GRANULARITY, NETWORK ASYMMETRY, AND AGGREGATE VOLATILITY

JESSE TWEEDLE

Abstract. I evaluate two competing theories for microfoundations of aggregate fluctuations. The network hypothesis suggests industry-level shocks propagate across the input-output (IO) network of the economy, resulting in aggregate fluctuations. The granular hypothesis suggests idiosyncratic shocks to very large firms result in aggregate fluctuations. My main contribution is to connect the two aggregate fluctuation hypotheses for the first time and theoretically and empirically quantify the contributions of each to volatility.

The network hypothesis depends crucially on certain plants being essential suppliers to the economy. However, they may be essential suppliers due to their productivity and not any underlying input-output requirements, which means productivity may be the source of both the granularity and network hypotheses. To disentangle these relationships, I document a plant-plant input-output network, then develop a model in which productivity and the exogenous IO network can vary independently and both combine to determine the observed IO network. Finally, I calibrate the model to uncover the underlying IO network and then investigate the empirical relationship between the uncovered IO network and aggregate volatility.

I find (i) the observed plant-plant IO network is very asymmetric, (ii) productivity doesn't vary enough to explain the observed IO network, (iii) and therefore the true underlying IO network explains the majority of the plant size distribution and 34% of aggregate volatility.

1. Introduction

Why are microeconomic shocks sources of aggregate volatility and how do they propagate across the economy? Are shocks transmitted across input-output linkages or not? The answer would seem to depend on two competing theories of microfoundations of aggregate fluctuations: the granularity hypothesis of Gabaix [18] and the unbalanced network hypothesis of Acemoglu, et al. [2]. If granular plants are the sources of aggregate fluctuations, then plants should account for the majority of fluctuations, independent of the input-output (IO) network. However, at the plant level, the two theories are intertwined—plants may be essential suppliers in the IO network because of their granularity. I aim to document and explore the relationship between granularity and IO networks and how they contribute to aggregate volatility.

There are several dimensions of interaction between the two theories. First, the conceptual difference between them depends on the reason for the shape of the individual size distribution, be they plants or sectors. The granularity hypothesis is agnostic about the underlying cause of the shape of the distribution, and one typically assumes that a fat-tailed productivity distribution is responsible (e.g., in a standard Melitz model). The network hypothesis, on the other hand, claims the fat-tailed size distribution is caused by an exogenous asymmetry in the IO network, so that certain sectors are very large because they supply an inordinately large portion of the economy. The insight I add is to let productivity and the IO network[1] vary independently at the plant level, and explain, with theory and data, how the size distribution is shaped by these two primary forces, and how that affects aggregate volatility.[2]

Second, plants and sectors are typically treated very differently in economic models and data. How does that affect the argument for microfoundations of aggregate fluctuations? Both the granular and network hypotheses require the number of microeconomic units to be very large (otherwise there would be no micro to provide foundations for). However, in many models and data, sectors are the only ones with IO networks, while plants are assumed to have differences in productivity but no variation in IO characteristics. This presents a problem for the unbalanced network hypothesis: plant networks can't be sources of aggregate fluctuations if there is no variation within sectors. On the other hand, sector-level models take expenditure shares as exogenous, implying no productivity or heterogeneity can affect the network itself.

In theory, there is no difference between sectors and plants—it is easy to include plant specific IO characteristics and reformulate a model with sectors into a model with plants with the same behaviour. This makes the distinction between plants and sectors an empirical one, in the sense that the deepest level we can study IO networks is the level at which data on inputs and outputs and transactions are recorded, which are typically industries. Put another way, the only way to explore the relationship between unbalanced IO networks and granularity at the plant level is to have data on inputs and outputs and trading relationships at the plant level. I combine theory and data on granularity and an asymmetric IO network at the plant level, and show how this translates into both hypotheses coexisting at the most disaggregated level in the economy.

Putting the pieces together, I compare the properties of each plant, granularity and the IO network, and explore empirically how these factors affect aggregate volatility.[3]

I use four crucial pieces of data and theory to explore these relationships. First, plant level data on commodity inputs and outputs to establish the unbalanced IO network at a much more disaggregated level than previous studies.[4] Second, I use theory to identify conditions under which productivity and the underlying IO network[5] determine the endogenous observed IO network. Third, I calibrate the model to match observed plant characteristics and uncover the underlying IO network. Fourth,

[1]And geography and trade costs.
[2]And geography and trade costs.
[3]And covariance centrality.
[4]And geography and trade costs and STF.
[5]And geography and trade costs.

I measure each plant's importance to the IO network measure and link it to its contribution to aggregate volatility.[6]

I[7] use the Annual Survey of Manufactures (ASM), a long-term establishment-level survey in Canada, covering 99% of output and value-added. The ASM comes with detailed data on commodity inputs and outputs for each plant, crucial to exploring the disaggregated IO network. Using these data, I construct plant-to-plant direct-requirements tables, in the tradition of industry-level input-output accounting at statistical agencies. The ASM has the relevant data on other plant-level characteristics, including industry, location, sales, value added, and employment.

To disentangle the two[8] forces shaping the observed unbalanced IO network—is a plant a central supplier beacuse it supplies an essential product or because it is so productive that every plant substitutes toward it? The endogenous IO network depends on those two primary factors, the productivities of individual plants and the unobserved plant-to-plant supply linkages. I extend the standard Cobb-Douglas input-output model to accomodate productivity differences and substitutability across plants (both within and across sector boundaries), which induces productive plants to become more central suppliers. The key to differentiating between productivity and network asymmetry is the behaviour of the tails of each distribution, and how they affect the tail of the size distribution. Recall that the argument for microfoundations of aggregate volatility depends on the fat tail of the size distribution as the number of plants in the economy gets very large. If productivity and network centrality are both distributed with power laws, I show that, as the number of plants gets large enough to apply the microfoundation argument, the fatter of the two tails will determine the tail of the size distribution. However, as in many applications of power laws, the empirics are more complex and both factors will matter.

[Way more here on model, geography, regions, etc? And another discussion of covariance matrix estimation and calculation of aggregate volatility, eigenvectors etc.]

Research on idiosyncratic shocks and aggregate volatility restarted in earnest when Gabaix [18] and Acemoglu et al. [2] revived the debate between Horvath [22, 23] and Dupor [13] on whether idiosyncratic shocks average out in aggregate. Gabaix [18] proposes that the largest, granular firms are so big that their idiosyncratic shocks do not average out at the aggregate level. Acemoglu et al. [2] suggest the reason for non-diversification of idiosyncratic shocks is an asymmetric input-output network, in which a shock to a sector that supplies a large number of other sectors propagates through the economy and generates aggregate fluctuations. I add an understanding of the connections between the two theories at an empirical level, specifically showing the complementarity between granularity and production networks and how idiosyncratic plant-level shocks rely on plant-level IO variation within industries. What really differentiates this work is that I explore the determinants of the observed IO network, whereas previous research assumes the network is exogenous.

[6]Centrality bit here.

[7]Way more here, about ASM and STF.

[8]three

The most direct predecessors of this paper are empirical studies of aggregate fluctuations. Starting with Shea [33], and continuing most recently with Di Giovanni, Levchenko and Méjean [11], Foerster, Sarte and Watson [15], Acemoglu et al. [1]. Foerster, Sarte and Watson [15] combined factor analysis with structural model of industrial production in the US, finding common shocks are the source of the majority of volatility, with idiosyncratic shocks becoming more important after the great moderation. Di Giovanni, Levchenko and Méjean [11] study fluctuations of French firm sales to individual countries and find idiosyncratic fluctuations account for the majority of aggregate volatility, and that much of it comes from covariances between firms. They suggest the firm covariances are due to firm-to-firm linkages, although they only observe industry-level IO data. In contrast to both papers, I use plant-level IO data to establish the determinants of plant covariances, using deeper levels of disaggregation to examine both covariances (firm level to plant level) and IO (industry level to plant level). As well, I study the determinants of the network itself, something taken as exogenous in previous empirical work.

Any study of granularity builds on a body of work on the determinants of firm size and the characteristics of its distribution, from specific applications in international trade [12, 9, 10], or studies on general characteristics and theories of the size distribution itself [30]. I add an endogenous network perspective to this research and use it to further explore the determinants of the plant size distribution and the sources of granularity. My work also fits naturally with Hottman et al. [24], who use detailed price and sales data on consumer non-durables to suggest 'firm appeal' is the dominant source of firm heterogeneity, accounting for $50 - 70\%$ of firm size. Holmes and Stevens [21] also provide evidence that demand characteristics are the main source of plant heterogeneity, in contrast to standard Melitz applications. In my case, the IO requirements of downstream plants translate into a dominant source of plant appeal, and therefore are a large determinant of plant size.

My argument is also related to recent work on customer-supplier relationships, especially Barrot and Sauvagnat [5], who study the disruption of production networks after natural disasters. In addition, research on customer-supplier relationships in Japan [6, 7] and the US [4] suggests larger plants have different input-output characteristics than smaller plants. Typically, customer-supplier relationship data only includes an indicator for whether a firm supplies another firm, not the strength of the relationship or the commodities made and used. In my case, I have measures of the strength of the interaction between plants. To this research, I add a characterization of the manufacturing IO network in Canada, focusing on differences across plants within industries.

