Scalable Expertise*

David Argente[†] Sara Moreira[‡] Ezra Oberfield[§] Venky Venkateswaran[¶]

Penn State Northwestern Princeton NYU Stern

July 2021

Abstract

We develop and test a theory of firm size and scope centered on the scalability of a firm's expertise, i.e. the extent to which its knowledge can be applied across multiple units (products, markets, or locations). In the model, heterogeneous firms choose scope along with the level and scalability of its expertise. A central mechanism of the model is the two-way feedback between scope and scalability, which amplifies the firm's response to changes in demand in a heterogeneous fashion. Firms whose expertise is more scalable have endogenously higher returns to scale, i.e. exhibit stronger responses to the same changes in demand. This also implies that when scalability and size are positively correlated, a rise in industry demand can increase concentration. Despite its parsimony, the model generates a rich set of testable predictions about which firms respond more to common shocks and the margins along which they adjust. We provide support for these predictions using data on multi-product and multi-establishment firms across several different contexts. We show that, in line with the theory, firms with higher size and scope exhibit stronger responses to demand changes. We also construct a novel measure of firm-level scalability using detailed product characteristics and use it to perform deeper and more direct tests of the theory. Finally, we show that innovations by scalable firms tend to diffuse more quickly to other firms, suggesting that changes in scalability can have larger effects on productivity beyond the firm.

Keywords: Firm size, multi-unit firms, scalability, returns to scale, market concentration, diffusion

*We thank Marios Angeletos, Lorenzo Caliendo, Pete Klenow, Giuseppe Moscarini, Esteban Rossi-Hansberg, Chad Syverson as well as numerous seminar and conference participants for useful comments and discussions. We would like to thank Lichen Zhang and Jiarui Cao for excellent research assistance. Empirical results are calculated based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are ours and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

[†]Email: dargente@psu.edu. Address: 403 Kern Building, University Park, PA 16801.

[‡]Email: sara.moreira@kellogg.northwestern.edu. Address: 2211 Campus Drive, Evanston, IL 60208.

[§]Email: edo@princeton.edu. Address: Julis Romo Rabinowitz Building, Princeton, NJ 08544.

[¶]Email: vvenkate@stern.nyu.edu. Address: 7-81 44 W 4th Street, New York, NY 10012.

1 Introduction

The cross-sectional distribution of firms shows enormous variation, not only in size but also in the extent and nature of product offerings. The largest firms in an industry are often several orders of magnitude larger than the median firm. Some firms offer a large number of relatively standardized products while others have more limited scope with products quite distinct from each other. What explains this type of variation? What considerations/factors shape a firm's decision about the scope and nature of its product offerings? Similar questions about the range of firms' activities arise in other contexts. For example, a firm has to decide how many and what types of markets (similar or unique) to enter. Similarly, firms can choose between R&D or advertising outlays that have broad impact (i.e. will be useful throughout the firm) and more targeted ones (e.g. ones that promote or improve a particular product line or unit). Some firms are managed centrally, whereas others provide affiliates with more autonomy. What are the implications of these decisions for the firm size distribution and industry concentration? How do these forces shape firms' responses to changes in the environment?

Motivated by these questions, we embed a novel theory of firm size and returns to scale into an otherwise canonical model of monopolistic competition in the tradition of Hopenhayn (1992) and Melitz (2003). Heterogeneous firms optimally choose *scope* (the number of 'units'), *expertise* (profitability-enhancing knowledge) and *scalability* (the extent to which the firm's expertise can be applied across its units). Higher productivity (or demand) sets off mutually reinforcing increases in scope and scalability, which amplify the usual direct effect of higher productivity/demand. Intuitively, a larger scope increases incentives to invest in scalable expertise. And if expertise becomes more scalable, adding units requires less of an investment in unit-specific expertise and therefore, becomes more attractive. Crucially, this mechanism is more potent for firms that rely more heavily on scalable expertise, implying that these firms have a more elastic scope margin. These more scalable firms tend to be larger when scalable and unit-specific expertise are gross substitutes, which we argue is the empirically relevant case.

The model makes a number of predictions. First, total revenue is a convex function of idiosyncratic productivity (both in logs). Without our mechanism, size is log-linear in productivity, as in Melitz (2003). The intuition stems from *marginal returns to scale* (MRTS), which captures how marginal costs change with scale. Here, MRTS is endogenous. For scalable firms, the ability to increase scope with a smaller 'cost' in terms of foregone expertise moderates the rate at which marginal costs rise. This induces them to make larger adjustments to overall size in response to a given change in demand and/or productivity. Under some conditions, amplification is so powerful that it can generate a power law distribution in size even when the distribution of exogenous productivity/demand is bounded.

The second prediction follows directly from this intuition – scalable firms exhibit stronger responses to common shocks. We show that, faced with a symmetric rise in industry demand, such firms increase their size, scope, and the scalability of their expertise by more than their less scalable counterparts. This result has important implications for the evolution of the firm size distribution in response to changes in the environment: to the extent that scalability and size are positively correlated in the cross-section, a rise in industry demand leads to an increase in the relative size of large firms, or equivalently in market concentration.

The key firm-level variable driving these results is scalability. In fact, scalability remains a sufficient statistic for marginal returns to scale and responsiveness to demand shocks even when we augment the model with additional dimensions of heterogeneity and curvature. In order to test the theory with 'standard' data, we also express the predictions in terms of more easily observed variables, i.e. size and scope, making use of the theory's predicted connection between these variables and scalability. The model predicts that firms with higher size or scope are the most responsive to common demand shocks.

We test these predictions using detailed establishment and product-level data from two distinct sources. We use the Kilts-Nielsen database covering firms and products in the consumer goods industry, and the National Establishment Time Series (NETS) covering firms and establishments for the entire economy. The definition of scope is different depending on the data source. In the Nielsen dataset it is defined as the number of distinct products that the firm sells and in the NETS it is defined as the number of establishments the firm operates. We construct multiple proxies for sectoral demand shocks. Our first proxy is based on the intensity of competition from Chinese imports, following Autor, Dorn and Hanson (2013). Our second proxy exploits variation in home prices across regions in the US during the housing boom and bust cycle in the 2000s, following Mian and Sufi (2011, 2014). In all these cases – across different interpretations of scope and various shifts in demand – we find support in the data for the central prediction of the theory: higher size or scope is associated with a larger elasticity of size and scope with respect to the demand shock.

Next, to perform a deeper and more direct test of the theory, we construct a measure of the scalability of expertise embedded in a firm's product portfolio, exploiting the detailed information about attributes of products in the Kilts-Nielsen dataset. Intuitively, our measure captures the extent to which the products in a firm's portfolio share common characteristics. We confront the model's predictions with respect to scalability, both in the cross-section and in response to shocks and show that they line up well with the observed patterns. In particular, we show that, as predicted by the theory, scalability is positively correlated with overall size and scope. Moreover, firms with high scalability exhibit a greater responsiveness to shocks across three margins – size, scope and scalability. Finally, we use this measure to test whether the pattern of scalability is informative about the effects of demand on market concentration. We show that, in line with the theory, sectors where scalability covaries more strongly with size experienced larger increases in concentration in response to rising demand.

Taken together, these results reveal the ability of the model, despite its parsimony, to successfully match a number of facts, both cross-sectional patterns as well as responses to shocks, across a range of settings. We also argue that the theory offers a unified framework to interpret patterns documented by other researchers in the firm dynamics and trade literature. For example, Aghion, Bloom, Lucking, Sadun and Van Reenen (2021) show that during the Great Recession, firms with more decentralized management outperformed their more centralized counterparts, even controlling for size. We view centralization of management as one manifestation of scalable, rather than unit-specific, expertise. Our theory then predicts that centralized firms would be more sensitive to changes in demand, in line with the observed patterns. Holmes and Stevens (2014) find that competition from China had a more muted impact on small, specialty firms. Our model suggests that one potential reason is that these firms use production techniques and/or knowledge that is less scalable, making them less sensitive to demand changes.[1]

Scalability also has implications from a policy perspective. First, the model predicts that, as with demand and productivity changes, scalable firms also respond more to taxes/subsidies. As such, measures of scalability can inform targeted policy interventions. Second, scalability can interact with, and amplify the effects of, distortions impeding resource allocation. This is particularly relevant when scalable firms face larger distortions (e.g. due to a size-dependent policy, which applies to larger firms but exempts smaller ones). Third, scalability of expertise can also be a mechanism that generates spillovers across markets. A change in one market that alters a firm's incentive to invest in scalable expertise would also affect its activities in other markets.

Our last empirical exercise documents an interesting connection between scalability and diffusion of knowledge. Specifically, we show that product attributes first introduced by scalable firms are more likely to subsequently appear in their competitors' product portfolios. This is consistent with a simple and intuitive idea: scalability is achieved by organizing and codifying knowledge in order to make it more readily usable across multiple products/locations/units within the firm. But, such codification also likely makes it easier for the knowledge to be used outside the boundaries of the firm where it originated. This positive association between scalability and diffusion has important implications, both normative and positive. First, may lead to an externality in the choice of scalability, if firms do not internalize benefits accruing to other firms from their expertise choices. Second, changes in the environment which affect firms' incentives to alter the mix of expertise (e.g. industry demand shocks) can have additional aggregate consequences.

We make several contributions. First, we uncover a novel mechanism – centered on the endogenously chosen scalability of expertise – and show that it implies heterogeneous returns to scale and responsiveness to shocks, even in a canonical model of monopolistic competition with CES demand and CRS production (at the product level). Second, we show that a parsimonious model augmented with this mechanism has the potential to explain a wide range of patterns in the data, both old and new. Third, on the empirical side, we document new facts about heterogeneity in long-run responses to changing demand conditions. We also devise a novel measure of scalability and demonstrate its connection to various margins of adjustment as well as diffusion of knowledge

[1] For more examples, see Section 5.1.

across firms. Our analysis points to the value of incorporating such measures when forecasting the heterogeneous effects of market-wide shocks.

Related Literature: Our analysis is related to a number of different strands in the literature on firm heterogeneity and dynamics. On the theory side, we contribute a novel theory of endogenous size, scope and returns to scale to the firm dynamics literature in the tradition of Hopenhayn (1992) and Melitz (2003). A strand of this literature studies models of endogenous scope, interpreted either as multiple product lines – as in e.g. Klette and Kortum (2004), Akcigit and Kerr (2018), Peters (2020), Garcia-Macia, Hsieh and Klenow (2019), Bernard, Redding and Schott (2011), Dhingra (2013), and Nocke and Yeaple (2014) – or multiple establishments, as in e.g. Luttmer (2011), Holmes (2011), and Cao, Hyatt, Mukoyama and Sager (2020). In some of these models (though not all), firms invest along the intensive margin (e.g. in improving the productivity/quality of each unit). We emphasize scalability of expertise and show that scalable firms are more responsive (on the margins of scope and overall size) to the same shock. We also show the interaction between scope and scalability amplifies the responses to demand shocks and present new evidence in support of this mechanism.

On the empirical front, we paint a rich and nuanced picture of heterogeneous responses to demand shocks, complementing the well-known studies of, e.g., Mian and Sufi (2011, 2014), and Autor, Dorn and Hanson (2013). Other papers studying the effects of demand shocks on firms include Aghion, Bloom, Lucking, Sadun and Van Reenen (2021), Mayer, Melitz and Ottaviano (2020), Hyun and Kim (2019), Lileeva and Trefler (2010), Bustos (2011), and Baldwin and Gu (2009). We contribute to this body of work both by documenting new facts on the heterogeneity in responses and by presenting a simple theory that can help explain a number of the empirical patterns, both existing and new.

Our interest in heterogeneous responses is shared by a few other papers including Moscarini and Postel-Vinay (2012), Fort, Haltiwanger, Jarmin and Miranda (2013), and Crouzet and Mehrotra (2020). The object of interest in these papers is the cyclicality of large versus small firms, i.e. in whether total employment of large firms exhibits a stronger correlation (usually, unconditional) with aggregate economic activity than that of small firms at business cycle frequencies. We, on the other hand, focus on medium- to long-run responses to identified demand shocks and show, both theoretically and empirically, how these responses vary by scalability, scope and size. Our finding that large, scalable firms are the most responsive echoes that of Holmes and Stevens (2014). We also connect to the recent literature on rising concentration in the US economy, e.g. Autor, Dorn, Katz, Patterson and Van Reenen (2020) and Barkai (2020). Rossi-Hansberg et al. (2021), Hsieh and Rossi-Hansberg (2020), and Benkard, Yurukoglu and Zhang (2021) show that this rise has been fueled by expansion of scope (across geographic or product markets) by the largest firms. Several recent papers argue that advancements in information and communication technologies (ICT) may have facilitated the rise of large firms, e.g. Hsieh and Rossi-Hansberg (2020), Aghion, Bergeaud, Boppart, Klenow and Li (2019), De Ridder (2019), Mariscal (2018), Rubinton (2020), and Lashkari, Bauer and Boussard (2018). Indeed, ICT is one example of scalable investment. While these papers have focused on the consequences of a reduction in the cost of ICT technologies, we highlight a different mechanism—that a common increase in demand can, on its own, increase concentration because more scalable firms respond more strongly to the increase in demand. Further, we view our notion of scalable expertise as capturing a broader variety of investments beyond ICT, including management practices, brand building, customer capital, or product features (as we pursue in our own empirical exploration). Our theory is tied concretely to scope, which helps to provide an intuitive way to map the theory to data, and allows us to test the theory directly and indirectly.

The rest of the paper is organized as follows. Section 2 presents our theoretical model and its predictions. We discuss the data and construction of key variables in Section 3 and confront the predictions of the theory with the data in Section 4. Section 5 discusses the role of the theory of scalable expertise within the literature studying firm heterogeneity and dynamics as well as its policy implications, and explores the relationship between scalability and diffusion. Section 6 concludes.

