NBER WORKING PAPER SERIES

FIRM VOLATILITY IN GRANULAR NETWORKS

Bernard Herskovic
Bryan Kelly
Hanno Lustig
Stijn Van Nieuwerburgh

Working Paper 19466
http://www.nber.org/papers/w19466

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2013

We thank Daron Acemoglu, Andy Atkeson, Steven Davis, Jan Eeckhout, Rob Engle, Xavier Gabaix, Stefano Giglio, Jakub Jurek, Matthias Kehrig, Toby Moskowitz, Philippe Mueller, Lubos Pastor, Jacopo Ponticelli, Alberto Rossi, Amit Seru, Chad Syverson, Allan Timmerman, Laura Veldkamp, Pietro Veronesi, Rob Vishny, and seminar participants at Berkeley Haas, Chicago Booth, KUL, CUNY, the UBC Winter Conference, USC, the Philadelphia Federal Reserve Bank annual conference, the Adam Smith conference at Oxford, NYU Stern, University of Amsterdam, London School of Economics, London Business School, the NBER Asset Pricing and Capital Markets conferences at the Summer Institute, SITE, the Becker-Friedman Institute conference on networks, the European Finance Association, the University of Maastricht, and MIT Sloan for comments and suggestions. We thank Lauren Cohen and Andrea Frazzini for generously sharing their data. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Firm Volatility in Granular Networks
Bernard Herskovic, Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh
NBER Working Paper No. 19466
September 2013, Revised June 2020
JEL No. E1,G10

ABSTRACT

Firm volatilities co-move strongly over time, and their common factor is the dispersion of the economy-wide firm size distribution. In the cross section, smaller firms and firms with a more concentrated customer base display higher volatility. Network effects are essential to explaining the joint evolution of the empirical firm size and firm volatility distributions. We propose and estimate a simple network model of firm volatility in which shocks to customers influence their suppliers. Larger suppliers have more customers and customer-supplier links depend on customers size. The model produces distributions of firm volatility, size, and customer concentration consistent with the data.

Bernard Herskovic
Anderson School
University of California at Los Angeles
bernard.herskovic@anderson.ucla.edu

Bryan Kelly
Yale School of Management
165 Whitney Ave.
New Haven, CT 06511
and NBER
bryan.kelly@yale.edu

Hanno Lustig
Stanford Graduate School of Business
655 Knight Way
Stanford, CA 94305
and NBER
hlustig@stanford.edu

Stijn Van Nieuwerburgh
Columbia University
Graduate school of Business
Uris Hall, office 809
3022 Broadway
New York, NY 10027
and NBER
svnieuwe@gsb.columbia.edu

1 Introduction

Recent research has explored how the firm size distribution and the network structure in an economy can influence aggregate volatility. Gabaix (2011) points out that "granularity," or extreme skewness in firm sizes, concentrates economic mass among a few very large firms, stifling diversification and increasing aggregate volatility. Acemoglu et al. (2012) and Carvalho (2010) show that sparsity of inter-sector linkages similarly inhibits diversification in an economy and raises aggregate volatility.

This research is silent about the impact of firm networks and firm size concentration on the volatility at the firm level. The volatility of firm-level stock returns and cash flows varies greatly over time (e.g. Lee and Engle, 1993) and across firms (e.g. Black, 1976; Christie, 1982). Firm-level fluctuations in uncertainty have important implications for investment and hiring decisions as well as firm value, as highlighted by Bloom (2009).[1] But the underlying determinants of firm volatility are poorly understood. In much of the work on volatility in economics and finance, firms are modeled to have heteroscedastic shocks without specifying the source of heteroscedasticity. Our goal is to understand, both theoretically and empirically, how inter-firm linkages and size distributions interact to endogenously produce heteroscedasticity at the firm level.

We propose a simple model in which firms are connected to other firms in a customer-supplier network.[2] It has three assumptions. Firms' growth rates are influenced by their own idiosyncratic shocks and by the growth rate of their customers. As a result, the firm-specific shocks propagate through the network via connected firms. The appendix provides a simple general equilibrium model with inter-connected firms and consumer demand shocks that delivers a structural interpretation. Second, the probability of a customer-supplier link

[1] See also Leahy and Whited (1996), Bloom, Bond, and Van Reenen (2007), Stokey (2016), Bloom et al. (2018) and the papers cited therein.

[2] A recent literature explores the role of production networks in macro-economics, trade, asset pricing, and banking. See Atalay et al. (2011), Foerster, Sarte, and Watson (2011), Acemoglu et al. (2012), Carvalho and Gabaix (2013), Ahern and Harford (2014), Acemoglu, Ozdaglar, and Tahbaz-Salehi (2015), Farboodi (2015), Herskovic (2018), Kramarz, Martin, and Mejean (2020), Stanton, Walden, and Wallace (2018). For reviews of the network literature, see Allen and Babus (2009), Jackson (2014), and the references therein.

depends on the size of the supplier so that large firms typically supply to a higher number of customers. Third, the importance of a customer-supplier link depends on the size of the customer. Large customers have a stronger connection with their suppliers, presumably because they account for a large fraction of their suppliers' sales. We provide microeconomic evidence for all three assumptions based on the observed customer-supplier networks among Compustat firms.[3] Differences in firms' network connections impart total firm volatility with cross-sectional heteroscedasticity.

Firms are aggregators of their own idiosyncratic shock and the shocks to connected firms. The sparsity and granularity of the firm's customer network, which in turn depend on the firm size distribution, determine the firm's volatility by affecting the diversification of the shocks that they are exposed to. Our model maps the firm size distribution to network formation, generating a rich set of testable implications for volatility in the cross-section and the time-series. We study data on firm-level sizes, volatilities, and customer-supplier linkages, establishing a new set of stylized facts about firm volatility and confirming the model's implications.

Firm-level volatilities exhibit a common factor structure tightly related to the economy-wide firm size dispersion. In the model, each supplier's network is a random draw from the entire population of firms, so that any firm's customer network inherits similar dispersion to that of the entire size distribution. An increase in size dispersion slows down every firm's shock diversification and increases their volatility. We find that firm volatilities possess a strong factor structure in the data.

The factor structure implies strong time series correlations between moments of the size and volatility distributions. An increase in size dispersion translates into higher average volatility among firms. It also raises the cross section dispersion in volatilities. In the time series, size dispersion has a 63% correlation with mean firm volatility and 78% with the

[3] Since we only have data for publicly listed firms, we cannot verify whether the same comovement between the firm size dispersion and the average and the dispersion of firm volatility extends to private firms. Davis et al. (2006) document a secular decline in the volatility of employment for private firms using annual LBD data from 1982 until 1997.

dispersion of firm volatility. Our paper is the first to document these facts and to provide an economic explanation for the factor structure in firm-level volatility by connecting it to firm size dispersion. A persistent widening in the firm size dispersion should lead to a persistent rise in mean firm volatility. We observe such a widening (increase in firm concentration) between the early 1960s and the late 1990s, providing a new explanation for the run-up in mean firm volatility studied by Campbell et al. (2001).

In the cross-section, differences in volatility across firms arise from two sources: differences in the number of customers and differences in customers' size dispersion. First, large firms are less volatile than small firms because they are connected to more customers, which improves diversification regardless of the size profile of its customer base. This effect also appears in the model's volatility factor structure. Smaller firms have larger exposures to the common volatility factor, implying that small firms have both higher levels of volatility and higher volatility of volatility. In the data, we find a strong negative correlation between firm size and variance, and small firms indeed have higher volatility factor exposures.

Second, holding the number of connections fixed, a supplier's customer network is less diversified if there is more dispersion in the size of its customers. Because customer size determines the strength of a link, severe customer size disparity means that shocks to the biggest customers exert an outsized influence on the supplier, raising the supplier's volatility. Differences in customer size disparity arise from probabilistic network formation; some suppliers will link to a very large or very small customer by chance alone. The data indeed show a strong positive correlation between a firm's out-Herfindahl, our measure of concentration in a firm's customer network, and its volatility. Firm size and firm out-Herfindahl remain the leading determinants of firm volatility after the inclusion of other determinants of volatility previously proposed in the empirical literature. Collectively, this evidence supports a network-based explanation of firm volatility.

To gauge its quantitative plausibility, we estimate our network model using data on the customer-supplier network among US firms. Our estimation targets moments and cross-moments of the distributions of firm size, firm variance, and inter-firm business linkages. The model can account for the large dispersion in firm volatility while respecting the evidence on the number and concentration of customers. The estimation reveals that this requires strong network effects: a firm's customer's customer shocks are nearly as important as the shocks that directly hit the firm's customer. Strong network effects are necessary but not sufficient for quantitatively matching observed firm volatilities. An internal diversification mechanism whereby larger firms have lower shock volatility complements the external diversification mechanism of the customer network. A statistical test fails to reject the null hypothesis that the moments in model and data are equal. One reason for the high estimate of the network parameter is underlying heterogeneity in the network parameter of firms that is correlated with network centrality.

The rest of the paper is organized as follows. Section 2 presents new empirical evidence on the link between the firm size and volatility distributions. Section 3 explains these links with a simple network model. Section 4 estimates the model and Section 5 tests two additional model predictions in the micro-data. One is on the network determinants of firm-level volatility and the other establishes that firm size dispersion is an important common factor driving firm volatilities. The structural model, the proofs of the theoretical results, and the auxiliary empirical evidence are relegated to the appendix.

2 Evidence on Firm Size Dispersion and Firm Variance

This section documents our main new empirical facts about the joint evolution of the firm size and firm variance distributions.

2.1 Data

We consider market-based and fundamentals-based measures of firm size and firm volatility. Both are calculated at the annual frequency. Our main measure of firm size is the equity market value at the end of the calendar year. The alternative fundamentals-based measure is total sales within the calendar year. All variables in our analysis are deflated by the consumer price index. Our main measure of firm variance is defined as the variance of daily stock returns during the calendar year. Fundamentals-based variance in year t is defined as the variance of quarterly sales growth (over the same quarter the previous year) within calendar years t to $t+4$.[4] The sample is the universe of publicly-listed firms. Stock market data are from CRSP for the period 1926-2016 and sales data are from the merged CRSP/Compustat file for the period 1952-2016. The cross-sectional size and variance distributions are well approximated by a lognormal distribution. As a result, each distribution may be summarized by two moments: the cross-sectional mean and cross-sectional variance of the log quantities. We use the term dispersion to denote the cross-sectional standard deviation.

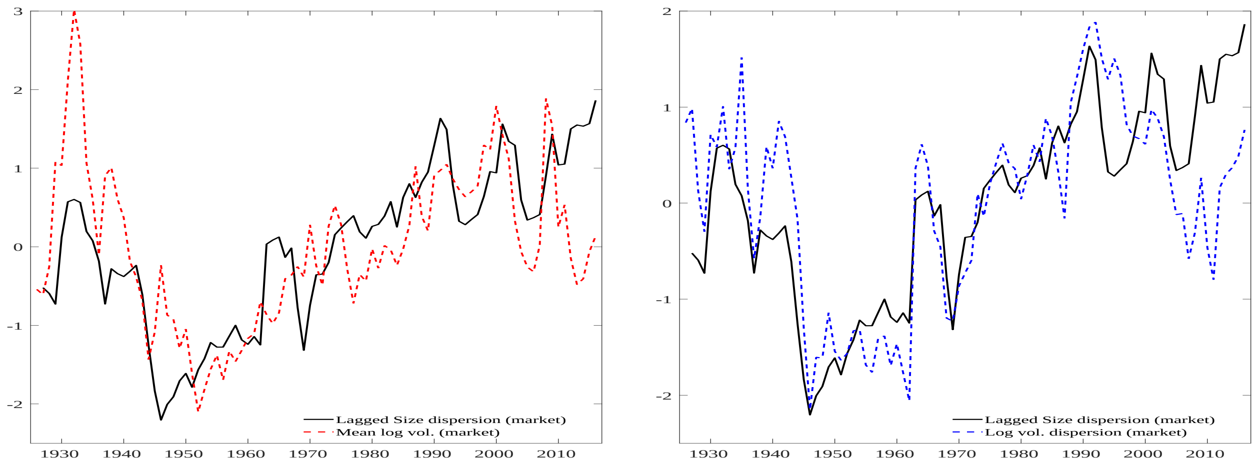
2.2 Comovement of Firm Size and Volatility Distributions

Panel A of Figure 1 plots the cross-sectional *average* of log firm variance, where variance is computed from daily stock returns, against lagged log firm size dispersion, where size is computed as the market capitalization. For ease of readability, both series are standardized to have mean zero and standard deviation one. The correlation between average firm variance and firm size dispersion is 63.3%. Mean firm variance experienced several large swings in the past century, especially in the 1920s and 1930s and again in the last two decades of the sample. These changes are preceded by similar dynamics in the cross-sectional dispersion of firm size. The high positive correlation between mean firm volatility and firm size dispersion is our first main new fact.

The second main stylized fact links the *dispersion* in firm variance to the dispersion in firm size. Panel B of Figure 1 shows a strong positive association between the cross-sectional dispersions of firm variance and firm size, based on the market measures. The correlation

[4] We also consider fundamental volatility measured by the standard deviation of quarterly sales growth within a single calendar year. The one- and five-year fundamental volatility estimates are qualitatively identical, though the one-year measure is noisier because it uses only four rather than twenty observations.

Figure 1: Market-based Dispersion in Firm Size and Firm Variance



(a) Panel A: Average Log Variance

(b) Panel B: Dispersion in Log Variance

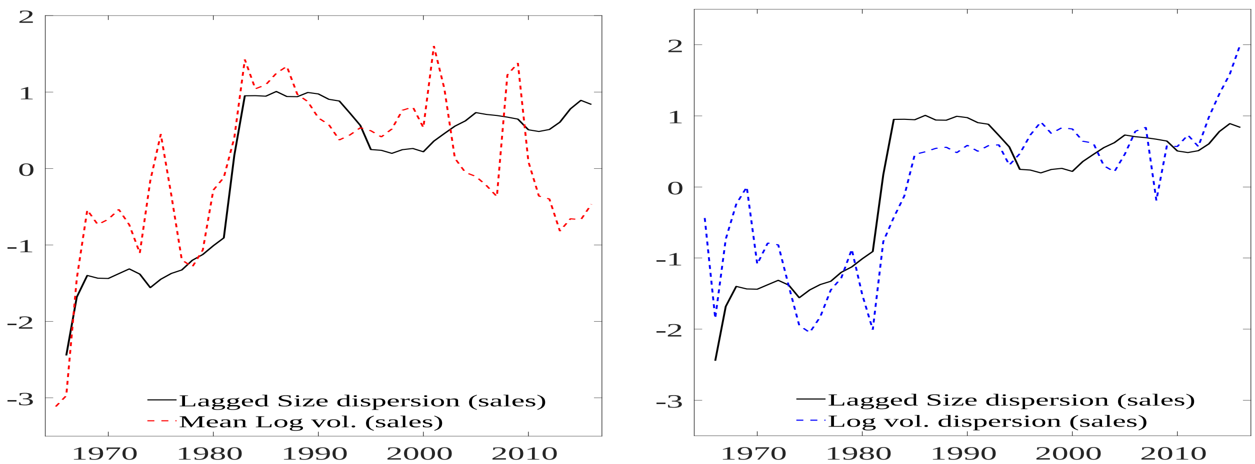
Notes: Panel A plots the lagged cross-sectional dispersion of log firm size (solid line) and the cross-sectional mean of the log variance distribution (dashed red line). Panel B plots the lagged cross-sectional dispersion of log firm size (solid line) and the cross-sectional standard deviation of the log variance distribution (dashed blue line). Firm size is measured as market value of equity; firm variance is measured as the variance of daily stock returns within the year. All series are rescaled for the figure to have mean zero and variance one. Sample is from 1926 to 2016 at annual frequency.

between the two time-series is 77.6%.

Appendix C.1 shows that the positive correlation between size dispersion and the first and second moments of the firm variance distribution is also present at business cycle frequencies; it uses the HP-filter to decompose the time series in trend and cycle components.

The same relationship between the moments of the firm size and variance distributions exists for our fundamentals-based measure. First, the correlation between average firm variance and lagged firm size dispersion is 68.7%. The left panel of Figure 2 shows the strong positive association. Because the sales-based data only start in 1965, their dynamics are more affected by the persistent increase in firm size dispersion and variance that took place between the 1960s and the 1990s. The right panel shows a strong positive correlation between the dispersion of variance and the dispersion of size for the sales-based measure. The correlation is 81.7%. Both facts corroborate the market-based evidence. Any explanation of these facts must confront the high degree of similarity between market volatilities and its

Figure 2: Fundamentals-based Dispersion in Firm Size and Firm Variance



(a) Average Log Variance

(b) Dispersion in Log Variance

Notes: Panel A plots the lagged cross-sectional dispersion of log firm size (solid line) and the cross-sectional mean of the log variance distribution (dashed red line). Panel B plots the lagged cross-sectional dispersion of log firm size (solid line) and the cross-sectional standard deviation of the log variance distribution (dashed blue line). Firm size is measured as firm sales; firm variance is measured based on 20 quarters of growth in firm sales. All series are rescaled for the figure to have mean zero and variance one. Sample is from 1965 to 2016 at annual frequency.

(more coarsely measured) fundamental counterpart.[5] This evidence suggests that financial explanations, such as discount rate or leverage effects (Black, 1976), are incomplete.[6]

2.3 Subsample Results

Table 1 establishes that the two key correlations between firm size dispersion and the average (column 3) and dispersion of firm variance (column 4) hold for various subsets of our data (listed in column 1). The second column reports the number of firms in each subsample. The first panel splits the sample of firms into three size terciles based on the market value of equity and calculates the moments within each group. Firms are resorted each year. Our main correlations are large and positive in all size groups.

[5] Market- and fundamentals-based measures of average log firm variance have an annual time-series correlation of 68.9%, while the two volatility dispersion measures have a correlation of 73.5%.

[6] We take the traditional perspective of asset pricing in endowment economies where firms' cash flow growth is taken as given. In richer settings, changes in discount rates may affect investment decisions and possibly input prices.

The second panel groups firms into nine industries. The two main correlations are large and positive in virtually all industries. The one exception is utilities, which only has a 17% correlation between size dispersion and average variance. For consumer durables, that same correlation is 35%. These two industries have the fewest firms. All correlations between size dispersion and variance dispersion are higher than 50%.

Third, we find virtually the same correlation for firms that have been public less than five years and firms that have been public for more than five years. Similar correlations are also found for firms that have been public for at least 10 and at least 25 years. This shows that our results are not driven by firms that recently went public, and whose size and volatility characteristics may be different from older firms or may have changed over time. We return to this composition hypothesis in Section 4.7.

Fourth, the results hold both in the first and in the second half of the sample. They are somewhat stronger in the first half of the sample. Again, this suggests that a changing composition of public firms, which applies to the second half of the sample, cannot be the sole driver of the results.

Fifth, the correlations between size dispersion and variance mean and dispersion are larger, but in the same ballpark, for firms listed on the NYSE and firms not listed on the NYSE.

To summarize, we observe a strong positive association of firm size dispersion with average firm variance and dispersion in firm variance, throughout the distribution of firm size, industry, age, and over time. The next section presents a network model that generates these associations.

3 A Network Model of Firm Growth and Volatility

This section develops a simple network model of connections between firms. It generates the positive correlations between firm size dispersion on the one hand and average firm volatility

Table 1: Composition

	# Firms	Correlation of Size Dispersion with Avg. Var.	Disp. Var
All Firms	3069.9	67.5	77.4
By size			
Smallest third	1023.4	65.8	59.9
Middle third	1023.0	59.3	67.2
Largest third	1023.5	55.0	66.4
By industry			
Consumer Non-Dur.	240.8	59.5	67.2
Consumer Durables	104.6	34.9	71.7
Manufacturing	516.4	51.7	77.3
Energy	150.2	67.7	57.4
Tech/Telecom	494.4	73.9	51.0
Retail	305.8	62.1	63.4
Healthcare	199.0	73.6	53.6
Utilities	111.5	16.9	58.9
Other	884.0	56.6	66.2
By Time Since Entry			
Less than 5 yrs.	905.0	74.0	79.3
At least 5 yrs.	2159.3	71.5	77.1
At least 10 yrs.	1447.5	72.9	76.4
At least 25 yrs.	489.0	46.9	77.9
By Subsample			
1926-1969	496.7	74.5	83.8
1970-2016	2573.1	55.2	42.7
By Exchange			
NYSE	1204.3	60.8	76.6
non-NYSE	1865.6	41.3	42.4

Notes: Annual data 1926-2016. We use the market-based volatility measure constructed from stock returns and the market-based measure of size (market equity). The second column reports the time series average of the cross section number of firms. Column 3 reports the correlation between average log volatility ($\mu_{\sigma,t}$) and lagged log size dispersion ($\sigma_{S,t-1}$). Column 4 reports the correlation between dispersion in log volatility ($\sigma_{\sigma,t}$) and lagged log size dispersion. For each grouping, we first compute the cross-sectional moments among the firms in the group, and then we compute the time-series correlation between the moments.

and the dispersion of firm volatility on the other hand. The reduced form network model of Section 3.1 is micro-founded by a structural production model in Section 3.2. Section 3.3 connects a firm's size to its position in the network. Section 3.4 presents the main result on firm variance. Section 3.5 extends the result to the case where the firm size distribution follows a power law, and characterizes comovement between firms' volatilities. Section 4 then uses size, volatility, and network moments to estimate the model presented here.