These papers are also part of a recent wave of interest in the formation and effects of social and economic networks. Carvalho and Voigtlander [8], Oberfield [32] and Jones [28] each apply these ideas specifically to production and growth, whereas other works focus on volatility and contagion in financial markets, such as Acemoglu et al. [3], Golub, Elliot, Jackson [14]. Other applications and background on several network measures used in this paper can be found in Jackson [26].

In Section 2, I present the plant-level volatility and IO data. I document an unbalanced IO network at a disaggregated level, with a few plants acting as central suppliers to the network.In Section 3, I present a simple, but necessary, extension to the IO model used in Acemoglu et al. [2] to allow plant IO characteristics to vary independently of productivity. The asymmetry of the network and the productivity distribution combine to determine plant sizes, which is the key to evaluating the granularity of the economy and its effect on aggregate volatility. In Section 4, I outline the asymptotic theory that gives a knife-edge prediction of the cause of granularity: the thicker tail of the distributions of productivity and network asymmetry are the sole cause of skewed firm size distribution.

In Section 5, I calibrate the model to uncover the underlying IO network from the endogenous, observed IO network and evaluate the competing theories of aggregate fluctuations. Previewing the main calibration results, the productivity distribution is not heterogenous enough to account for the asymmetry in the observed IO network. The majority of the observed IO network is due to the underlying IO network, consistent with results in Holmes and Stevens [21] that challenge the reliance of the plant size distribution on productivity alone.

In Section 6, I provide direct empirical support for the importance of both productivity and network asymmetry for determining granularity and aggregate volatility. An 10% increase in network centrality is associated with a 2.66% increase in plant size, and a 10% increase in productivity is associated with an 8% increase in plant size. Eliminating any asymmetry in the plant-plant IO network reduces aggregate volatility by 34%.

Section 7 concludes, and several Appendices follow, giving details on theory, measurement and development of the plant-plant IO network, and other sundry details.

2. Data

2.1. **Overview.** The data is from the Annual Survey of Manufactures (ASM), which covers 99% of industrial output in Canada. It is a long-running annual panel of manufacturing establishments, including information on all relevant industrial characteristics, including sales, value added, total intermediate inputs, location, employment, industry, and parent firm. I analyze the period from 1973 to 1999, covering several volatile periods in Canadian manufacturing, including recessions and recoveries in the 1980s and the early 1990s, as well as oil shocks in the 1970s. The average value-added of manufacturing over this period was approximately 50%, declining from 60% in 1973 to 40% in 1999. Aggregate volatility, measured by the standard deviation of the aggregate growth rate of total output, over this period was approximately 7% in manufacturing, slightly higher than the overall for Canada during the same period, around 5%.

The long-form survey, intended to provide additional detail for the biggest plants, covers approximately 18,000 plants and 92% of manufacturing output. It provides detailed data on commodity level inputs and outputs for each plant at the 9-digit Standard Classification of Goods (SCG) level. This commodity survey, which serves as part of the basis for Canada's input-output tables, provides the essential plant-level input-output data that allows me to investigate the plant-specific IO network and its effect on aggregate volatility. Just like industry-commodity level make-and-use tables in the national accounts, each plant can consume multiple inputs and produce multiple outputs, and each commodity may be produced or consumed by multiple plants, even across industries. Not surprisingly, there is considerable heterogeneity in input-output statistics at the plant-level, both within and across industries, and this heterogeneity plays a big part in the mechanisms outlined later in the paper.

2.2. **Sales and growth and volatility statistics.** In the Canadian manufacturing sector, a few industries play outsized roles in output, employment and value-added. Transportation equipment production alone accounted for 21.5% of total manufacturing output in Canada in 1997, suggesting that a shock to that industry will have significant effects on the economy as a whole. However, the top *ten* plants in that industry account for the vast majority of its output. Although industries look granular, plants within industries appear to have the same asymmetry and may themselves contribute to aggregate volatility.

Across the economy, the industry-level herfindahl is .159, showing the intense concentration of output in a few industries.[9] Within industries, output concentration at the plant level is even greater, with a mean herfindahl of .264 and a standard deviation of .144. Heterogeneity at both the plant and industry level is clear, and they contribute to an overall plant-level herfindahl of .0566. The potential for granularity is strong at both the industry and within-industry levels.

Table **??** also displays summary statistics of growth rates by plant, year and industry. The mean growth rate across all plants and years is .074 with a standard deviation of .329. The average aggregate growth rate is .075 with a standard deviation of .066. The mean industry growth rate is .067 with a standard deviation of .107. At more disaggregated levels, individual volatility is increasing. However, it is decreasing much slower than $1/\sqrt{N}$, which is a sign that shocks are not averaging out at either the industry or plant level. However, observed plant growth rates may be measuring aggregate shocks themselves and not idiosyncratic shocks.

2.3. **Input-output statistics.** Given two randomly chosen plants in Canadian manufacturing, how are they connected through their production processes? I attack

[9]The plant level herfindahl is year t is $h_t = \sqrt{\sum_{I \in \mathcal{I}} \sum_{i \in I} w_{it}^2}$. The industry level herfindahl in year t is $h_{\mathcal{I}t} = \sqrt{\sum_{I \in \mathcal{I}} w_{It}^2}$. I report the averages $\bar{h} = (1/T) \sum_{t=1}^{T} h_t$ and $\bar{h}_{\mathcal{I}} = (1/T) \sum_{t=1}^{T} h_{\mathcal{I}t}$. Within-industry herfindahls are defined similarly using the weights $w_{i/It}$. There is a possibility that survey design generates a mechanical relationship between herfindahl and weights, in which small industries (in terms of total output) have fewer plants in the survey, possibly because they are concentrated geographically and thus not many plants are required to estimate the total industry output for a province. In this case, there will be a negative relationship between the within-industry herfindahls and industry weights, because there are less firms to reduce the weight of sampled plants. To account for this, I also calculated the herfindahls using only the top 20 plants in each industry. The results are very similar, with $\bar{h} = .056$, $\bar{h}_{\mathcal{I}} = .187$, and $(1/N_{\mathcal{I}}) \sum_{I \in \mathcal{I}} \bar{h}_I = .270$.

the question in two ways: first, the industry-by-industry or plant-by-plant direct-requirements table, the empirical version of a matrix of input shares. This is one of the main IO measures developed by statistical agencies to measure direct and indirect input-output connections between industries, and is the basis of the empirical analysis in Acemoglu et al. [2]. Second, I create a new measure of input-output connections that measures the correlation of inputs, outputs, and direct linkages between two plants. The two measures complement each other and each have different advantages in different facets of the analysis.

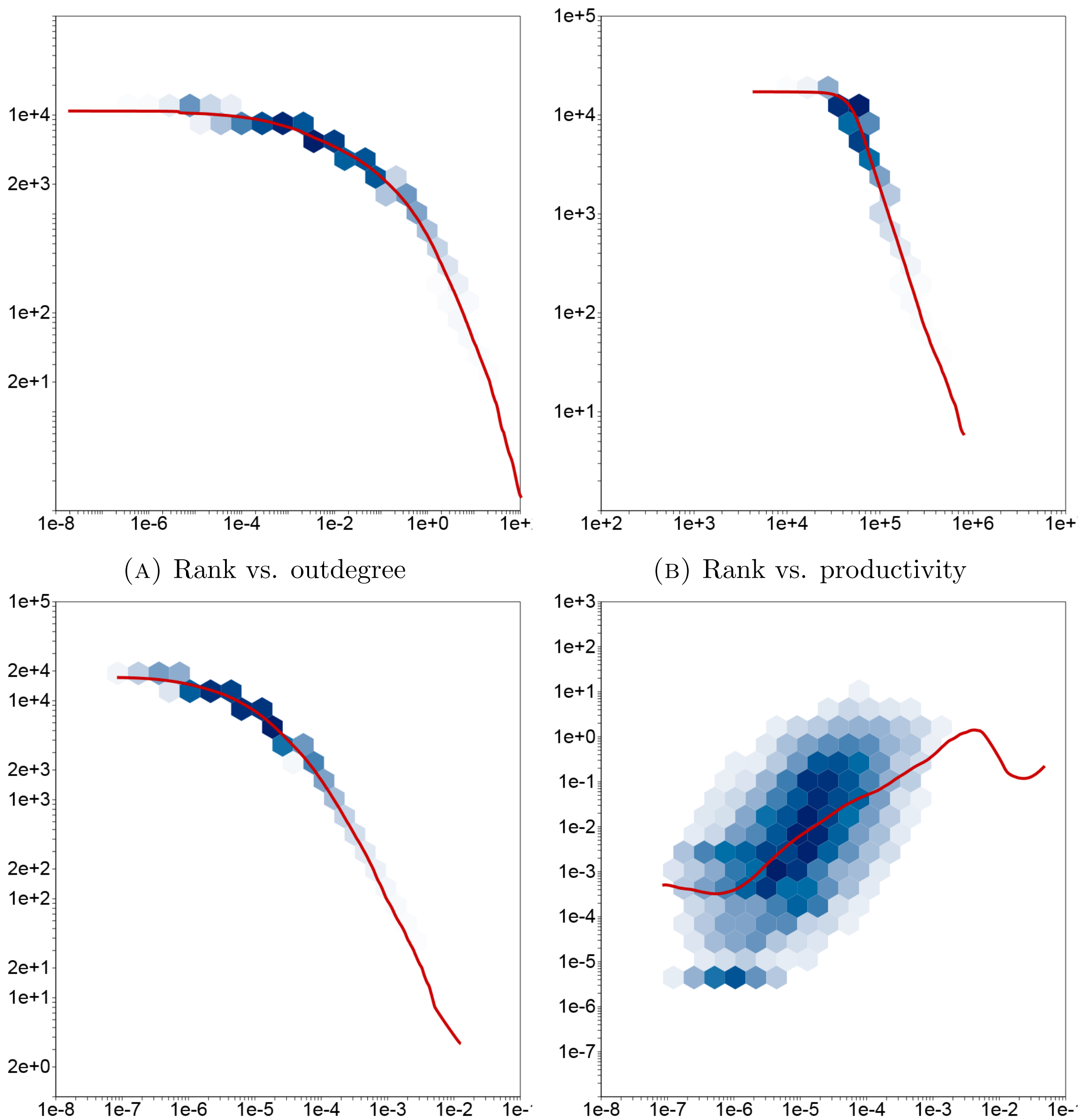
The main IO measure is the share of expenditure on inputs of plant j, g_{ij}. The goal of the empirical analysis is to construct the plant-level version of g_{ij} using the same method as the IO tables in statistical agencies. The ultimate result is a matrix , G of plant-plant input shares, where a typical element is g_{ij} and satisfies $\sum_{i\in N} g_{ij} = 1-\beta_i$, where β_i is the value added share of output of plant i (see Appendix 9 for the full derivation). If you arrange the plants in G by industry, it can be decomposed into industry-by-industry blocks. In addition, the weighted outdegree, d_i is a measure of the importance of plant i to the entire economy, measured by the sum of input shares across all other plants:

$$d_i \equiv \sum_{j\in J} g_{ji} \tag{1}$$

The summary statistics of the IO measures are displayed in Appendix **??**. The sparseness of the plant-level IO matrix is clear: of the 324 million possible connections between plants, less than 1% share any kind of link (in the sense of having strictly positive values of G), compared to 11% of industry pairs having connections. More disaggregated microeconomic levels become more and more sparse, suggesting there is significant heterogeneity in IO within sectors, because if every plant within an industry pair had IO characteristics that matched the aggregate level, the plant level connections would match the industry level. Instead, the IO network becomes more sparse as it becomes more disaggregated, so plants are only connected to certain other plants in another industry, and not all of them.

As shown by Acemoglu et al. [2], the increasing sparseness of the network does not necessarily mean idiosyncratic shocks are not important, as long as the network retains the asymmetric properties that represent the importance of a single plant or industry to the whole economy. In this case, the within-industry heterogeneity in IO suggests that there are important plants within important industries that are the source of aggregate fluctuations in the economy. The remaining question is whether the asymmetry in the observed IO network is truly due to asymmetry in the exogenous underlying IO network.

To illustrate the importance of asymmetry, Figure 1a plots the rank of d_i versus d_i itself, on a log-log scale (for the plants with a strictly positive d_i). The asymmetry and heavy-tailed distribution of outdegree is apparent—there are a few plants are very important to the network, and are significant suppliers of a large number of other plants. Figures 1b and 1c show similar relationships for labour productivity and plant size, respectively, suggesting all three have power law tails and that productivity and outdegree both determine the plant size distribution.



(A) Rank vs. outdegree (B) Rank vs. productivity

(C) Rank vs. size (D) Size vs. degree

FIGURE 1. Hexbin rank plots for plant characteristics, all with linear (power law) right tails. In panel (D), we can see that degree is positively correlated with plant size.[10]

Viewing plants as sectors, this confirms Acemoglu et al.'s [2] conjecture that the network asymmetry is preserved as the economy becomes more and more disaggregated. In addition, this result suggests that there is also significant plant-level variation of characteristics we normally only associate with industries. This is mainly because of lack of data at the plant or firm level, and so this paper provides the first evidence of IO network asymmetry at the plant level.

To confirm the power law behaviour of the sequence of plant outdegrees, I estimate the share parameter of the tail of the distribution. Following Gabaix and Ibragimov [19], I trim the distribution to the top 20$^{\text{th}}$ percentile of outdegree and estimate

$$\log(\text{rank}_i - 1/2) = \alpha - \beta \log d_i \tag{2}$$

The estimated shape parameter $\hat{\beta}$ is a measure of the strength of the asymmetry in the distribution—a shape parameter of 1 is Zipf's law.

The estimated parameter is $\hat{\beta} = 1.21$ (s.e. $= 0.011$), slightly lower (heavier tailed) than the sector level results from Acemoglu et al. [2], suggesting the plant level outdegree distribution is just as asymmetric as the industry level outdegree distribution in the US. Furthermore, plant size is positively correlated with outdegree in the hexbin plot in Figure 1(D). The elasticity of weight with respect to outdegree has an elasticity of .31 (s.e. .0049), suggesting a strong relationship between being an important plant to the economy (having a high outdegree) and being large (having a large influence vector).

There are two main takeaways from the IO data. First, there is incredibly asymmetry in plant-level IO network, suggesting certain plants are very important suppliers in the economy, especially within industries. Second, that asymmetry is positively associated with plant weight, confirming the relationship between the theoretical measure of asymmetry and the reduced form weight vector used for applying the granularity theory against diversification of idiosyncratic shocks.

However, keep in mind that these are observed IO statistics, not a true underlying IO network. Given the relationships between productivity and IO at the plant level, it is crucial to understand the data through the lens of an appropriate model.

3. Model

To study the relationships between volatility, endogenous unbalanced IO networks and the factors that determine them, I adapt the sectoral model of Acemoglu et al. [2], which is itself based on Long and Plosser [29]. There are three key additions.

First, I study individual plants and not sectors. Although technically easy (e.g., relabeling sectors as plants), it puts the focus on the determinants of granularity—is it the IO network or something else? This becomes crucial as we turn to the study of a very disaggregated economy, which is the primary reason for studying microfoundations of aggregate volatility. Second, I incorporate trade costs into the model by dividing the economy into multiple regions, each with their own individual demand and geographic characteristics.

[10] Hexbin plots have two advantages here. First, for confidentiality reasons I cannot display scatterplots because you can identify characteristics of each individual establishment. Second, scatterplots with thousands of points can be impossible to decipher. The hexbin plot is like a two-dimensional histogram, with the opacity of each hexagon representing the number of plants in that bin (representing the same thing as the length of a histogram bar).

Third, and most importantly, I relax the assumption that the IO network is exogenous. In my model, a plant may be a central supplier of the network because it is a required input in many other products (it has many high exogenous direct input coefficients) or because it is so productive that many other plants substitute toward it.

To introduce these features, I need a model in which productivity, geography and the exogenous IO network can vary independently to create an observed plant-level IO network that I can take to the data.

To start, there are R regions. A representative household in a specific region r inelastically supplies a labour L_r, and has Cobb-Douglas preferences over N different goods,

$$u_r(c_r) = \prod_{i \in N} c_{ri}^{\lambda^{ri}} \tag{3}$$

where c_{ri} is region r's consumption of good i. There is free migration between regions, so that the wage ω in equilibrium is constant across regions. Later, I normalize $\omega = 1$.

Each good is produced by a single plant using Cobb-Douglas combination of labour and a plant-specific intermediate input which is itself a CES aggregate of other products,

$$q_i = z_i l_i^{\beta} \left(\sum_{j \in N} \gamma_{ij}^{1/\sigma} q_{ij}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{(1-\beta)\sigma}{\sigma-1}} \tag{4}$$

where z_i is productivity, β is the labour share in production, q_{ij} is the quantity of plant j's product demanded by plant i, and σ is the elasticity of substitution between intermediates. The crucial part of production is γ_{ij}, which is the exogenous direct input coefficient. If γ_{ij} is high, then independent of plant j's productivity, plant i requires a lot of plant j's input to produce. If γ_{ij} is low but positive, then plant i may still demand a lot of q_{ij} if plant j is very productive. In this way, the endogenous IO network is determined jointly by productivity, substitutability and the exogenous IO network.

Trade between regions and plants comes at a cost. For instance, a plant, i, pays for quantity q_{ij}^s to be shipped, but only receives $q_{ij}^d = q_{ij}^s / \tau_{ij}$. In essence, the buyer pays a unit price $\tau_{ij} p_{ij}$, and the total expenditure that i spends on j is $\tau_{ij} p_{ij} q_{ij}^d$.[11]

With perfect competition, prices equal marginal costs for plant i, incorporating the iceberg trade costs,

$$p_i = C z_i^{-1} \left(\sum_{j \in N} \gamma_{ij} (\tau_{ij} p_j)^{1-\sigma} \right)^{\frac{1-\beta}{1-\sigma}} \tag{5}$$

where $C \equiv \beta^{-\beta}(1-\beta)^{\beta-1}\omega^{\beta}$ is independent of i.