2 Model

This section presents a simple theory in which heterogeneous firms choose the scope and size of their operations. The model makes a number of predictions about features of the cross-sectional distribution, how firms differ in their responses to shocks, and the relationship between scalability and concentration. In the following sections, we take these to data on multi-product and multi-establishment firms.

Our starting point is the canonical model of monopolistic competition by heterogeneous firms widely used in macroeconomics and trade. Firm i in sector j produces a composite output that is a CES aggregate of a continuum of products (index: u):

$$Q_{ij} \;\; = \;\; \left[\int_0^{N_{ij}} (Q_{uij})^{1-\frac{1}{\epsilon}}\, du\right]^{\frac{\epsilon}{\epsilon-1}} \;\; ,$$

where Q_{uij} denotes the quantity of product u, N_{ij} the (endogenous) measure of products and ϵ is the elasticity of substitution across products.

The composite output of sector j is also a CES aggregate of the firm-level composites, with elasticity of substitution θ. This structure implies that firm i faces the following demand function for product u:

$$Q_{uij} = \left(\frac{P_{uij}}{P_{ij}}\right)^{-\epsilon} \left(\frac{P_{ij}}{P_j}\right)^{-\theta} Q_j \;,$$

where Q_j denotes sector-wide output while P_j and P_{ij} are the ideal price indices for the sector and the firm respectively. Production is linear in labor input L_{uij}, i.e.

$$Q_{uij} = A_{ij} Z_{uij} L_{uij} \ .$$

The productivity of the firm has an exogenous (firm-specific) component, A_{ij} and an endogenous part, Z_{uij}, which we will refer to as expertise and describe in detail later. The firm also incurs a fixed cost of operating that depends on the measure of products it offers, $\mathcal{F}_j(N_{ij})$. Appendix A.1 shows that the profits of the firm (revenues net of wages) are given by:

$$\Pi_{ij} = G_j \quad \left(\int_0^{N_{ij}} (A_{ij} Z_{uij})^{\epsilon-1} du \right)^{\frac{\theta-1}{\epsilon-1}} - \mathcal{F}_j(N_{ij}) \tag{1}$$

where G_j is a common (i.e. sector-wide) equilibrium coefficient that scales firms' profits.

In what follows, we lighten notation by suppressing the sector subscript j.

Expertise We model expertise as a productivity shifter, but interpret it more broadly as capturing all forms of knowledge that allows the firm to operate more efficiently, i.e. extract more value from their inputs. Importantly, the expertise relevant to a particular product is a combination of two types of knowledge – *scalable* (or firm-wide) or *local* (u-specific). The former, denoted x_{ij}, is applicable to all the products of the firm, while the latter, denoted y_{ui}, reflects knowledge that is unique to a particular product.

Formally, $Z_{ui} = Z(x_i, y_{ui})$, where $Z(x, y)$ is increasing in both arguments. This flexible formulation can accommodate various interactions between the two forms of knowledge. One feature that will play a prominent role in our analysis is the extent to which the two are substitutable.[2] We impose that $Z(x, y)$ is homogeneous of degree 1, i.e. it can be expressed as $Z(x, y) = yz(k)$ where $k \equiv \frac{x}{y}$ and $z(k) \equiv Z(k, 1)$. We refer to the ratio k as a firm's *scalability* ratio.

Expertise is costly. In the next subsection, we consider a tractable specification for these costs – a capacity constraint – before turning to a more general formulation in Section 2.2.

2.1 Special Case

In this subsection, we make a number of simplifying assumptions that allow us to demonstrate the key economic forces at work in a clean and transparent fashion. In the next subsection, we will relax these assumptions and show that the key results obtain in a more general setting as well. First, we

[2] The formulation $Z(x, y)$ can also accommodate other interesting cases. For example, consider a setting where a firm's expertise is embodied in product characteristics. Scalable expertise might then correspond to characteristics that are common across products. In such a setting, if consumers value variety in product characteristics, they will place a lower value on a product portfolio weighted towards scalable expertise.

set $\theta = \epsilon = 2$ and $\mathcal{F}(N_i) = FN_i$, which simplifies the expression for profits in (1) to

$$\Pi_i = G \int_0^{N_i} A_i Z_{ui} du - FN_i \tag{2}$$

Next, the cost of expertise is assumed to take the form of a capacity constraint:

$$x_i + \int_0^{N_i} y_{ui}\, du \leq 1 \tag{3}$$

One interpretation of this formulation is that expertise requires managerial attention, which is limited. Finally, we assume that scalable and local components enter a firm's expertise with a constant elasticity:

Assumption 1 *Expertise is a constant elasticity function of scalable and local components*

$$Z_{ui} = Z(x_i, y_{ui}) = \left(x_i^{\frac{\sigma-1}{\sigma}} + y_{ui}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \qquad \sigma < 2\ .$$

The upper bound on σ is necessary for an interior solution.

The firm's problem Firm i's problem can thus be written as:

$$\max_{N_i, x_i, y_{ui}} \quad G \int_0^{N_i} A_i Z(x_i, y_{ui}) du - FN_i \qquad \text{s.t.} \qquad x_i + \int_0^{N_i} y_{ui}\, du \leq 1 \tag{4}$$

We begin by noting that, at the optimum, the firm will choose the same level of expertise for all its products, i.e. $Z_{ui} = Z_i$, or equivalently, $y_{ui} = y_i$, which simplifies the problem to

$$\max_{N_i, x_i, y_i} \quad G \;\; A_i N_i Z(x_i, y_i) du - FN_i \qquad \text{s.t.} \qquad x_i + N_i y_i \leq 1 \tag{5}$$

We next reduce the firm's problem as a choice over two dimensions. Since the constraint $x_i + N_i y_i = y_i(k_i + N_i) \leq 1$ will bind at the optimum, we can set $y_i = \frac{1}{k_i + N_i}$ where $k_i \equiv x_i/y_i$ is the scalability ratio. This implies that expertise can be expressed as $Z(x_i, y_i) = y_i z(k_i) = \frac{z(k_i)}{k_i + N_i}$. The firm's problem thus reduces to a choice over scope and the scalability ratio:

$$\max_{N_i, k_i} \quad G\ A_i N_i \frac{z(k_i)}{k_i + N_i} - FN_i \tag{6}$$

The solution is characterized by the following first-order conditions:

$$k_i : \qquad \frac{k_i z'(k_i)}{z(k_i)} = \frac{k_i}{k_i + N_i} \equiv S_i \tag{7}$$

$$N_i : \qquad GA_i \frac{z(k_i)}{k_i + N_i} - F = GA_i \frac{N_i}{k_i + N_i} \frac{z(k_i)}{k_i + N_i}\ , \tag{8}$$

Equation (7) equates the cost and benefits of making expertise more scalable (for a given scope N_i). The right hand side of the equation is the share of capacity devoted to scalable expertise, $\frac{k_i}{k_i+N_i} = \frac{x_i}{x_i+N_i y_i}$, which we denote as S_i. At the optimum, S_i is equated to the elasticity of expertise to scalable knowledge, $\frac{k_i z'(k_i)}{z(k_i)} = \frac{xZ_x}{Z} = \frac{k^{\frac{\sigma-1}{\sigma}}}{k^{\frac{\sigma-1}{\sigma}}+1}$. This elasticity is increasing (decreasing) in k_i if the elasticity of substitution between the two forms of expertise, σ, is larger (smaller) than 1. In other words, the share S_i and the scalability ratio k_i are positively (negatively) related when x and y are gross substitutes (complements).

Equation (7) can also be rearranged as a relationship between scope and the scalability ratio:

$$N_i \;=\; \frac{z(k_i) - k_i z'(k_i)}{z'(k_i)} \tag{9}$$

The expression on the right-hand side of (9) is the marginal rate of substitution $\frac{Z_y}{Z_x}$. Thus, at the optimum, the firm equates the marginal benefit of increasing the scalable component of expertise, NZ_x, to that of the non-scalable component $\frac{NZ_y}{N} = Z_y$. Note that the marginal rate of substitution $\frac{Z_y}{Z_x}$ is increasing in the scalability ratio k_i, so this equation describes a positive relationship between scope and scalability. Let $\mathbb{K}(N)$ denote the optimal choice of k for a given choice of scope. Note that $\frac{d\ln\mathbb{K}}{d\ln N} = \frac{d\ln k}{d\ln\frac{Z_y}{Z_x}} = \sigma$.

Given a scalability ratio, equation (8) determines the optimal scope. Here, the firm faces a tradeoff: larger scope means more units to bring in profit (net of fixed costs), but at the cost of lower expertise. Since capacity is fixed, raising scope (holding the scalability ratio constant) requires cutting back on expertise. This equation can be rearranged to yield the following relationship between scope and k_i:

$$N_i = \mathbb{N}(k_i; A_i) \equiv \; k_i\left[\left(\frac{GA_i}{F}\frac{z(k_i)}{k_i}\right)^{1/2} - 1\right]. \tag{10}$$

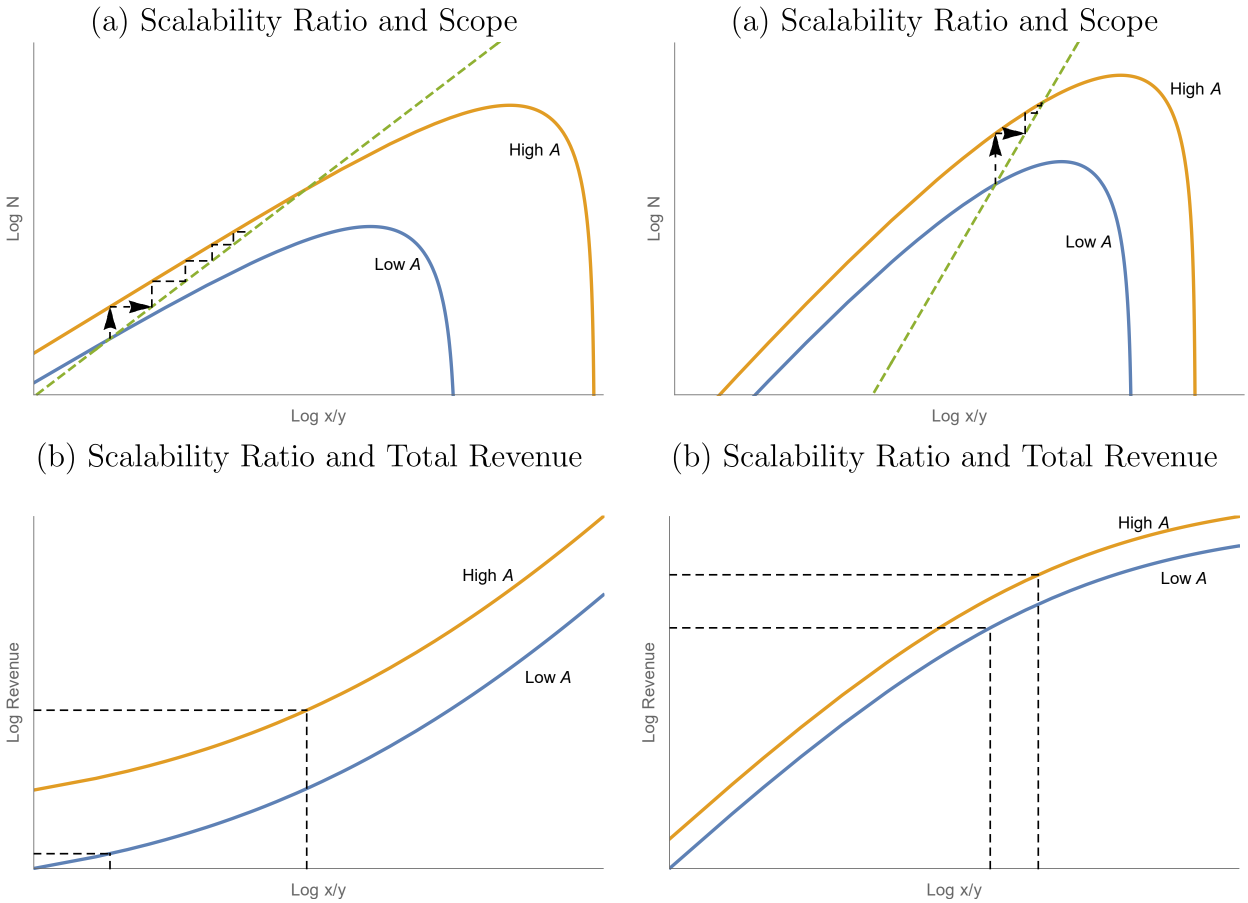
The left panel of Figure 1 shows the two curves, (9) and (10), on a log-log plot. The solid line displays (9), while the dashed ones show (10) for two levels of A_i. As mentioned earlier, the slope of the solid line is $\frac{1}{\sigma}$. The dashed line can be hump-shaped, but, an interior solution will always be in the upward sloping part.

The panel also depicts how the optimal choice of scalability and scope amplify exogenous differences in productivity (or equivalently, demand). A higher A_i leads the dashed line to shift upward (note that the solid line, which depicts equation (9), is independent of A_i). The arrows show successive rounds of adjustment. Higher productivity induces a larger scope, holding the scalability ratio fixed (the first vertical arrow). This in turn leads to an increase in the scalability ratio (the first horizontal arrow), which feeds back into further scope increases and so on.