3.1 Firm Growth

Define S_i as the size of firm $i = 1, \cdots, N$ with growth rate g_i. Firms are connected through a network. Firm i's growth rate depends on its own idiosyncratic shock and a weighted average of the growth rates of the firms j it is connected to:

$$g_i = \mu_g + \gamma_i \sum_{j=1}^{N} w_{i,j}\, g_j + \varepsilon_i. \tag{1}$$

The parameter $\gamma_i \in [0, 1)$ governs the rate of decay as a shock propagates through the network. A special case is $\gamma_i = \gamma, \forall i$. As discussed below, allowing for heterogeneity in γ_i strengthens the connection to the structural model and facilitates interpretation of the point estimate for γ. The network weight $w_{i,j}$ determines how strongly firm i's growth rate is influenced by the growth rate of firm j. If i and j are not connected, then $w_{i,j} = 0$. By convention, we set $w_{i,i} = 0$. We refer to the matrix of connection weights $\boldsymbol{W} = [w_{i,j}]$ as the network matrix. We assume that all rows of $\boldsymbol{W}$ sum to one. Thus, the largest eigenvalue of $\boldsymbol{W}$ equals one. In this directed network, connections are not necessarily symmetric. We identify firm i as the supplier and firm j as the customer. Firm j can be a customer of firm i without i being a customer of j.

Let $\boldsymbol{g}$ and $\boldsymbol{\varepsilon} \sim N\left(0, \sigma_\varepsilon^2 \boldsymbol{I}\right)$ be the $N \times 1$ vectors of growth rates and shocks, respectively. Let Γ be a diagonal matrix with generic element γ_i. For simplicity, set $\mu_g = 0$. The growth rate equation (1) for all firms can be written in vector form as:

$$\boldsymbol{g} = \Gamma \boldsymbol{W} \boldsymbol{g} + \boldsymbol{\varepsilon} = \left(\boldsymbol{I} - \Gamma \boldsymbol{W}\right)^{-1} \boldsymbol{\varepsilon}. \tag{2}$$

The Leontief inverse matrix $\left(\boldsymbol{I} - \Gamma \boldsymbol{W}\right)^{-1}$ is the key object describing the effects of network structure on the behavior of growth rates.

In this section, we purposely impose stark assumptions on the nature of the underlying innovations: each firm i experiences i.i.d. growth rate shocks $\varepsilon_i \sim N(0, \sigma_\varepsilon)$. In Section 4, we augment the model to consider heterogeneity in firms' idiosyncratic shock variances.

3.2 A Structural Interpretation

Appendix A derives equation (2) as the equilibrium outcome in a production economy. It is a constant returns-to-scale economy with N firms in the tradition of Long and Plosser (1983). The output of one firm is used as input in the production of another firm. A representative agent has Cobb-Douglas preferences defined over the N goods that are produced, with preference weights collected in the vector $\boldsymbol{\theta}$. The reduced-form shocks $\boldsymbol{\varepsilon}$ are consumer demand shocks for the various goods, $d\boldsymbol{\theta}$, scaled by the vector $\boldsymbol{\psi}$ whose i^{th} entry measures the ratio of output of firm i to economy-wide value added. These taste shocks travel upstream from customers to suppliers, i.e., from final goods producers to their suppliers, the producers of intermediate goods, and to their suppliers' suppliers, the producers of more basic inputs, etc., thereby affecting the growth in their real output $\boldsymbol{g} = dY/Y$. The parameter γ_i measures the share of output of firm i used by intermediate good producers. The entries of the network matrix W, w_{ij}, measure the sales of good i to intermediate goods firm j divided by the total output of firm i, and multiplied by γ_i. These entries are linked to the cost shares, which are primitives of the production function. The following proposition, proved in the appendix, shows that the structural model delivers the same "network equation" (3) as the reduced-form model's equation (2).

Proposition 1. *The responses of firm output to demand shocks $d\boldsymbol{\theta}$ equals:*

$$\frac{d\boldsymbol{Y}}{\boldsymbol{Y}} = (\boldsymbol{I} - \Gamma \boldsymbol{W})^{-1} \frac{d\boldsymbol{\theta}}{\boldsymbol{\psi}}. \tag{3}$$

Since the proof shows that product prices are unchanged, the response of firm sales (which is easy to measure) equals that of firm output (which is harder to measure). Appendix A works out an example with three layers of production that illustrates how demand shocks travel from the final good's firm to its supplier, the intermediate good's firm, and its supplier's supplier, the basic input firm. We provide a sufficient condition on structural parameters under which an increase in the first size dispersion, resulting from preference shocks, is associated with an increase in average firm variance. Simulation evidence for three-firm economies shows that this condition is satisfied for the vast majority of network parameters and preference shocks.

Shea (2002) and Kramarz, Martin, and Mejean (2020) consider different models where demand shocks travel upstream from customers to suppliers, like in our model. Acemoglu et al. (2012) solve a similar model, but emphasize productivity shocks that are transmitted downstream from suppliers to customers. In principle, shocks could travel both upstream and downstream. The empirical evidence in section 5.1 is consistent with upstream propagation.

3.3 Size Effects in Network Structure

The content of the spatial autoregression model (2) lies in the specification of the weighting matrix $\boldsymbol{W}$. Since we cannot directly measure w_{ij} in the data, we make two plausible assumptions on the probability and strength of connections between suppliers and their customers. These assumptions link the firm size distribution to the network structure of the economy. In Section 4, we confront the model with data on the production network in the US and validate these assumptions.

The firm size distribution determines the linkage structure between customers and suppliers. The existence of a link between supplier i and customer j is described by:

$$b_{i,j} = \begin{cases} 1 & \text{if } i \text{ connected to } j \\ 0 & \text{otherwise.} \end{cases}$$

Each element of the connections matrix, $\boldsymbol{B} = [b_{i,j}]$, is drawn from a Bernoulli distribution with probability $P(b_{i,j} = 1)$. This connection probability is assumed to be a function of the size of the supplier i:

$$P(b_{i,j} = 1) \equiv p_i = \frac{\tilde{S}_i}{Z} N^{-\zeta} \text{ (for } i \neq j), \tag{4}$$

where $\tilde{S}_i = S_i / E[S_i]$ is the relative size of firm i versus the population mean and Z is a scalar. While the functional form matters quantitatively, the crucial qualitative assumption is that the probability of a connection depends on the relative size of firm i. That is, larger firms have more connections on average. This is the model's first size effect.

Equation (4) also builds sparsity into the network. The sparsity parameter $\zeta \in (0, 1)$ governs the rate at which the likelihood of a connection decreases as the number of firms N grows large. It implies that the number of links (customers) in the system diverges as the number of firms goes to infinity, but that the probability of connecting to any single customer goes to zero. In a large economy, the expected number of customers for firm i, called the out-degree, is approximately:

$$N_i^{out} \approx N p_i = \frac{\tilde{S}_i}{Z} N^{1-\zeta}. \tag{5}$$

The number of linkages grows with the number of firms in the economy, but the rate of growth is slower when ζ is closer to 1.

The second key assumption on the network we make is that, conditional on a link existing between a supplier i and a customer j, the strength of that link depends on the size of the customer j:

$$w_{i,j} = \frac{b_{i,j} S_j}{\sum_{k=1}^{N} b_{i,k} S_k}, \quad \forall i, j. \tag{6}$$

One natural measure of size is sales. This weighting scheme then has the natural feature that customers j who represent a larger share of supplier i's sales have a stronger impact on the growth rate of i. This is the model's second size effect.

While there is obvious value in understanding more deeply why firms forge connections, modeling the endogenous choice over network linkages is notoriously difficult.[7] Our approach

[7]For some promising steps forward in this direction recently, see Acemoglu, Ozdaglar, and Tahbaz-Salehi

is to model two key features of the observed supplier-customer networks. Any network choice model would have to generate these features as an outcome. Our simpler approach suffices to study how the network structure among firms affects the link between the firm size and volatility distribution. This simplicity allows us prove theoretical results on the relationship between firm volatility and the structure of its customer network. Our paper also contributes new facts on the properties of production networks that the literature on endogenous network formation can target as outcomes. Also, our approach is what enables us to estimate the model in Section 4.

3.4 Firm Variance

Conditional on $\boldsymbol{W}$, the variance-covariance matrix of growth rates $\boldsymbol{g}$ is given by

$$\boldsymbol{V}(\boldsymbol{g}) = \sigma_\varepsilon^2 \left(\boldsymbol{I} - \Gamma \boldsymbol{W}\right)^{-1} \left(\boldsymbol{I} - \boldsymbol{W}'\Gamma\right)^{-1}. \tag{7}$$

The vector of firm volatilities is the square root of the diagonal of the variance-covariance matrix. In standard network settings, the Leontief inverse $\left(\boldsymbol{I} - \Gamma \boldsymbol{W}\right)^{-1}$ is an obstacle to deriving a tractable analytic characterization of volatility. Our model, in contrast, lends itself to a convenient variance representation when the number of firms in the economy becomes large.

Before deriving our main result, it is useful to build intuition for the behavior of variance by considering a simplified version of the network model. Suppose that growth rates follow the process:

$$\boldsymbol{g} = \left(\boldsymbol{I} + \Gamma \boldsymbol{W}\right)\left(\boldsymbol{\mu}_g + \boldsymbol{\varepsilon}\right). \tag{8}$$

In our full network model (1), a firm's growth rate is influenced by the growth rates of each of its connections. The latter are in turn influenced by their connections' growth rates, and

(2015), Farboodi (2015), Oberfield (2018), Stanton, Walden, and Wallace (2018), Herskovic and Ramos (Forthcoming), and Taschereau-Dumouchel (2018). The latter paper has the advantage that solving for the equilibrium production network is relatively straightforward even with a few thousand firms.

so on. The simplification in (8) differs from the full network in that idiosyncratic shocks only propagate one step in the supply chain and then die out. In fact, (8) is a first order approximation to (1), because:

$$(\boldsymbol{I} - \Gamma \boldsymbol{W})^{-1} \quad = \quad \boldsymbol{I} + \Gamma \boldsymbol{W} + (\Gamma \boldsymbol{W})^2 + (\Gamma \boldsymbol{W})^3 + \cdots \approx \boldsymbol{I} + \Gamma \boldsymbol{W},$$

under our maintained assumption that $\gamma_i \in [0, 1)$, $\forall i$. In this system, the variance of firm i's growth rate simplifies to:

$$V(g_i) = V\left(\gamma_i \sum_j w_{i,j}\varepsilon_j + \varepsilon_i\right) = \sigma_\varepsilon^2 \left(1 + \gamma_i^2 H_i^{out}\right), \tag{9}$$

where

$$H_i^{out} \equiv \sum_{j=1}^{N} w_{i,j}^2 \tag{10}$$

is the Herfindahl index of firm i's network of customers. We refer to H_i^{out} as the out-Herfindahl. Equation (9) shows that, to a first order approximation, the variance of a firm's growth rate is determined by its out-Herfindahl, the volatility of the underlying innovations σ_ε^2, and the strength of shock transmission in the network $\gamma_i, \forall i$.

The higher firm i's out-Herfindahl, the more concentrated is its network of connections. Standard diversification logic applies: a low degree of diversification in a firm's customer base raises its variance. The out-Herfindahl is driven by two characteristics: the number of customers and the size dispersion among its customers. The supplier is more diversified and has lower volatility when it has more customers and when the size dispersion of its customers is small.

Because all firms, large and small, draw their connections from the same economy-wide firm size distribution, all firms' customer networks have equal firm size dispersion *in expectation*. The expected customer size dispersion of any firm i is given by the *economy-wide* firm size dispersion. Thus, the economy-wide out-Herfindahl is a key determinant of the volatility of each firm. The economy-wide out-Herfindahl moves all firms' volatilities up and down together; it is a common factor in firms' volatilities. Differences between firms' volatilities arise from how many connections they draw from this economy-wide firm size distribution. Small firms draw fewer links on average than large firms and therefore are more volatile.

Our main theoretical proposition below formalizes the above intuition. It maps the variance of a firm to its size and to the concentration of firm sizes throughout the economy. It does so for the full network model, i.e., without the first-order approximation in (8). The proof is relegated to Appendix B as are the lemmas the proposition relies on.

Proposition 2. *Consider a sequence of economies indexed by the number of firms N. If the firm size distribution has finite variance, then the Leontief inverse has limiting behavior described by*

$$(I - \Gamma W)^{-1} \approx I + \Gamma W + \frac{\overline{\gamma}}{1-\overline{\gamma}}\Gamma \bar{W}.$$

where $\overline{\gamma} = \frac{1}{N}\sum_{i=1}^{N}\gamma_i$. Fixing the size of the i^{th} firm, S_i, volatility of firm i has limiting behavior described by

$$V(g_i|S_i) \approx \sigma_\varepsilon^2 \left[1 + \left(\frac{Z}{N^{1-\zeta}\tilde{S}_i} + \frac{\overline{\gamma}\,(2-\overline{\gamma})}{N(1-\overline{\gamma})^2}\right)\gamma_i^2 \frac{E[S^2]}{E[S]^2} + 2\gamma_i \frac{\overline{\gamma}}{1-\overline{\gamma}}\frac{\tilde{S}_i}{N}\right]. \tag{11}$$

This proposition highlights the determinants of firm-level growth rate variance in a large economy. A firm's variance depends on its own shock variance σ_ε^2 plus a term that reflects the network effect of interest. The latter consists of three terms.

First, firm variance depends on economy-wide firm size dispersion given by the ratio of the second non-central moment of the size distribution to the squared first moment $E[S^2]/E[S]^2$. In the special case where the firm size distribution is log-normal with variance σ_s^2, this ratio simplifies to $\exp(\sigma_s^2)$. This establishes a theoretical connection between firm size dispersion and the firm volatility distribution (and its mean and dispersion), just like we found empirically in Section 2. Economy-wide size dispersion is a common factor affecting all firms' volatilities, since more size dispersion makes every supplier less diversified. The firstround reduction in customer network diversification has ripple effects through the higher order terms: customers' customer networks are also less diversified, etc. As γ_i approach one, these higher order terms become quantitatively important as firms' idiosyncratic shocks have far reaching effects throughout the network. In the cross-section, firms with higher γ_i will propagate a given shock to firm size dispersion to their customers more strongly.

Second, relative firm size, $\tilde{S}_i$, affects volatility in two ways. Its primary role is captured by the first coefficient on $E[S^2]/E[S]^2$. Larger firms have a lower exposure to the common factor in volatility, economy-wide size dispersion, than do smaller firms because they typically connect to more customers and achieve better shock diversification. This diversification channel lowers the volatility of large firms. Smaller factor exposure also makes large firms less sensitive to fluctuations in the firm size dispersion. That is, they display lower volatility of volatility.

Third, the firm's relative size also appears in the numerator of the last term. Shocks to the largest firms are the most strongly propagated shocks in the model since these firms have the largest influence on their suppliers. Shocks to large firms ultimately "reflect" back, raising large firms' own volatility. In contrast, small firms' shocks die out relatively quickly in the network. Thus, the last term in (11) captures a countervailing increase in volatility for larger firms. In all of our numerical results, we find that the diversification channel of the first size term dominates the reflection channel of the last term. Only for the very largest firms does the reflection effect contribute meaningfully to firm variance.

Finally, we note that the strength of the diversification channel depends on the term $N^{1-\zeta}$. A high ζ means that there are relatively few linkages compared to the size of the economy. This network sparsity effect slows down the diversification benefits for all firms. Because the other two network terms in (2) decline at a faster rate N, the diversification channel becomes the dominant effect in an economy with many firms.[8]

[8]The same logic behind Proposition 2 allows us to describe the limiting behavior of covariances among firms' growth rates, $Cov(g_i, g_j)$. See Corollary 1 in Appendix B.

3.5 Additional Theoretical Results

We derive two more theoretical results with proofs in Appendix B. Proposition 3 characterizes firm variance when the firm size distribution follows a power law while Proposition 4 characterizes the comovement across firms' volatilities.

Proposition 3. *Consider a sequence of economies indexed by the number of firms N. If firm sizes are distributed according to a power law with exponent $\eta \in (1, 2]$, then firm variance decays at rate $N^{(1-\zeta)(2-2/\eta)}$.*

Gabaix (2011) emphasizes that extreme right skewness of firm sizes can also slow down volatility decay in large economies. We show that the firm-level network structure adds a complementary mechanism that slows down the volatility decay beyond Gabaix's granularity mechanism. In the absence of network effects, firm variance decays at rate $N^{2-2/\eta}$, where η is the power law coefficient of the firm size distribution. With network sparsity, the rate of decay is $N^{(2-2/\eta)(1-\zeta)}$, where ζ is our network sparsity parameter.

Proposition 4. *Consider a sequence of economies indexed by the number of firms N, and fix S_i and S_j. If the firm size distribution has finite fourth moment, then the covariance between the out-Herfindahls H_i and H_j has limiting behavior described by*

$$Cov(H_i, H_j | S_i, S_j) \approx \frac{1}{N^{1+2(1-\zeta)}} \frac{V(S^2)}{S_i S_j} \frac{Z^2}{E[S]^2}.$$

The covariance between $V(g_i|S_i)$ and $V(g_j|S_j)$ decays at the same rate.

Proposition 4 shows that comovement among firm variances decays at rate $N^{1+2(1-\zeta)}$ in a large economy. Intuitively, the volatility comovement among a pair of firms is lowest when both firms are large, since large firms have low exposure to overall size concentration. We test this prediction on the comovement between volatilities in the data in Section 5.2.

4 Simulated Method of Moments Estimation

Section 2 documents strong statistical relationships between the distributions of firm size and firm volatility. Section 3 provides a network-based foundation for these relationships. In this section we go one step further and ask whether the production network model can quantitatively account for the observed joint distribution of firm size, firm volatility, and inter-firm production network linkages. We estimate the key network parameters in a Simulated Method of Moments (SMM) framework. By insisting on matching data on customer-supplier relationships we provide evidence for the network assumptions made in the model.

4.1 Model Extensions

The stylized model of the previous section implicitly assumed that when two firms of equal size merge, the new firm is as volatile as each of the two original firms. A large literature on mergers suggests that many firms seek diversification benefits from mergers resulting in lower firm volatility of the combined entity. We label these gains from *internal diversification* to distinguish them from external diversification gains that accrue through the firm's network of customers. To capture internal diversification benefits in a simple way, we assume that the volatility of a firm's own fundamental shock depends negatively on its size. Specifically,

$$\sigma_{\varepsilon,i} = \sigma_\varepsilon + \lambda \log\left(1 + \frac{S_{median}}{S_i}\right), \tag{12}$$

where λ governs the sensitivity of fundamental volatility to firm size. The parameter σ_ε is the minimum shock volatility, after all possible internal diversification benefits have been exhausted. The larger the firm, the closer its shock volatility is to σ_ε. For the median firm, $S_i = S_{median}$, and the shock volatility equals $\sigma_\varepsilon + \lambda \log(2)$. This functional form keeps all firm variances positive. In Appendix C.7 we explore an alternative internal diversification function based on Stanley et al. (1996) and find similar results.

We also generalize the weighting function $w_{i,j}$, which governs the importance of customer

j in supplier i's network:

$$w_{i,j} = \frac{b_{i,j} S_j^{\psi}}{\sum_{k=1}^{N} b_{i,k} S_k^{\psi}}. \tag{13}$$

The sensitivity of the importance of the link to the size of the customer is governed by the curvature parameter ψ. When $\lambda = 0$ (and therefore $\sigma^2_{\varepsilon,i} = \sigma^2_{\varepsilon}$) and $\psi = 1$, we recover the simple model of Section 3.

4.2 Network Data, Selection, and Truncation

We use data on customer-supplier networks from the Compustat segment dataset. The sample is 1980-2012. If a customer represents more than 10% of its sellers' revenue, then the customer's name and sales amount are reported. Combining this information with the total sales, available in Compustat, we obtain the sales shares $w_{i,j}$. In a typical sample year, we have about 1,330 firms (suppliers) with non-missing customer information. We compute the cross-sectional distribution of the number of customers or out-degree (N_i^{out}), the dispersion of the customer network or out-Herfindahl (H_i^{out}), the number of suppliers or in-degree (N_i^{in}), and the dispersion of the supplier network (H_i^{in}). Key moments of these four distributions are reported in the first column of Table 3 and discussed further below. For consistency, the moments of the firm size and variance distributions, reported in that same table, are computed over the same sample period.

There are two data issues we address as part of our estimation algorithm. First, our network data as well as our firm size and volatility data cover a subset of firms. Only Compustat firms present in the customer segment data are included in the network statistics. Customer-supplier links between public and private companies are unobserved, as well as links of public firms with missing customer data. To capture this selection effect, we simulate the model for N firms, and allow all these firms to forge links with each other according to the network formation rules outlined above. However, we select only the largest N_{pub} firms to compute model-implied moments. Implicit is the assumption that large, listed firms are the most likely to have non-missing customer data. We set $N = 5,000$ for computational reasons; N_{pub} is a parameter estimated to match the average number of firms with non-missing network data.[9]

Second, firms in Compustat are only required to report customers that represent at least 10% of their sales. Despite some voluntary reporting of smaller customers, the vast majority of weights $w_{i,j}$ we observe exceed 10%. Our model accounts for all customers, with weights larger and smaller than 10%. To capture this truncation effect, we treat a link as unobserved whenever $w_{i,j}$ is below 10%. For consistency, we delete the few $w_{i,j}$ observations below 10% in the data as well. The procedure allows us to compare truncated moments in the model to truncated moments in the data. Since both truncated and untruncated moments are available in the model, we can make indirect inferences about the full network structure.

