based on both the firms' industry shares and firms' market shares. While firms endogenously choose the industries in which they operate, they do not endogenously choose their market shares in each industry. In particular, CoHHI is a measure of the commonality of market shares held by the same firm in two different industries, not just the choice to operate in both industries. In addition, our model generates predictions about the relationship between HHI and comovement, which are also supported in the empirical results. Because firms do not endogenously choose HHI, these results cannot be caused by the endogenous choice of firms.

Second, to the degree that the conglomerate network is endogenously determined, it is important to note that all of our results persist after controlling for industry-pair fixed effects. Thus, any time-invariant factor that leads industries to comove is absorbed by these fixed effects. In addition, we control for remaining time-series variation that is driven by major economic determinants of conglomeration, including vertical customer-supplier relationships and complimentary assets of Hoberg and Phillips. For our results to be spurious, there would need to be an underlying time-varying factor explaining all of the remaining time-series correlation that is both orthogonal to vertical relations and asset complementarity and also causes both comovement among industry growth rates and firms to increase their market shares within these industries at the same time.

Though the above arguments help to limit the magnitude of endogeneity concerns, as an additional analysis we study the effect of the United States granting Permanent Normal Trade Relations (PNTR) to China. PNTR was granted by Congress in October 2000 and became effective when China joined the World Trade Organization (WTO) at the end of 2001. Before the conferral of PNTR, the tariff rates of US imports from China required annual renewals, which had imposed a great amount

of uncertainty on the trade relations between China and US. Although PNTR did not change the import tariff rates that the United States actually applied to Chinese goods, it removed the uncertainty associated with these annual renewals. Without the yearly renewal of favorable rates, US import tariffs would have increased substantially.

Pierce and Schott (2016) show that granting PNTR to China caused declines in employment within US industries that were most at risk of higher tariffs without PNTR (exposed industries). We exploit the same shock to test whether employment growth declined for industries that were not directly affected by PNTR, but were connected to the exposed industries through the conglomerate network. Though this setting does not provide an exogenous change to the network, it does identify a specific exogenous shock that we can observe transmitting through the network.

Following Pierce and Schott (2016), we measure the NTR gap as the difference between the non-NTR rates to which tariffs would have risen in the industry if annual renewal had failed and the NTR tariff rates that were locked in by PNTR. This shock is time-invariant for each industry and therefore absorbed by industry fixed effects. However, as in Pierce and Schott, we can identify the effect of the NTR gap through its interaction with a dummy variable, $Post$, which is equal to one from 2001 onward to indicate years after the passage of PNTR.

To identify industries that could potentially receive the NTR shock from the exposed industries through the conglomerate network, we use the transmission network, $F'H$. To help address reverse causation, we use the 1999 network to ensure it is exogenous to the NTR shock in 2000. Thus, the results are not driven by conglomerates forming new industry linkages in response to the tariff shock.

For each industry k in the transmission network, we weight the NTR gap by the entries in the kth column of $F'H$. These entries represent the shocks from row industries that transmit to industry k. Thus, our weighting scheme provides for variation in an industry's exposure to the NTR gap based on the magnitude of the gap and the magnitude of the connection to the affected industry. We also normalize the measure by the sum of the column, excluding industry i. Therefore, Transmission NTR $\text{gap}_i = \frac{\sum_{k \neq i} \text{Transmission}_{k,i} \times \text{NTR gap}_k}{\sum_{k \neq i} \text{Transmission}_{k,i}}$. As above, this variable is identified through the interaction with the post dummy variable. If the employment shock transmits through the conglomerate network, we expect to find a negative coefficient on the interaction between the strength of the conglomerate network and the NTR gap. This reflects the change in employment for more exposed industries relative to less exposed industries.

We estimate the model using industry level employment data from the US Census Bureau's County Business Patterns (CBP) files from 1997 to 2018. These data offer the most detailed view of the United States' industrial structure available to the public. They provide annual data on employment at a detailed industry level, which covers nearly all establishments with paid employees in the private sector of the United States. Therefore, unlike the Compustat Segment data, the industry employment data in CBP files covers the universe of firms, both publicly listed and private. We calculate the strength of network linkages using data for all industries, but estimate regression coefficients using a sample of manufacturing industries, following Pierce and Schott.

Table VI presents the results of these tests. Column 1 replicates the findings in Pierce and Schott (2016). US industries directly exposed to the NTR shock experience a significant decline in employment. Column 2 shows that this shock transmits through the conglomerate network. Industries with greater network connections to the industries directly exposed to the NTR shock also experience declines in employment. These results provide further evidence that economic shocks transmit across industries through conglomerate firms.

In columns 3 and 4 of Table VI, we provide two more robustness checks. In column 3 we create a placebo variable identical to the interaction between the transmission network and the NTR gap by randomly assigning industries to the actual network weights in the transmission network, as above. We find that there is no correlation between this placebo network and employment growth. This implies that our main results are not caused by a general network-wide trend. Finally, in column 4, we include the strength of network connections from the 1997 input-output network. We find that downstream industries are affected by the shock, though upstream industries are not, consistent with Pierce and Schott (2016). However, we still find that the transmission network is significantly related to employment in connected industries.

These tests are important because they allow us to exploit the cross-sectional variation in the exposure to an identifiable shock. We also use the predetermined conglomerate network prior to the NTR shock to rule out reverse causality and control for input-output relations to control for alternative explanations. In addition, our dependent variable is not based on public firm filings in Compustat. Even after these controls, we still find that the conglomerate network serves as a conduit to spread economic shocks.