This amplification mechanism can turn explosive under some conditions – specifically, for high enough σ and/or A_i. In these cases, the firm finds it optimal to set $k_i, N_i \to \infty$. Intuitively, if the two types of expertise are very substitutable and the exogenous productivity is sufficiently high, the

Figure 1: Scalability and Scope

Substitutes ($\sigma > 1$) **Complements** ($\sigma < 1$)



Note: The dashed line in panel (a) plots (9), while the solid lines plot (10) for two values of A_i. The solid lines in (b) plot revenue $\mathbb{R}\left(k, \mathbb{N}(k; A_i); A_i\right)$, as a function of the scalability ratio for two different values of A_i. In the left two panels (Substitutes), we use $\sigma = 1.5$ and $\frac{GA_i}{F} = 0.2$ and $\frac{GA_i}{F} = 0.4$ for the two solid lines. For the right two panels, the corresponding values are $\sigma = 0.67$, $\frac{GA_i}{F} = 10$ and $\frac{GA_i}{F} = 20$.

incentives to substitute towards scalable expertise are very strong, to the point where we no longer have an interior solution. For now, we abstract from this possibility – in the more general version, this will be precluded with appropriate restrictions on primitives.

To show this amplification formally, we first express the optimal scope using the functions, $\mathbb{K}(\cdot)$ and $\mathbb{N}(\cdot)$, as $N_i = \mathbb{N}(\mathbb{K}(N_i); A_i)$. Differentiating implicitly with respect to A_i and rearranging gives:

$$\frac{d \ln N}{d \ln A} = \frac{\partial \ln \mathbb{N}}{\partial \ln A} + \frac{\partial \ln \mathbb{N}}{\partial \ln k} \frac{d \ln \mathbb{K}}{d \ln N} \frac{d \ln N}{d \ln A}$$

$$\Rightarrow \quad \frac{d \ln N}{d \ln A} = \frac{\frac{\partial \ln \mathbb{N}}{\partial \ln A}}{1 - \frac{\partial \ln \mathbb{N}}{\partial \ln k} \frac{d \ln \mathbb{K}}{d \ln N}} \tag{11}$$

The denominator in the expression in (11) shows the amplification. Since $\frac{d \ln \mathbb{K}}{d \ln N} > 0$ and, at any interior optimum, $\frac{\partial \mathbb{N}}{\partial k} > 0$, an increase in A_i sets off the sequence of mutually reinforcing changes we saw in the figure.

We can use (10) to derive[3]

$$\frac{\partial \ln \mathbb{N}}{\partial \ln k} = \frac{1}{2} \tag{12}$$
$$\frac{\partial \ln \mathbb{N}}{\partial \ln A} = \frac{1}{2}\frac{1}{1-S_i} . \tag{13}$$

Equation (13) shows that firms who use a relatively larger share of their capacity for scalable expertise (high S_i) adjust their scope by more in response to changes in A (holding the ratio k_i constant). Intuitively, increasing scope involves not only an additional fixed cost F but also requires an investment in unit-specific expertise. Since capacity is fixed, this leads to a trade-off between scope and expertise. When capacity is used mostly for scalable expertise (S_i is high), increasing scope exerts less pressure on the capacity constraint, i.e. induces a smaller reduction in expertise. In other words, high S_i firms face a lower effective cost of adding units and, therefore, have a more elastic scope margin. As we will show, this intuition extends to a more general setting: firms that devote a larger share to scalable expertise will adjust their scope by more in response to changes in demand.

Finally, we note that $\frac{d \ln \mathbb{K}}{d \ln N} = \sigma$. Using this, (12), and (13), (11) can be expressed as:

$$\frac{d \ln N}{d \ln A} = \frac{\frac{\partial \ln \mathbb{N}}{\partial \ln A}}{1-\frac{\sigma}{2}} = \frac{1}{2-\sigma}\left(\frac{1}{1-S_i}\right) \quad > \ 0 . \tag{14}$$

Thus, the elasticity of scope w.r.t productivity is increasing in the scalable share, S_i. Recall that when $\sigma > 1$, the scalable share S_i is positively related to the scalability ratio k_i. In this case, the empirically relevant one, the elasticity of scope is increasing in k_i.

This intuition directly extends to the responsiveness of the scalability ratio itself:

$$\frac{d \ln k}{d \ln A} = \sigma \frac{d \ln N}{d \ln A} = \frac{\sigma}{2-\sigma}\left(\frac{1}{1-S_i}\right) \quad > \ 0 .$$

Next, we turn to the effect of productivity on revenue $R_i = GA_i \frac{N_i z(k_i)}{k_i + N_i} \equiv \mathbb{R}(k_i, N_i; A_i)$. Changes

[3] To derive (12), re-arrange (8) as $\frac{(k+\mathbb{N}(k;A))^2}{kz(k)} = \frac{GA}{F}$, take logs and differentiate implicitly to obtain $2\frac{k+N\frac{\partial \ln \mathbb{N}}{\partial \ln k}}{k+N} - 1 - \frac{kz'(k)}{z(k)} = 0$. Solving for $\frac{\partial \ln \mathbb{N}}{\partial \ln k}$ and noting that $\frac{kz'(k)}{z(k)} = S$ from (7) yields

$$\frac{\partial \ln \mathbb{N}}{\partial \ln k} = \frac{\frac{1}{2}(1+S) - \frac{k}{k+N}}{\frac{N}{k+N}} = \frac{\frac{1}{2}(1+S) - S}{1-S} = \frac{1}{2}.$$

in productivity has both direct and indirect effects on this object:

$$\frac{d\ln R}{d\ln A} = \underbrace{\frac{\partial \ln \mathbb{R}}{\partial \ln A}}_{=1} + \underbrace{\frac{\partial \ln \mathbb{R}}{\partial \ln k}}_{=0}\underbrace{\frac{d\ln k}{d\ln A}}_{>0} + \underbrace{\frac{\partial \ln \mathbb{R}}{\partial \ln N}}_{>0}\underbrace{\frac{d\ln N}{d\ln A}}_{>0}$$

The first term is the direct effect, while the other two terms capture the indirect effects (arising through induced changes in scope and scalability). Since k is chosen to maximize $\frac{z(k)}{k+N}$, an envelope argument implies that, at the optimum, $\frac{\partial \ln \mathbb{R}}{\partial \ln k} = 0$.

The last term reflects the effect of a change in scope on revenue. Scope directly raises revenue, but also comes at the cost of lower expertise. The former dominates, so $\frac{\partial \ln \mathbb{R}}{\partial \ln N} > 0$. More importantly, the reduction in expertise from increased scope is smaller when the share of scalable expertise is high.

$$\frac{\partial \ln \mathbb{R}}{\partial \ln N} = 1 - \frac{N_i}{k_i + N_i} = S_i\ .$$

Along with (14), this implies the total effect of productivity on revenue is given by:

$$\frac{d\ln R}{d\ln A} = 1 + \frac{1}{2-\sigma}\frac{S_i}{1-S_i}\ \ >\ \ 0.$$

Thus, scalable firms, i.e. those with high S_i, adjust their overall size by more to a given change in productivity. This will play an important role in shaping the firm size distribution, as we will discuss in more detail below.

We summarize these results in the following proposition:

Proposition 1 *Firms with higher productivity have higher scalability, scope, and total revenue.*

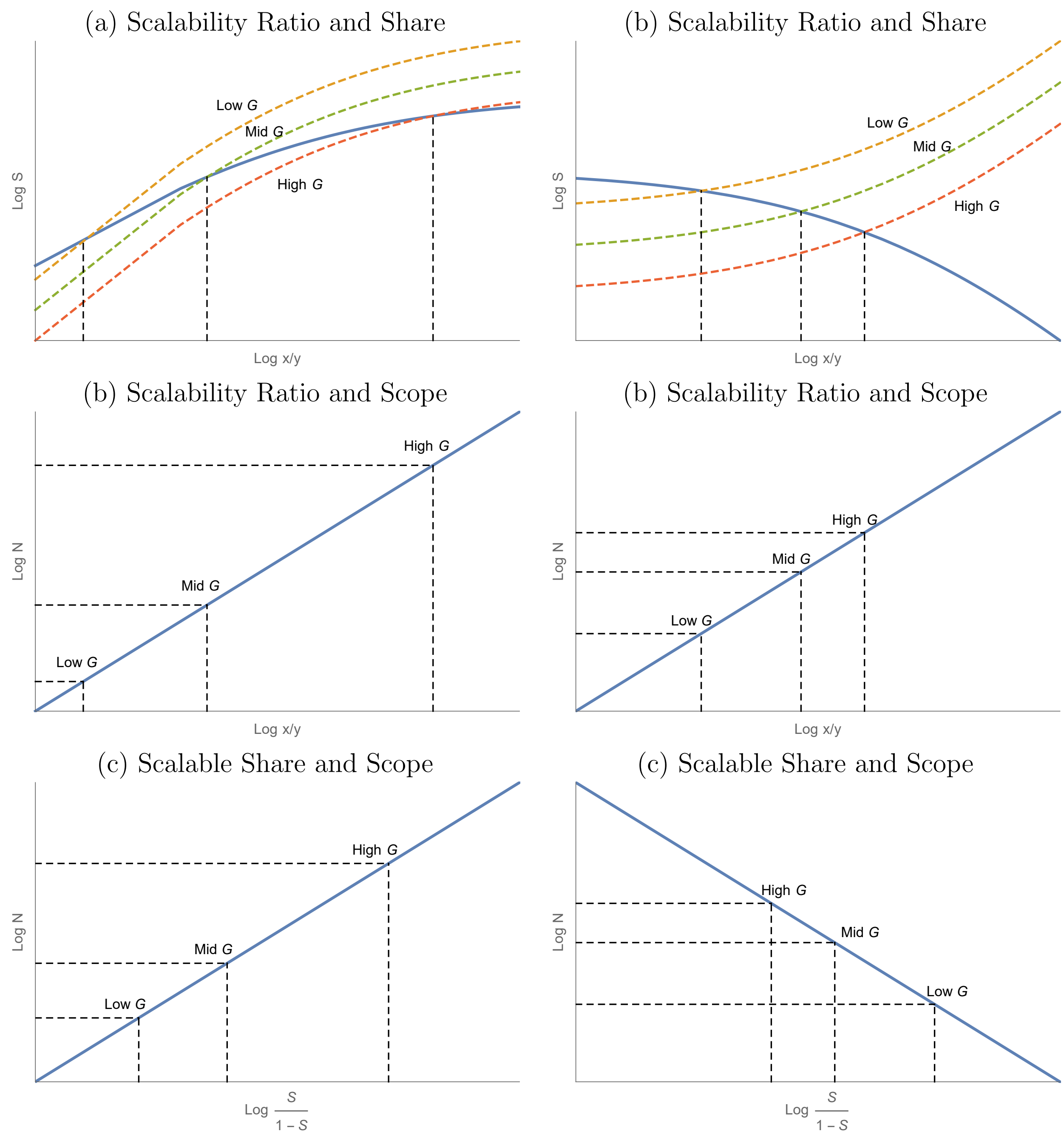
The impact of productivity on revenue *per unit* is more subtle. As with total revenue, there are direct and indirect effects. The direct effect is equal to $\frac{\partial \ln \mathbb{R}}{\partial \ln A} = 1$. The indirect effect (stemming from the change in scope) is negative. If $\sigma < 1$, the direct effect dominates so revenue per unit rises with productivity. On the other hand, if $\sigma > 1$, the feedback mechanism amplifying scope and scalability changes is so powerful that the indirect effect dominates. However, when we move to the more general model below, revenue per unit may rise even if $\sigma > 1$.

Response to Common Shocks

Next, we show that the forces discussed above also influence how firms respond to a common shock. Specifically we will analyze the effects of an increase in G, the sector-wide profit shifter. This can stem, for example, from an increase in industry demand (which is not entirely absorbed by the entry of new firms). We will show that firms differ systematically in their responses to this common shock, both in terms of the magnitude and the margins of the response.

Figure 2: Response to a Shock

Substitutes ($\sigma > 1$) **Complements** ($\sigma < 1$)



Note: The dashed lines in (a) plot (16), while the solid line depicts (15). Panel (b) shows (9) in logs, while the relationship in (c) is derived from (7). In the left panels (Substitutes), we use $\sigma = 1.5$ and $\frac{GA_i}{F} = 0.2$, $\frac{GA_i}{F} = 0.4$ and $\frac{GA_i}{F} = 0.8$ for the Low G, Mid G and High G lines. For the right panels, the corresponding values are $\sigma = 0.67$, $\frac{GA_i}{F} = 10$, $\frac{GA_i}{F} = 20$ and $\frac{GA_i}{F} = 40$.

Figure 2 shows the response of key variables to a change in G. The three left panels show the case where $\sigma > 1$, while the right panels show the same variables when $\sigma < 1$. In each panel, the dashed lines correspond to three equally spaced levels of (log) industry demand: going from the Low G line to the Mid G line represents the same change in $\ln G$ as going from Mid G to High G .[4] The horizontal axis of each panel is the scalability ratio, k.

The top panels show the effect of G on the (logs) of the scalability ratio and the scalable share S. The solid line shows combinations of k and S that are consistent with the optimal choice of scalability, i.e. equation (7). The dashed lines show the combinations of k and S consistent with the optimal choice of scope. Formally, the lines are obtained by expressing (7)-(8) as follows:

$$\ln S_i = \ln \frac{k_i z'(k_i)}{z(k_i)} \tag{15}$$

$$\ln S_i = -\frac{1}{2} \ln \left(\frac{z(k_i)}{k_i} A_i \right) - \frac{1}{2} \ln \left(\frac{G}{F} \right) , \tag{16}$$

The right hand side of (15) is the elasticity of expertise to the scalable component, i.e. $\frac{xZ_x}{Z}$. This is increasing (decreasing) in the scalable ratio k_i when the two types of expertise are relatively complementary (substitutable). In other words, (15) defines a positive (negative) relationship between k and S when σ is smaller (larger) than 1. These two cases are depicted in the two top panels of Figure 2. The right hand side of (16) is increasing in k_i for all σ, so the dashed lines are increasing in both panels. Their curvature does depend on σ – the curves are concave (convex) when σ is larger (smaller) than 1.