[11]Do I have to keep the supply/demand things here. It's really just iceberg trade costs.

$$p_{ri} = \prod_{i \in N} (\lambda_{ri}^{-1} \tau_{ri} p_i)^{\lambda_{ri}} \tag{6}$$

Remark 1. *Observed expenditure shares depend on productivity, geography and exogenous IO characteristics.*

The IO tables provided by statistical agencies gives an expenditure share of industry i on goods from industry j. The plant IO table I detail in the previous section is constructed the same way, an expenditure share of plant i on plant j. If we assume production is Cobb-Douglas, then the expenditure share parameter in production exactly determines the observed expenditure share. This is no longer true if the elasticity of substitution is not equal to 1. Define the observed expenditure share g_{ij},

$$g_{ij} = \frac{p_j q_{ij}}{p_i q_i} \tag{7}$$

In equilibrium, this simplifies to

$$g_{ij} = (1-\beta) \left(\frac{\gamma_{ij}(\tau_{ij} p_j)^{1-\sigma}}{\sum_{k \in N} \gamma_{ik}(\tau_{ik} p_k)^{1-\sigma}} \right) \tag{8}$$

If $\sigma = 1$, the observed expenditure share is exactly determined by the relative exogenous coefficient γ_{ij} (that is, if you rederive the solution starting with $\sigma = 1$ in the production function). However, it is clear that the observed expenditure shares are jointly determined by the vector of direct input coefficients $\gamma_{i\cdot}$ and the vector of prices and trade costs $\tau_{i\cdot}$, which are themselves determined by the vector of plant productivities (and more complex interconnections). Again, the observed IO network is endogenously determined by the vector of plant productivities, geography and the exogenous IO network.

Remark 2. *Expenditure shares still "determine" size, but they say nothing about the underlying determinants of the size distribution.*

In an important result, Acemoglu et al. [2] shows that the vector of industry sizes, normalized by total sales in the economy, which he calls the influence vector w (he uses the notation v, but I find using w to represent plant weight makes more sense), is the crucial link between the IO network and volatility. The influence vector determines the extent to which microeconomic shocks contribute to aggregate volatility, and the influence vector is determined by the characteristics of the exogenous IO network. Hence their claim that the IO network is the main determinant of aggregate volatility. Here I show that the same holds for the observed IO network. That is, an empirical association between the influence vector and the true IO network does not tell you the effect of the IO network on volatility, because the observed network may be entirely determined by productivity. Write the system of market clearing equations,

$$\sum_{r \in R} p_{ri} c_{ri} + \sum_{j \in N} p_i q_{ji} = p_i q_i, \text{ for } i \in N \tag{9}$$

And rewrite in terms of g_{ij} using (7),

$$\sum_{r \in R} p_{ri} c_{ri} + \sum_{j \in N} g_{ji} p_j q_j = p_i q_i, \text{ for } i \in N \tag{10}$$

Then a similar derivation to Acemoglu et al. [2] (see Appendix 8) gives you the influence vector as a function of the matrix of observed expenditure shares $G = [g_{ij}]$, observed demand shares $A = [a_{ri}]$, and regional labour $\vec{L}$,

$$w' = \left(\frac{\beta}{L}\right) \vec{L}' A (I - G)^{-1} \tag{11}$$

The influence vector, w, is always related to the observed IO network, but the observed IO network is endogenous. So observing the association between the influence vector and the IO network does not give you any information on the importance of the underlying IO network, $\Gamma = [\gamma_{ij}]$, or region demand characteristics, $\Lambda = [\lambda_{ri}]$.

Example 1. *Suppose $\gamma_{ij} = \tau_{ij} = 1$ for all $i, j \in N$. Then there is no exogenous IO network or geographic variation, and all of the observed IO characteristics are due to productivity.*

If $\gamma_{ij} = \tau_{ij} = 1$, then all plants use the same intermediate bundle and face the same intermediate input price. This means the expenditure share equation (7) reduces to

$$g_{ij} = (1 - \beta) \left(\frac{z_j^{\sigma - 1}}{\sum_{k \in N} z_k^{\sigma - 1}} \right) \tag{12}$$

Which is determined solely by relative productivities. In this case, if productivities are distributed with a power law, we will still observe an influence vector consistent with the unbalanced IO network, even though the underlying IO network is as balanced as possible.

Example 2. *Suppose $z_i = 1$ for all $i \in N$ and $\tau_{ij} = 1$ for all $i, j \in N$. Then there is no productivity or geography variation, and all of the observed IO characteristics are due to the exogenous IO network.*

When productivities and trade costs are identical across all plants, the expenditure share terms reduce to

$$g_{ij} = (1 - \beta) \left(\frac{\gamma_{ij}}{\sum_{k \in N} \gamma_{ik} (p_k / p_j)^{1 - \sigma}} \right) \tag{13}$$

where $(p_k/p_j)^{1-\sigma}$ terms can be written as a recursive function of relative prices and IO parameters, which implies the expenditure shares are determined only by IO parameters.

3.1. **Outdegree and unbalanced IO networks.** An unbalanced IO network is one in which individual plants are central suppliers to the entire economy. The easiest way to ask how central a plant is by adding up the expenditure shares of a plant's customers,

$$\hat{d}_i = \sum_{j \in N} g_{ji} \tag{14}$$

Example 3. *Suppose $\gamma_{ij} = d_j/N$, for $j \in N$, and $\tau_{ij} = 1$ for $i, j \in N$.*

Expenditure shares are

$$g_{ij} = (1-\beta)\left(\frac{d_j z_j^{\sigma-1}}{\sum_{k \in N} d_k z_k^{\sigma-1}}\right) \tag{15}$$

Observed outdegree is

$$\hat{d}_i = (1-\beta)\left(\frac{d_i z_i^{\sigma-1}}{(1/N)\sum_{k \in N} d_k z_k^{\sigma-1}}\right) \tag{16}$$

And, under a few more assumptions,[12] one element of the influence vector is

$$w_i = \mathbf{1}'\Lambda_{\cdot i}\beta + (1-\beta)\left(\frac{d_i z_i^{\sigma-1}}{\sum_{k \in N} d_k z_k^{\sigma-1}}\right) \tag{17}$$

This examples highlights the dependence of the influence vector on productivity and the unbalanced IO network—the distribution of w_i is determined by the distribution of $d_i z_i^{\sigma-1}$. Recall that the argument for microfoundations of aggregate shocks requires the distribution of w_i to have a thick tail even as the number of plants grows large. However, as the number of plants grows large, the thick tail of w_i will tend to be dominated by the thickest tail of the two distributions of outdegree and productivity.

3.2. **Geography.** Plants have a tendency to cluster in specific geographic areas, leaving groups of plants in close proximity (and low trade costs), while other plants are spread out over the rest of the regions in the economy. A plant that is closer to other plants, especially large ones, tend to be large themselves and also tend to purchase inputs from their neighbours.

Example 4. *Suppose $\tau_{ij} = \tau_j$, for $j \in N$, and that productivity and the exogenous IO parameters are constant, $z_i = 1$, for $i \in N$, and $\gamma_{ij} = 1$ for $i, j \in N$.*

The expenditure share depends solely on trade costs, and plant i buys the most from the plant with the lowest τ_j:

$$g_{ij} = (1-\beta)\left(\frac{\tau_j^{1-\sigma}}{\sum_{k \in N} \tau_k^{1-\sigma}}\right) \tag{18}$$

In addition, trade costs between plants are typically clustered so that groups of plants are closer to each other and all have low τ_{ij} values relative to plants in remote areas. This has a special effect on the size distribution of establishments. While productivity and demand characteristics usually have distributions with very skewed

[12]That plants are located so that

right tails, trade costs can have the opposite effect—there is a high number of plants concentrated in center areas of the economy, with other small establishments in remote areas. Since trade costs are bounded below by 1 and are inversely related to g_{ij}, trade costs have a much more significant effect at the middle and lower-end of the size distribution, not the right tail.

The next section pins down the theoretical basis for these concepts, and the following sections explore the empirical support for them. Since the effect of trade costs on size are bounded above, I focus the asymptotic theory on demand and productivity characteristics.

4. Asymptotic Theory

Asymptotic results are key to the arguments for and against the microfoundations of aggregate shocks.[13] The granular hypothesis relies on a thick tail of the size distribution. The unbalanced network hypothesis claims the reason *why* the size distribution has a thick tail is because of a thick tail of outdegree, a telling characteristic of an unbalanced IO network. Only by combining the two approaches can we understand the forces that shape the observed centrality and size distributions.

In what follows, I rely especially on the following property of power law distributions:

Remark 3. *Suppose the random variables X and Y follow power law distributions with parameters ζ_X and ζ_Y. Then the distribution of $X+Y$ and the distribution of XY both follow power laws with parameter* $\min\{\zeta_X, \zeta_Y\}$.

The same result follows for many similar combinations of power law random variables (see [17] or [27]). Using Remark 3, we are interested in explaining the tail parameter of the size distribution, β_v, given the tail parameters of the distributions of observed outdegree (ζ_d) and productivity (ζ_z).

Therefore, if the asymptotic results hold for this economy, network asymmetry cannot be the fundamental cause of the skewed plant size distribution because of the relative values of each tail parameter. But like so many other applications of power laws, the reality is not so black and white. In any case, we must understand the asymptotic argument first, and then ask if and when is it reasonable to apply it.

The network hypothesis relies on two sequential arguments. First, the tail of the distribution of the plant-level exogenous IO network characteristics must determine the tail of the distribution of the observed plant-level IO network characteristics. Second, the tail of the distribution of the observed IO network determines the tail of the plant size distribution. If either of these arguments fail, it is unlikely the underlying IO network is the cause of the skewed plant size distribution.