<div align="center">

Table VI

Transmission of Tariff Shocks Through the Conglomerate Network

</div>

This table presents coefficient estimates from panel regressions where the dependent variable is the industry growth rate of employment. NTR Gap is the difference between the non-Normal Trade Relations tariff rate and the NTR tariff rate. Post × Transmission NTR Gap$_i$ is the transmission gap weighted by the transmission matrix from the conglomerate network. IO Customer and IO Supplier are inter-industry connections from the input-output network. Coefficients and industry-clustered standard errors (in parentheses) are in percentages. Statistical significance is indicated by ***, **, and * for significance at 0.01, 0.05, and 0.10.

	Dependent variable: Employment growth			
	(1)	(2)	(3)	(4)
Post × NTR Gap$_i$	−7.420***	−5.756***	−7.428***	−2.250
	(1.457)	(1.674)	(1.453)	(1.999)
Post × Transmission NTR Gap$_i$		−4.640**		−4.091*
		(2.074)		(2.166)
Post × Placebo NTR Gap$_i$			0.460	
			(2.193)	
Post × IO Customer NTR Gap$_i$				−7.683**
				(3.207)
Post × IO Supplier NTR Gap$_i$				−2.060
				(4.976)
Industry and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.176	0.176	0.175	0.178
Observations	5,375	5,375	5,375	5,358

<div align="center">

VI. Robustness Tests

</div>

Though the above results show that inter-industry strength in the conglomerate network is correlated with the strength of the comovement of industry growth rates, there are potential concerns related to sampling bias, endogeneity, industry definitions, and mechanisms.

First, to address the concern that the dependent and explanatory variables are both derived from data restricted to publicly-traded firms, in Online Appendix Table I, we estimate the main regressions using industry level employment data from the US CBP data, as above. Consistent with the baseline results, a stronger CoHHI link between industries is associated with strong comovement in their growth rates of employment.

Second, to help address reverse causality concerns, we used lagged variables in the tariff shock tests in Table VI. To provide additional robustness, in Online Appendix Table II, we also re-estimate our main equations using lagged explanatory variables with industry-pair fixed effects. These tests control for any time-invariant determinant of industry comovement (industry-pair fixed effects) as well as reverse causation in the time-series, in which firms tend to diversify into industries that have recently experienced greater comovement (lagged explanatory variables). The results are qualitatively similar as the main results.

Third, to alleviate concerns that using large publicly traded firms to construct our network creates truncation bias in our results, we run a series of robustness tests. To test how changing the size threshold of the available Compustat data affects our results, we estimate the main tests using a sample that excludes observations of firms that are below the 25th percentile of sales in a given year. Online Appendix Table III shows that our results are nearly identical as in the main tests that use the full Compustat sample. In Online Appendix Table IV, we use a more extreme threshold of the 50th percentile of sales. The magnitude of the results are smaller, but they are still statistically significant. These tests suggest that our results are likely to be stronger if we had a larger sample that included smaller, private firms. Likewise, Online Appendix Table V shows that the results are nearly identical when we exclude foreign firms.

Fourth, in untabulated results, we compare Compustat data to the Economic Census data published by the US Census Bureau. For each industry, we calculate the ratio of Compustat sales to Census sales. We then divide industries into those that have above-median representation by Compustat firms and those with below-median representation. We find that the total Census-level sales of the average industry with above-median representation is statistically identical to the Census-level sales of the average industry with below-median representation. This means that large and small industries, as measured by Census data, are equally represented by Compustat data. This helps alleviate any concerns that Compustat data is biased towards large industries because it is biased towards large firms.

Fifth, we reconstruct our network measures using SIC 3-digit industry definitions, which are considerably more coarse than the NAICS codes we use in our main tests: the number of industry-pairs in the network based on 3-digit SIC codes is only about 25% of the number of industry-pairs in the main specifications. Online Appendix

Table VI shows that our main results are robust to constructing the network using these more coarse industry definitions.

Sixth, we use the tariff shock to provide direct evidence of the transmission of shocks from one segment to another within the same conglomerate firm. We create a variable that measures a firm-segment's exposure to the tariff shock from other segments within the same conglomerate, similar to our industry-level measure in Table VI. We find that a segment's sales growth is lower when other segments within the same firm have greater exposure to the tariff shock, consistent with prior literature on internal reallocation. The results and further details are in the Online Appendix.

Finally, we provide suggestive evidence that financial constraints influence the transmission of shocks within conglomerates. We define the financial constraints of an industry pair as the sales weighted average of the financial constraints index of Whited and Wu (2006) for the conglomerates that operate in both industries. If an industry pair is unconnected, we set its financial constraint to zero. In Online Appendix Table VIII, the interaction of CoHHI and the financial constraints variable shows that industry pairs with more financially constrained firms tend to comove less with each other. Consistent with Appendix B.1 of Matvos and Seru (2014), these results could imply that financially constrained firms allocate internal resources efficiently, causing less spillovers and thus, less comovement between industries. In contrast, unconstrained firms might engage in corporate socialism, causing more comovement of industry pairs. We caution that these results are speculative and a full investigation of the causes of internal redistribution is beyond the scope of this paper.

VII. Conclusion

To organize an analysis of the economy, researchers typically partition economic activity into a set of isolated industries, grouped together by common suppliers, production processes, technology, or customers. At the same time, economic activity is grouped together by common control derived from ownership, typically organized as firms. These two groupings create overlapping boundaries of economic activity, in which industries are groupings of firms, but at the same time, some firms are groupings of industries. In this paper, we organize these overlapping groupings into a unified network of industries and firms. Using this network perspective, we show that economic activity transmits across the economy through conglomerate firms that span multiple industries.