What happens when G rises? Equation (15) remains the same. Since G enters (16) log-linearly, increases in $\ln G$ induce parallel, downward shifts in the dashed line. In both the left and right panels, this raises the scalability ratio, working through the same channels we discussed in the context of changes in idiosyncratic productivity. However, the effect on the scalable share S_i depends on σ: when $\sigma > 1$ (the left panel), the incentives to shift capacity towards scalable expertise are very strong, resulting in their share rising with G. The opposite happens when the two types of expertise are relatively complementary ($\sigma < 1$, the right panel).

More importantly, the figure shows that these adjustments are systematically larger when the share S is high to begin with. This occurs in both panels – towards the right side of the graph, i.e. at higher levels of k in the substitutes case, and at lower k when the two forms of expertise are relatively complementary. The intuition is closely related to the one discussed above – a larger share of capacity being used for scalable expertise implies that increases in scope come at a relatively lower cost in terms of foregone expertise, which makes the scope margin relatively more elastic. This intuition is displayed in the remaining panels of the figure. The middle panel shows that the magnitude of the changes in scope are increasing (decreasing) in the scalability ratio k when the two types of expertise are substitutable (complementary). The bottom panel depicts the changes in

[4] The parameter values are the same as for Figure 1.

scope and a transformation of the share S. It shows that, for all σ, a higher share (or equivalently, a higher $\frac{S}{1-S} = (\sigma - 1)\, k$) is associated with larger percentage changes in scope (as well as in $\frac{S}{1-S}$).

The above discussion focuses on the differences between going from Low G to Mid G and from Mid G to High G. But, it is easy to see that the same logic also extends to the cross-section: these forces also induce heterogeneous responses to a common change in G. Specifically, firms with a high scalable share S_i make larger adjustments to scope and scalability ratio in response to common shocks. Formally,

$$\frac{d\ln N_i}{d\ln G} = \frac{1}{2-\sigma}\left(\frac{1}{1-S_i}\right) \qquad \frac{d\ln k_i}{d\ln G} = \frac{\sigma}{2-\sigma}\left(\frac{1}{1-S_i}\right) \tag{17}$$

Finally, we can also characterize the effects of the common shock on total size.

$$\frac{d\ln R_i}{d\ln G} = 1 + \frac{1}{2-\sigma}\left(\frac{S_i}{1-S_i}\right) \ . \tag{18}$$

Thus, the response of total revenue also rises with S_i. The intuition is similar to the effects of the firm-specific productivity shifter A_i. The following proposition collects these results.

Proposition 2 *Firms with a higher scalable share exhibit higher elasticities of scalability, scope, and size to common demand shocks. If $\sigma > 1$, these firms have higher k_i, so the elasticities of scalability, scope, and size are rising in scalability.*

Marginal Returns to scale

Underlying these results is the way in which scalable expertise determines a firm's *marginal returns to scale*. We define this as the negative of the elasticity of marginal cost with respect to output:

$$MRTS_i \equiv -\frac{Q_i c_i''(Q_i)}{c_i'(Q_i)} \ ,$$

where $c_i(Q_i)$ denotes firm i's cost function.

A higher $MRTS$ means that marginal cost rises more slowly with output, which in turn implies a larger response of output and revenue to demand. To see this formally, consider a firm that faces an iso-elastic demand function $Q_i = D_i P_i^{-\theta}$, where D_i is a demand shifter. The firm chooses output Q_i to maximize profit, $D_i^{\frac{1}{\theta}} Q_i^{\frac{\theta-1}{\theta}} - c_i(Q_i)$. The optimality condition is $\frac{\theta-1}{\theta} D_i^{\frac{1}{\theta}} Q_i^{-\frac{1}{\theta}} = c_i'(Q_i)$. The change in output in response to a change in the demand shifter is

$$\frac{d\ln Q_i}{d\ln D_i} = \frac{1}{1+\theta\frac{Q_i c_i''(Q_i)}{c_i'(Q_i)}} = \frac{1}{1-\theta\cdot MRTS_i} \ .$$

It follows that the response of total revenue to a change in D_i is given by

$$\frac{d\ln D_i^{\frac{1}{\theta}} Q_i^{\frac{\theta-1}{\theta}}}{d\ln D_i} = \frac{1}{\theta} + \frac{\theta-1}{\theta}\frac{d\ln Q_i}{d\ln D_i} = \frac{1}{\theta} + \frac{\theta-1}{\theta}\left(\frac{1}{1-\theta\cdot\ MRTS_i}\right) ,$$

That is, for a given the curvature of demand, returns to scale are sufficient to determine responses of output and revenue to a change in demand.[5]

In our setting, each firm has many products so output here refers to the composite good $Q_i = \left(\int_0^{N_i} Q_{iu}^{\frac{\epsilon-1}{\epsilon}} du\right)^{\frac{\epsilon}{\epsilon-1}}$. Noting the symmetry across products, the cost of producing Q_i solves:

$$c_i(Q_i) = \min_{x_i, y_i, N_i, L_i, Q_{iu}} \ N_i W L_i + F N_i ,$$

subject to the capacity constraint $x_i + N_i y_i \leq 1$, technology $A_i Z(x_i, y_i) L_i \leq Q_{iu}$, and within firm aggregation $Q_i \geq N_i^{\frac{\epsilon}{\epsilon-1}} Q_{iu}$. In this version of our model (with $\epsilon = 2$), the marginal returns to scale is given by (see Appendix A.3 for details):

$$MRTS_i = -\frac{Q_i c_i''(Q_i)}{c_i'(Q_i)} = \frac{S_i}{2-(1-S_i)\sigma} = \frac{1}{\frac{2-\sigma}{S_i}+\sigma} . \tag{19}$$

which is increasing in S_i. Intuitively, a higher S_i means the firm can increase scope with a smaller 'cost' in terms of foregone expertise. This moderates the rate at which its marginal costs rise with scale. This is stated formally in the following result:

Proposition 3 *Firms with higher scalable share of expertise have higher marginal returns to scale.*

The Size Distribution

The cross-sectional distribution of firms shows enormous variation in size. Understanding the sources of this variation is a classic question in economics, and has attracted renewed interest given recent increases in concentration.

In this subsection, we show how scalability of expertise shapes the firm size distribution. It essentially amplifies the effects of productivity on size: as productivity rises, firms endogenously change the composition of their expertise (as well as scope), which results in a more pronounced increase in size. This effect is more powerful at higher levels of productivity, inducing a convex relationship between size (total revenue) and productivity (both in logs). In contrast, the standard Hopenhayn-Melitz framework with iso-elastic demand leads to a log-linear relationship between total revenue and productivity. The following result states this formally.

[5]There are other notions of returns to scale in the literature. One common measure, used e.g. by Lashkari et al. (2018), is the ratio of average cost to marginal cost, or equivalently the reciprocal of the elasticity of total cost with respect to output. As discussed in Lashkari et al. (2018), this object, along with the markup, determines the revenue-to-cost ratio for the firm. Our notion of returns to scale, on the other hand, is the relevant one for characterizing responsiveness to demand shocks.

Proposition 4 *If $\sigma > 1$, size is strictly convex in productivity (both in logs), i.e.* $\frac{d^2 \ln R}{d(\ln A)^2} > 0$.

Under some conditions, this mechanism can even generate an unbounded size distribution – in particular, one following a power law – from a bounded productivity distribution. This requires additional assumptions on the productivity distribution. We begin with the following definition.

Definition 1 *A probability distribution function $H(A)$ has a constant right elasticity at upper bound $\bar{A}$ if* $\lim_{A \nearrow \bar{A}} \frac{\log[1-H(A)]}{\log\left[1-A/\bar{A}\right]} = \kappa$.

One example of a such a distribution function is the beta distribution with bounds $\underline{A}$ and $\bar{A}$ and parameters η and κ.[6] The next result states the conditions under which we obtain a power law distribution in size.

Proposition 5 *Suppose that $\sigma > 1$ and that A_i is distributed according to a distribution with a constant right elasticity at upper bound $\bar{A}$, where $\bar{A} \leq A^* \equiv \frac{F}{GZ_x(1,0)}$. Then,*

1. *If $\bar{A} < A^*$, then the distribution of revenue is bounded.*

2. *if $\bar{A} = A^*$, then the size distribution is unbounded and follows a power law:*

$$\lim_{R \to \infty} \frac{\log \Pr\left(Size > R\right)}{\log R} = -\kappa(\sigma - 1)$$

Thus, the amplification from the optimal choice of scalability and scope generates firms of unbounded sizes despite a finite upper bound for productivity. Of course, this result is somewhat special[7], but it serves as a stark illustration of the potential of scalable expertise to amplify small differences in productivity into enormous differences in firm size.

Growth and Concentration

There has been much interest in rising concentration of late, driven by recent trends in the US data over the past few decades. Here, we show how scalability of expertise induces a rise in concentration in response to a common increase in demand. This occurs because, when the two types of expertise are gross substitutes, the largest firms are also the most scalable and therefore, respond most to the higher demand and gain market share at the expense of their smaller, less scalable counterparts.

Proposition 6 *Let $\sigma > 1$. For a fixed set of firms, an increase in G leads to a first order stochastically dominant shift in the distribution of revenue shares among those firms.*

[6]The beta distribution has density $h(A) = \frac{\Gamma(\eta+\kappa)}{\Gamma(\eta)\Gamma(\kappa)} \frac{(A-\underline{A})^{\eta-1}(\bar{A}-A)^{\kappa-1}}{(\bar{A}-\underline{A})^{\eta+\kappa-1}}$.

[7]In the general model presented below, the power law result requires, in addition to the restriction on the upper bound $\bar{A}$, two knife-edge assumptions on curvature parameters, $\frac{\phi}{\omega} + \frac{\psi}{\mu\gamma} = 1$ and $\phi = \frac{\psi}{\mu}$. Nevertheless, the underlying force remains present in the more general setting, leading to larger size gaps among higher productivity firms than among lower productivity firms (for the same productivity gap).

This follows directly from two facts (i) firms with more revenue have expertise that is more scalable (Proposition 1) (ii) firms whose expertise is more scalable experience a larger increase in revenue in response to the common shock (proposition 2).

This naturally has important implications for concentration. In particular, it implies that a symmetric demand increase will raise concentration – more precisely, any measure of concentration that depends only on the distribution of sales shares and rises with a shift in sales from a lower-ranked firm to a higher-ranked firm. This includes several commonly used measures like the Gini coefficient, the Herfindahl-Hirschman Index, and concentration ratios.

Corollary 1 *Let $\sigma > 1$. Among a fixed set of firms, a common increase in G raises the concentration of sales among those firms.*

2.2 General Model

In this subsection, we analyze a more general version of the model. Specifically, we relax the assumptions on curvature and heterogeneity made in the previous subsection. The firm's problem is now given by:

$$\max_{x_i, y_i, N_i} \qquad G N_i^{\phi} \left(A_i Z(x_i, y_i)\right)^{\psi} - F N_i^{\omega} - H \cdot \left[x_i^{\mu} + N_i y_i^{\mu}\right]^{\gamma} \ . \tag{20}$$

There are several differences relative to the version analyzed in the previous subsection: (i) arbitrary elasticities of substitution both across products of a given firm as well as across composite goods of different firms within a sector (these are not necessarily equal to each other, so $\phi = \frac{\theta-1}{\varepsilon-1} \neq 1$)[8] (ii) a more flexible specification for the fixed operating costs, $\mathcal{F}(N_{ij})$ (iii) expertise is now subject to an explicit cost (instead of a capacity constraint) and (iv) the function $Z(x, y)$ is no longer restricted to be of the CES form (the only restriction is that it is CRS). Note that this more general formulation nests the previous one with the parameters $\phi = \psi = \omega = 1$ and $\gamma \to \infty$.

We begin by imposing regularity conditions. Let $\sigma(x/y)$ be the elasticity of substitution between x and y – note that for a general CRS function $Z(x, y)$, the elasticity is a function of the scalability ratio x/y. Let $\tilde{\sigma}\left(\tilde{x}/\tilde{y}\right)$ be the elasticity of substitution associated with the transformed CRS function $\tilde{Z}\left(\tilde{x}, \tilde{y}\right) \equiv Z\left(\tilde{x}^{1/\mu}, \tilde{y}^{1/\mu}\right)^{\mu}$. As we show in Appendix A.6, the two objects σ and $\tilde{\sigma}$ are related by $\frac{\tilde{\sigma}(k)-1}{\tilde{\sigma}(k)} = \frac{1}{\mu} \frac{\sigma\left(k^{1/\mu}\right)-1}{\sigma\left(k^{1/\mu}\right)}$. Note that $\tilde{\sigma}$ is between 1 and σ, and approaches 1 as μ grows large.

Assumption 2 *The parameters of the model satisfy*

(i) $1 > \frac{\phi}{\omega} + \frac{\psi}{\mu\gamma}$.

(ii) $\tilde{\sigma} \leq 1 + \omega$.

[8] In Appendix A.2, we present a version in which a firm's productivity varies exogenously across its products. When such variation follows a power law, consistent with the evidence in Bernard, Redding and Schott (2011), the firm's objective takes the same reduced form as (20).

(iii) If $\sigma > 1$ then $\phi \geq \psi/\mu$.