I approach the second part of the argument first. For the observed network to matter asymptotically, the outdegree distribution must have a thick tail. If not,

[13]In Appendix 8, I use Hulten's Theorem to show aggregate volatility depends on the herfindahl of the economy, and the herfindahl of the economy depends on the distribution of outdegree and productivity. These results are standard when applying the granular and network theories of aggregate fluctuations, so I omit them and focus on the new idea provided in this paper.

outdegree cannot be the ultimate source of the thick tail of the size distribution. If the outdegree distribution does have a thick tail, the parameter must match, or be "close" to matching (in a statistical sense) the tail of the size distribution. However, the measured tail parameter for the network is 1.21, about 20% higher than the plant size distribution's parameter of 1.04, which is consistent with a Zipf's law distribution of plant size.Therefore $\zeta_z < \zeta_d$ implies the degree distribution is dominated by some other plant characteristic, and thus does not determine plant size asymptotically or turn idiosyncratic shocks into aggregate fluctuations.

We can see this conclusion supported by prior research in different settings. A plethora of research on the firm size distribution conclude it is approximately described by Zipf's law in the upper tail (see [30] or [17], while Acemoglu et al. [2] measure the tail of the sector outdegree distribution at 1.38, much larger than the typical Zipf's law size distribution parameter of 1.

The first part of the argument, the required relationship between the observed and unobserved network characteristics is more problematic. The IO data are necessarily the observed shares, and so depend on both the underlying IO network and other plant characteristics, especially productivity. However, in absence of direct evidence for or against the underlying network, I suggest that, asymptotically, productivity is more likely to be the cause of the observed IO network, and possibly the final size distribution.

To establish this formally, I show that, under the assumptions of the model in the previous section, the tail of the size distribution is dominated by the thickest tail between productivity (adjusted for substitutability) and outdegree.

Proposition 4.1. *Suppose the distributions of outdegree and productivity both follow power laws with parameters ζ_d and ζ_z,*

$$P(d > x) = C_d x^{-\zeta_d} L_d(x), \tag{19}$$

$$P(z > x) = C_z x^{-\zeta_z} L_z(x) \tag{20}$$

Here, $L_d(x)$ and $L_z(x)$ are slowly varying functions, C_d and C_z are constants, and ζ_d and ζ_z are positive. Then the size distribution also follows a power law with parameter $\min\{\zeta_d, \zeta_z/(\sigma-1)\}$,

$$P(v > x) = C_v x^{-\min\left\{\zeta_d, \frac{\zeta_z}{(\sigma-1)}\right\}} L_v(x) \tag{21}$$

Proof. See Appendix 8. □

The distribution of labour productivity has a tail parameter of approximately 1.97, so for a suitable choice of σ, it is easy to match the empirical tail parameter of the plant size distribution. In particular, if $\sigma \approx 2.89$, the size distribution will approximately satisfy Zipf's law. It also could satisfy both, if substitutability for final goods is higher than for intermediates. Note that similar studies on productivity and size, especially ones focusing on international trade models, (e.g., see Appendix 9.2 for an extension of the model with monopolistic competition and plant entry and exit) gives the same result—firm size is determined by a combination of productivity and substitutability, with the size tail parameter being very close to 1 (see, e.g., a series of papers by di

Giovanni and Levchenko and their co-authors [12, 9, 10]). The difference here is that they observe the size distribution and assume it *must* be because of productivity. For more on power laws and the determination of firm size, see [30] or [17].

Although the asymptotic theory gives clear cut answers as to which factor is responsible for the shape of the size distribution, the empirical results suggest the truth is somewhere between the two extremes.

5. Calibration

In this section, I calibrate the model to match features of the data to further explore the relationships between the unbalanced IO network and volatility. Instead of applying asymptotic results directly to infer which tail, productivity or outdegree, dominates the size distribution, using the model described in Section 3, I use data on plant productivity $\vec{z}$, trade costs T, the observed region demand A, and the observed input share matrix G to solve for the unobserved region demand characteristics Λ and the unobserved technical requirement matrix Γ.

Although final demand didn't add to the explanation of the model and asymptotic theory, it is important empirically. Therefore, to match the data better, I change the consumer's utility function to a CES combination of each product,

$$u_r(c_{ri}) = \left(\sum_{i \in N} \lambda_{ri}^{\frac{1}{\sigma}} c_{ri}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{22}$$

Where c_{ri} is region r's consumption of plant i's output. Now the unobserved final demand characteristic λ_{ri} is similar to a γ_{ji} in firm j's production function, and the observed final demand share a_{ri} is similar to the observed expenditure share g_{ji}.

5.1. **Parameters.**

5.1.1. *Productivity.* First, I need a measure of productivity, z_i, for each establishment. There are several methods to estimate productivity, although they give similar results. First, I use the method of Gandhi, Navarro and Rivers [20] to estimate total-factor productivity (TFP). Second, labour productivity, either calculated by value-added per worker or value-added per dollar spent on wages (which accounts for labour and capital quality better than raw value-added per worker or hours worked).

The results are similar for all measures. Since revenue TFP varies less than labour productivity, and revenue TFP varies considerably more than physical TFP (consistent with Foster and Haltiwanger [16] and Huttman et al. [24]), revenue and physical productivity can explain even less of the observed IO network than labour productivity.

5.1.2. *Trade costs.* STF distance data, possible with some elasticity? Or σ gives the elasticity, really. T.

5.1.3. *Observed IO network.* Next, the matrix of input-shares G is described in Section **??** and Appendix **??**. In addition, the share of value-added in production, β_i, is calculated as exactly that. Next, the final demand parameter a_i (similar to the final demand category in the industry input-output tables), is calculated as that left over after all within-manufacturing production has been taken into account. Finally, I set the elasticity of substitution $\sigma = 2$. Later, I test the robustness of the model against changing the elasticity of substitution, since that directly affects the tail of the plant-size distribution as we saw in Section 4.

The model is simple enough to be solved directly, using the $N \times N$ observations in G, $(R+N) \times (R+N)$ observations in T, $R \times N$ final demand observations in A, and N productivity observations to solve for the $N \times N$ unknowns in Γ and $R \times N$ unknowns in Λ (along with the $R+N$ prices and normalized wage $\omega = 1$).

5.1.4. *Elasticities.* Could try same minimization program to get σ, etc.

5.2. **Results.** The main result is that productivity cannot explain much of the observed IO network; there is not enough heterogeneity in productivity to explain either the size distribution, the IO network or the final demand parameters. This is consistent with other recent work challening the dependence of the plant size distribution on productivity.

Trade costs explains x. Calculate gradient of plant size distribution log-normal tail parameters (actually, maybe I'll drop the power law stuff here and estimate log-normal distributions). Get some cool results. Simulate aggregate volatility later.

(To be completed pending confidentiality review.)

5.3. **Robustness.**

6. Volatility calibration and simulations

6.1. **Dynamic model.** To adapt the static model in Section **??** to include volatility, I use a strategy similar to Acemoglu [2]. In each period, plants receive idiosyncratic demand shocks γ'_{ijt} and λ'_{rit}, as well as productivity shocks z'_{it}. In each period, the equilibrium is equal to the static model with the new parameters $\gamma_{ijt} = \gamma_{ij}\gamma'_{ijt}$, $\lambda_{ijt} = \lambda_{ij}\lambda'_{ijt}$, and $z_{it} = z_i z'_{it}$.[14]

There are several important factors in the dynamic model that help us study the microfoundations of aggregate fluctuations, and the relative contributions of granularity, geography and exogenous IO characteristics to aggregate volatility.

Similar to the rest of the paper, the difference between the unobserved and observed parameters matters. The data are observed sales growth rates, but we would like to know the unobserved idiosyncratic shocks that gave rise to them. Furthermore, uncorrelated idiosyncratic shocks naturally result in correlated sales growth rates, depending on the linkages between plants and plants, and plants and regions.

Next, demand and productivity shocks may contribute differently to aggregate volatility. In previous work (see, e.g., Acemoglu NBER paper on macro things),

[14] Do I need more discussion or examples of how shocks are transmitted here? Will be tough with CES.

productivity shocks only propagate downstream, and demand shocks only propagate upstream. However, using a CES function, both types of shocks can propagate in both directions. For example, a positive productivity shock can propagate upstream because it affects downstream expenditure for the product (positively, if the elasticity of substitution is positive).

The distinction between demand and productivity is an important factor in the literature on the firm-size distribution (see *xyz* / Haltiwanger etc, and Section **??** above), so it's reasonable to expect the same pattern in volatility. Demand variation by plant contributes significantly more to the firm-size distribution and does variation in productivity. Similarly, idiosyncratic demand shocks may contribute significantly more to volatility than does idiosyncratic productivity shocks.

The last important note: idiosyncratic shocks may or may not be correlated. First, I attempt to match aggregate volatility by using uncorrelated shocks, but if the simulations can't match the data, I'll re-examine the assumptions, and see how far idiosyncratic shocks can go with reasonable parameter estimates.