The core of our model is an affiliation network in which industries are affiliated with firms and firms are affiliated with industries, but firms have no direct connections with other firms and industries have no direct connections with other industries. This creates a directed, weighted network in which shocks transmit from firms to industries in proportion to a firm's market share of total industry sales and from industries to firms in proportion to the fraction of a firm's sales attributed to the industry. From the affiliation network, we create three unique inter-industry networks that vary in their interpretation. In the first network, inter-industry connections represent the strength of the links from an industry through conglomerate firms, back to other industries. The connections in the second network represent two industries' commonality of shared in-links from overlapping conglomerate firms. The final network represents the strength of shared out-links from industries to common firms.

An important new perspective offered by our network is a microfoundation for the widely-used Herfindahl-Hirschman Index (HHI). We show that HHI is a special case of a more general measure we call CoHHI which represents the shared in-links of an industry. If the same firms command more similar market shares in two industries, then the industries have a higher CoHHI. This reflects a measure of cross-industry sales concentration through overlapping firms. Furthermore, just as statistical variance is a special case of covariance, HHI is a special case of CoHHI. This derivation also provides an economic rationale for HHI's specific formulation as the sum of squared market shares, which is an ad hoc choice made in the original papers by Hirschman and Herfindahl.

The conglomerate network also provides predictions on the covariance of growth rates across industries. Assuming idiosyncratic firm and industry shocks, we show that the volatility of an industry's growth equals a common volatility of industry growth rates plus firm-level volatility weighted by the HHI of the industry. The covariance of industry growth rates is equal to firm-level volatility weighted by the CoHHI between the two industries. Thus CoHHI describes the comovement of growth rates across the economy.

We test the predictions of our model using a panel of segment level data from 1997 to 2018. We show that the stronger is the connection between two industries in the conglomerate network, the stronger is the comovement of their growth rates of sales, asset levels, and employment. These results persist after controlling for industry-pair fixed effects, year fixed effects, changes in industry HHIs, customer-supplier links, and asset similarity measures. To help identify a causal relationship, we use lagged

values of network connections and find qualitatively similar results. We also exploit the cross-sectional variation in industries' exposure to tariff rate shocks following the granting of normal trade relations to China. We find that employment falls more in industries that have stronger connections to the industries directly affected by the tariff rate shock.

We believe our results have far-reaching implications. First, they help explain how idiosyncratic shocks aggregate to macroeconomic fluctuations. Second, they provide a new perspective on the incidence of diversified conglomerates across industries and time. Third, the conglomerate network generates a new measure of cross-industry concentration, CoHHI, and gives an economic microfoundation to HHI. Given the prevalence of HHI in academic research and among policy-makers, we believe this measure will be useful for understanding the organizational structure of economic activity within and across industries.

References

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, "The Network Origins of Aggregate Fluctuations," *Econometrica,* 80 (2012), 1977–2016.

Aggarwal, Rajesh K., and Andrew A. Samwick, "Why Do Managers Diversify Their Firms? Agency Reconsidered," *The Journal of Finance,* 58 (2003), 71–118.

Ahern, Kenneth R., and Jarrad Harford, "The Importance of Industry Links in Merger Waves," *Journal of Finance,* 69 (2014), 527–576.

Alchian, Armen A., and Harold Demsetz, "Production, Information Costs, and Economic Organization," *The American Economic Review,* 62 (1972), 777–795.

Antón, Miguel, Florian Ederer, Mireia Giné, and Martin Schmalz, "Common Ownership, Competition, and Top Management Incentives," *Working Paper,* (2018).

Atalay, Enghin, Ali Hortaçsu, James Roberts, and Chad Syverson, "Network structure of production," *Proceedings of the National Academy of Sciences of the United States of America,* 108 (2011), 5199–5202.

Azar, José, Martin Schmalz, and Isabel Tecu, "Anticompetitive Effects of Common Ownership," *Journal of Finance,* 73 (2018), 1513–1565.

Barabási, Albert László, *Network Science* (Cambridge University Press, 2016).

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

Berry, Steven, Martin Gaynor, and Fiona Scott Morton, "Do Increasing Markups Matter? Lessons from Empirical Industrial Organization," *Journal of Economic Perspectives*, 33 (2019), 44–68.

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

Carvalho, Vasco M., and Xavier Gabaix, "The Great Diversification and its Undoing," *American Economic Review*, 103 (2013), 1697–1727.

Clauset, Aaron, Cosma R. Shalizi, and M.E.J. Newman, "Power-Law Distributions in Empirical Data," *SIAM Review*, 51 (2009), 661–703.

di Giovanni, Julian, Andrei A. Levchenko, and Isabelle Mejean, "Firms, Destinations, and Aggregate Fluctuations," *Econometrica*, 82 (2014), 1303–1340.

Erdős, Paul, and Alfréd Rényi, "On Random Graphs I," *Publicationes Mathematicae*, 6 (1959), 290–297.

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

Giroud, Xavier, and Holger M. Mueller, "Firms' Internal Networks and Local Economic Shocks," *American Economic Review*, 109 (2019), 3617–3649.

Herfindahl, Orris Clemens, 1950, "Concentration in the Steel Industry," Ph.D. thesis Columbia University.

Herskovic, Bernard, Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh, "Firm Volatility in Granular Networks," *Journal of Political Economy*, 128 (2020), 4097–4162.

Hirschman, Albert O., *National Power and the Structure of Foreign Trade* (Berkeley and Los Angeles: University of California Press, 1945).

Hoberg, Gerard, and Gordon Phillips, "Text-Based Network Industries and Endogenous Product Differentiation," *Journal of Political Economy*, 124 (2016), 1423–1465.

Hoberg, Gerard, and Gordon M. Phillips, "Scope, Scale, and Competition: The 21st Century Firm," *Working paper*, (2021).