These restrictions are sufficient to guarantee there is a unique solution to the problem and that the solution is interior (see Appendix). Assumption 2*(i)* ensures that the curvature of cost of scope and expertise is larger than the curvature of the corresponding benefit. Assumption 2*(ii)* ensures that the feedback between increasing scope and increasing scalability is not explosive. Assumption 2*(iii)* ensures that the feedback between reducing scope and increasing local knowledge is not explosive.

As in the special case, firms with higher productivity or that face increased demand raise size, scope and scalability ratio. Whether they raise the share of scalable expertise, $S_i \equiv \frac{x_i^\mu}{x_i^\mu + N_i y_i^\mu}$ depends on whether scalable and local expertise are gross substitutes or complements.

Proposition 7 *Suppose Assumption 2 holds. Then*

1. *Firms with higher A_i have higher size, scope, and scalability ratio. They also have higher scalable shares if $\sigma > 1$.*
2. *Firms respond to an increase in demand by raising size, scope, and scalability ratio: $\frac{d\ln R}{d\ln G} > 0$, $\frac{d\ln N}{d\ln G} > 0$, and $\frac{d\ln x/y}{d\ln G} > 0$. The scalable share S also rises if $\sigma > 1$.*

Note that the primitive heterogeneity is in the form of a 'level' shifter, A_i, which interacts with the various sources of curvature (which are assumed to be homogeneous across firms) to generate endogenous heterogeneity in effective curvature (rather than imposed exogenously). Moreover, even with all these rich interactions, responses to shocks are governed by a single variable, the scalable share of expertise. That is, a firm's scalable share of expertise is a sufficient statistic for how its size, scope, and scalability change with demand.[9]

Proposition 8 *Scalability ratio is a sufficient statistic for the elasticities of scope, size, size per unit, scalability ratio, and the scalable share of expertise to demand.*

We now restrict attention to the empirically relevant region of the parameter space

Assumption 3 *We assume that $1 < \tilde{\sigma} < 1 + \frac{\phi}{1 - \frac{\psi}{\mu\gamma}}$, and that σ is non-decreasing in scalability ratio.*

Proposition 9 *Suppose Assumptions 2 and 3 hold. In response to the same change in G, firms with higher scalability ratio (or, equivalently, higher scalable share) have larger changes in size, scope, scalability ratio, and the scalable share:*

$$\frac{d}{d(x/y)}\left(\frac{d\ln R}{d\ln G}\right) > 0,\ \frac{d}{d(x/y)}\left(\frac{d\ln N}{d\ln G}\right) > 0,\ \frac{d}{d(x/y)}\left(\frac{d\ln x/y}{d\ln G}\right) > 0,\ \frac{d}{d(x/y)}\left(\frac{d\ln \frac{S}{1-S}}{d\ln G}\right) > 0$$

$$\frac{d}{dS}\left(\frac{d\ln R}{d\ln G}\right) > 0,\ \frac{d}{dS}\left(\frac{d\ln N}{d\ln G}\right) > 0,\ \frac{d}{dS}\left(\frac{d\ln x/y}{d\ln G}\right) > 0,\ \frac{d}{dS}\left(\frac{d\ln \frac{S}{1-S}}{d\ln G}\right) > 0$$

[9] As we will show in Section 2.3, this property turns out to be robust to richer heterogeneity in level shifters.

In many settings, it may not be possible to directly observe a firm's scalability ratio or scalable share. Nevertheless, if $\sigma > 1$, then firms with higher scalability also have higher size and scope. A simple consequence of Proposition 9 is that firms with higher size or scope will exhibit stronger responses to the same increase in demand.

Corollary 2 *Suppose Assumptions 2 and 3 hold. In response to the same change in G, firms with higher size or higher scope have larger changes in size, scope, scalability ratio, and scalable share of knowledge:*

$$\frac{d}{dR}\left(\frac{d\ln R}{d\ln G}\right) > 0,\ \frac{d}{dR}\left(\frac{d\ln N}{d\ln G}\right) > 0,\ \frac{d}{dR}\left(\frac{d\ln x/y}{d\ln G}\right) > 0,\ \frac{d}{dR}\left(\frac{d\ln \frac{S}{1-S}}{d\ln G}\right) > 0$$
$$\frac{d}{dN}\left(\frac{d\ln R}{d\ln G}\right) > 0,\ \frac{d}{dN}\left(\frac{d\ln N}{d\ln G}\right) > 0,\ \frac{d}{dN}\left(\frac{d\ln x/y}{d\ln G}\right) > 0,\ \frac{d}{dN}\left(\frac{d\ln \frac{S}{1-S}}{d\ln G}\right) > 0\ .$$

2.3 Adding More Heterogeneity

In this subsection, we allow for richer heterogeneity across firms. The firm's objective is now assumed to be given by:

$$\pi_i = \max_{N,x,y}\ GN^{\phi}(A_i Z(x,y))^{\psi} - F_i N^{\omega} - H_i\left[\left(\frac{x}{a_i^x}\right)^{\mu} + N\left(\frac{y}{a_i^y}\right)^{\mu}\right]^{\gamma}$$

Relative to the baseline version in Section 2.1-2.2, now firms are also allowed to vary in the cost of introducing additional units, F_i, in the overall cost of expertise, H_i, as well as in the cost of each type of expertise, through the parameters a_i^x and a_i^y.

These additional dimensions of heterogeneity break the one-to-one mapping between size, scope, and scalability in the cross-section. Nevertheless, the respective elasticities of size, scope, scalability ratio, and the scalable share with respect to demand are exactly the same as in Section 2.2. In particular, the scalable share of expertise (or equivalently scalability ratio) remains a sufficient statistic for these elasticities. Under Assumptions 2 and 3, in response to the same change in demand, firms with higher share of scalable expertise (or equivalently more scalable firms) have larger increases in size, scope, and scalability ratio. That is, Propositions 8 and 9 continue to hold.

We next turn to the implications for concentration of rising industry demand. Recall that, in the general model with $\sigma > 1$, a common rise in demand had a disproportionate effect on the largest firms because they were the most scalable. With additional dimensions of heterogeneity, scalability still drives responsiveness to demand but is not necessarily tightly linked to size. Nevertheless, how scalability covaries with size carries information about the effect of demand on concentration. Loosely speaking, the stronger the positive association between scalability and size, the larger is the response of concentration to a common demand shock.

To state this formally, we first define a partial ordering of how scalability covaries with size. We will say that scalability covaries more strongly with size in industry 1 compared to industry 2 if the marginal distributions of size and of scalability are the same, and if for any level of size, R, the distribution of scalability among firms in industry 2 with size weakly less than R first order stochastically dominates the distribution of scalability among such firms in industry 1. The ordering is strict if the stochastic dominance is strict for some R. Intuitively, this ordering captures the effect of moving scalability from smaller to larger firms holding the marginal distributions fixed.

Proposition 10 *For any admissible measure of concentration of size, the same increase in G will raise concentration of size (among a fixed set of firms) by more in the industry in which the scalable share of expertise covaries more strongly with size.*

3 Data

We use product and establishment-level data to validate the theory. In this section, we provide details on the datasets, the various measures that characterize firms' heterogeneity, and the construction of measures of common shocks.

3.1 Datasets

3.1.1 Product data

We use comprehensive data for firms and products sold in the consumer goods industry from 2006 to 2015 collected by the Nielsen Retail Measurement Services (RMS) and provided by the Kilts-Nielsen Data Center at the University of Chicago Booth School of Business. This data set is collected from point-of-sale systems in grocery, drug, and general-merchandise stores. The consumer product goods industry accounts for 14% of the total consumption of goods in the U.S.[10] The Nielsen RMS data set covers about 40% of the sales of the industry, and nearly the universe of firms and products in the industry.

We link products to firms by using information obtained from GS1 US, the single official source of barcodes. Appendix B provides detailed information on the Nielsen RMS and GS1 US datasets. The combined data set includes detailed information about the characteristics of each firms' product; most notably, it includes information on the product's attributes, well-measured quantities, and revenue. The original RMS data set consists of more than one million distinct products identified by barcodes, making it possible to track product sales and prices over time and at a very detailed level. Barcodes are organized into a hierarchical structure. Each barcode is classified into one of the 1070 product modules, that are organized into 104 product groups. We follow Hottman,

[10]This industry includes non-durables (also known as consumer packaged goods) and semi-durable goods. It excludes consumer durables, producer intermediates, and producer capital.

Redding and Weinstein (2016) and Argente, Lee and Moreira (2020) and define sectors based on the classification of product group, which is more detailed than a 4-digit level of Standard Industrial Classification (SIC). We use the detailed product attributes, such as information about a product's size, packaging, formula, and flavor to measure the degree of scalable expertise across firms.

For our analysis, we explore changes in size over time at the firm-level within sector. Thus, for each firm-sector-year combination, we compute total sales, total products and average sales per product. We also make use of various measures capturing entry and exit of firms and products, as well as data on location of stores selling the products when creating measures of shocks.

3.1.2 Establishment data

We use data covering firms and establishments covering all sectors from 1990 to 2016 obtained from the National Establishment Time Series (NETS). The NETS data set consists of longitudinally linked Dun & Bradstreet establishment-level data.[11] NETS provides yearly employment and sales information for 'lines of business' in a specific location (similar to the definition of an establishment).

Each establishment is assigned a data universal numbering system identifier that makes it possible to track its sales and employment over time. For each establishment, we know location, industry classification, and parent company. We use parent company information as our definition of firms. In our main analysis we characterize the industry of establishments at the 4-digit level of SIC and the locations are mapped to Metropolitan Statistical Areas (MSAs). We apply some sample restrictions and robustness exercise that makes us confident that our results are representative of U.S. business activity in the static cross section.[12]

As with the product data, we mostly explore changes in size over time at the firm-level within sector. To do so, we create for each firm-sector variables capturing total employment (or total sales), the number of establishments, and the average employment (sales) per location, and study how these variables evolve over time. We also use geographic information when creating measures of shocks and various measures capturing entry and exit of firms and establishments.

3.2 Sectoral shocks

The theory of scalable expertise has predictions on how different firms respond to common shocks, both in terms of the magnitude of the response and the margins they adjust. We explore two distinct shocks to the industry demand. Our first shock is sector-specific and capture a sector's exposure to Chinese import competition. Our second shock is firm and sector-specific and exploits

[11]The data is provided by Walls & Associates. The dataset is the result of a large-scale data collection effort involving interviews, mailings, private third-party sources, and public data sources. Appendix B provides detailed information on the data.

[12]We follow Crane and Decker (2019), who recently compared NETS to sources based on administrative records. Their paper describes potential limitations with regard to the inclusion of non-employee firms on NETS and imputed observations, and shows that applying some sample restrictions leads to high correlations in the distributions for the level of employment between NETS data and U.S. Census Longitudinal Business Dataset.

firms' heterogeneous exposure to regional variation in home price appreciation during the housing boom and bust cycle in the 2000s.

3.2.1 China import penetration shock

Our measures of trade exposure follow closely Autor, Dorn and Hanson (2013) and Acemoglu, Autor, Dorn, Hanson and Price (2016) by capturing the change in the import penetration ratio. We collect data on trade between China and the US from UN Comtrade, NBER-CES and UNIDO. Our baseline measure of the change in the Chinese import penetration ratio for a given sector j over the period 2006–2015, is defined as

$$\Delta IP_{j,06-15} = \frac{M_{j,15} - M_{j,06}}{Y_{j,06} + M_{j,06} - E_{j,06}} \times 100 \tag{21}$$

where $M_{j,t}$ is the imports from China into the U.S., $E_{j,t}$ is the exports, and $Y_{j,t}$ is the industry shipments. We use 2006 as the baseline year since it is the earliest period for which we have simultaneously trade, product and establishments data. Intuitively, the measure $\Delta IP_{j,06-15}$ captures group-level changes in imports from China. In order to address endogeneity concerns, we also work with an alternative measure that uses imports in other high income countries as an instrument, following Acemoglu, Autor, Dorn, Hanson and Price (2016). The motivation for the instrument is that high-income economies are similarly exposed to a Chinese supply shock, but are unaffected by US-specific shocks that affect US import demand.[13]

We map the import penetration measures to changes in residual demand for equivalent products produced by US firms. Thus, we define the shocks of the different sectors as the additive inverse of the import penetration measures to proxy a positive demand shock: $\Delta G_j = -\Delta IP_{j,06-15}$. For the establishments data, we use standard SIC mapping to trade data to create sectoral shocks at the 4-digit SIC level. For the product data, we use the concordance developed by Bai and Stumpner (2019) to map Nielsen's product groups to trade data, and obtain shocks at the product group level.

This shock allows us to use variation across manufacturing sectors affected by trade. There is substantial heterogeneity in import penetration across sectors. As of 2006, imports from China were already important for some sectors. Nevertheless, between 2006 and 2015 there was substantial acceleration in many sectors. Figure B.2.1 in Appendix B.2 shows the measure of Chinese import penetration by sector. There is a substantial amount of heterogeneity within the consumer product industry, and within manufacturing more generally. As expected, on average, sectors that produce semi-durables and durable products were more affected by import penetration of China products and thus have more negative demand shocks than sectors related to food products.

[13] See Appendix B.2.1 for more details on the instrument $\Delta IP_{j,06-15}$ and measures of import penetration for other time periods.

3.2.2 House price shock

Our measure of house price shock allows us to construct a firm-sector specific demand shock G_{ij}, by exploring geographical variation in consumer demand arising from the evolution in the housing market and firm-sector specific pre-existing exposure to different locations. The inspiration for this shock builds on the work of Mian and Sufi (2011, 2014) that document the role of the housing net worth channel in suppressing consumer demand either through a direct wealth effect or through tighter borrowing constraints driven by the fall in collateral value. We build two shocks covering the housing boom and bust cycle in the 2000s.