6.2. **Calibration strategy.** I now describe the strategy to estimate the parameters of the idiosyncratic shock processes $\vec{\sigma}$. First, start with the empirical variance-covariance matrix of sales growth rates, $V = [\text{cov}(g_{it}, g_{jt})]$. The empirical data to match is then $X = \text{vec}(V)$. Next, guess the potential parameters, $\vec{\sigma} = (\sigma_\gamma, \sigma_\lambda, \sigma_z)$, and simulate the dynamic version of the model described in Section **??**, and calculate the simulated variance-covariance matrix of sales growth $\hat{V}(\vec{\sigma})$, which results in the predicted data, $\hat{X}(\vec{\sigma})$. The optimal parameters $\vec{\sigma}^*$ are the solution to the minimization program

$$\arg\min_{\vec{\sigma}} ||\hat{X}(\vec{\sigma}) - X||' W ||\hat{X}(\vec{\sigma}) - X||, \tag{23}$$

where the weight is chosen to similar to weighted-least-squares estimators. For instance, larger plants will typically have more reliable sales growth statistics, so one might choose weights to be equal to the empirical plant weight matrix $W = [w_i w_j]$.[15] This choice also has the benefit of attempting to match aggregate volatility at the same time as matching each individual covariance estimate, because aggregate volatility is the weighted sum of the individual covariances (under some assumptions). Another choice is to weight the covariance estimates based on the variance of the covariance estimate itself (which is decreasing in T), or simply to use $W = [1]$. I use simulated annealing to solve the minimization program.

I wish to approximate certain aspects of the variance-covariance matrix. First, the contribution of covariance to aggregate volatility. We know the majority of observed volatility is due to covariances between plants, not the variances of individual plants. Therefore, appropriate estimates of $\vec{\sigma}$ will produce simulated covariances that account for the majority of aggregate fluctuations. Second, the predicted aggregate volatility should match empirical aggregate volatility. Although, with a weight matrix $W =$

[15] X is $N^2 \times 1$, so W must be $N^2 \times N^2$.

$[w_i w_j]$, the minimization program attempts to match aggregate volatility exactly, there's no guarantee it will actually work. The benefit of using this calibration scheme is that I can numerically evaluate the performance of the calibration: the resulting R^2 gives an indication of how close the estimation gets.

In addition, I estimate standard errors of the parameters by bootstrapping equation (**??**).

6.3. Results.

6.3.1. *Parameter estimates.* Tables of stuff, with confidence intervals.

6.3.2. *Contributions to aggregate volatility.* Demand, productivity shocks. How close does it come to data? R^2, etc. I'm guessing it doesn't come close. Add aggregate shocks.

6.4. Counterfactuals.

6.4.1. *Geography and trade costs.* Eliminate trade costs, keep sigmas as given, calculate volatility.

6.4.2. *Productivity.* Eliminate productivity variation, keep sigmas, calculate volatility.

6.4.3. *Demand / IO Parameters.* Eliminate demand variation, keep sigmas, calculate volatility.

6.4.4. *Shocks and levels?* Do both at once?

7. Conclusion

I investigate the relationship between idiosyncratic shocks, unbalanced input-output (IO) networks and aggregate volatility. Using detailed data on commodity inputs and outputs in Canadian manufacturing, I study a plant-level IO network and its effect on aggregate volatility. My main contribution is to account for the endogenous observed IO network and quantify the separate effects of productivity and the underlying IO network on plant size and aggregate volatility.

To differentiate between the granular and network hypotheses of aggregate fluctuations, I use a model in which productivity and the underlying IO network vary independently and use the plant-plant IO network data to uncover the model parameters. I find that productivity cannot explain the asymmetry in the observed IO network and that the majority of the variation in plant size, and therefore aggregate volatility, is caused by the underlying IO network.

I compare the properties of the IO network to each plant's contribution to aggregate volatility, and confirm that more central plants matter more for aggregate volatility. Specifically, a 10% increase in a plant's outdegree is associated with a 2.66% increase in size, while a 10% increase in labour productivity is associated with an 8% larger plant. The asymmetry of the IO network contributes 34% to aggregate volatility in Canadian manufacturing.

In conclusion, to investigate the propagation of idiosyncratic shocks, I acknowledge and investigate the endogeneity of the observed IO network and find the underlying IO network does account for a sizable proportion of aggregate volatility. Future research can extend this work in several ways: using the plant-plant IO network to directly investigate the propagation mechanism of idiosyncratic shocks, adding in financial linkages between establishments within firms, or identifying supply chains across the economy, instead of just manufacturing. Doing so will increase our knowledge of the complex linkages that underpin our economy.

References

[1] Daron Acemoglu, Ufuk Akcigit, and William Kerr. Networks and the Macroeconomy: An Empirical Exploration. 2015.

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

[3] Daron Acemoglu, Azuman Ozdaglar, and Alireza Tahbaz-Salehi. Systemic Risk and Stability in Financial Networks. 2013.

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

[5] Jean-Noel Barrot and Julien Sauvagnat. Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks. 2015.

[6] Andrew B. Bernard, Andreas Moxnes, and Yukiko U. Saito. Production Networks, Geography and Firm Performance. 2015.

[7] Vasco Carvalho, Nirei Makoto, and Yukiko U. Saito. Supply Chain Disruptions: Evidence from the Great East Japan Earthquake. *RIETI Discussion Paper Series*, (14-E-035), 2014.

[8] Vasco Carvalho and Nico Voigtländer. Input Diffusion and the Evolution of Production Networks. 2014.

[9] Julian Di Giovanni and Andrei Levchenko. Country size, international trade, and aggregate fluctuations in granular economies. *Journal of Political Economy*, 120(6):1083–1132, 2012.

[10] Julian Di Giovanni and Andrei Levchenko. Firm Entry, Trade, and Welfare in Zipf's World. *Journal of International Economics*, 89(2):283–296, 2013.

[11] Julian Di Giovanni, Andrei Levchenko, and Isabelle Méjean. Firms, Destinations, and Aggregate Fluctuations. *Econometrica*, 82(4):1303–1340, 2014.

[12] Julian Di Giovanni, Andrei Levchenko, and Romain Rancière. Power Laws in Firm Size and Openness to Trade: Measurement and Implications. *Journal of International Economics*, 85(1):42–52, 2011.

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

[14] Matthew Elliott, Benjamin Golub, and Matthew O. Jackson. *Financial Networks and Contagion*, volume 104. 2014.

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

[16] Lucia Foster, John Haltiwanger, and Chad Syverson. Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review*, 98(1):394–425, 2008.

[17] Xavier Gabaix. Power Laws in Economics and Finance. *Annual Review of Economics*, 1(1):255–294, 2009.

[18] Xavier Gabaix. The Granular Origins of Aggregate Fluctuations. *Econometrica*, 79(3):733–772, 2011.
[19] Xavier Gabaix and Rustam Ibragimov. Rank 1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents. *Journal of Business & Economic Statistics*, 29(1):24–39, 2011.
[20] Amit Gandhi, Salvador Navarro, and David Rivers. On the Identification of Production Functions: How Heterogeneous is Productivity? 2013.
[21] Thomas J Holmes and John J Stevens. An Alternative Theory of the Plant Size Distribution, with Geography and Intra- and International Trade. *Journal of Political Economy*, 122(2), 2014.
[22] Michael Horvath. Cyclicality and Sectoral Linkages: Aggregate Fluctuations from Independent Sectoral Shocks. *Review of Economic Dynamics*, 1(4):781–808, 1998.
[23] Michael Horvath. Sectoral shocks and aggregate fluctuations. *Journal of Monetary Economics*, 45(1):69–106, 2000.
[24] Colin Hottman, Stephen J. Redding, and David E. Weinstein. Quantifying the Sources of Firm Heterogeneity. 1, 2014.
[25] CR Hulten. Growth Accounting with Intermediate Inputs. *The Review of Economic Studies*, 45(3):511–518, 1978.
[26] Matthew O. Jackson. *Social and Economic Networks*. Princeton University Press, 2010.
[27] Anders Hedegaard Jessen and Thomas Mikosch. Regularly Varying Functions. *Publications de L'Institut Mathematique*, 80(94):171–192, 2006.
[28] Charles Jones. Intermediate Goods and Weak Links in the Theory of Economic Development. *American Economic Journal: Macroeconomics*, 3(2):1–28, 2011.
[29] John B. Long and Charles I. Plosser. Real Business Cycles. *Journal of Political Economy*, 91(1):39—-69, 1983.
[30] Erzo G. J. Luttmer. Selection, Growth, and the Size Distribution of Firms. *Quarterly Journal of Economics*, (August):1103–1144, 2007.
[31] Marc J. Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695—-1725, 2003.
[32] Ezra Oberfield. Business networks, production chains, and productivity: A theory of input-output architecture. 2011.
[33] John Shea. Complementarities and comovements. *Journal of Money, Credit, and Banking*, 34(2):412–433, 2002.

8. Appendix: Theory

See important model notation in Table 1.