Jensen, Michael C., "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers," *American Economic Review,* 76 (1986), 323–329.

Kleinert, Jörn, Julien Martin, and Farid Toubal, "The Few Leading the Many: Foreign Affiliates and Business Cycle Comovement," *American Economic Journal: Macroeconomics,* 7 (2015), 134–159.

Lamont, Owen, "Cash Flow and Investment: Evidence from Internal Capital Markets," *Journal of Finance,* 52 (1997), 83–109.

Lewellen, Wilbur G., "A Pure Financial Rationale for the Conglomerate Merger," *Journal of Finance,* 26 (1971), 521–537.

Maksimovic, Vojislav, and Gordon Phillips, "Do Conglomerate Firms Allocate Resources Inefficiently Across Industries? Theory and Evidence," *The Journal of Finance,* 57 (2002), 721–767.

Matsusaka, John G., and Vikram Nanda, "Internal Capital Markets and Corporate Refocusing," *Journal of Financial Intermediation,* 11 (2002), 176–211.

Matvos, Gregor, and Amit Seru, "Resource Allocation within Firms and Financial Market Dislocation: Evidence from Diversified Conglomerates," *Review of Financial Studies,* 27 (2014), 1143–1189.

Pierce, Justin R., and Peter K. Schott, "The Surprisingly Swift Decline of US Manufacturing Employment," *American Economic Review,* 106 (2016), 1632–1662.

Scharfstein, David S., and Jeremy C. Stein, "The Dark Side of Internal Capital Markets: Divisional Rent-Seeking and Inefficient Investment," *Journal of Finance,* 55 (2000), 2537–2564.

Seru, Amit, "Firm Boundaries Matter: Evidence From Conglomerates and R&D Activity," *Journal of Financial Economics,* 111 (2014), 381–405.

Sharifkhani, Ali, and Mikhail Simutin, "Feedback Loops in Industry Trade Networks and the Term Structure of Momentum Profits," *Journal of Financial Economics, forthcoming,* (2021).

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

Shin, Hyun-Han, and René M. Stulz, "Are Internal Capital Markets Efficient?," *Quarterly Journal of Economics,* 113 (1998), 531–552.

Stein, Jeremy C., "Internal Capital Markets and the Competition for Corporate Resources," *Journal of Finance,* 52 (1997), 111–133.

Stigler, George J., "A Theory of Oligopoly," *Journal of Political Economy,* 72 (1964), 44–61.

Tate, Geoffrey, and Liu Yang, "The Bright Side of Corporate Diversification: Evidence From Internal Labor Markets," *Review of Financial Studies,* 28 (2015), 2203–2249.

Teece, David J., "Economies of scope and the scope of the enterprise," *Journal of Economic Behavior & Organization,* 1 (1980), 223–247.

Villalonga, Belén, "Does Diversification Cause the Diversification Discount?," *Financial Management,* 33 (2004), 5–27.

Whited, Toni M., and Guojun Wu, "Financial Constraints Risk," *Review of Financial Studies,* 19 (2006), 531–559.

Williamson, Oliver E., *Markets and Hierarchies: Analysis and Antitrust Implications* (New York: Free Press, 1975).

Online Appendix

"The Conglomerate Network"

## Kenneth R. Ahern, Lei Kong, and Xinyan Yan

I. Network Visualizations

To help visualize the conglomerate network, Online Appendix Figure I presents the CoHHI network for manufacturing industries in 2015. Each inter-industry connection represents a CoHHI score above a minimum threshold. Industries listed in boxes are aggregated to more coarse definitions for brevity. To give further intuition for the structure of the network, Online Appendix Figure II provides a detailed representation of the links between the paper, chemicals, and plastic industries. The firms listed are those firms that operate in at least two of the three industries. The CoHHI of the paper and chemicals industries is driven by their common exposure to the same firms, with Procter & Gamble as a key conduit. Likewise, the chemicals and plastics industries are connected through common exposure to conglomerate firms, with Bayer as the strongest connection.

II. Direct Evidence of the Transmission of Shocks within a Conglomerate

To examine how shocks propagate from one segment to another in a conglomerate, for each segment with the conglomerate, we create a companion NTR gap in addition to its own NTR gap. Specifically, the companion NTR gap using 1999 conglomerate structure is defined as *Other NTR* $Gap_{i,j,k} = \sum_{l \neq k} \left(\frac{Sales_{j,l,1999}}{\sum_{l \neq k} Sales_{j,l,1999}} \times NTRGAP_l \right)$, where i represents the segment, j represents the parental firm of the segment, k represents the industry of segment i, t represents year, $Sales_{j,l,1999}$ is firm j's sales in industry l in 1999, and NTR GAP$_l$ represents the NTR gap of industry l. To ensure that what matters is that segments are linked to other industries in which the parent firm is operating and not other industries in general, we create a placebo conglomerate structure in 1999. Specifically, for each segment, we replace all industries to which the segment is currently linked with randomly assigned industries. In other words, while

holding a segment's own NTR gap unchanged, we replace NTR GAP$_l$ with randomly assigned industries when calculating the companion NTR gap of this segment.