We start by measuring changes in house prices across-locations and instrumenting for them using variation in housing supply elasticity proposed by Saiz (2010). The intuition for this instrument is that for a fixed housing demand shock during the housing boom, house prices should rise more in areas where housing supply is less elastic. During the housing bust, it is then precisely those areas where house prices rose most that see the largest declines in house prices.[14] For each MSA, denoted by m, we compute changes in house prices indices (HP) and estimate the following equation separately for the housing boom (2001-2006) and bust (2006-2011), denoted as τ:

$$\Delta \log(\mathrm{HP})^{\tau}_{m} = \rho \, \mathrm{SupplyElasticity}_{m} + \delta X_{m} + \varepsilon_{m} \tag{22}$$

where X_m is a vector of controls that control for other changes in the local economy as in Stroebel and Vavra (2019). We obtain house price indices at both the MSA level from the FHFA House Price Index. Table A.II in Appendix presents results from the first-stage regression and various robustness. The instrument is highly predictive of house price changes over both periods, with low-elasticity MSAs experiencing larger house price gains during the housing boom, and larger house price drops during the housing bust.

After retrieving the predicted change in price index across different locations $\Delta \widehat{\log(\mathrm{HP})}^{\tau}_{m}$, we build firm-specific shocks separately using information on the location of firms. In particular, we measure housing price shocks faced by firm i in sector j in either the housing boom or bust as follows:

$$\Delta D^{\tau}_{ij} = \sum_{m=1}^{M^{\tau_0}_{ij}} \frac{\mathrm{Size}^{\tau_0}_{ijm}}{\sum_{m=1}^{M^{\tau_0}_{ij}} \mathrm{Size}^{\tau_0}_{ijm}} \Delta \widehat{\log(\mathrm{HP})}^{\tau}_{m} \tag{23}$$

where the variable Size_{ijts} refers to measures of size in the baseline years τ_0 (2001 for boom period, and 2006 for bust period). In the establishment NETS data, size is measured as total employment of establishments in a particular location (with robustness with sales measures), and we are able to cover both the housing boom and bust. Our baseline analysis includes firms in sectors where the location of the establishment likely reflects where the demand occurs. For this reason, in our

[14] Saiz (2010) uses information on the geography of a metropolitan area to measure the ease with which new housing can be constructed. The index assigns a high elasticity to areas with a flat topology without many water bodies, such as lakes and oceans.

baseline sample, we exclude manufacturing businesses, where the location of a plant likely reflects where the production occurs, and not necessarily the demand.[15] In the product Nielsen data, size is measured as total sales in stores in a particular MSA, and thus a location is directly associated with the demand channel. The product data, however, is only available starting in 2006, and thus we are only able to cover the bust period. Throughout the paper, we report the results for each period and dataset independently.[16]

3.3 Mapping Model to Data

3.3.1 Measuring size and scope

In the establishment NETS data, we have variables capturing size and the number of establishments. Our baseline measure of size is defined as the total employment a firm i in sector j in year t, and our measure of scope is the total number of distinct establishments. We use employment as baseline because it is better measured on than revenue on NETS, but we also use total revenue as robustness.[17] The measure of scope based on number of establishments aligns well with the more traditional measures capturing the intensive and extensive growth of firms as in Cao, Hyatt, Mukoyama and Sager (2020).

The product Nielsen data allows us to define alternative measures using more detailed data. We define *size* as the total revenue of a firm i in sector j in year t, and *scope* as the total number of distinct barcodes sold across all stores. These measures are similar to the ones used in Hottman, Redding and Weinstein (2016). Barcodes are by design unique to every product – changes in any attribute of a good (e.g. forms, sizes, package, formula) results in a new barcode, and thus allows us to have a close map to the theory.

3.3.2 Measuring scalability

In this section, we propose a measure that captures the extent to which knowledge can be applied across multiple products. In particular, we leverage the detailed information on product-level characteristics in the Kilts-Nielsen data to construct measures of scalability. The underlying intuition of our measure is that products are bundles of characteristics, and specific characteristics can be shared across multiple products. Thus, firms that have products that have many common characteristics, are more likely to have more scalable knowledge.

We build the measure using detailed descriptions of products. The original data classifies products into modules and within each module products are described using various "attributes" vari-

[15] We consider several alternative samples, including using all sectors to evaluate the sensitivity to this baseline sample. In Appendix, we show that the results are qualitatively similar and are not dependent on these assumptions.

[16] In Appendix B.2, we provide more details on the distribution of these shocks.

[17] In Appendix D.1 we show that using total revenue rather than employment does not affect our findings. This is not surprising given that revenue is, for the most part, imputed from employment.

ables, such as package, size, flavor, formula.[18] These attributes can take different values, which we term "characteristics". Consider the product module: razor blades. The products in this module have the following five attributes: form, consumer type, scent, skin condition, and generic. The attribute "form" can take the following characteristics: "adjustable", "assorted", "injector", "moving", "pivoting" etc. The product module of powder detergents, on the other hand, is described by the following four attributes: form, container, type, and generic. The attribute "form" for detergents could be "pack", "pod", "refill", "table".

With these data, we are able to describe products as a bundle of characteristics, some of which it shares across multiple products, while others are only used by one or a small number. The distinguishing feature of scalable expertise is its applicability to multiple products, in contrast to 'local' or product-specific expertise. In line with this interpretation, we map the scalability of expertise of firm i to the fraction of common attributes across its product portfolio. In particular, we propose a scalability index of firm i in sector j in product module m at time t is defined as follows:

$$\mathcal{SI}_{mijt} \equiv 1 - \frac{\text{Unique}_{mijt}}{\text{Scope}_{mijt} \times \text{NumAttributes}_{mjt}} \ ,$$

where the variable Unique_{mijt} counts the number of distinct characteristics in the portfolio of products of the firm in module m. This is normalized by the total number of attribute-cells to be filled, i.e. module-level scope Scope_{mijt} times the number of attributes for each product in that module $\text{NumAttributes}_{mjt}$. The index captures the share of common characteristics within the portfolio of products of the firms, i.e. the likelihood that the characteristics of products are shared within the portfolio of a given firm. If no characteristic is repeated across the products of a firm, the scalability index equals 0. For example, a single-product firm will have a number of different attributes equal to the number of distinct characteristics. By contrast, when products of a single firm share many characteristics, the scalability index converges to 1. Note that the index is a relative measure, intended to measure the *composition* of expertise, not its *level*. This leads us to the following mapping between the index (or more precisely, a simple transformation thereof) to scalability in the theory:

$$\frac{x_{mijt}}{y_{mijt}} \equiv \frac{\mathcal{SI}_{mijt}}{1 - \mathcal{SI}_{mijt}} \ .$$

We aggregate this measure to the firm-sector level using revenue-weights, to be at the same aggregation level of our size and scope measures.

A potential concern is that the scalability index might increase mechanically as firms add products and, thus, biased towards 1 for firms with a large number of them. We address this using a bootstrap procedure. For a firm with N products in a module, we compute what its scalability index would be if its portfolio had N randomly selected products from that module. The measure of scalability that we use in the regressions is a firm's scalability relative to the scalability of a random

[18] This information is available for approximately 61% of all barcodes in the data. We use a total of 20 distinct attributes. Each product module has between 4 and 8 active attributes.

portfolio of the same size. In Appendix C, we provide additional details on the construction of this measure, including examples and robustness.

4 Validating the Theory

Our theory has multiple testable implications. We divide them into two sets. The first set uses only "standard" measures: size and scope. Second, we test predictions that incorporate scalability directly. We evaluate how scalability covaries with size and scope in the cross-section, assess whether scalable firms are more sensitive to shocks, and how the distribution of scalability affects changes in concentration.

4.1 Testable implications with size and scope measures

The theory describes how firms differ systematically in their responses to a common shock, both in terms of the magnitude and the margins of the response. Corollary 2 implies that in response to the same change in G, firms with higher size or higher scope have larger changes in size and scope. In this section, we develop statistical tests to evaluate if the predictions fit well the data using establishment and product data, and two distinct shocks. Our first shock is sector-specific and captures different exposure to the intensity of competition from Chinese imports. We also use a second shock that exploits cross-regional variation in home prices during the housing boom and bust cycle in the 2000s, and the pre-existing exposure of firms to different locations.

We first test whether firms with higher size or higher scope are more sensitive to changes in demand, using the following regression specifications:

$$\begin{aligned}
\Delta\text{size}_{ij} &= \alpha_{RR}\,\Delta G_{ij} + \beta_{RR}\,(\Delta G_{ij} \times \text{size}_{ij}) + \gamma_{RR}\,\text{size}_{ij} + \psi_{RRj} + \epsilon_{RRij} \\
\Delta\text{scope}_{ij} &= \alpha_{NR}\,\Delta G_{ij} + \beta_{NR}\,(\Delta G_{ij} \times \text{size}_{ij}) + \gamma_{NR}\,\text{size}_{ij} + \psi_{NRj} + \epsilon_{NRij} \\
\Delta\text{size}_{ij} &= \alpha_{RN}\,\Delta G_{ij} + \beta_{RN}\,(\Delta G_{ij} \times \text{scope}_{ij}) + \gamma_{RN}\,\text{scope}_{ij} + \psi_{RNj} + \epsilon_{RNij} \\
\Delta\text{scope}_{ij} &= \alpha_{NN}\,\Delta G_{ij} + \beta_{NN}\,(\Delta G_{ij} \times \text{scope}_{ij}) + \gamma_{NN}\,\text{scope}_{ij} + \psi_{NNj} + \epsilon_{NNij}
\end{aligned} \tag{24}$$

where Δsize_{ij} and Δscope_{ij} refers to difference of log size and scope during the period of the shock, respectively. ΔG_{ij} refers to either the Chinese import penetration, or the housing price shocks.[19] The control variables size_{ij} and scope_{ij} are log of total sales/employment and scope measured before the shock, and they allow us to control for potential systematic associations between size and growth, as well as mean reversion effects. We also account for differences in average sensitivity to shocks across sectors by using sector-specific fixed-effects (ψ_j). We standardize the variables size_{ij} and scope_{ij} within sector j, to help compare the magnitudes across regressions and samples.

[19]Note that in the case of the China shock, the shock is sector specific, ΔG_j, and thus, when we use these specifications with sector fixed effects, we cannot identify the coefficients α on the shock.

The key parameters of equations 24 are the interaction terms β_{RR}, β_{NR}, β_{RN}, and β_{NN}. These coefficients measure differences in degree of sensitivity to shocks. A key prediction of the model is that all these coefficients should be positive.

Table 1 presents our baseline results for the estimates of β_{RR} (columns 1, 3, 5, 7, 9), β_{NR} (columns 2, 4, 6, 8, 10), β_{RN} (columns 11, 13, 15, 17, 19), and β_{NN} (columns 12, 14, 16, 18, 20), across five different versions. These five different versions correspond to different combinations of datasets and shocks: (i) establishment data with China import penetration shock (2006-2015); (ii) establishment data with housing price shock during boom (2001-2006); (iii) establishment data with housing price shock during bust (2007-2011); (iv) product data with China import penetration shock (2006-2015); (v) product data with housing price shock during bust (2007-2011). In regressions using establishment data, we use employment as our benchmark measure of size, and scope refers to the number of establishments of a firm. In regressions using product data our measure of size is the total sales of a firm-sector, and scope is measured as the number of distinct barcodes.

Our results show that larger size firms exhibit greater size sensitivity to common shocks (β_{RR}), and also greater scope sensitivity (β_{NR}). Likewise, firms that operate multiple units are more likely to make larger changes to size (β_{RN}) and scope (β_{NN}) when affected by common shocks. These results indicate that the predictions of the theory align remarkably well with our results, independently of the datasets and nature of shocks.

Table 1: Heterogeneous Response to Common Shocks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Δ size	Δ scope	Δ size	Δ scope	Δ size	Δ scope	Δ size	Δ scope	Δ size	Δ scope
$\Delta G \times$ size	0.042*	0.023**	0.021***	0.002***	0.003**	0.001**	0.123***	0.010***	0.666***	0.066*
	(0.024)	(0.012)	(0.002)	(0.000)	(0.001)	(0.001)	(0.017)	(0.004)	(0.161)	(0.037)
Obs.	321,525	321,525	5,933,769	5,933,769	6,968,641	6,968,641	24,146	24,146	27,930	28,137
R-squared	0.047	0.023	0.053	0.006	0.036	0.011	0.154	0.069	0.088	0.031
Sector	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
	Δ size	Δ scope	Δ size	Δ scope	Δ size	Δ scope	Δ size	Δ scope	Δ size	Δ scope
$\Delta G \times$ scope	0.116***	0.059***	0.004*	0.000	0.006**	0.004**	0.057***	0.030***	0.246*	0.147***
	(0.023)	(0.018)	(0.002)	(0.002)	(0.002)	(0.002)	(0.016)	(0.004)	(0.143)	(0.038)
Obs.	321,115	321,115	5,933,769	5,933,769	6,968,641	6,968,641	24,146	24,146	27,930	28,137
R-squared	0.043	0.242	0.008	0.071	0.015	0.091	0.073	0.136	0.055	0.077
Sector	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sample	China	China	Housing	Housing	Housing	Housing	China	China	Housing	Housing
Period	06-15	06-15	01-06	01-06	06-11	06-11	06-15	06-15	06-11	06-11
Data	NETS	NETS	NETS	NETS	NETS	NETS	Nielsen	Nielsen	Nielsen	Nielsen

Note: The table reports the results of estimating equation 24. The dependent variable is the log change in the total employment/sales of firm i in sector j in the period of the shock or the change in the number of plants/products of firm i in sector j in the period of the shock. The reported coefficients are β_{RR} (columns 1, 3, 5, 7, 9), β_{NR} (columns 2, 4, 6, 8, 10), β_{RN} (columns 11, 13, 15, 17, 19), and β_{NN} (columns 12, 14, 16, 18, 20), across five different versions. Columns 1/2/11/12 use the NETS data and the China import penetration shock 2006-2016. Columns 3/4/13/14 use NETS data and the housing price shock from 2001-2006. Columns 5/6/15/16 use NETS data and the housing price shock from 2006-2011. Columns 7/8/17/18 use Nielsen data and the China import penetration shock 2006-2016. Columns 9/10/19/20 use Nielsen data and the housing price shock from 2006-2011. All the specifications include sector effects and robust standard errors.