8.1. **Full model.**

8.1.1. *Consumers.* There are R regions, with a representative consumer in each with utility function $u_r(c_r)$,

$$u_r(c_r) = \left(\sum_{i \in N} \lambda_{ri}^{\frac{1}{\sigma}} c_{ri}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{24}$$

Labour is inelastically supplied given the stock of labour in region r, L_r. Consumer r's problem is

$$\max_{c_r} u_r(c_r) \text{ s.t. } \sum_{i \in N} p_{ri} c_{ri} \leq w_r L_r \tag{25}$$

Table 1. Table of Notation

R	$\triangleq$	Set of regions. Abusing notation, R is also the number of regions.
N	$\triangleq$	Set of plants. Abusing notation, N is also the number of plants.
G	$\triangleq$	$N \times N$ matrix of observed plant input shares. An element g_{ij} is the share of plant j's input in plant i's sales.
Γ	$\triangleq$	$N \times N$ matrix of exogenous plant input demand characteristics. An element γ_{ij} enters plant i's demand for plant j's output.
A	$\triangleq$	$R \times N$ matrix of observed region-plant demand shares. An element a_{ri} is the share of region r's total expenditure on plant i's output.
Λ	$\triangleq$	$R \times N$ matrix of exogenous region input demand characteristics. An element λ_{ri} enters region r's demand for plant i's output.
T	$\triangleq$	$(R+N) \times (R+N)$ matrix of trade costs. An element τ_{ri} is the cost of trade between region r and i, and an element τ_{ij} is the cost of trade between plants i and j.
z_i	$\triangleq$	Productivity of plant i.
σ	$\triangleq$	Elasticity of substitution.
β_i	$\triangleq$	Share of value-added in plant i's production.

Consumer r must pay a trade cost τ_{ri} to buy from plant i, so that

$$p_{ri} = \tau_{ri} p_i \tag{26}$$

The solution gives r's price index

$$p_r = \left(\sum_{i \in N} \lambda_{ri} (\tau_{ri} p_i)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{27}$$

8.1.2. *Producers.* There are N producers. Producer i's production function is

$$f_i(l_i, q_{i1}, \ldots, q_{iN}) = z_i l_i^{\beta} \left(\sum_{j \in N} \gamma_{ij}^{\frac{1}{\sigma}} q_{ij}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{(1-\beta)\sigma}{\sigma-1}} \tag{28}$$

Producer i's problem is to minimize cost

$$\min_{(l_i, q_{i1}, \ldots, q_{iN})} \sum_{i \in N} p_{ij} q_{ij} \text{ s.t. } f_i \geq \bar{q}_i \tag{29}$$

Producer i's input cost for one unit of intermediate input is

$$\eta_i = \beta^{-\beta}(1-\beta)^{\beta-1}\left(\sum_{j\in N}\gamma_{ij}(\tau_{ij}p_j)^{1-\sigma}\right)^{\frac{1-\beta}{1-\sigma}} \tag{30}$$

Given perfect competition, plant i's price is

$$p_i = z_i^{-1}\eta_i \tag{31}$$

8.1.3. *Market clearing.* Labour is free to migrate between regions. Total labour in the economy is

$$\sum_{r\in R} L_r = L \tag{32}$$

Now, each plant i is in a region r, and the total value added produced by those plants in r add up to total income in that region,

$$\sum_{i\in r}\beta s_i = \omega L_r \tag{33}$$

For goods, producer i supplies the other producers $j \in N$, and each region $r \in R$, giving market clearing

$$\sum_{r\in R} c^s_{ri} + \sum_{j\in N} q^s_{ji} = q^s_i, \text{ for } i \in N \tag{34}$$

Iceberg trade costs mean producer i ships $c^s_{ri} = \tau_{ri}c_{ri}$ to region r and $q^s_{ji} = \tau_{ji}q_{ji}$. Replacing those terms and multiplying all terms by p_i,

$$\sum_{r\in R} p_i\tau_{ri}c_{ri} + \sum_{j\in N} p_i\tau_{ji}q_{ji} = p_i q^s_i, \text{ for } i \in N \tag{35}$$

8.1.4. *Equilibrium.* Equilibrium in the economy means two sets of prices $\{p_r : r \in R\}$, $\{p_i : i \in N\}$, wage ω normalized to 1, and labour stocks by region $\{L_r : r \in R\}$, that solve the consumer's and producer's problems for each region and producer, and the labour and goods markets clear.

8.1.5. *Solving the model given data.* Given data on T, G, A, w, β, solve for Λ and Γ. Must also solve for p_r and p_i, and normalize $\omega = 1$. Anything else? L_r?

Need to make assumptions about the elasticity σ, then solve.

Have price equations:

$$p_r = \left(\sum_{i\in N} \lambda_{ri}(\tau_{ri}p_i)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}, \text{ for } r \in R \tag{36}$$

$$\eta_i = \beta_i^{-\beta_i}(1-\beta_i)^{\beta_i-1}\omega^{\beta_i}\left(\sum_{j\in N} \gamma_{ij}(\tau_{ij}p_j)^{1-\sigma}\right)^{\frac{1-\beta_i}{1-\sigma}}, \text{ for } i \in N \tag{37}$$

$$p_i = z_i^{-1}\eta_i, \text{ for } i \in N \tag{38}$$

And share equations:

$$a_{ri} = \lambda_{ri}\tau_{ri}^{1-\sigma}\left(\frac{p_i}{p_r}\right)^{1-\sigma}, \text{ for } r \in R, i \in N \tag{39}$$

$$g_{ij} = (1-\beta_i)\lambda_{ij}\tau_{ij}^{1-\sigma}\left(\frac{p_j}{\eta_i}\right)^{1-\sigma}, \text{ for } i \in N, j \in N \tag{40}$$

$$\lambda_{ri} = a_{ri}\tau_{ri}^{\sigma-1}\left(\frac{p_i}{p_r}\right)^{\sigma-1}, \text{ for } r \in R, i \in N \tag{41}$$

$$\gamma_{ij} = (1-\beta_i)^{-1}g_{ij}\tau_{ij}^{\sigma-1}\left(\frac{p_j}{\eta_i}\right)^{\sigma-1}, \text{ for } i \in N, j \in N \tag{42}$$

And region income equations,

$$\beta_i s_i = wl_i \tag{43}$$

$$\sum_{i\in R} \beta_i s_i = \omega L_r \tag{44}$$

$$\sum_{r\in R} L_r = L \tag{45}$$

And finally, sizes:

$$\omega A'\vec{L} + G's = s, \text{ or} \tag{46}$$

$$s = \omega(I - G')^{-1}A'\vec{L} \tag{47}$$

How many unknowns? $p_r \to R$, $p_i, \eta_i \to 2N$, $\Lambda \to RN$, $\Gamma \to N^2$, $s \to N$, $L_r \to R$, ω. So $R + 2N + RN + N^2 + N + R + 1$. How many equations? $R + 2N + RN + N^2 + R + 1 + N$. Done, and also $\omega = 1$? Where does that come in? Maybe sizes also data.

8.1.6. *Solving the model given parameters.* Do the opposite. Same equations, start with z, β, T (data), σ (maximization), Γ, Λ, and so on. Solve the same equations for s, all p, A, G.

But strategy is different.

8.2. **Derivation of influence vector.** Using the definition of observed expenditure shares,

$$g_{ji} = \frac{\tau_{ji} p_i q_{ji}}{p_j q_j} \tag{48}$$

Rewrite the system of market clearing equations

$$\sum_{r \in R} \tau_{ri} p_i c_{ri} + \sum_{j \in N} \tau_{ji} p_i q_{ji} = p_i q_i, \text{ for } i \in N \tag{49}$$

as

$$\sum_{r \in R} \tau_{ri} p_i c_{ri} + \sum_{j \in N} g_{ji} p_j q_j = p_i q_i, \text{ for } i \in N \tag{50}$$

Then replace $\tau_{ri} p_i c_{ri} = a_{ri} p_r c_r = a_{ri} \omega L_r$ and define total sales as $s_i = p_i q_i$,

$$\sum_{r \in R} a_{ri} \omega L_r + \sum_{j \in N} g_{ji} s_j = s_i, \text{ for } i \in N \tag{51}$$

Rewrite in vector form, using $\vec{L} = (L_1, \ldots, L_R)$, write $a_{\cdot i}$ as the i-th column of A and $g_{\cdot i}$ as the i-th column of G,

$$\omega a'_{\cdot i} \vec{L} + g'_{\cdot i} s = s_i, \text{ for } i \in N \tag{52}$$

Now stack those N equations on top of each other, which stacks the vectors $g'_{\cdot i}$ (now the *row* vectors of G'), which gives

$$\omega A' \vec{L} + G' s = s \tag{53}$$

Rearrange and factor out s,

$$s - G' s = \omega A' \vec{L} \tag{54}$$

$$(I - G') s = \omega A' \vec{L} \tag{55}$$

Then pre-multiply by the Leontief matrix, the inverse of $(I - G')$,

$$s = \omega (I - G')^{-1} A' \vec{L} \tag{56}$$

To get the influence vector, use $\omega L = \beta \sum_{i \in N} s_i$ and $w_i = s_i / \left(\sum_{j \in N} s_j \right)$, and finally normalize wages to 1 ($\omega = 1$) and take the transpose of both sides:

$$w' = \left(\frac{\beta}{L} \right) \vec{L}' A (I - G)^{-1} \tag{57}$$

If value-added varies across plants, the relevant equation is

$$A' \overrightarrow{(\beta' w)_r} + G' w = w \tag{58}$$

Or,

$$A'\overrightarrow{(\beta' w)}_r = (I - G')^{-1} w \tag{59}$$

8.3. **Aggregate volatility depends on the product of the distributions of outdegree and productivity.** Aggregate volatility scales according to $||w||_2$ (see Hulten's Theorem [25] and Theorem 1 of Acemoglu et al. [2]). To add to those results, I characterize the behaviour of $||w||_2$ in terms of the distributions of outdegree and productivity.