In Online Appendix Table VII, we show the results using the 1999 conglomerate structure in the first two columns and the placebo results in the last two columns. The companion NTR gap has a negative and significant impact on segment sales growth whereas the placebo tests reveal tiny and insignificant impacts.
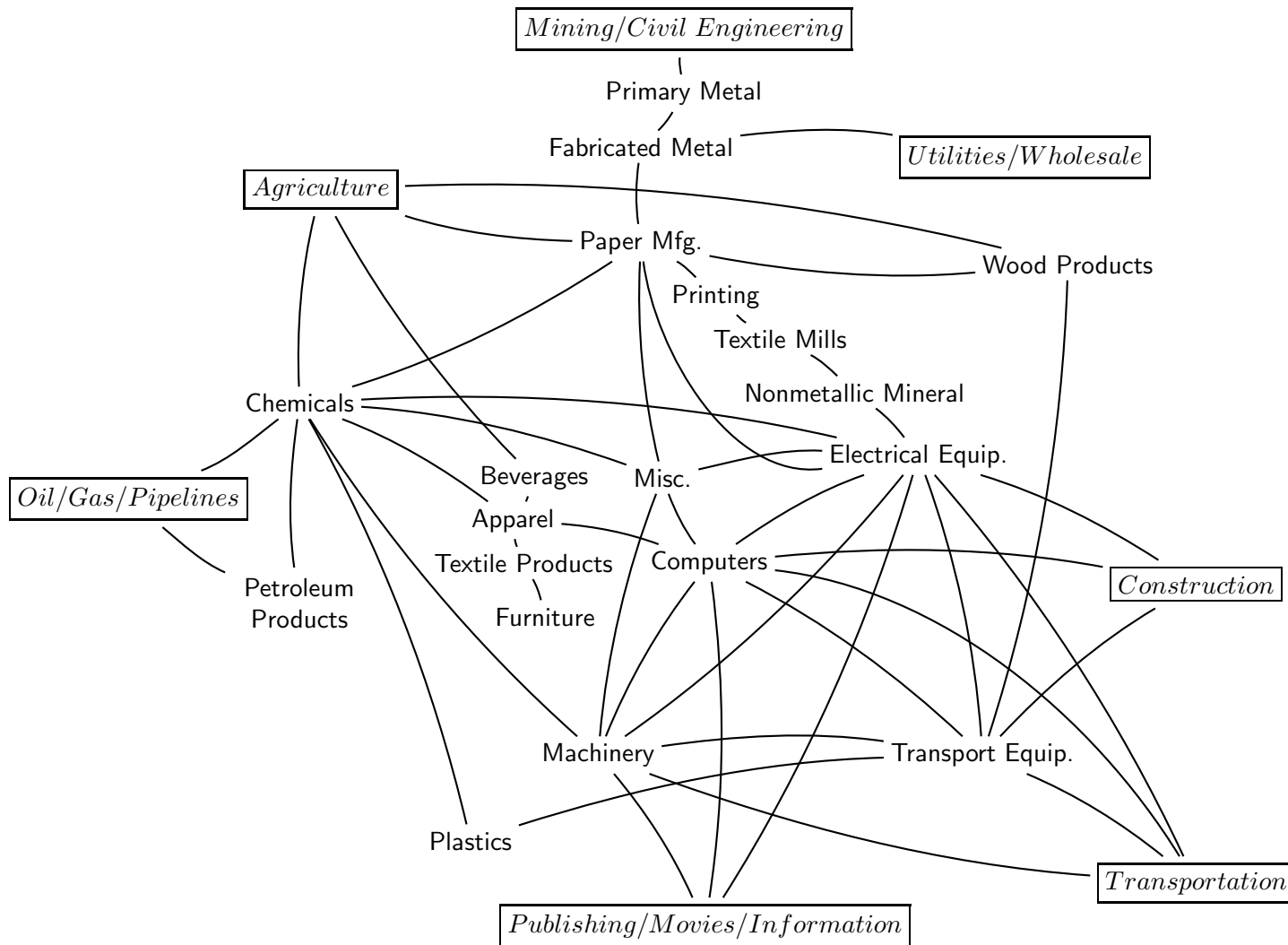
ONLINE APPENDIX FIGURE I

Conglomerate Network in Manufacturing in 2015

Each node in the network is a three-digit manufacturing industry, except non-manufacturing industries, which are aggregated at a higher level and presented in boxes. Lines between industries represent Co-HHI measures above a minimum threshold. Because many of the manufacturing industries are connected to credit and securities industries in the Co-HHI network, we omit these links for ease of viewing the other Co-HHI connections. Data are from Compustat.
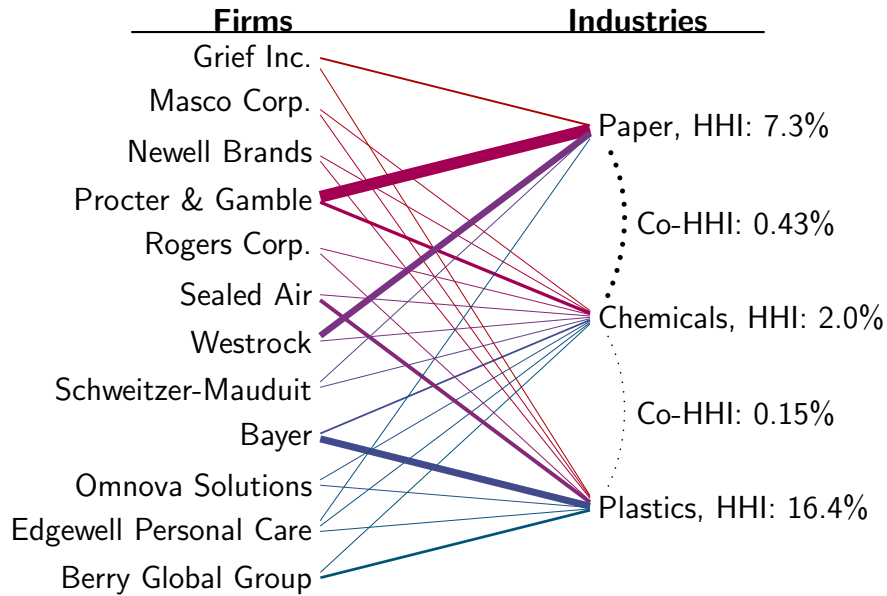
Online Appendix Figure II

The Affiliation Network of the Paper, Chemicals, and Plastics Industries
This figure represents the affiliation network of three manufacturing industries, defined at the three-digit NAICS level: Paper, Chemicals, and Plastics. The listed firms are those that operate in at least two of the three industries. Each industry contains additional firms not represented in the figure that do not operate segments in at least two of these three industries. The widths of the lines are scaled by firms' market shares in each industry. There is a weak Co-HHI relationship between Plastics and Paper not shown in the figure. Data are from Compustat for year 2015.