4.3 Parameters

Our empirical approach is to estimate the key parameters that govern the network by SMM. These parameters are listed in Panel A of Table 2. While all parameters jointly determine all moments, we nevertheless provide some intuition for which moments most directly identify each parameter. First, we assume that all firms have the same network parameter $\gamma_i = \gamma$. The parameter γ governs the strength of the network. When $\gamma = 0$, there are no network effects. A value of γ close to 1 implies strong higher-order effects, or equivalently, slow spatial decay in the network. Second, the fundamental volatility of the innovations σ_ε governs the average level of firm volatility. Third, the probability of forming a supplier-customer connection in (4) depends on the parameter Z. This parameter affects the average number of customers, the average out-degree. Fourth, as just discussed, λ affects the relative volatility of large and small firms (over and above the differences generated by the network) and ψ affects the relative importance of large and small customers in determining a supplier's growth rate. The average number of firms with network data pins down N_{pub}. Finally, the

[9] One additional argument for using public firms only in estimation, besides data availability, is that firm variance is hard to measure for private firms.

Table 2: Parameter Estimates

		(1) Benchmark	(2) No Network	(3) No ID
γ	*Network propogation effect*	0.918	0.000	0.918
		(0.026)	–	–
σ_ε	*Fundamental shock volatility*	0.298	0.298	0.388
		(0.035)	–	–
λ	*Internal diversification effect*	0.131	0.131	0.000
		(0.025)	–	–
Z	*New connections*	0.003	0.003	0.003
		(0.001)	–	–
ψ	*Sensitivity of connection weight to customer size*	0.184	0.184	0.184
		(0.003)	–	–
μ_s	*Mean of initial log size distribution*	13.412	13.412	13.412
		(0.243)	–	–
σ_s	*Standard deviation of initial log size distribution*	5.104	5.104	5.104
		(0.102)	–	–
N_{pub}	*Number of public firms*	1332	1332	1332
		(47.7)	–	–

Notes: This table report model parameters. Column 1 reports the estimated parameters using Simulate Method of Moments estimation, along with the parameters' standard errors in parenthesis. Columns 2 and 3 report the parameters from the benchmark estimation but without network effects (i.e. $\gamma = 0$) and without internal diversification (i.e. $\lambda = 0$), respectively.

mean and variance of the observed log firm size distribution identify the first two moments of the log-normally distributed firm size distribution, μ_s and σ_s, in the model. We collect the parameters to be estimated in the vector Θ.

Two parameters are determined outside the estimation. Mean firm growth rate μ_g is set equal to zero, matching the observed full-sample growth rate in real market capitalization. Second, $1 - \zeta$ is the elasticity of the average number of customers to the number of firms in the economy; see equation (5). We set $\zeta = 0.87$ to match the estimated elasticity in our network data.[10]

[10] We estimate a time-series regression of the log average number of connections on a constant and the log number of firms. The slope of this regression, which corresponds to $1-\zeta$, is estimated at 0.13 with t-statistic of 2.85. The estimation sample is the same sample of Compustat firms for which we have non-missing customer information (1980-2012). The parameter ζ is difficult to identify since other things change in the time series.

4.4 Simulated Method of Moments

Our SMM estimation is standard; see Gourieroux, Monfort, and Renault (1993) and Duffie and Singleton (1993). The estimation chooses the parameter vector Θ which minimizes the distance between the data moments, collected in the $1 \times K$ vector $\mathcal{G}$, and the corresponding moments obtained from a simulation of the network model, collected in $\hat{\mathcal{G}}(\Theta)$:

$$\mathcal{F} = \min_{\Theta} \mathbb{E}\left[\left(\mathcal{G} - \hat{\mathcal{G}}(\Theta)\right) \mathcal{W} \left(\mathcal{G} - \hat{\mathcal{G}}(\Theta)\right)'\right].$$

We target $K = 28$ moments in the estimation, listed in Table 3. They consist of four moments of the firm size distribution, four moments of the firm volatility distribution, twelve moments of the network distribution, seven correlation moments between size, variance, and network moments, and one moment related to the number of firms with network data. All moments are expressed as logs or log differences and are of similar magnitude. For simplicity and robustness we use the $K \times K$ identity weighting matrix for $\mathcal{W}$. We draw an initial size distribution for $N = 5,000$ firms and then simulate the network links (the $b_{i,j}$ shocks) and fundamental shock volatilities (the ε shocks) 100 times. The moment function $\hat{\mathcal{G}}(\Theta)$ is the average over the 100 draws. The calculation of standard errors is detailed in Appendix C.2. We also provide a Wald test of the null hypothesis that all targeted moments are the same in model and data. The p-value calculation is also in Appendix C.2.

4.5 Estimation Results

The first column of Table 2 reports the SMM point estimates and standard errors for the parameters in Θ. The corresponding moments are in the second column of Table 3. To disentangle the separate roles of network effects and internal diversification, we consider two restricted versions of the benchmark model. Model "No Network" in column 2 of Table 2 is a model without network effects; it has the same parameters as the benchmark except that it sets $\gamma = 0$. Model "No ID" in column 3 of Table 2 shuts down internal diversification by setting $\lambda = 0$ and σ_ε equal to the median volatility of the benchmark model in (12). All other parameters are held at benchmark values. The resulting model fits are reported in the last two columns of Table 3.

Firm Size Panel A shows that the benchmark model closely matches the moments of the firm size distribution. It generates not only the correct average firm size and size dispersion, but also large enough differences between the median firm and the smallest firms and between the median firm and the largest firms. This is mostly by virtue of the calibration. The log normality assumption on firm size fits the data very well. Since firm size is pre-determined, the restricted models fit the size distribution equally well.

Firm Variance The network effects reveal themselves in the distribution of firm variance, reported in Panel A of Table 3. The benchmark model matches the mean firm variance. Fundamental shock volatility is estimated just below 30% for the least volatile firms. The median firm has fundamental shock volatility of 38.9% ($0.298 + 0.131 \log(2)$). More importantly, the model generates most of the observed dispersion in firm variance (86% vs 105% in the data). The cross-sectional differences in firm variance are driven by differences in firms' sizes and customer networks. Some firms have a poorly diversified customer base resulting in high variance, while others achieve a high degree of external diversification resulting in low variance. The point estimate for γ is 0.918 and is precisely estimated. A γ close to 1 points to strong network effects. The null hypothesis of no network effects ($H_0 : \gamma = 0$) is strongly rejected. In sum, the cross-sectional dispersion of firm variance is what identifies the parameter γ.[11]

The "No Network" model in column 3 of Table 3, generated under the null of no network effects, shows that the dispersion in firm volatility is very small. The model with no network

[11] In a dynamic model, time series variation in this cross-sectional dispersion would additionally aid in identification of the network parameters. Cyclical variation in the cross-sectional dispersion in firm-level TFP could also help explain time series variation in firm-level variance (Midrigan and Xu, 2014; Kehrig, 2015).

effects still has an internal diversification mechanism which creates a negative cross-sectional correlation between size and variance. The first row of Panel D shows that the No Network model matches that correlation closely. But the internal diversification mechanism generates far too little volatility dispersion.[12] Network effects are necessary. However, they are not sufficient. The "No ID" model in column 4 of Table 3 generates about half of the observed dispersion in volatility of the data and of the benchmark model that has both network effects and diversification. Appendix C.4 shows that these conclusions regarding the model without network effects and the model without ID effects are robust when we re-estimate all other parameters of the model freely.

Network Moments A key question is whether the network effects that are necessary to deliver the dispersion in firm variance are consistent with the observed properties of the network. Panel B of Table 3 shows that this is indeed the case. The median firm in the model has 1.12 customers whose sales represent at least 10% of that firm's total sales, compared to 1.0 in the data. There are large differences in the cross-section. The firms with the most customers (at the 90th percentile) have 83% more customers than the median firm. The model produces a similar difference at 102%.[13]

[12]While internal diversification generates substantial variation in firm variance, most of this variation occurs between very small firms. The dispersion among the largest N_{pub} firms, reported in the table, is very small.

[13]The restricted models have the same network moments as the benchmark model since they have the same size distribution.

Table 3: Size, Variance, and Network Moments

	(1) Data	(2) Model Bench	(3) Model No Network	(4) Model No ID
Panel A: Size and Volatility Distributions				
log Size				
Average	19.83	19.83	19.83	19.83
Stadard Deviation	2.56	2.53	2.53	2.53
prc 50 - prc 10	2.91	2.37	2.37	2.37
prc 90 - prc 50	3.88	3.91	3.91	3.91
log Variance				
Average	−1.40	−1.40	−2.42	−1.39
Stadard Deviation	1.05	0.86	0.01	0.46
prc 50 - prc 10	1.38	0.40	0.00	0.24
prc 90 - prc 50	1.31	1.69	0.02	0.93
Panel B: Network Moments				
N^{out}				
Median	1.00	1.12	1.12	1.12
log prc 50 - log prc 10	0.00	0.08	0.08	0.08
log prc 90 - log prc 50	0.83	1.02	1.02	1.02
H^{out}				
Median	0.05	0.18	0.18	0.18
log prc 50 - log prc 10	1.27	1.88	1.88	1.88
log prc 90 - log prc 50	1.85	1.71	1.71	1.71
N^{in}				
Median	1.00	1.00	1.00	1.00
log prc 50 - log prc 10	0.00	0.00	0.00	0.00
log prc 90 - log prc 50	1.69	1.09	1.09	1.09
H^{in}				
Median	1.00	1.00	1.00	1.00
log prc 50 - log prc 10	0.95	0.89	0.89	0.89
log prc 90 - log prc 50	0.00	0.00	0.00	0.00
Panel C: Cross-sectional Correlations				
Corr(log Size,log Var)	−0.64	−0.61	−0.67	−0.69
Corr(log N^{out},log Size)	−0.00	0.31	0.31	0.31
Corr(log H^{out},log Size)	−0.28	−0.88	−0.88	−0.88
Corr(log H^{out},log Var)	0.22	0.51	0.68	0.75
Corr(log N^{in},log Size)	0.53	0.34	0.34	0.34
Corr(log H^{in},log Size)	−0.47	−0.34	−0.34	−0.34
Corr(log H^{in},log Var)	0.13	0.19	0.19	0.21
Panel D: Firms with Network Data				
log number of public firms	7.15	6.95	6.95	6.95
Panel E: Goodness of Fit Statistics				
FVAL	–	2.88	7.53	3.72
Wald	–	24.76	–	–
p-value	–	0.64	–	–

Notes: This table reports different size, variance, and network moments both from the data and from our simulation. We use Compustat and CRSP data from 1980 to 2012, and the moments reported are time-series averages of cross-sectional moments. Column 2 reports the moments from a SMM estimation of the benchmark model. Columns 3 and 4 report the SMM estimation of a model without network effects (i.e. $\gamma = 0$) and without internal diversification effects (i.e. $\lambda = 0$), respectively. Panel A reports average, standard deviation, and percentile spreads for both log size and log variance. Panel B reports median and percentilae spreads for four distinct network moments: out-degree or number of customers (N^{out}), out-Herfindahl (H^{out}), in-degree of number of suppliers (N^{in}), and in-Herfindahl (H^{in}). Panel C reports correlations between log size, log variance, and network moments. Panel D reports the log number of firms with network data avaiable. Finally, in Panel E, we report the distance function minimized in the SSM and a Wald test for the null hypotheses that all moments in the data (Column 1) and in the estimated model (Column 2) are equal.

Customer network concentration (out-Herfindahl), constructed from network weights as in equation (10), is also similar in model and data. We recall that the out-Herfindahl directly affects firm variance per equation (9). Dispersion in out-Herfindahl directly maps into dispersion in variance. The model generates the same amount of dispersion in out-Herfindahls as in the data. The firms with the most sparse customer networks (90th percentile) have 185% higher out-Herfindahls than the median firm in the data and 171% in the model.

As Panel C shows, these high-H^{out} firms tend to have higher variance. The cross-sectional correlation is 51% in the model, compared to 22% in the data. Firms with under-diversified (more concentrated) customer networks also tend to be smaller. The correlation is -88% in the model and -28% in the data. The reason these model correlations are not perfect is, first, because higher-order network effects affect variance as well (equation 11 versus equation 9), and second, because internal diversification provides an independent source of dispersion in variance. The model without internal diversification produces a correlation between out-Herfindahl and variance that is indeed higher at 75%. The random nature of the network accounts for the less than perfect correlation between size and out-Herfindahl.

Our network model assumes that larger firms have more connections, on average. Panel C bears this out with a 31% correlation between size and number of customers N^{out}. In the data the correlation is 0%, suggesting little empirical support for the assumption. However, that correlation is subject to a large truncation bias. The correlation between untruncated N^{out} and size is 69% in the model, implying a 38% bias. Applying that bias to the data would suggest a 38% correlation, indicating stronger support for the assumption. Appendix C.3 discusses biases arising from truncation in more detail. It also uses industry-level input-output data from the Bureau of Economic Analysis (BEA) which do not suffer from truncation. The BEA data show a 58% correlation between size and untruncated N^{out}. The BEA data imply an even larger truncation bias. Appendix C.3 also provides evidence for the second size assumption underlying the model, that larger customers represent a larger share of a firm's sales.

We also study the implications of our model for two in-degree moments: the number of suppliers a firm has (N^{in}) and the concentration of the supplier network (H^{in}). Because the Compustat segment data do not contain information on a firm's suppliers, we must use the customer data to infer supplier links. Since we measure suppliers the exact same way in the model, the selection issues that this data structure creates are reproduced in the model. Panel B of Table 3 shows that the model matches the median number of (truncated) suppliers and its dispersion, as well as the (truncated) supplier concentration and its dispersion. The model also fits the cross-sectional correlations between in-degree moments and firm size and variance quite well. The positive correlation between size and number of suppliers is entirely caused by truncation, since the model assumes no relationship. Subtracting the 34% upward bias from the 53% correlation in the data reduces the empirical estimate to below 20%.

We conclude that our simple model is consistent with the key moments of the cross-sectional distribution of firm size, firm variance, and the number and concentration of customer-supplier relationships.

Panel E of Table 3 reports the SMM function value, which is the sum of the squared distances between model and data moments. The restricted versions of the model in the last two columns have a substantially worse fit. Removing internal diversification increases the L_2 norm from 2.88 to 3.72. Removing the network mechanism makes for a far greater deterioration in fit, from 2.88 to 7.53. The panel also reports a Wald statistic for the null hypothesis that all moments (in Table 3) are equal in model and data. We fail to reject the null hypothesis, providing further support for the network model.[14]

[14]According to the same Wald statistic, the $\gamma = 0$ model also cannot be rejected either. However, this result needs to be interpreted cautiously. The Wald statistic is poorly behaved for the $\gamma = 0$ model since that model cannot generate any meaningful variation in the dispersion of firm variance. The columns of the Hessian matrix corresponding to these moments are very close to zero which leads to invertibility problems for the matrices that enter in the computation of the Wald statistic. This issue does not arise for the benchmark model since those corresponding columns of the Hessian matrix are not close to zero.

4.6 Interpreting the Network Effects

The estimation lends strong support for the presence of network effects between firms. When interpreted through the lens of the structural model in Appendix A, a point estimate for γ of 0.918 suggests that most of firms' output serves as input into other firms' production and only 8.2% is final goods consumption. Such a low estimate for the consumption share is belied by sectoral input-output data, which suggest a final goods share around 37%.[15] One possibility is that the estimation uncovers links between firms that go beyond input-output linkages. Financial ties such as trade credit could strengthen network effects between firms (e.g. Raddatz, 2010; Bams, Bos, and Pisa, 2016). Another possibility is that a common value of γ fails to provide an accurate representation of network transmission effects. Indeed, if network transmission is heterogeneous across firms, for example with each firm possessing its own γ_i as in equation (1), then the estimated common γ parameter will be a biased estimate of the average network effect. To investigate, we consider a model under the benchmark parameters except that γ_i can take on one of three values that are centered around 0.6: $\gamma_i \in \{.22, .60, .98\}$. We choose the central value of 0.6 as a reasonable number for the average share of value-added that is bought by other firms. We assume a positive cross-sectional correlation between γ_i and the firm's eigenvector centrality. The top-third of firms with the highest centrality measure is assigned the highest value for γ_i, the middle third the middle value, and the bottom third the lowest value. This positive correlation seems natural in a production economy; firms that sell more to other firms are more central in the production network. After simulating data from this heterogeneous-γ model, we estimate a homogenous-γ model. We estimate $\gamma = 0.93$, very close to the point estimated of 0.918 presented above. The estimate far exceeds the mean and median value for γ_i of 0.6. The intuition for the "bias" is that heterogeneity in γ_i contributes to the cross-sectional dispersion in firm variance. The econometrician who assumes a constant-γ model needs to rely on stronger network effects

[15]Using 2007 BEA data from table "The Use of Commodities by Industries," the ratio of final goods consumption to total industry output is 0.373.

instead to generate the same cross-sectional difference in firm variances. When heterogeneity in γ_i aligns with eigenvector centrality, this force is particularly strong because firms that strongly propagate idiosyncratic shocks (i.e. high-γ_i firms) are also centrally located in the network, playing a key role in the cross-sectional distribution of volatility.

4.7 Composition Effects

An interesting question, left for future research, is whether the comovement of firm size dispersion on the one hand and average firm volatility and the dispersion of firm volatility on the other hand extends to the full cross-section of public and private firms. Answering this question requires access to panel data on sales or TFP for private firms. For example, Bloom et al. (2018) use such data for the manufacturing sector. One issue is that manufacturing is a rapidly shrinking part of the economy. An additional challenge is that measuring firm-level volatility requires a sufficiently long time-series of at least quarterly frequency, which may bias the sample of private firms that can be studied.

A related question is to what extent our results are driven by changes over time in the nature of public firms. Davis et al. (2006) show that private firms started going public earlier in their life cycle in the 1990s, at a time that they were both smaller and more volatile than the average public firm. Doidge et al. (2018) show that this process went into reverse after the year 2000, with fewer private firms going public and more firms delisting. The fraction of small firms with market capitalization below \$100 million in real 2015 dollars has fallen from 60% in the 1980s to 40% in 1997 to 22% in 2016. The changing composition of the universe of public firms could in principle generate the positive comovement between firm size dispersion and firm volatility.

We have already presented empirical evidence in Figure 1 and in Table 1 that speaks against the composition hypothesis. In Appendix C.6, we provide additional evidence from the model. Specifically, we show that a model without network effects in which the changes in the firm size dispersion over time are caused by a changing composition of public firms (and match the observed changes in the firm size dispersion) cannot generate the observed changes in the mean and dispersion of firm volatility. That said, with better data, future work could extend our analysis to the universe of public and private firms.

5 Additional Testable Implications

The network model suggests two additional testable predictions. First, a firm's volatility should depend on its size and on its out-Herfindahl. Second, all firms' volatilities should comove because they are driven by a common factor: the economy-wide dispersion in firm size. We find empirical support for both predictions.

5.1 Determinants of Firm Variance

A large literature has examined the determinants of firm variance. Black (1976) proposed that differences in leverage drive heterogeneity in firm variance. Comin and Philippon (2005) study the role of industry competition and R&D intensity. Davis et al. (2006) emphasize age effects. Brandt et al. (2010) argue that institutional ownership is related to firm variance. Our model predicts a negative correlation between firm variance and firm size and a positive correlation between firm variance and customer network concentration (out-Herfindahl), controlling for other firm characteristics.

Table 4 reports panel regressions of firm-level log annual return variance on log size and log out-Herfindahl, controlling for a range of firm characteristics including log age, leverage, industry concentration, as well as industry and cohort fixed effects.[16] Consistent with our model, we find that size and out-Herfindahl are important determinants of firm variance. The elasticity of firm variance to firm size is around -0.15 and precisely estimated in all specifications. Decreasing the log firm size from the 90^{th} percentile to the median increases firm variance by 58% (-3.88 $\times$ -0.15), which is more than one-half of a standard

[16] Cohorts are defined by the year in which a firm first appears in the CRSP/Compustat data set.

deviation in variance. The elasticity of firm variance to customer concentration is 0.05 and also significant in all specifications. Increasing the log out-Herfindahl from the median to the 90^{th} percentile increases firm variance by 9.3% (1.85×0.05). As we showed in Table 3, log size and log out-Herfindahl are negatively correlated in both network model and data. Network concentration conveys similar information since size determines network structure. Given that concentration in the customer network is measured with noise, size likely captures an important part of the true network concentration effect. Nevertheless, the table provides strong evidence that network concentration matters separately for firm variance and survives the inclusion of other well-known volatility determinants such as age and leverage.

As an aside, when we use the in-Herfindahl instead of the out-Herfindahl, we find that it does not enter significantly either in isolation or after controlling for size. This suggests that the firm variance data are more consistent with upstream than downstream shock propagation.