The magnitudes of the estimates differ across the five different combinations of datasets and shocks. To interpret the magnitude, consider a positive demand shock induced by a decline in China's import penetration. A firm that is one standard deviation larger than average in its sector increases its size by 4.2 percentage points and scope by 2.3 percentage points more than the average firm in its sector. Notice that the implied changes in total size are larger than the changes in scope. This indicates that larger firms also exhibit greater sensitivity of size per unit to demand. The magnitudes tend to be larger when we interpret scope as number of products (as opposed to number of establishments) and measure size using revenue (as opposed to employment).

Robustness – We evaluate the robustness of our results with respect to the samples and measures of growth. In our baseline results, we measure growth as log changes among the sample of surviving firms. In particular, we do not take into account the effect of differences in the cyclical sensitivities of the rate of entry and exit of small and large size/scope firms. Thus our results should be perceived as capturing the intensive margin differences. We re-estimate the equations above using a sample that also includes exits by using the bounded growth rates of Davis and Haltiwanger (1992), our results are all qualitatively similar.[20]

We evaluate the impact of the inclusion of different sectors in the results. For example, in our baseline results with establishment data and the house price shocks, we use the sample of non-manufacturing firms. We consider several alternative samples, including using all sectors and more restrictive samples of establishments to evaluate the sensitivity to this baseline sample. In Appendix, we show that the results are qualitatively similar and do not depend on these assumptions.[21]

We also evaluate how our results are affected when we use different samples of firms. One potential limitation of the results with establishment data is that they may be affected by the large number of non-employer single-unit establishments (in line with evidence in Crane and Decker (2019)). We study how our results are affected when we exclude these establishments and focus on variation coming from medium-sized and large firms. In Appendix, we show that most of the results are qualitatively similar.[22]

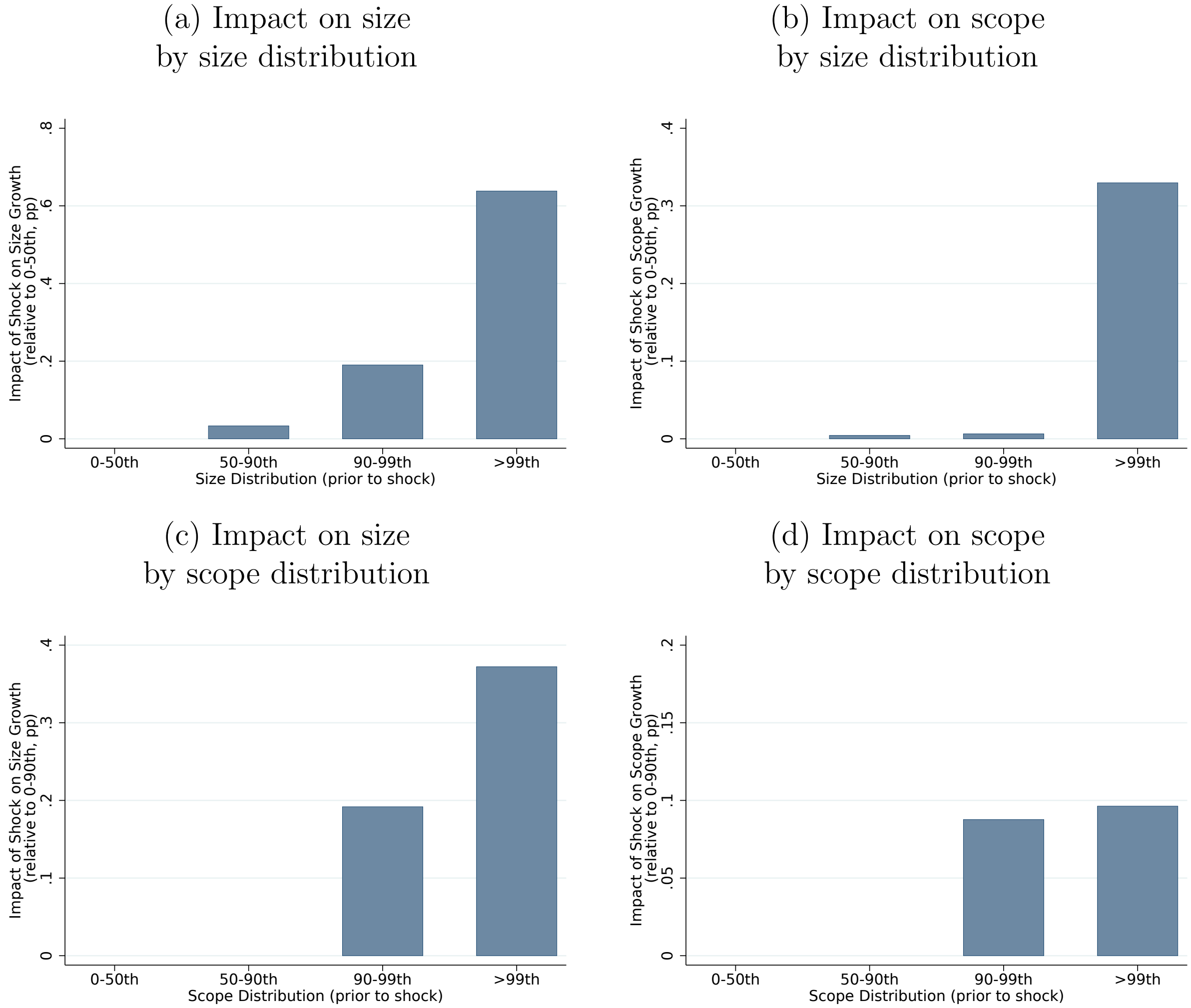
Finally, we also explore alternative specifications to equation 24. We consider a non-parametric specification in which we study the differential impact of the shock in different parts of the size and scope distribution. Figure 3 presents the results for the establishment data, when we estimate the impact for the China shock on four size/scope quantiles: below median, 50-90 percentile, 90-99 percentile, and top 1%. The results indicate monotonicity on the impact of the shock aligned with the results of Table 1. For example, results studying impact on size as a function of size indicate that larger firms tend to respond more strongly than small and medium firms.

[20]Tables D3, D7 and D12 in Appendix D.1 provides the results for the establishment-level data, and Table D16 in Appendix D.2 provides the results for the product-level data.

[21]Tables D9 and D14 in Appendix D.1 provides the results for the establishment-level data.

[22]Tables D10 and D15 in Appendix D.1 provides the results for the establishment-level data.

Figure 3: Heterogeneous response to China import penetration shock: non-parametric specification



Note: The figure shows the impact of shocks for different size/scope quartiles. We plot $\hat{\beta}_k$'s from estimating the specification $\Delta \mathrm{Y}_{ij} = \alpha + \sum_k \beta_k \left(\Delta G_{ij} \times \mathrm{d}_{k,ij}\right) + \sum_k \gamma_k \left(\mathrm{d}_{k,ij}\right) + \epsilon_{ij}$, where $\mathrm{d}_{k,ij}$ are dummies for four size/scope quantiles: below median, 50-90 percentile, 90-99 percentile, and top 1% (for the scope quantiles, we aggregate the first two dummies because more than 90 percent of firms have only one establishment). The dependent variable $\Delta \mathrm{Y}_{ij}$ refers to changes in log size/scope. Panel (a) plots the results of changes in size as a function of size quantiles, comparable to β_{RR}. Likewise, panel (b), (c) and (d) plots the results comparable to β_{NR}, β_{NR}, and β_{NN}, respectively. The estimates are computed using establishment-level data and the China import penetration shock.

Related literature – There is a more established literature studying whether the growth of large or small firms covary more with business cycles. For example, Crouzet and Mehrotra (2020) recently documented that the largest firms are less cyclical than other firms. There are two important differences between our results and this body of work. First, we are interested in firms' responses to changes in demand, either idiosyncratic or common. As such, we exploit variation in demand that is external to the sector. The literature on size cyclicality, in contrast, has focused on unconditional covariation with aggregate conditions.[23] Second, we focus on firms' longer-run responses to shocks (10 years for the China shock and 5 years for the housing shocks), rather than adjustments at business cycle frequencies (typically quarterly or annual changes).

4.2 Testable implications with scalability measures

4.2.1 Correlation between scalability, scope and size

Given the central role played by scalability in the theory, it is natural to ask: how well are the predictions with respect to scalability borne out by the data? In particular, Proposition 7 predicts that scalability is positively related to size and scope in the cross-section. We evaluate if this relationship is consistent with data by estimating:

$$\begin{aligned} \ln\left(\frac{x_{ijt}}{y_{ijt}}\right) &= \alpha_R + \beta_R \text{ size}_{ijt} + \lambda_{Ri} + \Gamma_{Rjt} + \varepsilon_{Rijt} \\ \ln\left(\frac{x_{ijt}}{y_{ijt}}\right) &= \alpha_N + \beta_N \text{ scope}_{ijt} + \lambda_{Ni} + \Gamma_{Njt} + \varepsilon_{Nijt} \end{aligned} \tag{25}$$

where $\frac{x_{ijt}}{y_{ijt}}$ is scalability as defined in Section 3.3, and size_{ij} and scope_{ij} are log of total revenue and scope of firm i in sector j at time t. We control for firm effects (λ_i) and a set of sector $\times$ time fixed effects (Γ_{jt}).

The results are presented in Table 2. In line with Proposition 7, columns (1) and (3) show that $\beta_R > 0$ and $\beta_N > 0$, which indicates that large size and large scope firms have relatively more scalable knowledge. Columns (2) and (4) indicate that these findings are robust to including firm effects, consistent with the fact that as firms increase their size and scope, they also increase the scalability of their product portfolio.

We also evaluate how the results compare across different types of consumer product sectors. Table D17 in the Appendix shows that these results are robust across different categories in the Nielsen data and to disaggregating sectors further.

[23] An exception is Park (2020), who also uses establishment NETS data and studies changes in China import penetration between 1991-2007. His results are consistent with our estimates of the size sensitivity captured by a positive estimate of the coefficient β_{RR} in equation 24.

Table 2: **Cross-Sectional Relationship: Scalability, Scope and Size**

	(1)	(2)	(3)	(4)
size	0.145***	0.172***		
	(0.021)	(0.013)		
scope			0.258***	0.240***
			(0.027)	(0.018)
Observations	235,674	233,004	235,671	233,001
R-squared	0.021	0.405	0.054	0.417
Firm	N	Y	N	Y
Period-Sector	Y	Y	Y	Y

Note: The table shows the results estimating equation 25 in the product-level data. The dependent variable is (the log of) scalability from (25) and the independent variables are (the logs of) size (revenue) and scope. All the variables are standardize relative to the mean of the sector and time period (year). Scalability index is adjusted relative to the alternative (bootstrapped) index.

4.2.2 Heterogeneous response to symmetric shocks with scalability measures

Next, we explore how scalability affects how firms respond to the same sectoral shock. Proposition 9 predicts that in response to the same shock, firms with higher scalability have larger changes in size, scope, scalability. We study these prediction using the following specifications:

$$\begin{aligned}
\Delta\ \text{size}_{ij} &= \alpha_{RS}\ \Delta G_j + \beta_{RS}\ \left(\Delta G_j \times \ln \frac{x_{ij}}{y_{ij}}\right) + \gamma_{RS} \ln\left(\frac{x_{ij}}{y_{ij}}\right) + \psi_{RSj} + \epsilon_{RSij} \\
\Delta\ \text{scope}_{ij} &= \alpha_{NS}\ \Delta G_j + \beta_{NS}\ \left(\Delta G_j \times \ln \frac{x_{ij}}{y_{ij}}\right) + \gamma_{NS} \ln\left(\frac{x_{ij}}{y_{ij}}\right) + \psi_{NSj} + \epsilon_{NSij} \\
\Delta \ln\left(\frac{x_{ij}}{y_{ij}}\right) &= \alpha_{SS}\ \Delta G_j + \beta_{SS}\ \left(\Delta G_j \times \ln \frac{x_{ij}}{y_{ij}}\right) + \gamma_{SS} \ln\left(\frac{x_{ij}}{y_{ij}}\right) + \psi_{SSj} + \epsilon_{SSij}
\end{aligned} \tag{26}$$

where the dependent variables Δ size$_{ij}$, Δ scope$_{ij}$, or $\Delta \ln\left(\frac{x_{ij}}{y_{ij}}\right)$ are changes in size, scope and scalability, respectively. G_j is the China import penetration shock from 2006 to 2015 for sector j.[24] We control for sector fixed effects Γ_j to account for differences in response that are heterogeneous across sectors. The coefficient of interest is β_{RS}, β_{NS}, and β_{SS} measures the heterogeneous response of firms with different scalability – the theory predicts all three coefficients should be positive.

The results are presented in Table 3. Column (1) shows that, in response to common shocks, firms with higher scalability adjust size by more (relative to the average firm, whose response is now picked up by the sector fixed effects). The sign of the coefficient is consistent with the theory although it lacks statistical power in this specification. Column (2) shows that firms with higher

[24] For the housing shock, these specifications lack statistical power and we are unable to sign the impact of scalability on the response to shocks. We present these results in Table D18 in Appendix D.4.