Comovement of Employment Growth and Network Links

This table presents coefficient estimates from panel regressions where the dependent variable is $(g_k - g_j)^2$, where g_i is the growth rate of industry i employment. Variable definitions are provided in the text. All regressions include industry-pair fixed effects and year fixed effects. Coefficients and industry-pair standard errors (in parentheses) are in percentages. Statistical significance is indicated by ***, **, and * for significance at 0.01, 0.05, and 0.10.

	Dependent variable: $(g_k - g_j)^2$					
	(1)	(2)	(3)	(4)	(5)	(6)
Shared in-links (Co-HHI)	−0.542**		−0.569**			
	(0.231)		(0.232)			
Shared out-links		0.093	0.137			
		(0.114)	(0.114)			
Shared links dummy				−0.047**		
				(0.020)		
Transmission					−0.150	
					(0.145)	
Transmission dummy						−0.047***
						(0.014)
Sum of HHI	0.036***	0.036***	0.036***	0.035***	0.036***	0.035***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.007)	(0.007)
Input-Output link	0.279***	0.278***	0.279***	0.279***	0.287***	0.287***
	(0.036)	(0.036)	(0.036)	(0.036)	(0.025)	(0.025)
Hoberg-Phillips similarity	0.583	0.589	0.581	0.589	0.580*	0.582*
	(0.478)	(0.478)	(0.478)	(0.478)	(0.338)	(0.338)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.169	0.169	0.169	0.169	0.169	0.169
Observations	3,264,827	3,264,827	3,264,827	3,264,827	6,529,654	6,529,654

Comovement of Industry Growth and Lagged Shared Network Links
This table presents coefficient estimates from panel regressions where the dependent variable is $(g_k - g_j)^2$, where g_i is the growth rate of industry i for sales (Panel A) and assets (Panel B). Explanatory variables are lagged by one year. Variable definitions are provided in the text. All regressions include industry-pair fixed effects and year fixed effects. Coefficients and industry-pair clustered standard errors (in parentheses) are in percentages. Statistical significance is indicated by ***, **, and * for significance at 0.01, 0.05, and 0.10.

	Dependent variable: $(g_k - g_j)^2$			
	(1)	(2)	(3)	(4)
Panel A: Sales growth				
Shared in-links (Co-HHI)	−13.645***		−13.254***	
	(2.897)		(2.911)	
Shared out-links		−2.607**	−1.789	
		(1.107)	(1.095)	
Shared links dummy				−0.620***
				(0.225)
Sum of HHI	0.136	0.132	0.135	0.129
	(0.120)	(0.120)	(0.120)	(0.120)
Input-Output link	−0.006	−0.007	−0.005	−0.004
	(0.342)	(0.342)	(0.342)	(0.342)
Hoberg-Phillips similarity	−13.092**	−12.924**	−13.044**	−12.951**
	(5.489)	(5.483)	(5.487)	(5.483)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.066	0.066	0.066	0.066
Observations	3,278,544	3,278,544	3,278,544	3,278,544

	Dependent variable: $(g_k - g_j)^2$			
	(1)	(2)	(3)	(4)
Panel B: Asset growth				
Shared in-links (Co-HHI)	−16.219***		−15.881***	
	(2.556)		(2.561)	
Shared out-links		−2.528**	−1.544	
		(1.169)	(1.132)	
Shared links dummy				−0.661***
				(0.236)
Sum of HHI	−1.230***	−1.234***	−1.231***	−1.237***
	(0.119)	(0.119)	(0.119)	(0.119)
Input-Output link	0.010	0.007	0.011	0.010
	(0.358)	(0.358)	(0.358)	(0.358)
Hoberg-Phillips similarity	−7.429	−7.257	−7.389	−7.283
	(5.486)	(5.484)	(5.485)	(5.484)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.056	0.056	0.056	0.056
Observations	3,202,103	3,202,103	3,202,103	3,202,103

ONLINE APPENDIX TABLE III

Comovement of Industry Growth and Shared Network Links
Robustness to Excluding Firms Below the 25th Percentile of Sales

This table replicates Table III of the main paper, but uses observations from networks that exclude firms with sales below the 25th percentile of sales per year. The table presents coefficient estimates from panel regressions where the dependent variable is $(g_k - g_j)^2$, where g_i is the growth rate of industry i for sales (Panel A) and assets (Panel B). Variable definitions are provided in the text. All regressions include industry-pair fixed effects and year fixed effects. Coefficients and industry-pair clustered standard errors (in parentheses) are in percentages. Statistical significance is indicated by ***, **, and * for significance at 0.01, 0.05, and 0.10.