Table 4: Determinants of Firm-Level Volatility

	Dependent Variable: Log Firm Volatility											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log Size	-0.16		-0.15	-0.15		-0.15	-0.14		-0.15	-0.15		-0.15
	-16.98		*-15.67*	*-28.68*		*-18.38*	*-19.45*		*-19.02*	*-20.15*		*-19.60*
Log H^{out}		0.06	0.03		0.05	0.02		0.05	0.02		0.05	0.02
		4.22	*2.65*		*4.46*	*2.17*		*4.44*	*1.96*		*18.67*	*6.29*
Log Age				-0.09	-0.20	-0.14						
				-2.89	*-9.93*	*-8.15*						
Leverage				0.14	0.09	0.02	0.16	0.10	0.04			
				1.49	*1.24*	*0.40*	*1.73*	*1.43*	*0.79*			
Ind. Conc.				0.53	0.02	0.13	0.59	0.04	0.16			
				1.70	*0.04*	*0.42*	*2.02*	*0.11*	*0.50*			
Constant	-1.00	-0.37	-0.80	-0.92	0.00	-0.55	-1.15	-0.46	-0.89	-0.98	-0.41	-0.84
	-19.70	*-8.82*	*-11.81*	*-5.88*	*0.05*	*-4.56*	*-10.89*	*-4.46*	*-8.48*	*-26.85*	*-10.66*	*-18.94*
FE	None	None	None	None	None	None	Cohort	Cohort	Cohort	Ind.	Ind.	Ind.
Obs.	167,440	28,718	28,232	133,026	22,970	22,949	136,879	23,084	23,063	167,440	28,718	28,232
Adj. R^2	0.283	0.028	0.299	0.338	0.133	0.373	0.358	0.145	0.396	0.397	0.137	0.379

Notes: This table reports panel regressions of firms' return volatility on a range of characteristics including log size, log customer network Herfindahl (H^{out}), log age, book leverage, competition measured by industry size concentration, as well as cohort and industry fixed effects in certain specifications. Size is defined as the lagged market equity relative to the cross-sectional average market equity. Book leverage is defined as debt in current liabilities plus total long-term debt divided by total assets. Industries are defined as 4-digit GICS codes and industry concentration is measure by the Herfindahl index of sales. Cohorts are defined by the year in which a firm first appears in the CRSP/Compustat data set. Standard errors are clustered at the industry level. Sample is at annual frequency from 1980 to 2012.

5.2 Comovement in Firm Variances

Recent research has documented a puzzling degree of common variation in the panel of firm-level volatilities. Herskovic et al. (2016) show that a single common factor explains roughly one-third of the variation in log firm variance for the panel of CRSP stocks.[17] They also show that this strong factor structure is not only a feature of return variance, but also of sales growth variance. The puzzling aspect of this result is that the factor structure remains nearly completely intact after removing all common variation in returns or sales growth rates. Hence, common volatility dynamics are unlikely to be driven by an omitted common return or sales growth factor. That paper raises the question of what the common factor in (idiosyncratic) firm volatility might be.

Our granular network model suggests an answer. It predicts a high degree of comovement in firm volatility arising from concentration in the firm size distribution. Proposition 2 implies that the dispersion of the firm size distribution is the common factor. Proposition 4 further clarifies that large firms have low exposure to this size concentration factor. In other words, fluctuations in firm size dispersion are an important determinant of fluctuations in firm-level volatility, and more so for small than for large firms. Furthermore, if the true data generating process is a network model, then factor model residual volatilities will possess a similar degree of comovement as total volatilities, despite residual growth rates themselves being nearly uncorrelated. To understand this, consider that in the literature, idiosyncratic volatility is typically constructed by first removing the aggregate component of growth rates with a statistical procedure such as principal component analysis, then calculating the volatilities of the residuals. In a granular network model like ours, such a factor regression approach is misspecified. There is no dimension-reducing common factor that fully captures growth rate comovement since, by virtue of the network, *every* firm's shock may be systematic. A

[17] Similarly, Engle and Figlewski (2015) document a common factor in option-implied volatilities since 1996, and Barigozzi et al. (2014) examine the factor structure in realized volatilities of intra-daily returns since 2001. Bloom et al. (2018) show that firm-specific sales growth and productivity exhibit cross-sectional dispersion that fluctuates with the macro economy.

sign of the misspecification of the factor model is that the residuals exhibit a volatility factor structure that looks very similar to the factor structure for total firm volatility.

Table 5 provides empirical support for the fact that the common factor in firm volatilities is the firm size dispersion. It contains panel regressions for firm variance using three different factor models. In column 1, the factor is the lagged dispersion in log firm size (market-based). In column 2 the factor is the lagged cross-sectional average volatility. Since this is essentially the first principal component of firm volatilities, it is a natural yardstick for any one-factor model. The third column instead uses the contemporaneous average volatility as a factor. Because it uses finer conditioning information, it can be considered an upper bound on the explanatory power of a single factor. The main point of Table 5 is that the explanatory power of firm size dispersion (R^2 of 15.65%) is nearly as high as that of mean volatility (R^2 of 20.94%), and about half as large as the one-factor upper bound (R^2 of 33.65%). Firm size dispersion is a powerful factor.

Columns 4-6 repeat the exercise on model-generated data. For the model calculations, we match the observed cross-sectional average and standard deviation of the log size distribution, year by year, from 1980 to 2012. The model implies a firm variance distribution in each year. We use the model-implied time series for firm size dispersion and mean firm variance to estimate the factor model. Similar to the data, lagged firm size dispersion is about as strong a predictor as lagged average firm variance. The R^2 is about half that of contemporaneous mean variance.[18]

Panel B confirms the model prediction that large firms, in quintile 5 (Q5) of the firm size distribution have lower exposure to the common size dispersion factor than small firms in quintile 1 (Q1). Finally, the data in Panel A show that the size dispersion factor explains a smaller fraction of the variation for large firms than for small firms, as predicted by Proposition 4.

[18] In unreported results, we confirmed these results for fundamental variance, as well as for residual market and fundamental variances. Residual variances are obtained by orthogonalizing firm returns or sales growth rates to a common factor in returns or sales growth rates, and then taking a variance.

Table 5: Comovement in Firm Variances

	Data			Model		
	(1)	(2)	(3)	(4)	(5)	(6)
	Factors			Factors		
	$\sigma_{S,t-1}$	$\mu_{\sigma,t-1}$	$\mu_{\sigma,t}$	$\sigma_{S,t-1}$	$\mu_{\sigma,t-1}$	$\mu_{\sigma,t}$
Panel A: Return Volatility R^2 by Size Quintile						
All firms	15.65	20.94	33.65	52.18	55.92	93.22
Q1	19.44	22.85	29.37	51.99	54.97	84.14
Q5	11.06	17.62	36.55	51.91	56.10	98.48
Panel B: Return Volatility Loadings by Size Quintile						
All firms	0.55	0.48	0.73	1.31	0.67	1.06
Q1	0.71	0.57	0.73	1.12	0.57	1.15
Q5	0.37	0.43	0.76	0.31	0.16	0.24

Notes: Each factor model regression is a time-series regression of log total volatility on a factor. Total volatility is measured as variance of daily returns within the calendar year, and size is measured as market equity. All volatility factor regressions take the form $\log \sigma_{i,t} = a_i + b_i factor_t + e_{i,t}$. For each stock i, we estimate the factor model and report the cross-sectional average of the R^2 in Panel A and cross-sectional average of the slopes in Panel B. We require a minimum of 25 observation to run the regression. In each panel, we report the full sample average and the average within the first and last quintile of the time-series average size distribution. We deflate size by CPI when constructing size quintiles. We consider three different volatility factors. The first, motivated by our network model, is the lagged cross section standard deviation of log market equity, $\sigma_{S,t-1}$ (Columns 1 and 4). The second and third factors we consider are the lagged and contemporaneous cross-sectional average log volatility, $\mu_{\sigma,t-1}$ (Columns 2 and 5) and $\mu_{\sigma,t}$ (Columns 3 and 6). All three factors use the full sample in the cross section. Columns 1-3 report the estimation results for the data at annual frequency from 1926 to 2016. We estimate the same factor structure for the model as well in Columns 4-6.

6 Conclusion

We document new features of the joint evolution of the firm size and firm volatility distribution and propose a new model to account for these features. In the model, shocks are transmitted from customers to suppliers in a production network. The larger the supplier, the more customer connections it has, the better diversified its customer base, and the lower its volatility. Large customers have a strong influence on their suppliers, so shocks to large firms have an important effect throughout the economy. A equilibrium model with consumer demand shocks and multiple inter-connected firms delivers a structural interpretation of the reduced-form network model.

When the firm size dispersion increases in this economy, large firms become more important, and many customer networks become less diversified. In those times, average firm volatility is higher as is the cross-sectional dispersion of volatility. Because the underlying innovations are i.i.d. over time, the model endogenously generates "uncertainty shocks." The model quantitatively replicates the most salient features of the firm size and the volatility distributions, while being consistent with data on customer network linkages. The estimation reveals the importance of strong network effects, without which the model cannot account for the large dispersion in firm volatilities.

Future work could explore in more depth the various sources of network linkages to fully account for the strong complementarity in firm-level volatilities.

References

Acemoglu, Daron, Ufuk Akcigit, and William Kerr. 2016. "Networks and the macroeconomy: An empirical exploration." *NBER Macroeconomics Annual* 30 (1):273–335.

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. 2012. "The Network Origins of Aggregate Fluctuations." *Econometrica* 80 (5):1977–2016.

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. 2015. "Systemic Risk and Stability in Financial Networks." *American Economic Review* 105 (2):564–608.

Ahern, Kenneth and Jarrad Harford. 2014. "The Importance of Industry Links in Merger Waves." *Journal of Finance* 69 (2):527–576.

Allen, Franklin and Ana Babus. 2009. "Networks in finance." In *The network challenge: strategy, profit, and risk in an interlinked world*. Wharton School Publishing, 367 – 382.

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

Bams, Dennis, Jaap Bos, and Magdalena Pisa. 2016. "Trade credit: Elusive insurance of firm growth." Research Memorandum 029, Maastricht University, Graduate School of Business and Economics (GSBE).

Barigozzi, Matteo, Christian T. Brownlees, Giampiero M. Gallo, and David Veredas. 2014. "Disentangling systematic and idiosyncratic dynamics in panels of volatility measures." *Journal of Econometrics* 182 (2):364–384.

Black, Fischer. 1976. "Studies in Stock Price Volatility Changes." *Proceedings of the 1976 meetings of the Business and Economics Statistics Section* :171–181.

Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77 (3):623–685.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2018. "Really Uncertain Business Cycles." *Econometrica* 86 (3):1031–1065.

Bloom, Nick, Stephen Bond, and John Van Reenen. 2007. "Uncertainty and Investment Dynamics." *The Review of Economic Studies* 74 (2):391–415.

Brandt, M.W., A. Brav, J.R. Graham, and A. Kumar. 2010. "The idiosyncratic volatility puzzle: Time trend or speculative episodes?" *Review of Financial Studies* 23 (2):863–899.

Campbell, John Y., Martin Lettau, Burton G. Malkiel, and Yexiao Xu. 2001. "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk." *The Journal of Finance* 56 (1):1–43.

Carvalho, Vasco M. 2010. "Aggregate Fluctuations and the Network Structure of Intersectoral Trade." Working Paper.

Carvalho, Vasco M and Xavier Gabaix. 2013. "The Great Diversification and its Undoing." *American Economic Review* 103 (5):1697–1727.

Christie, Andrew A. 1982. "The stochastic behavior of common stock variances: Value, leverage and interest rate effects." *Journal of Financial Economics* 10 (4):407 – 432.

Cohen, Lauren and Andrea Frazzini. 2008. "Economic Links and Predictable Returns." *The Journal of Finance* 63 (4):1977–2011.

Comin, Diego A. and Thomas Philippon. 2005. "The Rise in Firm-Level Volatility: Causes and Consequences." In *NBER Macroeconomics Annual 2005*, vol. 20. MIT Press, 167 – 228.

Davis, Steven J., John Haltiwanger, Ron Jarmin, and Javier Miranda. 2006. "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms." In *NBER Macroeconomics Annual 2006*, vol. 21. MIT Press, 107–180.

Doidge, Craig, Kathleen Kahle, Andrew Karolyi, and Rene Stulz. 2018. "Eclipse of the Public Corporation or Eclipse of the Public Markets?" *Journal of Applied Corporate Finance* 30 (1).

Duffie, Darrell and Kenneth Singleton. 1993. "Simulated Moments Estimation of Markov Models of Asset Prices." *Econometrica* 61 (4):929–952.

Engle, Robert F. and Stephen Figlewski. 2015. "Modeling the Dynamics of Correlations Among Implied Volatilities." *Review of Finance* 19 (3):991–1018.

Farboodi, Maryam. 2015. "Intermediation and Voluntary Exposure to Counterparty Risk." Working Paper.

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

Gabaix, Xavier. 2011. "The Granular Origins of Aggregate Fluctuations." *Econometrica* 79 (3):733–772.

———. 2016. "Technical Note on Demand shocks and Upstream Propagation." Tech. rep., New York University.

Gourieroux, Christian, Alain Monfort, and Eric Renault. 1993. "Indirect Inference." *Journal of Applied Econometrics* 8:85–118.

Herskovic, Bernard. 2018. "Networks in Production: Asset Pricing Implications." *Journal of Finance* 73:1785–1818.

Herskovic, Bernard, Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh. 2016. "The Common Factor in Idiosyncratic Volatility." *Journal of Financial Economics* 119 (2):249–283.

Herskovic, Bernard and Joao Ramos. Forthcoming. "Acquiring Information Through Peers." *American Economic Review* .

Jackson, Matthew O. 2014. "Networks in the Understanding of Economic Behaviors." *Journal of Economic Perspectives* 28 (4):3–22.

Kehrig, Matthias. 2015. "The cyclical nature of the productivity distribution." *Earlier version: US Census Bureau Center for Economic Studies Paper No. CES-WP-11-15* .

Kramarz, Francis, Julien Martin, and Isabelle Mejean. 2020. "Volatility in the small and in the large: The lack of diversification in international trade." *Journal of International Economics* 122:103276.

Leahy, John V. and Toni M. Whited. 1996. "The Effect of Uncertainty on Investment: Some Stylized Facts." *Journal of Money, Credit and Banking* 28 (1):64–83.

Lee, Gary G.J. and Robert F. Engle. 1993. "A Permanent and Transitory Component Model of Stock Return Volatility." UCSD Working Paper.

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

Midrigan, Virgiliu and Daniel Xu. 2014. "Finance and Misallocation: Evidence from Plant-Level Data." *American Economic Review* 104 (2):422–458.

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

Raddatz, Claudio. 2010. "Credit chains and sectoral comovement: Does the use of trade credit amplify sectoral shocks?" *The Review of Economics and Statistics* 92:985–1003.

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

Stanley, Michael HR, LuÝs AN Amaral, Sergey V Buldyrev, Shlomo Havlin, Heiko Leschhorn, Philipp Maass, Michael A Salinger, and H Eugene Stanley. 1996. "Scaling behaviour in the growth of companies." *Nature* 379 (6568):804–806.

Stanton, Richard, Johan Walden, and Nancy Wallace. 2018. "Mortgage loan flow networks and financial norms." *The Review of Financial Studies* 31 (9):3595–3642.

Stokey, Nancy L. 2016. "Wait-and-see: Investment options under policy uncertainty." *Review of Economic Dynamics* 21:246–265.

Taschereau-Dumouchel, Mathieu. 2018. "Cascades and Fluctuations in an Economy with an Endogenous Production Network." Working Paper.

White, Halbert. 2014. *Asymptotic theory for econometricians.* Academic press.

A Structural Model of Upstream Shock Transmission

This appendix sets up a simple structural model that delivers upstream transmission of shocks, and that delivers the network equation (1). It is a simple version of the canonical Long and Plosser (1983) model, applied to firms rather than to industries. Our derivation relies heavily on teaching notes that Xavier Gabaix graciously shared with us (Gabaix (2016)). The result generalizes Acemoglu, Akcigit, and Kerr (2016).

A.1 Setup

Households supply labor $\overline{L}$ inelastically. They derive utility from consuming each of the N goods:

$$U = \sum_{i=1}^{N} \theta_i \log c_i,$$

where θ_i is a taste shifter for good i. Without loss of generality, $\sum_{i=1}^{N} \theta_i = 1$. The household chooses its consumption basket to maximize total utility.

Good i is produced using labor as well as all of the other commodities:

$$Y_i = \exp(z_i) L_i^{b_i} \prod_{j=1}^{N} X_{ij}^{a_{ij}},$$

where the coefficient satisfy $\sum_{j=1}^{N} a_{ij} + b_i = 1, i = 1, 2, \ldots, N$, implying constant returns to scale.

A.2 Characterizing Equilibrium

Market clearing for goods implies that, for each good j:

$$C_j + \sum_{i=1}^{N} X_{ij} = Y_j \tag{A1}$$

Market clearing in the labor market implies that $\sum_{i=1}^{N} L_i = \overline{L}$.

Each firm maximizes profits. The first order conditions for profit maximization dictate that, for each good i, the demand for labor inputs and other goods satisfy:

$$\pi_j X_{ij} = a_{ij} \pi_i Y_i, \tag{A2}$$
$$w L_i = b_i \pi_i Y_i. \tag{A3}$$

where π_i is the price of good i.

Define economy-wide value added G as $G = \sum_{i=1}^{N} \pi_i C_i = w\overline{L}$. Define the ratio of firm i's output to total value added as ψ_i:

$$\psi_i = \frac{\pi_i Y_i}{G}.$$

We can restate equation (A2) as:

$$a_{ij} = \frac{\pi_j X_{ij}}{\pi_i Y_i} = \frac{\pi_j X_{ij}}{\psi_i G}. \tag{A4}$$

This shows that a_{ij} is the cost of input j in the value of the output i.

The household's first order condition for good j is given by:

$$\pi_j C_j = \theta_j G$$

Substituting this equation and equation (A4) into the market clearing condition (A1), we obtain:

$$\begin{aligned}
\pi_j Y_j &= \pi_j C_j + \sum_i \pi_j X_{ij}, \\
\psi_j G &= \left(\theta_j + \sum_i a_{ij}\psi_i\right) G, \\
\psi_j &= \theta_j + \sum_i a_{ij}\psi_i.
\end{aligned}$$

If firm i is the supplier of good i while firm j denotes a customer who uses good i as an input, then we get:

$$\psi_i = \theta_i + \sum_{j=1}^{N} a_{ji}\psi_j \tag{A5}$$

We use $\boldsymbol{A}$ to denote an $N \times N$ matrix with a_{ij} as the $(i,j)^{th}$ element of the matrix. We use $\boldsymbol{\psi}$ to denote the vector of firm shares with ψ_i as the i^{th} element, and we use $\boldsymbol{\theta}$ to denote the vector of preference parameters θ_i. Using matrix notation, equation (A5) can be stated as follows:

$$\boldsymbol{\psi} = \boldsymbol{\theta} + \boldsymbol{A}'\psi, \tag{A6}$$

which in turn implies that the firm shares can be stated as follows:

$$\boldsymbol{\psi} = (\boldsymbol{I} - \boldsymbol{A}')^{-1}\boldsymbol{\theta}.$$

The firm shares do not depend on the productivity shocks, but are only determined by preferences $\boldsymbol{\theta}$ and costs shares $\boldsymbol{A}$.

A.3 Link with W Matrix

Define $\tilde{w}_{ij}$ as the ratio of sales of supplier i to customer j, divided by the output of good i:

$$\tilde{w}_{ij} = \frac{\pi_i X_{ji}}{\pi_i Y_i} = \frac{\pi_i X_{ji}}{\pi_j Y_j}\frac{\pi_j Y_j}{\pi_i Y_i} = a_{ji}\frac{\psi_j}{\psi_i} \tag{A7}$$

Let $\Psi = diag(\psi)$ be the $N \times N$ matrix with the vector ψ on its diagonal. Then (A7) can be written in matrix notation as:

$$\tilde{\boldsymbol{W}} = \boldsymbol{\Psi}^{-1}\boldsymbol{A}'\boldsymbol{\Psi}.$$

Note that the rows of the $\tilde{\boldsymbol{W}}$ sum to less than one:

$$\begin{aligned}
\sum_{j=1}^{N} \tilde{w}_{ij} &= \sum_{j=1}^{N} a_{ji}\frac{\psi_j}{\psi_i}, \\
&= \sum_{j=1}^{N} \frac{\pi_i X_{ji}}{\pi_i Y_i}, \\
&= \frac{\pi_i Y_i - \pi_i C_i}{\pi_i Y_i} \equiv \gamma_i < 1,
\end{aligned}$$

unless good i is only used as an intermediary good but not used for final consumption. We call the fraction of total output of good i used as inputs for other firms, γ_i.

In our empirical implementation, we cannot directly measure the expenditure share matrix $\boldsymbol{A}$ for the U.S economy. However, we make two plausible assumptions. First, we assume that $\tilde{w}_{ij}$ is more likely to be non-zero when firm i is larger. Second, we assume that $\tilde{w}_{ij}$ is more likely to be larger when firm j is larger. Sales is our proxy for size.