Write an element of the influence vector w_i as

$$w_i = \frac{\beta}{N} + (1-\beta)\left(\frac{d_i z_i^{\sigma-1}}{\sum_{k\in N} d_k z_k^{\sigma-1}}\right) \tag{60}$$

Then the Euclidean norm of v can be written

$$||w||_2 = \sqrt{\sum_{i\in N}\left[\frac{\beta^2}{N^2} + (1-\beta)^2\left(\frac{d_i z_i^{\sigma-1}}{\sum_{k\in N} d_k z_k^{\sigma-1}}\right)^2 + 2(1-\beta)\left(\frac{\beta}{N}\right)\left(\frac{d_i z_i^{\sigma-1}}{\sum_{k\in N} d_k z_k^{\sigma-1}}\right)\right]} \tag{61}$$

$$||w||_2 = \sqrt{\frac{\beta^2}{N} + (1-\beta)^2\sum_{i\in N}\left(\frac{d_i z_i^{\sigma-1}}{\sum_{k\in N} d_k z_k^{\sigma-1}}\right)^2 + 2(1-\beta)\left(\frac{\beta}{N}\right)\sum_{i\in N}\left(\frac{d_i z_i^{\sigma-1}}{\sum_{k\in N} d_k z_k^{\sigma-1}}\right)} \tag{62}$$

Rewrite slightly,

$$||w||_2^2 = \frac{\beta^2}{N} + (1-\beta)^2\sum_{i\in N}\left(\frac{d_i z_i^{\sigma-1}}{\sum_{k\in N} d_k z_k^{\sigma-1}}\right)^2 + 2(1-\beta)\left(\frac{\beta}{N}\right) \tag{63}$$

$$||w||_2^2 = \frac{\beta^2}{N} + 2(1-\beta)\left(\frac{\beta}{N}\right) + (1-\beta)^2 h_g^2 \tag{64}$$

$$||w||_2^2 = \frac{\beta(2-\beta)}{N} + (1-\beta)^2 h_g^2 \tag{65}$$

$$||w||_2^2 \geq (1-\beta)^2 h_g^2 \tag{66}$$

Implying $||w||_2^2 = \Omega\left(h_g^2\right)$.

In addition, $||w||_2^2 = \mathcal{O}\left(h_g^2\right)$. To see this, first note

$$h_g^2 \geq \frac{1}{N}\left(\sum_{i\in N}\frac{d_i z_i^{\sigma-1}}{\sum_{k\in N} d_k z_k^{\sigma-1}}\right)^2 = \frac{1}{N} \tag{67}$$

which we can rearrange to get $1/(N h_g^2) \leq 1$.

$$||w||_2^2/h_g^2 = \frac{\beta(2-\beta)}{N h_g^2} + (1-\beta)^2 \tag{68}$$

Meaning

$$\limsup_{N\to\infty} \frac{||w||_2^2}{h_g^2} = \limsup_{N\to\infty} \left[\frac{\beta(2-\beta)}{Nh_g^2} + (1-\beta)^2\right] \tag{69}$$

Using the result that $(Nh_g^2)^{-1}$ is bounded above by 1,

$$\limsup_{N\to\infty} \frac{||w||_2^2}{h_g^2} \leq \limsup_{N\to\infty} \left[\beta(2-\beta) + (1-\beta)^2\right] \tag{70}$$

$$\limsup_{N\to\infty} \frac{||w||_2^2}{h_g^2} \leq \beta(2-\beta) + (1-\beta)^2 < \infty \tag{71}$$

So $||w||_2^2 = \mathcal{O}\left(h_g^2\right)$, which combined with the Big-Ω result gives

$$||w||_2 = \Theta\left(h_g\right) \tag{72}$$

8.4. **Proof of Proposition 4.1.**

Proof. One element of the influence vector, w_i, is

$$w_i = \frac{\beta}{N} + (1-\beta)\left(\frac{d_i z_i^{\sigma-1}}{\sum_{k\in N} d_k z_k^{\sigma-1}}\right) \tag{73}$$

As $N \to \infty$, the first term approaches zero, and the distribution of w is determined by the relative product term $d_i z_i^{\sigma-1}$, which means

$$w_i \to \chi d_i z_i^{\sigma-1} \tag{74}$$

$$F_v(x) = F_v\left(\chi d_i z_i^{\sigma-1}\right) \tag{75}$$

$$P(v > x) \to P(\chi d z^{\sigma-1} > x) \tag{76}$$

$$= P(d z^{\sigma-1} > \chi^{-1} x) \tag{77}$$

$$P(v > x) = P(d z^{\sigma-1} > \chi^{-1} x) \tag{78}$$

$$= \int_{\underline{d}}^{\infty} P\left(z > \left[\frac{x}{\chi d}\right]^{1/(\sigma-1)}\right) \mathrm{d}F_d(d) \tag{79}$$

$$= \int_{\underline{d}}^{\infty} C_z \left[\frac{x}{\chi d}\right]^{-\zeta_z/(\sigma-1)} \mathrm{d}F_d(d) \tag{80}$$

$$= \chi^{\zeta_z/(\sigma-1)} C_z x^{-\zeta_z/(\sigma-1)} \int_{\underline{d}}^{\infty} d^{\zeta_z/(\sigma-1)} \mathrm{d}F_d(d) \tag{81}$$

For the integral to exist, we need $\zeta_z/(\sigma-1) < \zeta_d$. If so, it is a constant (independent of x), so combine the other constants into $C_v = \chi^{\zeta_z/(\sigma-1)} C_z \int_{\underline{d}}^{\infty} d^{\zeta_z/(\sigma-1)} \mathrm{d}F_d(d)$, and write

$$P(v > x) = C_v x^{-\zeta_z/(\sigma-1)} \tag{82}$$

TABLE 2. Summary statistics: growth rates and herfindahls from aggregate, industry, and plant levels. The sample runs from 1973-1999, so each measure has 26 years of observations (growth rates from 1974-1999, herfindahls from 1973-1998). Labour productivity statistics aregiven for 1992.

		Obs.	Mean	S.D.
Aggregate	Growth rate	26	.075	.066
	Herfindahl	26	.0566	.0071
Industries	Growth rate	232×26	.067	.107
	Herfindahl	26	.159	.0129
	(Within-I) herfindahl	232×26	.264	.144
Plants	Growth rate	306146	.0739	.329
	Residual growth rate	306146	$-.0697$	.383
	Labour productivity	32710	51771	87407

So v has a power law distribution with parameter $\zeta_z/(\sigma-1)$. If $\zeta_z/(\sigma-1) > \zeta_d$, we need to derive it the other way, and end up with a power law distribution with parameter ζ_d. Therefore the distribution can be expressed by

$$P(v > x) = C_v x^{-\min\{\zeta_d, \zeta_z/(\sigma-1)\}} \tag{83}$$

Or,

$$\log P(v > x) = \log C_v - \min\{\zeta_d, \zeta_z/(\sigma-1)\} \log x \tag{84}$$

□

9. APPENDIX: DATA AND EMPIRICS

9.1. **Growth and herfindahl statistics.** See Table 2.

9.2. **Intensive and Extensive Margins of Volatility.** In the main text, I assume there is no extensive margin of volatility. One may wonder how the results change if I allow for plant entry and exit. To test this empirically, I use a similar decomposition to Di Giovanni, Levchenko and Méjean [11].

First, write sales of plant i at year t as s_{it}. Let I_t be the set of plants operating in year t, and $I_{t/t-1}$ be the set of plants operating in both years t and $t-1$. Then the

Table 3. Intensive vs. Extensive Margin Volatility

	S.D.	Relative S.D.
Aggregate Volatility, $\tilde{\sigma}_A$	.065	1.00
Intensive Volatility, σ_A	.066	1.02
Extensive Volatility, σ_ν^2	.009	.138

log-difference aggregate growth rate of sales is

$$\tilde{g}_{At} \equiv \ln\left(\sum_{i\in I_t} x_{it}\right) - \ln\left(\sum_{i\in I_{t-1}} x_{it-1}\right) \tag{85}$$

$$= \ln\left(\frac{\sum_{i\in I_{t/t-1}} x_{it}}{\sum_{i\in I_{t/t-1}} x_{it-1}}\right) - \left[\ln\left(\frac{\sum_{i\in I_{t/t-1}} x_{it}}{\sum_{i\in I_t} x_{it}}\right) - \ln\left(\frac{\sum_{i\in I_{t/t-1}} x_{it-1}}{\sum_{i\in I_{t-1}} x_{it-1}}\right)\right] \tag{86}$$

$$= g_{At} - \ln\left(\frac{\nu_{t,t}}{\nu_{t,t-1}}\right) \tag{87}$$

where g_{At} is the intensive margin of growth and the other term is the extensive margin of growth. Now aggregate volatility is

$$\tilde{\sigma}_A^2 = \sigma_A^2 + \sigma_\nu^2 - 2\mathrm{Cov}(g_{At}, g_\nu) \tag{88}$$

Calculating each of these in the data, we see that the extensive margin matters little (consistent with the results in Di Giovanni, Levchenko and Méjean [11].