	Dependent variable: $(g_k - g_j)^2$			
	(1)	(2)	(3)	(4)
Panel A: Sales growth				
Shared in-links (Co-HHI)	−18.747***		−17.932***	
	(2.615)		(2.629)	
Shared out-links		−5.021***	−3.727***	
		(1.310)	(1.244)	
Shared links dummy				−1.295***
				(0.221)
Sum of HHI	2.511***	2.505***	2.508***	2.498***
	(0.111)	(0.111)	(0.111)	(0.111)
Input-Output link	0.465	0.462	0.466	0.470
	(0.322)	(0.322)	(0.322)	(0.322)
Hoberg-Phillips similarity	−16.455***	−16.170***	−16.376***	−16.216***
	(4.279)	(4.271)	(4.277)	(4.273)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.061	0.061	0.061	0.061
Observations	3,252,468	3,252,468	3,252,468	3,252,468

continued on next page

	Dependent variable: $(g_k - g_j)^2$			
	(1)	(2)	(3)	(4)
Panel B: Asset growth				
Shared in-links (Co-HHI)	−18.059***		−17.229***	
	(2.609)		(2.617)	
Shared out-links		−5.062***	−3.812***	
		(1.474)	(1.410)	
Shared links dummy				−1.441***
				(0.233)
Sum of HHI	2.657***	2.651***	2.654***	2.643***
	(0.114)	(0.114)	(0.114)	(0.114)
Input-Output link	−0.150	−0.154	−0.149	−0.145
	(0.326)	(0.325)	(0.326)	(0.325)
Hoberg-Phillips similarity	−20.197***	−19.936***	−20.124***	−19.975***
	(5.538)	(5.527)	(5.534)	(5.527)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.052	0.052	0.052	0.052
Observations	3,161,000	3,161,000	3,161,000	3,161,000

ONLINE APPENDIX TABLE IV

Comovement of Industry Growth and Shared Network Links
Robustness to Excluding Firms Below the Median Size

This table replicates Table III of the main paper, but uses observations from networks that exclude firms with sales below the median sales level per year. The table presents coefficient estimates from panel regressions where the dependent variable is $(g_k - g_j)^2$, where g_i is the growth rate of industry i for sales (Panel A) and assets (Panel B). Variable definitions are provided in the text. All regressions include industry-pair fixed effects and year fixed effects. Coefficients and industry-pair clustered standard errors (in parentheses) are in percentages. Statistical significance is indicated by ***, **, and * for significance at 0.01, 0.05, and 0.10.

	Dependent variable: $(g_k - g_j)^2$			
	(1)	(2)	(3)	(4)
Panel A: Sales growth				
Shared in-links (Co-HHI)	−9.299***		−8.279***	
	(2.825)		(2.847)	
Shared out-links		−6.341***	−5.525***	
		(1.643)	(1.592)	
Shared links dummy				−0.963***
				(0.241)
Sum of HHI	2.429***	2.421***	2.424***	2.417***
	(0.119)	(0.119)	(0.119)	(0.119)
Input-Output link	1.321***	1.327***	1.328***	1.332***
	(0.365)	(0.365)	(0.365)	(0.365)
Hoberg-Phillips similarity	0.618	0.851	0.723	0.785
	(4.989)	(4.983)	(4.986)	(4.987)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.063	0.063	0.063	0.063
Observations	2,593,920	2,593,920	2,593,920	2,593,920

	Dependent variable: $(g_k - g_j)^2$			
	(1)	(2)	(3)	(4)
Panel B: Asset growth				
Shared in-links (Co-HHI)	−10.991***		−9.946***	
	(2.671)		(2.686)	
Shared out-links		−6.566***	−5.591***	
		(1.684)	(1.624)	
Shared links dummy				−1.180***
				(0.258)
Sum of HHI	3.702***	3.693***	3.697***	3.688***
	(0.119)	(0.119)	(0.119)	(0.119)
Input-Output link	0.544	0.549	0.550	0.557
	(0.370)	(0.370)	(0.370)	(0.370)
Hoberg-Phillips similarity	−15.681***	−15.428***	−15.581***	−15.498***
	(5.282)	(5.275)	(5.279)	(5.276)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.049	0.049	0.049	0.049
Observations	2,509,535	2,509,535	2,509,535	2,509,535

ONLINE APPENDIX TABLE V

Comovement of Industry Growth and Shared Network Links
Robustness to Excluding Foreign Firms

This table replicates Table III of the main paper, but uses observations from networks that exclude firms incorporated outside of the U.S. The table presents coefficient estimates from panel regressions where the dependent variable is $(g_k - g_j)^2$, where g_i is the growth rate of industry i for sales (Panel A) and assets (Panel B). Variable definitions are provided in the text. All regressions include industry-pair fixed effects and year fixed effects. Coefficients and industry-pair clustered standard errors (in parentheses) are in percentages. Statistical significance is indicated by ***, **, and * for significance at 0.01, 0.05, and 0.10.

	Dependent variable: $(g_k - g_j)^2$			
	(1)	(2)	(3)	(4)
Panel A: Sales growth				
Shared in-links (Co-HHI)	−21.811***		−20.918***	
	(2.448)		(2.471)	
Shared out-links		−6.244***	−4.365***	
		(1.362)	(1.284)	
Shared links dummy				−2.050***
				(0.236)
Sum of HHI	1.800***	1.793***	1.797***	1.784***
	(0.113)	(0.113)	(0.113)	(0.113)
Input-Output link	0.490	0.488	0.494	0.498
	(0.329)	(0.330)	(0.330)	(0.330)
Hoberg-Phillips similarity	−16.236***	−15.930***	−16.157***	−15.973***
	(4.323)	(4.310)	(4.320)	(4.311)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.070	0.070	0.070	0.070
Observations	3,371,361	3,371,361	3,371,361	3,371,361

continued on next page

	Dependent variable: $(g_k - g_j)^2$			
	(1)	(2)	(3)	(4)
Panel B: Asset growth				
Shared in-links (Co-HHI)	−22.676***		−22.294***	
	(2.223)		(2.235)	
Shared out-links		−3.901**	−1.875	
		(1.579)	(1.479)	
Shared links dummy				−2.084***
				(0.237)
Sum of HHI	1.588***	1.583***	1.586***	1.571***
	(0.112)	(0.112)	(0.112)	(0.112)
Input-Output link	−0.817**	−0.820**	−0.815**	−0.808**
	(0.340)	(0.340)	(0.340)	(0.340)
Hoberg-Phillips similarity	−17.680***	−17.447***	−17.644***	−17.428***
	(5.050)	(5.036)	(5.048)	(5.036)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.058	0.058	0.058	0.058
Observations	3,289,294	3,289,294	3,289,294	3,289,294