Define $w_{ij} = \tilde{w}_{ij}/\gamma_i$ and collect the γ_i in a diagonal matrix Γ. The matrix $\boldsymbol{W}$ is:

$$\boldsymbol{W} = \Gamma^{-1}\tilde{\boldsymbol{W}}$$

We note that all the rows of $\boldsymbol{W}$ sum to one: $\sum_{j=1}^{N} w_{ij} = 1$.

A.4 Upstream Transmission of Preference Shocks

We are now ready to state the main result, Proposition 1, which delivers the network equation of the reduced-form model in the main text from the structural model described in this appendix.

We note at the outset that because the level of wages and prices is not determined, we choose the wage as the numéraire and express all good prices relative to the price of labor.

Proposition. *The responses of firm output to taste shocks $d\boldsymbol{\theta}$ approximately equals:*

$$\frac{d\boldsymbol{Y}}{\boldsymbol{Y}} = (\boldsymbol{I} - \Gamma\boldsymbol{W})^{-1}\frac{d\boldsymbol{\theta}}{\boldsymbol{\psi}}. \tag{A8}$$

Proof. By differentiation of equation (A5), we obtain the following system of equations:

$$d\psi_i = d\theta_i + \sum_{j=1}^{N} a_{ji} d\psi_j, i = 1, \ldots, N.$$

This system of equations implies the following system for growth rates in firm shares:

$$\begin{aligned} \frac{d\psi_i}{\psi_i} &= \frac{d\theta_i}{\psi_i} + \sum_{j=1}^{N} a_{ji}\frac{d\psi_j}{\psi_i}, \\ \frac{d\psi_i}{\psi_i} &= \frac{d\theta_i}{\psi_i} + \sum_{j=1}^{N} \gamma_i w_{ij}\frac{d\psi_j}{\psi_j}, \end{aligned}$$

where we have used the mapping from the cost shares to the expenditure shares in (A7) and the link between w_{ij} and $\tilde{w}_{ij}$. Hence, in matrix notation, we get the following result:

$$\frac{d\boldsymbol{\psi}}{\boldsymbol{\psi}} = \frac{d\boldsymbol{\theta}}{\boldsymbol{\psi}} + \Gamma\boldsymbol{W}\frac{d\boldsymbol{\psi}}{\boldsymbol{\psi}}.$$

This in turn implies that:

$$\frac{d\boldsymbol{\psi}}{\boldsymbol{\psi}} = (\boldsymbol{I} - \Gamma\boldsymbol{W})^{-1}\frac{d\boldsymbol{\theta}}{\boldsymbol{\psi}}$$

Let $\widehat{x}$ denote dx/x. To complete the proof, we need to show that $\widehat{\boldsymbol{Y}} = \widehat{\boldsymbol{\psi}}$.
We start with the goods market. First, note that the firm share $\psi_i = \frac{\pi_i Y_i}{G}$. Hence, the percentage changes satisfy: $\widehat{\pi}_i + \widehat{Y}_i = \widehat{\psi}_i$, because $\widehat{G} = 0$. The latter is because aggregate labor supply $\overline{L}$ is fixed and $\widehat{w} = 0$, a normalization. Similarly, for firm j: $\widehat{\pi}_j + \widehat{Y}_j = \widehat{\psi}_j$. Second, equation (A4) implies that:

$$\widehat{\psi}_i = \widehat{\pi}_j + \widehat{X}_{ij}.$$

By combining the two preceding equations, we get the following expression for the growth rate of intermediate inputs:

$$\widehat{X}_{ij} = \widehat{\psi}_i - \widehat{\psi}_j + \widehat{Y}_j. \tag{A9}$$

Next, we turn to the labor market where:

$$wL_i = b_i\pi_i Y_i = \psi_i b_i G.$$

In growth rates, this implies that:

$$\widehat{L}_i = \widehat{\psi}_i. \tag{A10}$$

Finally, we consider the change in output in firm i, which is given by:

$$\widehat{Y}_i = \widehat{z}_i + b_i \widehat{L}_i + \sum_j a_{ij} \widehat{X}_{ij}$$

Because there are only preference shocks and no productivity shocks, $\widehat{\boldsymbol{z}} = 0$. After substituting expressions (A9) and (A10), we obtain:

$$\widehat{Y}_i = b_i \widehat{\psi}_i + \sum_j a_{ij} \left(\widehat{\psi}_i - \widehat{\psi}_j + \widehat{Y}_j \right).$$

This can be simplified to yield:

$$\widehat{Y}_i = \widehat{\psi}_i + \sum_j a_{ij} \left(-\widehat{\psi}_j + \widehat{Y}_j \right),$$

where we have used the constant returns to scale assumption $b_i + \sum_j a_{ij} = 1$. In matrix notation, this implies that:

$$\widehat{\boldsymbol{Y}} - \boldsymbol{A}\widehat{\boldsymbol{Y}} = \widehat{\boldsymbol{\psi}} - \boldsymbol{A}\widehat{\boldsymbol{\psi}} \Rightarrow \widehat{\boldsymbol{Y}} = \widehat{\boldsymbol{\psi}}.$$

The proposition implies that all prices are constant, i.e., unaffected by taste shocks: $\widehat{\pi}_i = 0$. □

By setting the reduced-form shocks equal to the scaled taste shocks $\boldsymbol{\varepsilon} = d\boldsymbol{\theta}/\boldsymbol{\psi}$ and $\mu_g = 0$ (an assumption we make in our estimation exercise), equation (A8) is identical to the network equation (2) in the main text.

The structural model provides an interpretation of the shocks as consumer preference shocks. It also provides an interpretation of the network coefficients γ_i, which modulate the strength of the network effects. In the structural model, γ_i is the share of output of firm i used as inputs in other firms. Intuitively, shocks to one firm do not propagate if other firms do not use that firm's good as an input. Finally, the structural model makes clear that demand shocks propagate upstream.

Since we cannot accurately measure γ_i for different goods produced by different firms, in our empirical work, we impose that $\gamma_i = \gamma, \forall i$. Then $1 - \gamma$ denotes the average share of output that accrues to final consumption. That is, $1 - \gamma$ measures the leakage in the production network due to final goods consumption.

A Simple Example with Two Firms A simple example where good 1 is the intermediate good and good 2 is the final good illustrates the upward shock propagation:

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 \\ \alpha & 0 \end{bmatrix}$$

Then

$$\tilde{\boldsymbol{W}} = \Gamma \boldsymbol{W} = \begin{bmatrix} 0 & \alpha \frac{\psi_2}{\psi_1} \\ 0 & 0 \end{bmatrix}$$

and

$$(\boldsymbol{I} - \Gamma \boldsymbol{W})^{-1} = \begin{bmatrix} 1 & \alpha \frac{\psi_2}{\psi_1} \\ 0 & 1 \end{bmatrix}$$

Using our expression, this implies that the growth rates can be expressed as:

$$\frac{d\psi_1}{\psi_1} = \frac{d\theta_1}{\psi_1} + \alpha \frac{\psi_2}{\psi_1} \frac{d\theta_2}{\psi_2}. \tag{A11}$$

$$\frac{d\psi_2}{\psi_2} = \frac{d\theta_2}{\psi_2}. \tag{A12}$$

Generally, if $\boldsymbol{A}$ is lower triangular, which means that i uses only inputs from $j < i$, then $(\boldsymbol{I} - \Gamma \boldsymbol{W})^{-1}$ is upper triangular. Hence, demand shocks are transmitted upstream, from the final goods firm (firm 2) to the intermediate goods firm (firm 1).

Example with Three Firms A simple three-firm example where good 1 is the basic input , good 2 is intermediate good firm and firm 3 produces the final good illustrates taste shocks affect the distribution of firm volatility even hough they leave aggregate value added unchanged ($\widehat{G} = 0$). Under upstream shock transmission, the $\boldsymbol{A}$ matrix is lower triangular (i uses only inputs from $j < i$):

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 & 0 \\ \alpha & 0 & 0. \\ \beta & \nu & 0. \end{bmatrix}$$

Then

$$\tilde{\boldsymbol{W}} = \Gamma \boldsymbol{W} = \begin{bmatrix} 0 & \alpha\frac{\psi_2}{\psi_1} & \beta\frac{\psi_3}{\psi_1} \\ 0 & 0 & \nu\frac{\psi_3}{\psi_2} \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$(\boldsymbol{I} - \Gamma \boldsymbol{W}) = \begin{bmatrix} 1 & -\alpha\frac{\psi_2}{\psi_1} & -\beta\frac{\psi_3}{\psi_1} \\ 0 & 1 & -\nu\ \frac{\psi_3}{\psi_2} \\ 0 & 0 & 1 \end{bmatrix}$$

Then the matrix that governs the transmission of shocks is upper triangular:

$$(\boldsymbol{I} - \Gamma \boldsymbol{W})^{-1} = \begin{bmatrix} 1 & \alpha\frac{\psi_2}{\psi_1} & \alpha\frac{\psi_2}{\psi_1}\nu\frac{\psi_3}{\psi_2} + \beta\frac{\psi_3}{\psi_1} \\ 0 & 1 & \nu\frac{\psi_3}{\psi_2} \\ 0 & 0 & 1 \end{bmatrix}$$

Using our expression from equation (A8), this implies that the growth rates of each firm can be expressed as:

$$\begin{aligned} \frac{d\psi_1}{\psi_1} &= \frac{d\theta_1}{\psi_1} + \alpha\frac{\psi_2}{\psi_1}\frac{d\theta_2}{\psi_2} + (\alpha\nu + \beta)\frac{\psi_3}{\psi_1}\frac{d\theta_3}{\psi_3}. \\ \frac{d\psi_2}{\psi_2} &= \frac{d\theta_2}{\psi_2} + \nu\frac{\psi_3}{\psi_2}\frac{d\theta_3}{\psi_3}. \\ \frac{d\psi_3}{\psi_3} &= \frac{d\theta_3}{\psi_3}. \end{aligned}$$

Demand shocks are transmitted upstream, from the final goods firm (firm 3) to the raw inputs firm (firm 1). The latter is exposed to firm 3's shocks via firm 2 as well as directly. To see that, it helps to restate these expressions as:

$$\begin{aligned} \frac{d\psi_1}{\psi_1} &= \frac{d\theta_1}{\psi_1} + \alpha\frac{\psi_2}{\psi_1}(\nu\frac{\psi_3}{\psi_2}\frac{d\theta_3}{\psi_3} + \frac{d\theta_2}{\psi_2}) + \beta\frac{\psi_3}{\psi_1}\frac{d\theta_3}{\psi_3} \\ \frac{d\psi_2}{\psi_2} &= \nu\frac{\psi_3}{\psi_2}\frac{d\theta_3}{\psi_3} + \frac{d\theta_2}{\psi_2} \\ \frac{d\psi_3}{\psi_3} &= \frac{d\theta_3}{\psi_3} \end{aligned}$$

The first term in the first equation above measures the direct impact of firm 1's demand shocks on firm 1. The second term measures the impact of firm 2's shocks on firm 1. The third term measures the direct impact of firm 3's demand shock on firm 1. Note that the second term also comprises the impact of firm 3's shock on firm 2. Hence, firm 1 is exposed to demand shocks for final goods directly and via firm 2.

A.5 Size Dispersion and Average Variance in Example Economies

We provide a sufficient condition under which an increase in firm size dispersion coincides with an increase in average firm variance, as well as simulation evidence for three-firm economies that shows that the sufficient condition is likely to be satisfied.

To show the implications for firm variance, we are interested in $\text{Var}\left(\frac{dY}{Y}\right) = \text{Var}\left(\frac{d\psi}{\psi}\right)$, and, from the model, we know that changes in firm size relate to preference shocks in the following way:

$$\frac{d\psi}{\psi} = \left(I - \tilde{W}\right)^{-1}\frac{d\theta}{\psi} = \left(I - \Psi^{-1}A'\Psi\right)^{-1}\frac{d\theta}{\psi},$$

where $\frac{d\theta}{\psi} = \left[\frac{d\theta_1}{\psi_1}, \frac{d\theta_2}{\psi_2}, \frac{d\theta_3}{\psi_3}, \ldots\right]'$, $\psi = [\psi_1, \psi_2, \psi_3, \ldots]'$, and $\Psi = \text{diag}(\psi)$ is a diagonal matrix with the vector ψ on the main diagonal. Hence, the variance-covariance of firms' growth rates is given by:

$$\begin{aligned}
\text{Var}\left(\frac{d\psi}{\psi}\right) &= \left(I - \Psi^{-1}A'\Psi\right)^{-1}\text{Var}\left(\frac{d\theta}{\psi}\right)\left(I - \Psi A\Psi^{-1}\right)^{-1} \\
&= \left(I - \Psi^{-1}A'\Psi\right)^{-1}\Psi^{-1}\text{Var}\left(d\theta\right)\Psi^{-1}\left(I - \Psi A\Psi^{-1}\right)^{-1} \\
&= \left(\Psi - A'\Psi\right)^{-1}\Sigma_0\left(\Psi - \Psi A\right)^{-1} \\
&= \Psi^{-1}\left(I - A'\right)^{-1}\Sigma_0\left(I - A\right)^{-1}\Psi^{-1} \\
&= \Psi^{-1}M\Psi^{-1}
\end{aligned}$$

where $\Sigma_0 \equiv \text{Var}(d\theta)$, and $M = (I - A')^{-1}\Sigma_0(I - A)^{-1}$. The variance of firm i is given by $(M)_{ii}\,\psi_i^{-2}$, where M_{ii} is the i^{th} diagonal element of M, which consists only of model primitives. Average firm variance is the average of the diagonal elements of the variance matrix:

$$\text{Average Firm Variance} = \frac{1}{n}\sum_{i=1}^{n}\frac{(M)_{ii}}{\psi_i^2} \tag{A13}$$

When comparing two economies with different size dispersion, it is desirable to hold the size of the final consumption goods sector constant. The latter is given in equilibrium by $\frac{1}{\sum_{i=1}^{n}\psi_i}$. Without such an assumption, one could engineer a higher average firm variance by lowering all firms' ψ_i's. But the resulting increase in variance would mechanically be coming from an increase in the consumption share. Henceforth, we study changes in the firm size distribution that keep the size of the final goods sector constant.

The next proposition shows a sufficient condition under which an increase in size dispersion is accompanied by an increase in the average firm variance.

Proposition 5. *An economy with a uniform firm size distribution (i.e. firms of equal size ψ_i) has lower average firm variance than an economy with a non-uniform size distribution (i.e. firms of unequal size), holding fixed the consumption share, provided that the cross-sectional covariance between $1/\psi_i$ and $(M)_{ii}$ satisfies:*

$$2\overline{\psi^{-1}}\sum_i (M)_{ii}\left(\psi_i^{-1} - \overline{\psi^{-1}}\right) > \underbrace{-\sum_i (M)_{ii}\left(\psi_i^{-1} - \overline{\psi^{-1}}\right)^2}_{<0} \tag{A14}$$

where $\overline{\psi^{-1}} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\psi_i}$.

Proof. Let AvgVar $_{\text{Equal Size}}$ (AvgVar $_{\text{Unequal Size}}$) be the average firm variance under an (un)equal firm size distribution:

$$\begin{aligned}
\text{AvgVar}_{\text{ Unequal Size}} &= \frac{1}{n}\sum_i (M)_{ii}\,\psi_i^{-2} = \frac{1}{n}\sum_i (M)_{ii}\left[\overline{\psi^{-1}} + d_i\right]^2 \\
&= \frac{1}{n}\sum_i (M)_{ii}\left[\left(\overline{\psi^{-1}}\right)^2 + d_i^2 + 2\overline{\psi^{-1}}d_i\right] \\
&> \frac{1}{n}\sum_i (M)_{ii}\left(\overline{\psi^{-1}}\right)^2 > \frac{1}{n}\sum_i (M)_{ii}\,\overline{\psi}^{-2} = \text{AvgVar}_{\text{ Equal Size}}
\end{aligned}$$

where $d_i = \psi_i^{-1} - \overline{\psi^{-1}}$. The first inequality holds by the covariance bound, and the last inequality holds because of Jensen's inequality. Finally, the last equality holds because under equal size and holding consumption share fixed we have $\psi_i = \overline{\psi}$ for every i. □

The sufficient condition provided by equation (A14) requires the cross-sectional covariance between the inverse of firm-size-to-value-added ratio, i.e. ψ_i^{-1}, and network-implied volatility to be greater than a negative lower bound. An implication of this result is that more dispersion in firm size distribution leads to higher average firm volatility whenever firm size is orthogonal to the network structure (e.g. $\psi_i \perp (M)_{ii}$).

To gain insight in how likely the positive relationship between size dispersion and average variance is to arise, we resort to a simulation exercise. To avoid that the specific correlation structure of preference shocks drives our results, we assume a symmetric correlation structure with $\text{var}\,(d\theta_i) = \sigma^2$ for every $i = 1, \ldots, n$, and $\text{cov}\,(d\theta_i, d\theta_j) = \rho\sigma^2$ for every $i \neq j$:

$$\Sigma_0 = \text{Var}\,(d\theta) = \begin{bmatrix} \sigma^2 & \rho\sigma^2 & \cdots & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \cdots & \rho\sigma^2 \\ \vdots & & \ddots & \vdots \\ \rho\sigma^2 & \rho\sigma^2 & \ldots & \sigma^2 \end{bmatrix}.$$

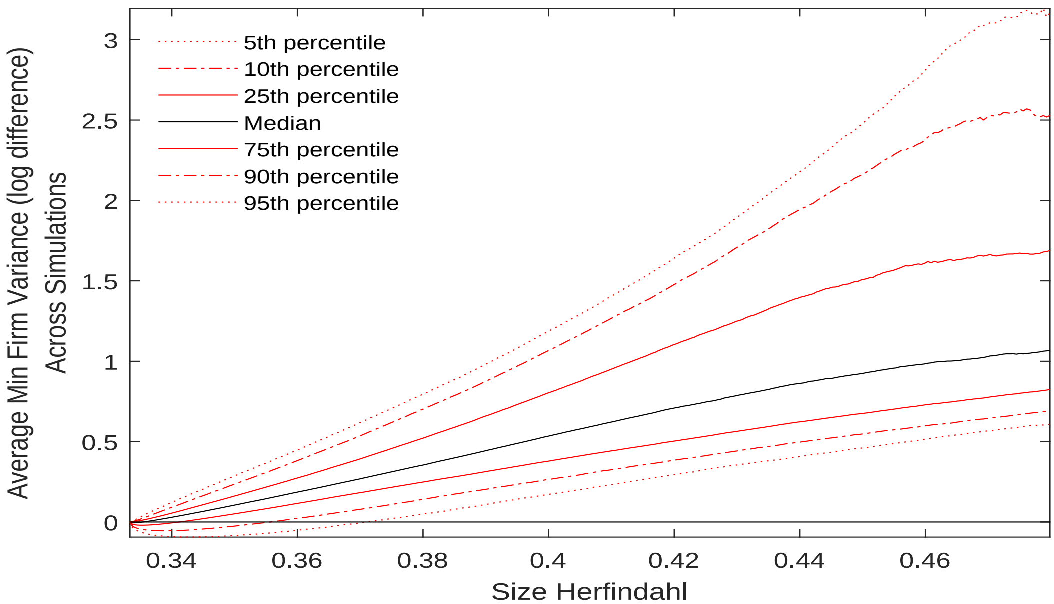
By the nature of preference shocks, the last shock has to be implied by $d\theta_n = 1 - \sum_{i=1}^{n-1} d\theta_i$. This relation pins down the value for ρ based either on the variance of the last shock, i.e. $\sigma^2 = (n-1)\sigma^2 + (n-1)(n-2)\rho\sigma^2$, or on the covariance of the last shock with the other shocks, i.e. $\rho\sigma^2 = -\sigma^2 - (n-2)\rho\sigma^2$. Either one delivers a preference shock correlation given by:

$$\rho = -\frac{1}{n-1}.$$

We specialize to a three-firm setting. We simulate 50,000 distinct networks, i.e., entries of matrix $\boldsymbol{A}$ of 3 firms. Each entry of the network matrix A_{ij} is drawn from a uniform distribution in the unit interval. We focus on changes in preference parameters that are (i) feasible, i.e. $\theta_i \geq 0$ for every i and $\sum_i \theta_i = 1$; and (iii) do not change consumption share, i.e. $\frac{1}{\sum_{i=1}^{n} \psi_i}$ remains unchanged, where $\psi = (I - A')^{-1}\theta$ is implied by equilibrium conditions. We require each simulated economy to satisfy two conditions: (i) each row of A have to sum to at most one, (ii) there exist feasible preference parameter values such that firms having the same size is an equilibrium. The first condition is a network-parameter restriction. We impose the second condition so that we can increase firm size dispersion, starting from an equilibrium where firms have the same size. For each of the 50,000 simulated economies, we compute an underlying feasible preference shock vector $\boldsymbol{d\theta}$ to deliver the minimum average firm variance for each possible level of model-implied size Herfindahl. Choosing the minimum variance is a conservative choice when establishing a rise in average variance.

Figure A1 plots the change in average average firm variance relative to the equal-size economy across all 50,000 models (solid line), as well as percentiles of the distribution of average variance across simulations, at each level of size Herfindahl (dashed lines). Overall, there is clear positive relation between size dispersion and average firm variance. This is not only the case for the average model, but for the vast majority of models. There is a tail of models (about 5-10%) for which increases in Herfindahl below 0.02 are associated with declines in average firm variance. But even for this set of models, the relationship turns positive for larger increases in size dispersion. In sum, for the vast majority of network models and preference shocks, the relationship between size dispersion and average variance is positive.