ONLINE APPENDIX TABLE VI

Comovement of Industry Growth and Shared Network Links
Robustness to Coarse Industry Definitions

This table replicates Table III of the main paper, but uses observations from networks based on 3-Digit SIC codes. The table presents coefficient estimates from panel regressions where the dependent variable is $(g_k - g_j)^2$, where g_i is the growth rate of industry i for sales (Panel A) and assets (Panel B). Variable definitions are provided in the text. All regressions include industry-pair fixed effects and year fixed effects. Coefficients and industry-pair clustered standard errors (in parentheses) are in percentages. Statistical significance is indicated by ***, **, and * for significance at 0.01, 0.05, and 0.10.

	Dependent variable: $(g_k - g_j)^2$			
	(1)	(2)	(3)	(4)
Panel A: Sales growth				
Shared in-links (Co-HHI)	−18.393***		−16.487***	
	(5.047)		(4.984)	
Shared out-links		−4.906***	−4.582***	
		(0.919)	(0.909)	
Shared links dummy				−1.041***
				(0.207)
Sum of HHI	2.570***	2.562***	2.568***	2.555***
	(0.231)	(0.231)	(0.231)	(0.231)
Input-Output link	−0.171***	−0.170***	−0.170***	−0.169***
	(0.053)	(0.053)	(0.053)	(0.053)
Hoberg-Phillips similarity	−1.504	−1.309	−1.403	−1.270
	(4.592)	(4.596)	(4.592)	(4.574)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.066	0.066	0.066	0.066
Observations	845,868	845,868	845,868	845,868

	Dependent variable: $(g_k - g_j)^2$			
	(1)	(2)	(3)	(4)
Panel B: Asset growth				
Shared in-links (Co-HHI)	−13.935***		−11.982**	
	(4.709)		(4.726)	
Shared out-links		−5.010***	−4.774***	
		(0.946)	(0.941)	
Shared links dummy				−1.160***
				(0.223)
Sum of HHI	0.019	0.013	0.018	0.006
	(0.237)	(0.237)	(0.237)	(0.237)
Input-Output link	−0.303***	−0.302***	−0.302***	−0.301***
	(0.057)	(0.056)	(0.056)	(0.056)
Hoberg-Phillips similarity	−22.009***	−21.804***	−21.876***	−21.797***
	(4.361)	(4.354)	(4.357)	(4.375)
Industry-pair and year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.053	0.053	0.053	0.053
Observations	821,164	821,164	821,164	821,164

ONLINE APPENDIX TABLE VII

Transmission of Tariff Shocks Within Conglomerates

This table presents coefficient estimates from panel regressions where the dependent variable is the industry growth rate of sales. NTR Gap is the difference between the non-Normal Trade Relations tariff rate and the NTR tariff rate. Other NTR Gap is based on other segments within the conglomerate firm. Firm-clustered standard errors are in parentheses. Statistical significance is indicated by ***, **, and * for significance at 0.01, 0.05, and 0.10.

	Dependent variable: Sales growth			
	1999 Fixed Network		Placebo Network	
	(1)	(2)	(3)	(4)
Post $\times$ NTR Gap$_i$	0.002	0.002	-0.054	-0.050
	(0.063)	(0.063)	(0.057)	(0.057)
NTR Gap$_i$	-0.859		0.576*	
	(0.883)		(0.313)	
Post $\times$ Other NTR Gap$_i$	-0.197**	-0.194**	0.013	0.018
	(0.079)	(0.079)	(0.062)	(0.062)
Other NTR Gap$_i$	-1.857*		-0.201	
	(1.122)		(0.153)	
Segment fixed effects	Yes		Yes	
Segment-industry fixed effects		Yes		Yes
Year fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.268	0.270	0.269	0.271
Observations	24,143	24,095	24,143	24,095

ONLINE APPENDIX TABLE VIII

Financial Constraints

This table presents coefficient estimates from panel regressions where the dependent variable is $(g_k - g_j)^2$, where g_i is the growth rate of industry i for sales in column 1 and assets in columns 2. Industry-pair financial constraint is the Whited-Wu index of financial constraints for an industry-pair-year observation. All regressions include industry-pair and year fixed effects. Coefficients and industry-pair clustered standard errors (in parentheses) are in percentages. Statistical significance indicated by ***, **, and * for significance at 0.01, 0.05, and 0.10.

	Sales growth	Asset growth
Industry-pair financial constraint	0.979	0.831
	(0.733)	(0.879)
Shared in-links (CoHHI)	−0.923	10.068
	(10.278)	(10.757)
CoHHI×Financial constraint	51.493**	77.677***
	(24.514)	(24.269)
Sum of HHI	2.511***	1.690***
	(0.112)	(0.116)
Input-Output link	1.038***	0.193
	(0.362)	(0.367)
Hoberg-Phillips similarity	−20.760***	−29.298***
	(4.365)	(5.637)
Industry-pair and year fixed effects	Yes	Yes
Adjusted R^2	0.068	0.054
Observations	3,560,590	3,471,255