Figure A1: Minimum Average Firm Variance by Size Herfindahl for 50,000 Network Models



Notes: The figure plots the average firm variance as size Herfindahl increases for 50,000 simulated networks. For each simulation, we find the feasible firm size distribution that delivers the size Herfindahl displayed in the x-axis and the minimum average firm variance. Specifically, in each simulation, we assume that there are three firms and we randomly draw the entries of the matrix A from a uniform distribution in the unit interval. We require each simulated network to satisfy the two properties: (i) all rows in A sum to less than or equal to one, (ii) there has to be a distribution of preference parameters such that firms having equal shares is an equilibrium. Once we draw a feasible network, we then solve for different preference parameters that minimize the average firm variance and deliver the Herfindahl target. Finally, we compute the log difference between the simulated economy relative to the same economy when firm have equal size (i.e. minimum size Herfindahl).

B Proofs of Propositions in Main Text

Returning to the full network specification in (1), the following results formalize the preceding intuition in a large N economy. First, we provide a limiting description of each supplier's customer Herfindahl. In this appendix, we use the simplified notation H to represent the object H^{out} from the main text. Throughout, we use asymptotic equivalence notation $x(N) \sim y(N)$ to denote that $x(N)/y(N) \to 1$ almost surely as $N \to \infty$. We use $x(N) \approx a + b(N) + c(N)$ to denote that $x(N) = a + b(N) + c(N) + o(\min[a, b(N), c(N)])$. Where there is no ambiguity, we suppress time and/or firm subscripts. We start by stating and proving two helper lemmas.

B.1 Lemma 1

The following lemma highlights the common structure in customer Herfindahls across suppliers. The ratio of the second non-central moment of the size distribution to the squared first moment captures the degree of concentration in the entire firm size distribution. Economy-wide firm size concentration affects the customer network concentration of all firms. Differences in customer Herfindahl across suppliers are inversely related to supplier size, capturing the model feature that larger firms are connected to more firms on average. Under lognormality, $E[S^2]/E[S]^2$ equals the (exponentiated) cross-sectional standard deviation of log firm size. The lemma applies more generally to the case where the size distribution has finite variance, and below we analyze firm volatility decay in the case of power law size distributions with infinite variance but finite mean.

Lemma 1. *Fix the size of the i^{th} firm, S_i, and consider a sequence of economies indexed by the number of firms N. If the firm size distribution has finite variance, then*

$$\sum_{k=1}^{N} b_{i,k} S_k^q \sim \frac{N^{(1-\zeta)}}{Z} \tilde{S}_i E[S^q], \quad q = 1, 2. \tag{B1}$$

Therefore,

$$H_i \sim \frac{1}{N^{(1-\zeta)}} \frac{Z}{\tilde{S}_i} \frac{E[S^2]}{E[S]^2}. \tag{B2}$$

In particular, if $\log(S)$ *is normal with variance σ_s^2, then*

$$H_i \sim \frac{1}{N^{1-\zeta}} \frac{Z}{\tilde{S}_i} \exp(\sigma_s^2).$$

Proof. Given S_i, and for fixed N, the expected value of $N^{-1} \sum_k b_{i,k} S_k^q$ $(q = 1, 2)$ is

$$E\left[\frac{1}{N}\sum_k b_{ik} S_k^q | S_i\right] = E\left[\frac{1}{N}\sum_k p_i S_k^q | S_i\right] = \frac{p_i}{N} E\left[\sum_k S_k^q | S_i\right] \to p_i E\left[S^q\right] = \frac{N^{-\zeta}}{Z} \tilde{S}_i E[S^q].$$

The first equality follows from the fact that all $b_{ik}, k = 1, ..., N$ are iid with expectation (conditional on S_i) of p_i. Thus, the stated convergence is given by the law of large numbers and this, together with the definition of p_i in equation (4) of the main text, establishes the asymptotic relation in (B1).

Next, we analyze H_i. Because $b_{i,k}^2 = b_{i,k}$, we have

$$H_i = \sum_k w_{i,k}^2 = \frac{\sum_k b_{i,k} S_k^2}{\left(\sum_k b_{i,k} S_k\right)^2}.$$

Applying (B1) in the numerator and denominator and appealing to Slutsky's theorem delivers (B2).

Lastly, in the special case of lognormality, $E[S^2]/E[S]^2 = e^{\sigma_s^2}$. □

B.2 Lemma 2

As described above, H_i is the first order determinant of firm volatility in our model. The following intermediate lemma allows us to capture higher order network effects that contribute to a firm's variance.

Lemma 2. *Consider a sequence of economies indexed by the number of firms N.*

(i) Fixing S_j, $E[w_{i,j}|S_j] \sim \tilde{S}_j/N$.

(ii) Fixing S_i, $\sum_{k=1}^{N} \frac{\gamma_i \gamma_k b_{i,k} b_{k,j} S_k}{\sum_{l=1}^{N} b_{k,l} S_l} \sim p_i \gamma_i \overline{\gamma}$, where $\overline{\gamma} = \frac{1}{N}\sum_{i=1}^{N} \gamma_i$.

(iii) Fixing S_i, and defining the matrix $\bar{W}$ as $[\bar{W}]_{i,j} = \tilde{S}_j/N$, we have

$$[W^q]_{i,j} \sim \frac{\tilde{S}_j}{N} = [\bar{W}]_{i,j}, \quad \textit{for } q \geq 2,$$

and, more generally,

$$[(\Gamma W)^q]_{i,j} \sim \overline{\gamma}^{q-1} \frac{\gamma_i \tilde{S}_j}{N} = \overline{\gamma}^{q-1} [\Gamma \bar{W}]_{i,j}, \quad \textit{for } q \geq 2.$$

(iv) Fixing S_i,

$$[W\bar{W}']_{i,j} \sim \frac{E[S^2]}{NE[S]^2}, \quad [\bar{W}\bar{W}']_{i,j} \sim \frac{E[S^2]}{NE[S]^2}, \quad \textit{and} \quad [WW']_{i,i} = H_i, \quad \textit{for all } i,j.$$

Additionally fixing S_j,

$$[WW']_{i,j} \sim \frac{E[S^2]}{NE[S]^2}, \quad \textit{for all } i \neq j.$$

Proof.

(i) Fix S_j and recall $w_{i,j} = \frac{b_{ij} S_j}{\sum_k b_{ik} S_k}$. Then, applying Lemma 1 to the denominator,

$$E\left[w_{i,j}|S_j\right] \sim S_j E\left[\frac{b_{ij}}{\frac{N^{1-\zeta}}{Z} S_i}\right] = S_j E\left[E\left[\frac{b_{ij}}{\frac{N^{1-\zeta}}{Z} S_i}|S_i\right]\right] = S_j E\left[\frac{p_i}{\frac{N^{1-\zeta}}{Z} S_i}\right] = \frac{\tilde{S}_j}{N}.$$

(ii) Define $Z_k \equiv \frac{\gamma_i \gamma_k b_{i,k} b_{k,j} S_k}{\sum_{l=1}^{N} b_{k,l} S_l}$. Then $\mu_k \equiv E[Z_k|S_i] \sim \frac{p_i}{N}\gamma_i \overline{\gamma}$ because

$$\begin{aligned} \mu_k &= E\left[\frac{\gamma_i \gamma_k b_{i,k} b_{k,j} S_k}{\sum_{l=1}^{N} b_{k,l} S_l}|S_i\right] = p_i E\left[\frac{b_{k,j} S_k}{\sum_{l=1}^{N} b_{k,l} S_l}|S_i\right] \gamma_i \gamma_k && \text{(by } b_{i,k} \perp) \\ &\sim p_i E\left[\frac{b_{k,j} S_k}{\frac{N^{1-\zeta}}{Z} S_k}\right] \gamma_i \gamma_k && \text{(by Lemma 1)} \\ &= \frac{Z}{N^{1-\zeta}} p_i E\left[p_k\right] \gamma_i \gamma_k = \frac{p_i}{N} \gamma_i \gamma_k. && \text{(since } E\left[p_k\right] = N^{-\zeta}/Z) \end{aligned}$$

Note that μ_k is $O(N^{-(1+\zeta)})$, so $\sum_k \mu_k$ converges. By the LLN for heterogeneous and dependent variables,[19]

$$N^{-1}\sum_{k=1}^{N} Z_k - N^{-1}\sum_{k=1}^{N} \mu_k \xrightarrow[N\to\infty]{a.s.} 0$$

[19] E.g., Theorem 3.47 of White (2001).

and therefore

$$\sum_{k=1}^{N} \frac{\gamma_i \gamma_k b_{i,k} b_{k,j} S_k}{\sum_{l=1}^{N} b_{k,l} S_l} = \sum_{k=1}^{N} Z_k \sim \sum_{k=1}^{N} \mu_k = p_i \gamma_i \overline{\gamma}.$$

(iii) We begin with the case $q = 2$.

$$[(\Gamma W)^2]_{i,j} = \sum_k \gamma_i w_{i,k} w_{k,j} \gamma_k = \sum_k \left(\frac{\gamma_i b_{i,k} S_k}{\sum_m b_{i,m} S_m} \frac{\gamma_k b_{k,j} S_j}{\sum_l b_{k,l} S_l} \right) = \left(\frac{S_j}{\sum_m b_{i,m} S_m} \right) \left(\sum_k \frac{\gamma_i \gamma_k b_{i,k} b_{k,j} S_k}{\sum_l b_{k,l} S_l} \right).$$

Applying Lemma 1 to the denominator of the first term, we have

$$\frac{S_j}{\sum_m b_{i,m} S_m} \sim \frac{S_j}{S_i} \frac{Z}{N^{1-\zeta}}.$$

The asymptotic behavior of the second term is given by part (ii) of this lemma along with the definition of p_i. Combining the limits for the first and second terms, we have

$$[W^2]_{i,j} \sim \left(\frac{S_j}{S_i} \frac{Z}{N^{1-\zeta}} \right) p_i \gamma_i \overline{\gamma} = \frac{\tilde{S}_j}{N} \gamma_i \overline{\gamma}.$$

For induction, assume that $[(\Gamma W)^q]_{i,j} \sim \overline{\gamma}^{q-1} \gamma_i \tilde{S}_j / N$ for $q > 2$. Then

$$[\Gamma W \, (\Gamma W)^q]_{i,j} \sim \sum_{k=1}^{N} \frac{\gamma_i \gamma_k \overline{\gamma}^{q-1} b_{i,k} S_k \tilde{S}_j}{N \left(\sum_l b_{i,l} S_l \right)} = \gamma_i \overline{\gamma}^{q-1} \frac{\tilde{S}_j}{N} \frac{\sum_k b_{i,k} S_k \gamma_k}{\sum_l b_{i,l} S_l} \sim \frac{\tilde{S}_j}{N} \gamma_i \overline{\gamma}^q.$$

The first equivalence is from matrix multiplication of W and the asymptotic equivalent of W^q (from the induction assumption), and the equalities are immediate.

(iv) Subsequent results involve the matrix $\bar{W}$, where $[\bar{W}]_{i,j} = \tilde{S}_j / N$. First, we have

$$[W \bar{W}']_{i,j} = \sum_k \frac{b_{i,k} S_k^2}{\left(\sum_l b_{i,l} S_l \right) N E[S]} \sim \frac{E[S^2]}{N E[S]^2},$$

which follows the same argument structure as part (ii) of this lemma. Next,

$$[\bar{W} \bar{W}']_{i,j} = \sum_k \frac{\tilde{S}_k^2}{N^2} \sim \frac{E[S^2]}{N E[S]^2},$$

which relies on the LLN for $N^{-1} \sum_k S_k^q$ for $q = 1, 2$. Note that $H_i = [W W']_{ii}$, where the behavior of H_i is given in Lemma 1.

For off-diagonal elements of WW', fix S_i and S_j. Then

$$[W W']_{i,j} = \frac{\sum_k b_{i,k} b_{j,k} S_k^2}{\left(\sum_l b_{i,l} S_l \right) \left(\sum_l b_{j,l} S_l \right)} \sim \frac{E[S^2]}{N E[S]^2},$$

which follows from preceding results. In particular, noting that $E\left[\sum_k b_{i,k} b_{j,k} S_k^2 | S_i, S_j \right] = p_i p_j E\left[\sum_k S_k^2 | S_i, S_j \right]$, the numerator is asymptotically equivalent to $\tilde{S}_i \tilde{S}_j \frac{N^{1-2\zeta}}{Z^2} E[S]^2$, and the limits of denominator sums are given in Lemma 1.

□

B.3 Main Result on Firm Volatility: Proposition 2

Our main theoretical proposition connects the variance of a firm to its size and to the concentration of firm sizes throughout the economy.

Proposition. *Consider a sequence of economies indexed by the number of firms N. If the firm size distribution has finite variance, then the Leontief inverse has limiting behavior described by*

$$(I-\Gamma W)^{-1} \approx I + \Gamma W + \frac{\overline{\gamma}}{1-\overline{\gamma}}\Gamma \bar{W}.$$

where $\overline{\gamma} = \frac{1}{N}\sum_{i=1}^{N}\gamma_i$. Fixing the size of the i^{th} firm, S_i, volatility of firm i has limiting behavior described by

$$V(g_i|S_i) \approx \sigma_\varepsilon^2\left[1 + \left(\frac{Z}{N^{1-\zeta}\tilde{S}_i} + \frac{\overline{\gamma}\,(2-\overline{\gamma})}{N(1-\overline{\gamma})^2}\right)\gamma_i^2\frac{E[S^2]}{E[S]^2} + 2\gamma_i\frac{\overline{\gamma}}{1-\overline{\gamma}}\frac{\tilde{S}_i}{N}\right].$$

Proof. Because $V(g) = \sigma_\varepsilon^2(I-\Gamma W)^{-1}(I-W'\Gamma)^{-1}$, we study the behavior of $(I-\Gamma W)^{-1}$ as the number of firms N becomes large, and holding S_i fixed. The Leontief inverse can be rewritten as an infinite sum, $(I-\Gamma W)^{-1} = I + \Gamma W + (\Gamma W)^2 + ...$ From Lemma 2, $(\Gamma W)^q \sim \overline{\gamma}^{q-1}\Gamma\bar{W}$ and $[(\Gamma W)^q]_{i,j} \sim \overline{\gamma}^{q-1}\gamma_i\tilde{S}_j/N$ for $q > 2$, which therefore implies that

$$\begin{aligned}(I-\Gamma W)^{-1} &= I + \Gamma W + (\Gamma W)^2 + (\Gamma W)^3 + \dots \\ &\approx I + \Gamma W + \Gamma\bar{W}\left(\overline{\gamma} + \overline{\gamma}^2 + \overline{\gamma}^3 + \dots\right) \\ &= I + \Gamma W + \frac{\overline{\gamma}}{1-\overline{\gamma}}\Gamma\bar{W}.\end{aligned}$$

By Slutsky's theorem, the growth rate variance is determined in the limit by the outer product of $I + \Gamma W + \frac{\overline{\gamma}}{1-\overline{\gamma}}\Gamma\bar{W}$. Expanding out this product produces

$$\begin{aligned}(I-\Gamma W)^{-1}(I-W'\Gamma)^{-1} \approx{}& I + \Gamma W + W'\Gamma + \Gamma W W'\Gamma + \frac{\overline{\gamma}}{1-\overline{\gamma}}\Gamma\bar{W} + \frac{\overline{\gamma}}{1-\overline{\gamma}}\bar{W}'\Gamma \\ &+ \frac{\overline{\gamma}}{1-\overline{\gamma}}\Gamma W\bar{W}'\Gamma + \frac{\overline{\gamma}}{1-\overline{\gamma}}\Gamma\bar{W}W'\Gamma + \left(\frac{\overline{\gamma}}{1-\overline{\gamma}}\right)^2\Gamma\bar{W}\bar{W}'\Gamma.\end{aligned} \tag{B3}$$

The i^{th} firm's variance conditional on its size, $V(g_i|S_i)$, is asymptotically equivalent to the i^{th} diagonal element of (B3). Diagonal elements of W are zero by definition. Limits of the remaining terms ($\bar{W}$, WW', $W\bar{W}'$, etc.) are given explicitly in Lemmas 1 and 2. Therefore,

$$\begin{aligned}V(g_i|S_i) \approx \sigma_\varepsilon^2\Bigg[&1 + 0 + 0 + \gamma_i^2\frac{1}{N^{1-\zeta}}\frac{Z}{\tilde{S}_i}\frac{E[S^2]}{E[S]^2} + 2\gamma_i\frac{\overline{\gamma}}{1-\overline{\gamma}}\frac{\tilde{S}_i}{N} + 2\gamma_i^2\frac{\overline{\gamma}}{1-\overline{\gamma}}\frac{E[S^2]}{NE[S]^2} \\ &+\gamma_i^2\left(\frac{\overline{\gamma}}{1-\overline{\gamma}}\right)^2\frac{E[S^2]}{NE[S]^2}\Bigg].\end{aligned} \tag{B4}$$

Note that the fourth term on the right side, $\gamma_i^2\frac{1}{N^{1-\zeta}}\frac{Z}{\tilde{S}_i}\frac{E[S^2]}{E[S]^2}$, is the asymptotic equivalent of $\gamma_i^2 H_i$, as given in Lemma 1. Rearranging yields the result. □

Corollary 1. *As a corollary to Proposition 2, fixing S_i and S_j, the limiting behavior of the covariance of growth rates for firms i and j is*

$$Cov(g_i, g_j|S_i, S_j) \approx \sigma_\varepsilon^2\left[\gamma_i w_{i,j} + \gamma_j w_{j,i} + \frac{\overline{\gamma}}{N(1-\overline{\gamma})}(\gamma_i\tilde{S}_i + \gamma_j\tilde{S}_j) + \frac{\gamma_i\gamma_j}{N(1-\overline{\gamma})^2}\frac{E[S^2]}{E[S]^2}\right].$$

Proof. Immediate from equation (B3) and Lemma 2. □

B.4 Firm Size Follows Power Law: Proposition 3

Thus far we have assumed that the firm size distribution has finite variance, thus the slow volatility decay in Proposition 2 arises only due to network sparsity. Gabaix (2011) emphasizes that extreme right skewness of firm sizes can also slow down volatility decay in large economies. In the next result, we show that the firm-level network structure adds a mechanism to further slow down volatility decay beyond Gabaix's (2011) granularity mechanism, which depends on the power law behavior of the size distribution. In the absence of network effects, power law-based firm sizes would imply that firm variance decays at rate $N^{2-2/\eta}$. For any given rate of decay determined by the power law, network sparsity further lowers the decay rate by ζ.

Proposition. *Consider a sequence of economies indexed by the number of firms N. If firm sizes are distributed according to a power law with exponent $\eta \in (1, 2]$, then firm variance decays at rate $N^{(1-\zeta)(2-2/\eta)}$.*

Proof. Note that in equation (B4), the term that converges the slowest is the one involving firm i's Herfindahl, H_i (the i^{th} diagonal element of WW'). Therefore, H_i determines the rate of convergence for firm variance.

Recall the expression for a firm's Herfindahl:

$$H_i = \sum_k \frac{b_{i,k} S_k^2}{\left(\sum_l b_{i,l} S_l\right)^2} = \frac{N_i^{2/\eta} N_i^{-2/\eta} \sum_k b_{i,k} S_k^2}{N_i^2 \left(N_i^{-1} \sum_l b_{i,l} S_l\right)^2},$$

where $N_i = \sum_k B_{i,k}$ describes firm i's customer count.

From Gabaix (2011, Proposition 2) in the case $\eta \in (1, 2]$, we have that

$$N_i^{-2/\eta} \sum_k b_{i,k} S_k^2 \xrightarrow{d} u,$$

where u is a Levy-distributed random variable. Because $\eta > 1$, mean size is finite and so the LLN implies $N_i^{-1} \sum_l b_{i,l} S_l \xrightarrow{a.s.} E[S]$. Therefore,

$$N_i^{2(1-1/\eta)} H_i = \frac{N_i^{-2/\eta} \sum_k b_{i,k} S_k^2}{N_i^{-1} \sum_k b_{i,k} S_k} \xrightarrow{d} \frac{u}{E[S]}.$$

Finally, note that holding S_i fixed,

$$N_i \sim N p_i = N^{1-\zeta} \frac{\tilde{S}_i}{Z}$$

due to the fact that $E[N_i | S_i] = N p_i$. Combining these facts gives the result:

$$N^{(1-\zeta)(2-2/\eta)} H_i \xrightarrow{d} \frac{u}{E[S]} \left(\frac{Z}{\tilde{S}_i}\right)^{2-2/\eta}.$$

□

B.5 Comovement in Firm Volatilities: Proposition 4

We have referred to the volatility structure described in Proposition 2 as a factor model. The next result characterizes comovement among firms' volatilities when N is large. Because H_i determines the rate of convergence for firm variance, we may understand how firm variances comove in a large economy by studying the asymptotic covariance among H_i and H_j. As the number of firms grows, not only does the level of volatility decay, but so does its variance and covariance between the volatilities of different firms. Proposition 4 shows that comovement among firm variances decays at rate $N^{1+2(1-\zeta)}$. Intuitively, covariance among a pair of firms is lowest when both firms are large, since large firms have low exposure to overall size concentration.

Proposition. *Consider a sequence of economies indexed by the number of firms N, and fix S_i and S_j. If the firm size distribution has finite fourth moment (for example, if it is lognormal), then the covariance between*

H_i and H_j has limiting behavior described by

$$Cov(H_i, H_j|S_i, S_j) \approx \frac{1}{N^{1+2(1-\zeta)}} \frac{V(S^2)}{S_i S_j} \frac{Z^2}{E[S]^2}.$$

The covariance between $V(g_i|S_i)$ and $V(g_j|S_j)$ decays at the same rate.

Proof. Because H_i determines the rate of convergence for firm variance, we may understand how firm variances covary in a large economy by studying the asymptotic covariance among H_i and H_j.

Define $\hat{E}_i[S^2] = N_i^{-1} \sum_k b_{i,k} S_k^2$, where $N_i = \sum_k B_{i,k}$. We first characterize the asymptotic behavior of

$$\begin{aligned} Cov\left(\hat{E}_i[S^2], \hat{E}_j[S^2]|S_i, S_j\right) &= E\left[N_i^{-1} N_j^{-1} \sum_k \sum_l b_{i,k} b_{j,l} S_k^2 S_l^2 |S_i, S_j\right] \\ &\quad -E\left[N_i^{-1} \sum_k b_{i,k} S_k^2 |S_i, S_j\right] E\left[N_j^{-1} \sum_k b_{j,k} S_k^2 |S_i, S_j\right] \end{aligned}$$

Because size has finite fourth moment, the LLN implies that (holding S_i, S_j fixed)

$$N^{-(1-2\zeta)} \sum_k b_{i,k} b_{j,k} S_k^4 \overset{a.s.}{\to} E[S^4] \tilde{S}_i \tilde{S}_j / Z^2,$$

$$N^{-(1-2\zeta)} (N-1)^{-1} \sum_{k \neq l} b_{i,k} b_{j,l} S_k^2 S_l^2 \overset{a.s.}{\to} E[S^2]^2 \tilde{S}_i \tilde{S}_j / Z^2,$$

and

$$N^{-(1-\zeta)} \sum_k b_{i,k} S_k^2 \overset{a.s.}{\to} E[S^2] \tilde{S}_i / Z.$$

These, together with $N_i \sim N p_i$, imply that

$$\begin{aligned} E\left[N_i^{-1} N_j^{-1} \sum_k \sum_l b_{i,k} b_{j,l} S_k^2 S_l^2 |S_i, S_j\right] &\approx N^{1-2\zeta} \frac{E[S^4] \tilde{S}_i \tilde{S}_j}{N^2 p_i p_j Z^2} + (N-1) N^{1-2\zeta} + \frac{E[S^2]^2 \tilde{S}_i \tilde{S}_j}{N^2 p_i p_j Z^2} \\ &= \frac{1}{N} E[S^4] + \frac{N-1}{N} E[S^2]^2 \end{aligned}$$

and

$$E\left[N_i^{-1} \sum_k b_{i,k} S_k^2 |S_i, S_j\right] E\left[N_j^{-1} \sum_k b_{j,k} S_k^2 |S_i, S_j\right] \sim N^{2(1-\zeta)} \frac{E[S^2]^2 \tilde{S}_i \tilde{S}_j}{N^2 p_i p_j Z^2} = E[S^2]^2$$

so that

$$Cov\left(\hat{E}_i[S^2], \hat{E}_j[S^2]|S_i, S_j\right) \approx N^{-1} V(S^2).$$

Since $H_i = N_i^{-1} \hat{E}_i[S_k^2] / \left(N_i^{-1} \sum_k b_{i,k} S_k^2\right)$, we have

$$Cov(H_i, H_j|S_i, S_j) \approx \frac{1}{N^{2(1-\zeta)}} \frac{Z^2}{\tilde{S}_i \tilde{S}_j E[S]^4} Cov\left(\hat{E}_i[S^2], \hat{E}_j[S^2]|S_i, S_j\right),$$

which delivers the stated asymptotic approximation. □

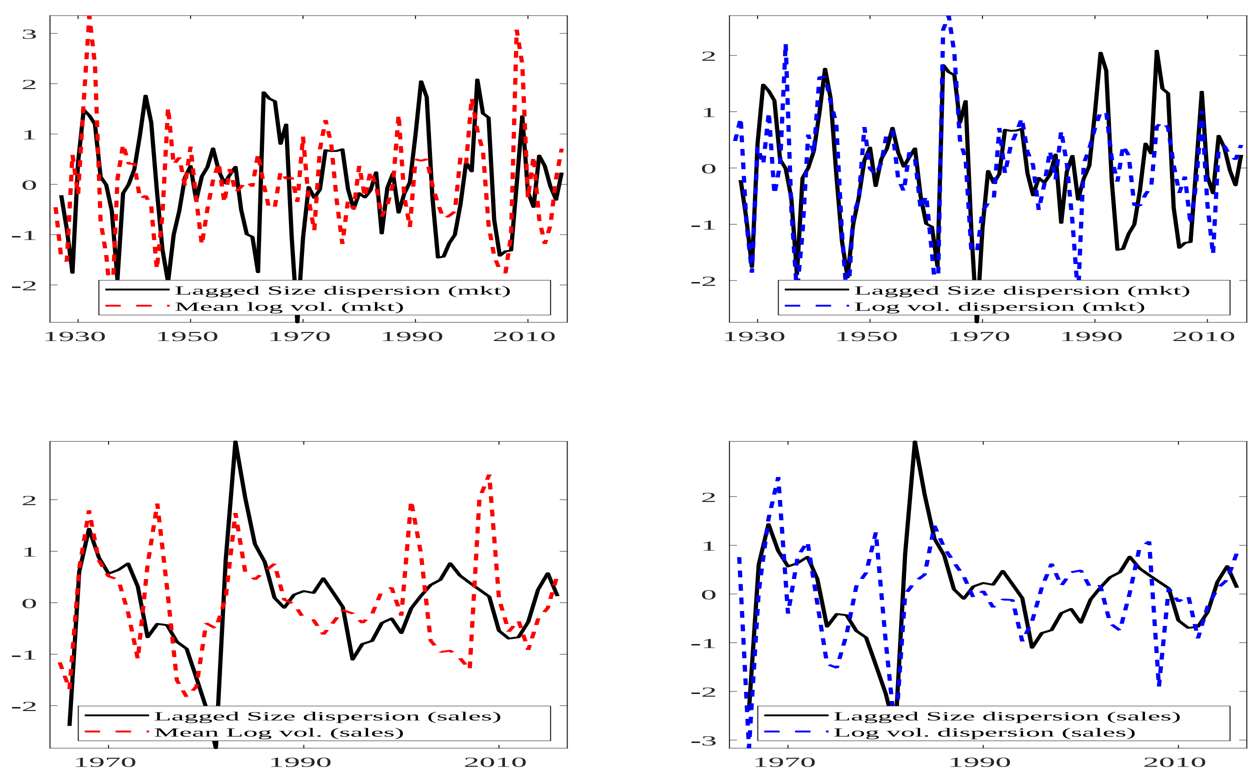
C Empirical Appendix

This appendix discusses several additional empirical results.

C.1 Frequency Decomposition

To study the trend and cycle in the size and volatility moments, we apply the Hodrick-Prescott filter with a smoothing parameter of 50. Figure C1 reports HP-detrended moments. The top-left panel shows firm size dispersion and mean firm variance based on market capitalization and return volatility. The top right panel reports size dispersion and variance dispersion, also based on market data. The bottom two panels are the counter-parts where size and variance are based on sales data. The correlations between the cyclical component in average variance and size dispersion are 25.8% for the market-based measure and 42.0% for the fundamentals-based measure. The correlations between the cyclical component in variance dispersion and size dispersion are 72.9% for the market-based measure and 54.1% for the sales-based measure. These results suggest that comovement between the dispersion in the firm size distribution and moments of the firm variance distribution occur both at cyclical and at low frequencies.

Figure C1: Detrended Size and Variance Moments



Notes: The figure plots HP-detrended time series moments of size and volatility distributions using smoothing parameter of 50.

C.2 Simulated Method of Moments Estimation Details

Objective Function Our estimation chooses the parameter vector Θ which minimizes the distance between the data moments, collected in the $1 \times K$ vector $\mathcal{G}$, and the corresponding moments obtained from a simulation of the network model, collected in $\hat{\mathcal{G}}(\Theta)$:

$$\mathcal{F} = \min_{\Theta} \mathbb{E}\left[\left(\mathcal{G} - \hat{\mathcal{G}}(\Theta)\right) \mathcal{W} \left(\mathcal{G} - \hat{\mathcal{G}}(\Theta)\right)'\right],$$

where the moment function $\hat{\mathcal{G}}(\Theta)$ is the average over the 100 draws, while $\mathcal{G}$ is a vector with the time-series average of the cross-sectional moments from the data.

Formally, these two moments functions are defined as

$$\begin{aligned} \mathcal{G} =& \frac{1}{T}\sum_{t=1}^{T} g_t \\ \hat{\mathcal{G}}(\Theta) =& \frac{1}{\mathcal{T}(T)} \sum_{t=1}^{\mathcal{T}(T)} \hat{g}_t\left(\Theta\right), \end{aligned}$$

where T is the number of time-series observations (33 for our sample), $\mathcal{T}(T)$ is the number of simulations (100 for our model). The function g_t is a $K \times 1$ vector of cross-sectional moments from the data for year t, while the function $\hat{g}_t\left(\Theta\right)$ is a $K \times 1$ vector of cross-sectional moments from the simulation t and parameter vector Θ. The weighting matrix $\mathcal{W}$ is the identity matrix. All moments are expressed in logs or are log differences so that they have the same order of magnitude.

Standard Error Calculation To derive the estimator asymptotics, we assume that $\frac{T}{\mathcal{T}(T)} \to \tau$ with τ being a finite positive number. Under sufficient regularity conditions, the asymptotic distribution of Θ is given by:

$$\sqrt{T}\left(\hat{\Theta} - \Theta_0\right) \longrightarrow \mathcal{N}\left(0, V_0\right),$$

where

$$\begin{aligned} V_0 =& (1+\tau)\left(G_0'\mathcal{W}G_0\right)^{-1} G_0'\mathcal{W}\Omega_0\mathcal{W}G_0\left(G_0'\mathcal{W}G_0\right)^{-1}, \\ G_0 =& \mathbb{E}\left[\frac{\partial}{\partial\Theta'}\hat{g}_t\left(\Theta\right)\Big|_{\Theta=\Theta_0}\right], \\ \Omega_0 =& \sum_{j=-\infty}^{\infty} \mathbb{E}\left[\left(g_t - \mathbb{E}(g_t)\right)\left(g_t - \mathbb{E}(g_t)\right)'\right]. \end{aligned}$$

To compute standard errors of the estimated parameters, we estimate G_0 as the numerical derivative of $\hat{\mathcal{G}}(\Theta)$ evaluated at the estimated parameters. We compute numerical derivatives by central difference approximation and by changing parameters by one percent. We estimate Ω_0 as the variance-covariance matrix of the moments from the data. For the estimated terms $\hat{G}_0$ and $\hat{\Omega}_0$ and assuming $\tau = \frac{33}{100}$, the asymptotic standard errors are given by the square root of the diagonal elements of the following variance covariance matrix:

$$\frac{1}{T}\hat{V}_0 = \left(1 + \frac{T}{\mathcal{T}(T)}\right)\frac{1}{T}\left(\hat{G}_0'\mathcal{W}\hat{G}_0\right)^{-1}\hat{G}_0'\mathcal{W}\hat{\Omega}_0\mathcal{W}\hat{G}_0\left(\hat{G}_0'\mathcal{W}\hat{G}_0\right)^{-1}.$$

Wald statistic To test whether some moments are collectively statistically equal to zero, we conduct the following hypothesis testing:

$$H_0 : h(\Theta_0) = 0 \qquad \text{v.s.} \qquad H_1 : h(\Theta_0) \neq 0,$$

where $h(\Theta)$ is a vector of moments to be tested. In our case, $h(\Theta)$ includes all the estimation moments.

Let $H(\Theta)$ be a matrix of the partial derivatives of $h(\Theta)$ with respect to Θ, then the Wald statistic is given by

$$W = Th(\hat{\Theta})' \left[H(\hat{\Theta})\hat{V}_0 H(\hat{\Theta})' \right]^{-1} h(\hat{\Theta}) \sim \chi^2_J$$

where J is the number of moments being tested.

C.3 Network Data and Truncation

Compustat Segment Data Our data for annual firm-level linkages come from Compustat. It includes the fraction of a firm's dollar sales to each of its major customers. Firms are required to supply customer information in accordance with Financial Accounting Standards Rule No. 131, in which a major customer is defined as any firm that is responsible for more than 10% of the reporting seller's revenue. Firms have discretion in reporting relationships with customers that account for less than 10% of their sales, and this is occasionally observed. In our data, 23% of firms report at least one customer that accounts for less that 10% of its sales. The Compustat data have been carefully linked to CRSP market equity data by Cohen and Frazzini (2008). This link allows us to associate information on firms' network connectivity with their market equity value and return volatility. We update the Cohen and Frazzini data to 2012. Cohen and Frazzini (2008) used the data to show that news about business partners does not immediately get reflected into stock prices. Atalay et al. (2011) and Herskovic (2018) also use Compustat sales linkage data to develop a model of customer-supplier networks.

BEA Data To gather additional evidence, we study a second data source which is at the industry level. Industry input-output data are from the Bureau of Economic Analysis (BEA). Because industry definitions vary quite dramatically over time, we focus on a set of 65 industries we can track consistently over time between 1997 and 2015. These are the input-output use tables, after redefinitions at producers' prices. In related work, Ahern and Harford (2014) use the network topography implied by the BEA industry data to show that the properties of these networks have a bearing on the incidence of cross-industry mergers. Herskovic (2018) also uses the BEA data to study the asset pricing effects of aggregate shocks to network moments.

Truncation To evaluate the effect of the 10% customer share truncation as well as the effect of selection, we conduct two exercises. First, we use the model. Since the model specifies the full network before imposing truncation, we can compute both truncated and untruncated moments in the model. If the model is the true data generating process, the difference between untruncated and truncated moments is the truncation bias. Second, we use the BEA data, which do not suffer from truncation. Again, we can artificially impose truncation on these data to study the effect.

Table C1 reports our findings for several key correlations of interest. Panel A reports truncated moments, including the Compustat moments reported in the main text, Panel B reports untruncated moments. The model and BEA data both show large truncation biases. This is especially true for moments that involve the number of customers (N^{out}) or the number of suppliers (N^{in}). The second row shows that $Corr(N_i^{out}, \log S_i)$ has a massive downward bias of 38 percentage points in the model and 65 percentage points in the BEA data. The fourth row shows that $Corr(N_i^{in}, \log S_i)$ has an upward bias of 35 percentage points in the model and 17 percentage points in the BEA data. Since Herfindahl moments are based on network weights squared, they are dominated by the larger weights. Large customers or suppliers are less likely to be missing or truncated. The Herfindahl correlations in rows 3 and 5 are nearly unbiased in the BEA data. The model shows a modest bias for row 3 but a sizeable bias for row 5.

Empirical Evidence for Model Assumptions The model makes two key assumptions on the network. First, large suppliers are more likely to have more customers: $Corr(S_i, N_i^{out}) > 0$. Second, the importance of a customer is higher the larger that customer is: $Corr(w_{ij,S_j} > 0$. These are assumptions on the full untruncated network of firms. The untruncated BEA data in column (4) of Table C1 provide empirical support for both assumptions in rows 1 and 2. The former correlation is 58% and the latter

Table C1: Truncation Analysis

	Panel A: Truncated			Panel B: Untruncated	
	(1)	(2)	(3)	(4)	(5)
	Compustat	BEA	Model	BEA	Model
$Corr\,(w_{i,j}, \log S_j)$	0.19	0.55	1.00	0.40	1.00
$Corr\,\left(N_i^{out}, \log S_i\right)$	-0.00	-0.07	0.31	0.58	0.69
$Corr\,\left(H_i^{out}, \log S_i\right)$	-0.28	-0.39	-0.88	-0.38	-0.67
$Corr\,\left(N_i^{in}, \log S_i\right)$	0.53	0.34	0.34	0.17	-0.01
$Corr\,\left(H_i^{in}, \log S_i\right)$	-0.47	0.10	-0.34	0.10	0.14

Notes: The table reports time-series averages of cross-sectional correlations between various features of customer-supplier networks with size and volatility. We report the following correlation in rows 1-5: correlation between the log network weight and log customer size, correlation between the out-degree (number of customers) and log supplier size, correlation between the out-Herfindahl and log supplier size, correlation between the in-degree (number of suppliers) and log supplier size, correlation between the in-Herfindahl and log supplier size, and correlation between the network weight and log size. Column (1) reports the correlation using Compustat sample for firms with network data available. Compustat sample is at firm-level and annual frequency for the period 1980-2012. Columns (2) and (4) are based on annual industry-level Bureau of Economic Analysis data for a set of 65 consistently measured industries for the period 1997-2015. Column (2) reports the correlation using BEA data assuming and artificially imposing 10% truncation on the network weights, which implies that we discard all customer-supplier pairs that represent less than 10% of supplier sales. Column (4) is not truncated. Columns (3) and (5) report the correlations for the truncated and untrucated model, respectively.

correlation is 40%. The former correlation is severly downward biased, as we discussed above. Adding either the model-implied bias or the BEA data-implied bias to the Compustat estimate raises the "untruncated" Compustat to a substantial positive number. The latter correlation is already positive in the Compustat data and suffers much less from bias. We conclude that the empirical evidence provides support for the two main assumptions in the model.

C.4 Model Estimation without Network or without Internal Diversification Effects

To highlight the importance of network effects for the model's ability to simultaneously match size, variance and network moments, we re-estimate two versions of the model. One version does not feature network effects, i.e., $\gamma = 0$. The other version does not feature internal diversification (ID), i.e., $\lambda = 0$. In each case, we allow the estimation to freely choose all other parameters. The estimated parameters and model-implied moments are in Tables C2 and C3. Both estimations make clear that the combination of network effects and internal diversification is necessary to successfully match the cross-sectional distribution of log variance.

Although the average log firm variance is similar to our benchmark, the estimated model without network effects generates almost no cross-sectional dispersion in variance. See column (3) of Table C3. It features fundamental shock volatility of 49.6%, which is significantly greater than our benchmark estimation of 29.8%; see column (2) of Table C2. If $\gamma > 0$, then the growth rate of a firm depends on the growth rate of other firms, which generates more dispersion in variance and amplifies firm-level volatility. In our estimation with $\gamma = 0$, a higher fundamental shock volatility is necessary to compensate for the lack of network effects and to match the average log variance. Although the estimation with $\gamma = 0$ matches the average variance, the only source of variance heterogeneity is the internal diversification channel embedded in $\sigma_{i,t}$. This channel is insufficient to match the cross-sectional dispersion in volatility. These results shows that the network effects are crucial to generate variance dispersion.

The model estimated without internal diversification, i.e., $\lambda = 0$, also falls short in terms of generating cross-sectional dispersion in variances. See column (4) of Table C3. It generates only half of the dispersion

Table C2: Parameter Estimates

		(1) Benchmark	(2) No Network	(3) No ID
γ	*Network propogation effect*	0.918	0.000	0.949
		(0.026)	–	(0.005)
σ_ε	*Fundamental shock volatility*	0.298	0.496	0.335
		(0.035)	(0.025)	(0.010)
λ	*Internal diversification effect*	0.131	0.090	0.000
		(0.025)	(3.674)	–
Z	*New connections*	0.003	0.002	0.004
		(0.001)	(0.000)	(0.000)
ψ	*Sensitivity of connection weight to customer size*	0.184	0.189	0.208
		(0.003)	(0.005)	(0.002)
μ_s	*Mean of initial log size distribution*	13.412	13.192	13.384
		(0.243)	(0.157)	(0.166)
σ_s	*Standard deviation of initial log size distribution*	5.104	5.291	5.152
		(0.102)	(0.035)	(0.052)
N_{pub}	*Number of public firms*	1332	1337	1388
		(47.7)	(40.8)	(46.3)

Notes: This table reports model parameters. Column 1 reports the estimated parameters from the SMM. Columns 2 and 3 report the estimated parameters assuming no network effects (i.e. $\gamma = 0$) and not internal diversification (i.e. $\lambda = 0$), respectively. Compared to the exercise in the main text, all model parameters are re-estimated in Columns 2 and 3. Estimates' standard errors are in parenthesis.

observed in the data—a standard deviation of 0.55 out of the 1.05 observed in the data. The estimation with $\lambda = 0$ has significantly stronger network effects with γ estimated at 0.949; see column (3) of Table C2. The estimation compensates the lack of internal diversification with stronger network effects. Although this is an improvement relative to the model without network effects, this estimation generates far less dispersion in variance than what we observe in the data.

Table C3: Size, Variance, and Network Moments

	(1) Data	(2) Model Bench	(3) Model No Network	(4) Model No ID
Panel A: Size and Volatility Distributions				
log Size				
Average	19.83	19.83	19.86	19.83
Standard Deviation	2.56	2.53	2.64	2.60
prc 50 - prc 10	2.91	2.37	2.49	2.56
prc 90 - prc 50	3.88	3.91	4.05	3.89
log Variance				
Average	−1.40	−1.40	−1.40	−1.40
Standard Deviation	1.05	0.86	0.00	0.55
prc 50 - prc 10	1.38	0.40	0.00	0.34
prc 90 - prc 50	1.31	1.69	0.01	1.07
Panel B: Network Moments				
N^{out}				
Median	1.00	1.12	1.12	1.00
log prc 50 - log prc 10	0.00	0.08	0.08	0.00
log prc 90 - log prc 50	0.83	1.02	1.02	1.10
H^{out}				
Median	0.05	0.18	0.19	0.24
log prc 50 - log prc 10	1.27	1.88	1.91	2.00
log prc 90 - log prc 50	1.85	1.71	1.66	1.44
N^{in}				
Median	1.00	1.00	1.00	1.00
log prc 50 - log prc 10	0.00	0.00	0.00	0.00
log prc 90 - log prc 50	1.69	1.09	1.10	1.09
H^{in}				
Median	1.00	1.00	1.00	1.00
log prc 50 - log prc 10	0.95	0.89	0.90	0.84
log prc 90 - log prc 50	0.00	0.00	0.00	0.00
Panel C: Cross-sectional Correlations				
Corr(log Size,log Var)	−0.64	−0.61	−0.66	−0.70
Corr(log N^{out},log Size)	−0.00	0.31	0.31	0.35
Corr(log H^{out},log Size)	−0.28	−0.88	−0.89	−0.89
Corr(log H^{out},log Var)	0.22	0.51	0.66	0.69
Corr(log N^{in},log Size)	0.53	0.34	0.36	0.36
Corr(log H^{in},log Size)	−0.47	−0.34	−0.37	−0.36
Corr(log H^{in},log Var)	0.13	0.19	0.20	0.22
Panel D: Firms with Network Data				
log number of public firms	7.15	6.95	6.95	6.98
Panel E: Goodness of Fit Statistics				
FVAL	–	2.88	6.50	3.45
Wald	–	24.76	26.80	27.33
p-value	–	0.64	0.53	0.50

Notes: This table reports different size, variance, and network moments both from the data and from our simulation. It takes the same structure as Table 3 in the main text. The only difference is that all model parameters are re-estimated in Columns 3 and 4.

C.5 Model with Aggregate Shocks

The model can be extended to allow for aggregate shocks. Firm growth in the benchmark model is given by:

$$g_i = \mu_g + \gamma_i \sum_{j=1}^{N} w_{i,j}\, g_j + \varepsilon_i.$$

In order to incorporate correlation in the ε shocks, we can assume one additional source of aggregate risk:

$$g_i = \mu_g + \gamma_i \sum_{j=1}^{N} w_{i,j}\, g_j + \nu_{agg} + \varepsilon_i.$$

where ν_{agg} has mean zero and variance σ^2_{agg}, and is independent from ε_i. In matrix notation, the growth rates become

$$\boldsymbol{g} = \Gamma \boldsymbol{W} \boldsymbol{g} + \boldsymbol{\varepsilon} + \boldsymbol{\nu} = \left(\boldsymbol{I} - \Gamma \boldsymbol{W}\right)^{-1} \left[\boldsymbol{\varepsilon} + \boldsymbol{\nu}\right],$$

and the variance-covariance matrix becomes

$$\boldsymbol{V}\left(\boldsymbol{g}\right) = \left(\boldsymbol{I} - \Gamma \boldsymbol{W}\right)^{-1} \left[\Sigma_{agg} + \Sigma_{idio}\right] \left(\boldsymbol{I} - \Gamma \boldsymbol{W}\right)^{-1},$$

where Σ_{agg} is an $N \times N$ matrix with σ^2_{agg} in each entry and Σ_{idio} is an $N \times N$ diagonal matrix with $\sigma^2_{\varepsilon,i}$ in the main diagonal.

We can numerically solve this version of the model with aggregate shocks and vary the parameter σ_{agg} from 0, which is our benchmark model, to 0.20, which delivers 20% volatility solely from aggregate shocks. These results are in Table C4. As aggregate volatility increases, average firm variance naturally increases, but firms also become more homogeneous in terms of volatility and cross-sectional dispersion in variance declines. As a result, there is also a lower correlation between log size and log variance. These features move the model with high aggregate risk farther from the data. The network moments as well as the size distribution remain unchanged in this exercise. The size distribution is driven by μ_s and σ_s, which are unchanged. The number of connections, who connects with whom, and the intensity of the network linkages all depend on the size distribution and other parameters that are kept at their benchmark values.

A second exercise which studies the effects of aggregate volatility while keeping average firm volatility the same as in the benchmark. The results are similar to the previous exercise, and are available from the authors upon request. The fit deteriorates in terms of the cross-sectional distribution of variances as well as the correlation between size and variance.

Table C4: Size, Variance, and Network Moments by Varying Aggregate Volatility

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Aggregate Volatility (σ_{agg})	–	0.00	0.01	0.02	0.03	0.04	0.05	0.10	0.20
Fundamental Shock Volatility (σ_{ε})	–	0.298	0.298	0.298	0.298	0.298	0.298	0.298	0.298
Panel A: Firm Size and Volatility Distribution									
log Size									
Average	19.83	19.83	19.83	19.83	19.83	19.83	19.83	19.83	19.83
Standard Deviation	2.56	2.53	2.53	2.53	2.53	2.53	2.53	2.53	2.53
prc 50 - prc 10	2.91	2.37	2.37	2.37	2.37	2.37	2.37	2.37	2.37
prc 90 - prc 50	3.88	3.91	3.91	3.91	3.91	3.91	3.91	3.91	3.91
log Variance									
Average	−1.40	−1.40	−1.33	−1.14	−0.89	−0.64	−0.39	0.62	1.86
Standard Deviation	1.05	0.86	0.82	0.74	0.65	0.56	0.48	0.25	0.10
prc 50 - prc 10	1.38	0.40	0.36	0.28	0.21	0.15	0.11	0.04	0.01
prc 90 - prc 50	1.31	1.69	1.63	1.47	1.26	1.07	0.89	0.39	0.12
Panel B: Network Moments									
N^{out}									
Median	1.00	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12
log prc 50 - log prc 10	0.00	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
log prc 90 - log prc 50	0.83	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
H^{out}									
Median	0.05	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18
log prc 50 - log prc 10	1.27	1.88	1.88	1.88	1.88	1.88	1.88	1.88	1.88
log prc 90 - log prc 50	1.85	1.71	1.71	1.71	1.71	1.71	1.71	1.71	1.71
N^{in}									
Median	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
log prc 50 - log prc 10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
log prc 90 - log prc 50	1.69	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09
H^{in}									
Median	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
log prc 50 - log prc 10	0.95	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89
log prc 90 - log prc 50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Panel C: Cross-sectional Correlations									
Corr(log Size,log Var)	−0.64	−0.61	−0.60	−0.58	−0.56	−0.53	−0.51	−0.45	−0.40
Corr(log N^{out},log Size)	−0.00	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31
Corr(log H^{out},log Size)	−0.28	−0.88	−0.88	−0.88	−0.88	−0.88	−0.88	−0.88	−0.88
Corr(log H^{out},log Var)	0.22	0.51	0.50	0.49	0.47	0.45	0.43	0.38	0.35
Corr(log N^{in},log Size)	0.53	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34
Corr(log H^{in},log Size)	−0.47	−0.34	−0.34	−0.34	−0.34	−0.34	−0.34	−0.34	−0.34
Corr(log H^{in},log Var)	0.13	0.19	0.19	0.18	0.17	0.17	0.16	0.14	0.12
Panel D: Firms with Network Data									
log number of public firms	7.15	6.95	6.95	6.95	6.95	6.95	6.95	6.95	6.95
Panel E: Goodness of Fit Statistics									
FVAL	–	2.88	2.93	3.11	3.50	4.09	4.82	9.05	16.50
Wald	–	24.76	–	–	–	–	–	–	–
p-value	–	0.64	–	–	–	–	–	–	–

Notes: This table reports different size, variance, and network moments both from the data and from our model simulation. It takes the same structure as Table 3 in the main text. Column 1 reports the data. Column 2 repeats the estimation results for the benchmark model. The benchmark model has zero aggregate shock volatility ($\sigma_{agg} = 0.00$). Columns 3–9 report simulations for a model with increasing aggregate volatility, keeping idiosyncratic shock volatility fixed at its benchmark value.

C.6 Composition Effect

The evidence in Davis et al. (2006) and Doidge et al. (2018) suggests that there may be a sample composition explanation whereby firm size dispersion and firm volatility first rise together (before 1997) and then fall (after 1997) as firms decide to go public earlier (before 1997) and again later (after 1997). We argue in this appendix that this alternative explanation cannot account for the facts on firm volatility.

First, we note that the dynamics of firm size dispersion and average firm volatility in Figure 1 are not obviously consistent with a composition explanation. Specifically, firm size dispersion among public firms rises pre-1997, but does not fall afterwards. It continues to fluctuate in the last 20 years of the sample. Firm volatility rises strongly until 1997, but then fluctuates afterwards.

Second, Table 1 provides empirical evidence against the sample composition hypothesis by studying sample splits based on firm size, on industry, on vintage (how long firms have been publicly traded), on historical period, and on exchange on which the firm is listed. It shows that the correlation between firm size dispersion and firm-level volatility (both average vol and the dispersion of vol) is found in all size, industry vintage, and exchange of listing groups. The correlations are lower for the non-NYSE group, for the second half of the sample, and for the Tech industry. If the sample composition story were driving the results, we would expect stronger results exactly in those three subsamples.

Third, we can use the model to ask whether the dynamic selection effect is quantitatively strong enough to explain the observed comovement between the firm size dispersion and average firm-level volatility and the dispersion of firm-level volatility. Simulation-based evidence from the model shows that the "changing composition of the public firms" explanation falls far short in explaining the firm volatility data, in the absence of network effects. The details of our simulation exercise are as follows:

- We define six sub-periods in the data in which there are noticeable changes in the firm size dispersion. These subperiods are chosen to capture the changing composition of public firms, and are listed in column (1) of Panel A of Table C5. They are 1981-1987, 1987-1992, 1992-2000, 2000-2004, 2004-2007, 2007-2012. The percentage change in the firm size dispersion from the start date to the end date of each subperiod in the data is listed in column (3) of Table C5. Column (2) reports the percentage change in the number of publicly listed firms in each of these subperiods. The subperiods alternate between a growing and shrinking number of public firms. When the number of public firms grows, the firm size dispersion tends to grow, and vice versa.

- Recall that the model contains both private and public firms. We call firms above a size threshold $\underline{S}$ public and those below it private. In the baseline model, this threshold is constant and chosen to match the observed number of public firms in the data. For this choice of $\underline{S}$, we match the 1981 level of firm size dispersion.

- We simulate six alternative economies, where we choose a different threshold $\underline{S}$ to match the observed firm size dispersion at the other turning point dates 1987, 1992, 2000, 2004, 2007, and 2012. Column (7) shows that the model indeed matches the observed changes in firm size dispersion in each subperiod exactly. The six alternative economies have the same parameters as the benchmark model except that we shut down the network effects ($\gamma_i = \gamma = 0$, $\forall i$).

- Column (6) reports the change in the number of public firms in the model over each period, induced by the different listing thresholds. The change in the number of firms over these sub-periods is substantial. But so are the changes in the number of public firms in the data, listed in column (2). The model does a reasonable job matching these numbers, even though they were not targeted in the exercise. Put differently, the composition hypothesis can generate substantial fluctuations in the firm size dispersion.

- In the model, lowering the listing threshold results in smaller firms being public. Because of the internal diversification force in the model, these firms have higher volatility. Column (8) reports the average firm volatility. It rises when more firms go public and falls when fewer firms are public. Column (9) reports the dispersion in firm volatility.

- The main finding from this exercise is that the quantitative changes in average firm volatility and volatility dispersion that are induced by changes in the firm size dispersion (which itself is of the

Table C5: Composition Effect Among Public Firms

	Data				Model			
	Public Firms	Public Firms with Network Data			Public Firms	Public Firms with Network Data		
Year (1)	# of Firms (2)	Size Disp (3)	Avg Vol (4)	Vol Disp (5)	# of Firms (6)	Size Disp (7)	Avg Vol (8)	Vol Disp (9)
Panel A: No network effects								
1987-1981	0.28	0.21	0.18	0.24	0.27	0.21	0.01	0.02
1992-1987	−0.02	−0.11	0.27	0.14	−0.14	−0.11	−0.01	−0.01
2000-1992	0.16	0.14	0.59	−0.24	0.16	0.14	0.01	0.02
2004-2000	−0.19	−0.28	−1.20	−0.06	−0.37	−0.28	−0.02	−0.03
2007-2004	0.04	0.02	1.34	−0.17	0.03	0.02	0.00	0.00
2012-2007	−0.06	0.01	−1.69	0.13	0.02	0.01	0.00	0.00
Panel B: Different internal diversification function								
1987-1981	0.28	0.21	0.18	0.24	0.28	0.21	0.16	0.04
1992-1987	−0.02	−0.11	0.27	0.14	−0.14	−0.11	−0.09	−0.02
2000-1992	0.16	0.14	0.59	−0.24	0.16	0.14	0.10	0.03
2004-2000	−0.19	−0.28	−1.20	−0.06	−0.38	−0.28	−0.22	−0.05
2007-2004	0.04	0.02	1.34	−0.17	0.03	0.02	0.01	0.00
2012-2007	−0.06	0.01	−1.69	0.13	0.02	0.01	0.01	0.00

Notes: Columns (1) to (5) report data moments, while Columns (6) to (9) report their model counterparts. All moments are percentage changes over the period listed in the first column. We report three moments both from the data and model: size dispersion, average volatility, and volatility dispersion. We compute them using only public firms with network data available. Panel A is for our benchmark model. Panel B uses an alternative internal diversification function.

identical magnitude as in the data) are small. In particular, they are much smaller than the observed changes in average firm volatility and volatility dispersion reported in columns (4) and (5).

- This shows that the composition hypothesis is quantitatively far too weak to produce the observed fluctuations in the firm volatility distribution.

One downside of the baseline model with $\gamma = 0$ is that it generates low levels of volatility dispersion. To give the composition hypothesis a better shot, we redo the simulation exercise with a different internal diversification function, taken from Stanley et al. (1996), that generates about the right baseline level of volatility dispersion (absent network effects). Panel B of Table C5 presents the results. The changes in average firm volatility in column (8) and the dispersion in firm volatility in column (9) are similar to those in Panel A, and far from the data.

We conclude that compositional changes to the distribution of public firms, of the kind documented by Davis et al. for the late 1990s and by Karolyi et al. for the period since then, cannot explain the quantitative comovement of firm size dispersion, average firm volatility, and volatility dispersion, absent network effects. We have included a discussion of the composition effect

C.7 Alternative Internal Diversification Function

This appendix considers an alternative functional form for the internal diversification effect, due to Stanley et al. (1996). Table C7 compares the data (column 1), our benchmark model with $\sigma_{\varepsilon,i} = \sigma_\varepsilon + \lambda \log\left(1 + \frac{S_{median}}{S_i}\right)$ (column 2), the proposed alternative specification for firm volatility, $\sigma_{\varepsilon,i} = 6.66 \times S_i^{-0.15}$, with the parameters taken directly from Stanley et al. (Alternative 1, column 3), and the same specification $\sigma_{\varepsilon,i} = \kappa \times S_i^{\alpha}$, where κ and α are estimated on our data (Alternative 2, column 4). For the specification where we estimate κ and α, we also re-estimate all other parameters of the model to give that alternative model the best shot. The estimated parameters are reported in column 3 of Table C6. When we fix κ and α in Alternative 1, we also fix the other parameters of the model (column 2 of Table C6).

Table C6: Parameter Estimates with Alternative Internal Diversification Function

		(1) Bench	(2) Alternative 1	(3) Alternative 2
γ	*Network propogation effect*	0.918	0.918	0.913
		(0.026)	–	(0.011)
σ_ε	*Fundamental shock volatility*	0.298	–	–
		(0.035)	–	–
λ	*Internal diversification effect*	0.131	–	–
		(0.025)	–	–
Z	*New connections*	0.003	0.003	0.003
		(0.001)	–	(0.000)
ψ	*Sensitivity of connection weight to customer size*	0.184	0.184	0.179
		(0.003)	–	(0.011)
μ_s	*Mean of initial log size distribution*	13.412	13.412	13.456
		(0.243)	–	(0.144)
σ_s	*Standard deviation of initial log size distribution*	5.104	5.104	5.089
		(0.102)	–	(0.086)
N_{pub}	*Number of public firms*	1332	1332	1321
		(47.7)	–	(27.0)
κ		–	6.660	2.249
		–	–	(0.300)
α		–	0.150	0.097
		–	–	(0.005)

Notes: This table reports model parameters. Column 1 reports the estimated parameters from the SMM in the benchmark model. Columns 2 holds the parameters constant at their benchmark values but uses an alternative internal diversification function due to Stanley et al. (1996). It takes the values for the parameters κ and α of that function directly from Stanley et al. (1996). Column 3 re-estimates all parameters, including those of the internal diversification function.

Turning to the results in Table C7, two main observations stand out. First, the alternative functional form for the internal diversification effect explains the data really poorly when the parameters are taken from the Nature paper. Average firm variance in column 3 is far too high and there is too much dispersion in firm variance. Since the size distribution is not affected, neither are the network moments. The poor fit results in a large distance between model and data (FVAL of 3.76). Likely this is because the Stanley et al. paper fits data on total firm volatility rather than fundamental volatility. The additional amplification from network effects in our model leads to excessive firm variance.

Second, the alternative model fares better when the parameters are re-estimated (in column 4 of Table C7). Mean firm variance is somewhat too high, unlike the benchmark parameterization. The 90-50 percentile difference in firm variance also deteriorates. Finally, the correlation between firm size and firm variance deteriorates meaningfully compared to the benchmark. However, both the standard deviation of firm variance and the 50-10 percentile improve. Overall the distance between this model and the data is 2.29, which is slightly lower than the benchmark. The estimated network strength $\hat{\gamma}$ is nearly unaffected for the alternative functional form for the diversification effect. In sum, our results are robust to using this alternative internal diversification function.

Table C7: Size, Variance, and Network Moments with Alternative Internal Diversification Function

	Data	Benchmark Model	Alt ID No Reestimation	Alt ID Reestimation
	(1)	(2)	(3)	(4)
Panel A: Firm Size and Volatility Distribution				
log Size				
Average	19.83	19.83	19.83	19.86
Standard Deviation	2.56	2.53	2.53	2.52
prc 50 - prc 10	2.91	2.37	2.37	2.33
prc 90 - prc 50	3.88	3.91	3.91	3.90
log Variance				
Average	−1.40	−1.40	−0.38	−1.31
Standard Deviation	1.05	0.86	1.18	1.01
prc 50 - prc 10	1.38	0.40	0.88	0.93
prc 90 - prc 50	1.31	1.69	2.17	1.76
Panel B: Network Moments				
N^{out}				
Median	1.00	1.12	1.12	1.14
log prc 50 - log prc 10	0.00	0.08	0.08	0.10
log prc 90 - log prc 50	0.83	1.02	1.02	1.00
H^{out}				
Median	0.05	0.18	0.18	0.18
log prc 50 - log prc 10	1.27	1.88	1.88	1.87
log prc 90 - log prc 50	1.85	1.71	1.71	1.74
N^{in}				
Median	1.00	1.00	1.00	1.00
log prc 50 - log prc 10	0.00	0.00	0.00	0.00
log prc 90 - log prc 50	1.69	1.09	1.09	1.09
H^{in}				
Median	1.00	1.00	1.00	1.00
log prc 50 - log prc 10	0.95	0.89	0.89	0.88
log prc 90 - log prc 50	0.00	0.00	0.00	0.00
Panel C: Cross-sectional Correlations				
Corr(log Size,log Var)	−0.64	−0.61	−0.69	−0.78
Corr(log N^{out},log Size)	−0.00	0.31	0.31	0.31
Corr(log H^{out},log Size)	−0.28	−0.88	−0.88	−0.88
Corr(log H^{out},log Var)	0.22	0.51	0.46	0.61
Corr(log N^{in},log Size)	0.53	0.34	0.34	0.33
Corr(log H^{in},log Size)	−0.47	−0.34	−0.34	−0.33
Corr(log H^{in},log Var)	0.13	0.19	0.22	0.25
Panel D: Firms with Network Data				
log number of public firms	7.15	6.95	6.95	6.94
Panel E: Goodness of Fit Statistics				
FVAL	–	2.88	3.76	2.29
Wald	–	24.76	–	25.13
p-value	–	0.64	–	0.62

Notes: This table reports different size, variance, and network moments both from the data and from our simulation. The structure is the same as Table 3 in the main text. Column 1 is for the data. Column 2 is for the benchmark model estimated by SMM. Columns 3 and 4 are for the alternative internal diversification (ID) function based on Stanley et al. (1996). Column 3 takes the values for the ID function parameters κ and α of that function directly from Stanley et al. (1996). Column 4 re-estimates all parameters, including those of the ID function.