Table 3: Scalability and Responsiveness to Shocks

	(1) Δ size	(2) Δ scope	(3) $\Delta \ln\left(\frac{x}{y}\right)$	(4) Δ size	(5) Δ scope	(6) $\Delta \ln\left(\frac{x}{y}\right)$
$\Delta G \times \ln\left(\frac{x}{y}\right)$	0.0187	0.0187***	0.0119***	0.0073	0.0099*	0.0117***
	(0.014)	(0.006)	(0.003)	(0.014)	(0.006)	(0.003)
$\Delta G \times$ scope				0.0329***	0.0242***	0.0016
				(0.012)	(0.005)	(0.003)
Observations	14,186	14,186	13,467	14,186	14,186	13,467
R-squared	0.049	0.044	0.256	0.063	0.097	0.256
Sector	Y	Y	Y	Y	Y	Y
Shock	China	China	China	China	China	China
Period	2006-2015	2006-2015	2006-2015	2006-2015	2006-2015	2006-2015

Note: The table shows the results of estimating equation 26. The dependent variable in Columns (1) and (4) is the change in the (log of) size (revenue) of firm i in sector j. The dependent variable in Columns (2) and (5) is the change in scope and in Columns (3) and (6) is the change in scalability of firm i in sector j. The independent variable is the demand shock interacted with the baseline level of scalability in 2006 or the same shock interacted with the levels of scope in 2006. In all specifications we use the China import penetration shock from 2006-2015. All specifications include sector fixed effects.

scalability do indeed adjust their scope by more in response to a positive import penetration shock. Lastly, our theory also predicts that, in response to common shocks, firms with higher scalability adjust scalability by more. Column (3) shows that scalability responds positively to changes in demand captured by the China import penetration shock. Columns (4)-(6) show that our measure of scalability still captures sensitivity of shocks when controlling for firm scope.[25]

4.2.3 Concentration

Lastly, we show suggestive evidence that the covariance between scalability and size carries information about the response of concentration among a fixed set of firms. Recall that Proposition 10 indicates that the same increase in G will raise concentration of size (among a fixed set of firms) by more in the sector in which scalability covaries more strongly with size. In order to determine the industries where scalability is more concentrated, for each sector j, we calculate the covariance between scalability and size in 2006 ($W_{j,2006}$), while focusing on firms active in the period 2006-2015.[26]

[25] While the theory implies that scalability is a sufficient statistic for a firm's sensitivity to shocks, we find that scope still explains some variation in the size and scope regressions. A possible reason that the scope still has predictive power is that our proxy measures true scalability with noise, and that our bootstrap procedure removes some of the variation in scalability. Nevertheless, we find it remarkable that our measure of scalability still carries information about sensitivity to shocks even controlling for scope.

[26] More specifically, we estimate $\ln\left(\frac{x_{ij}}{y_{ij}}\right) = \lambda_j + W_j \text{ scope}_{ij} + \varepsilon_{ij}$ for each sector, where $\frac{x_{ijt}}{y_{ijt}}$ is scalability and scope_{ij} is log of number of products of firm i in sector j in 2006. For over 80% of the sectors, the coefficient of a

Table 4: Market Size and Concentration

	(1)	(2)	(3)
	Δ HHI	Δ Share Top 5	-Δ ln # firms 80% share
$\Delta G \times \hat{W}_{2006}$	0.0063*	0.0096**	0.0547***
	(0.003)	(0.005)	(0.018)
Observations	112	112	110
R-squared	0.065	0.080	0.099

Note: The table reports the results of estimating equation 27. The dependent variable in Column (1) is change in the Herfindahl index in sector j from 2006-2015. The dependent variable in Column (2) is the change in the market share held by the top 5 firms in the sector and in Column (3) is the log change in the total number of firms with 80% of the market share times minus one. $\hat{W}_{j,2006}$ is the covariance between scalability and size in 2006 estimated using a reduced form approach for each sector separately and considering firms with positive sales both in 2006 and 2015. Both columns use the China import penetration shock. The data used is the Nielsen data.

To test Proposition 10, we then estimate the following specification:

$$\Delta Y_j = \alpha\ \Delta G_j + \beta\ \left(\Delta G_j \times \hat{W}_j\right) + \gamma \hat{W}_j + \epsilon_j \tag{27}$$

where ΔY_{jt} is an outcome variable measuring the change concentration in sector j between 2006 and 2015. We use the Chinese import penetration shock, ΔG_j, which is common across firms in a sector, to capture long term differences in market concentration. The coefficient of interest is β which measures if market concentration changes more in sectors in which scalability is more concentrated.

Table 4 assesses this prediction using several different measures of concentration. Column (1) measures concentration as Herfindahl Index in the sector, and the results indicate that concentration increased in sectors where the covariance of scalability and size is larger. Column (2) measures concentration as the share of revenue held by the top 5 firms in the sector (the concentration ratio of the top 5 firms) and also finds results consistent with the fact that concentration also increased more in sectors where the covariance is higher. Lastly, Column (3) measures concentration by the additive inverse of the total number of firms with 80% of the market. Consistent with the previous columns, using this measure we find an increase in concentration. Overall the table provides suggestive evidence that the covariance between scalability and size carries relevant information about the response of concentration to a common demand shock.

regression of scalability on size is positive.

5 Discussion

5.1 Other Evidence

The theory in this paper also offers a useful, unified framework to interpret a number of facts – both cross-sectional and in response to changes in demand – documented by other papers studying firm heterogeneity and dynamics. In this subsection, we discuss a few of them and relate them to the effects of scalable expertise.

Holmes and Stevens (2014) find that Chinese import competition had a disproportionate impact on larger plants relative to smaller, more 'specialty' plants with more knowledge about local customer needs. Through the lens of our model, specialty plants can be thought of as having relatively lower scalability.[27] The theory predicts that this will make them less responsive to demand conditions. Aghion, Bloom, Lucking, Sadun and Van Reenen (2021) show that during the Great Recession, firms with more decentralized management outperformed their more centralized counterparts, even controlling for size. It seems natural to conjecture that centralized firms have more of their knowledge in a scalable form, which makes them more sensitive to turbulence.

Hyun and Kim (2019) also look at spillovers within firms during the Great Recession. Specifically, they find that when a firm faces reduced demand in one location (e.g. due to a local housing shock), its sales fall in other locations as well. This is suggestive of scalable investments: a demand decrease in one location reduces the firm's incentives to invest and, to the extent that such investment is scalable, the effects (e.g. a decline in sales) are seen across the firm (i.e. in other locations as well). In line with our results, they also find that large firms respond more than small firms.

Lileeva and Trefler (2010), Bustos (2011) and Lashkari, Bauer and Boussard (2018) show that firms facing higher demand (e.g. due to an foreign tariff cut, a domestic decline in input tariffs or increased export demand) were more likely to improve productivity by adopting technologies like manufacturing information systems and information and communication technologies, which are suggestive of increased scalability, consistent with the predictions of our model. Bustos (2011) further shows that such adoption was accompanied by larger increases in product and process innovation. Mayer, Melitz and Ottaviano (2020) also find that export demand shocks induce increases in product scope (in addition to sales per product).

Cao, Hyatt, Mukoyama and Sager (2020) find that changes in scope (number of plants) are more important for large firms compared to small firms. Rossi-Hansberg, Sarte and Trachter (2021), Hsieh and Rossi-Hansberg (2020), Smith and Ocampo (2020) and Benkard, Yurukoglu and Zhang (2021) also find that increases in scope (number of locations or product markets served by a firm) are important for understanding the recent increase in national concentration. This is consistent with larger firms having higher scalability and, therefore, a more elastic scope margin in response to demand or productivity.

[27] In our data as well, we find that smaller firms tend to have lower scalability.

5.2 Policy

Our analysis also yields insights that are relevant for policy. First, the economic forces that make scalable firms more responsive to demand and productivity changes also make them more sensitive to policy interventions (such as revenue taxes or subsidies). As such, measures of scalability can inform targeted policy interventions.

A second insight pertains to the effect of distortions that affect allocation resources across firms. These are widely regarded as an important contributor to allocative inefficiencies and lower aggregate productivity, especially in developing countries – see, e.g., Hsieh and Klenow (2009). Scalability can interact with, and amplify the effects of, such distortions.[28] Scalable firms will respond more strongly, which could result in a more acute allocative inefficiency (and a larger drop in aggregate TFP) relative to the one that might be inferred using a model with homogeneous returns to scale. In other words, when assessing the aggregate effects of a distortion, it is important to account for its covariance with returns to scale (or equivalently, with scalability). This becomes even more salient when the underlying distortions are size-dependent.[29] Specifically, if larger firms are more scalable and subject to greater restrictions, the misspecification is especially acute.

Finally, scalability of expertise can also be a mechanism that generates spillovers across markets. A change in one market that alters a firm's incentive to invest in scalable expertise would also affect its activities in other markets.

5.3 Diffusion

Lastly, in this subsection, we document an interesting link between scalability and diffusion of knowledge across firms. The underlying hypothesis is an intuitive one: scalable expertise, i.e. knowledge that can be more readily used across different products of a given firm is also more likely to be useful to other firms in the industry. One interpretation of this assumption is that scalability is the result of practices that enhance the applicability of knowledge to different products (e.g. standardization or codification of procedures). These practices also will make it easier for such knowledge to be used outside the firm. Alternatively, one could think of scalable expertise as innovation along dimensions that are fundamentally more attractive for other firms to learn.

To do this, we first construct a novel measure of diffusion across firms. This leverages the same attribute-level information from the Nielsen data. We start from the introduction of a new characteristic by a firm and then count the number of products by other firms that share that characteristic and were introduced after that date. This is a simple, *ex-post* measure of how broad-based a given characteristic becomes within the market after its introduction. Formally, for a

[28] Endogenous scope and scalability also complicate measurement of distortions: inference using a model with homogeneous returns to scale à la Hsieh and Klenow (2009) may be misleading.

[29] See Restuccia and Rogerson (2008) and Guner, Ventura and Xu (2008). Aghion et al. (2008) and Garcia-Santana and Pijoan-Mas (2014) provide evidence on industrial licensing and small scale reservation laws in India. Gourio and Roys (2014) and Garicano, Lelarge and Van Reenen (2016) provide evidence from France on regulation that applies to larger firms but exempts smaller ones.

characteristic c introduced by firm i, we define diffusion as follows:

$$\mathbb{D}_{c,m,i,t,\tau} = \frac{\text{Num. of products with } c \text{ introduced by firm } -i \text{ between } t \text{ and } t+\tau}{\text{Num. of products introduced by firm } -i \text{ between } t \text{ and } t+\tau}$$

where the numerator counts the number of times a characteristic c introduced by firm i in module m is observed in products introduced by other firms in the same module between time t and time $t+\tau$. The denominator counts the total amount of products introduced by other firms between n t and $t+\tau$. Clearly, $D_{c,m,i,t,\tau}$ is always between 0 and 1.

Our main empirical specification takes the form:

$$\mathbb{D}_{cmit\tau} = \alpha + \beta\, \mathcal{SI}_{amit-1} + \gamma\, scope_{mit} + \lambda_{amt\tau} + \theta_{ami} + \epsilon_{amit\tau} \tag{28}$$

where c, a, m, i, and t denote the characteristic, attribute, product module, firm and time period, respectively. Thus, the dependent variable $\mathbb{D}_{cmit\tau}$ is the diffusion measure described above, while $\mathcal{SI}_{amit-1}$ is our Scalability Index measured for a given attribute a in firm i and module m in period $t-1$. Our coefficient of interest is β: in other words, we are interested in how the diffusion of a particular feature introduced by firm i over $(t, t+\tau)$ relates to the scalability of that firm in $t-1$. We lag scalability to mitigate potential endogeneity concerns. We estimate this relationships controlling for the total number of products of the firm in the same module, $N_{m,i,t}$, and under several specifications of fixed effects; the most saturated one has both attribute $\times$ module $\times$ time $\times$ age, $\lambda_{a,m,t,\tau}$, as well as firm $\times$ attribute $\times$ module effects, $\theta_{a,m,i}$.

Our results, presented in columns (1)-(2) Table 5, show that the level of scalability is indeed positively associated with the level of diffusion. In particular, column (2) shows that the relationship is strong and significant after controlling for attribute $\times$ module $\times$ time $\times$ age effects.

Our strategy for measuring diffusion also points to an alternative way of defining scalability as a 'forward-looking' measure: specifically, by treating scalability as diffusion within a firm, we can construct an *ex-post* measure of how scalable a new feature turns out to be. Formally, this alternative Scalability Index for a characteristic c introduced by firm i is defined as:

$$\widetilde{\mathcal{SI}}_{cmit\tau} = \frac{\text{Num. of products with } c \text{ introduced by firm } i \text{ between } t \text{ and } t+\tau}{\text{Num. of products introduced by firm } i \text{ between } t \text{ and } t+\tau}$$

We repeat our earlier analysis with this alternative measure of scalability, i.e. estimate the following specification:

$$\mathbb{D}_{cmit\tau} = \alpha + \beta\, \widetilde{\mathcal{SI}}_{amit-1\tau} + \gamma\, scope_{mit} + \lambda_{amt\tau} + \theta_{ami} + \epsilon_{amit\tau} \tag{29}$$

where, as before, the scalability measure $\widetilde{\mathcal{SI}}$ is aggregated to the attribute level and lagged. Again, as before, we include the total number of products sold by the firm, $N_{m,i,t}$ and various fixed effects as controls. The results are shown in columns (3) and (4) in Table 5 and confirm that the strong

Table 5: **Diffusion and Scalability**

Diffusion	(1)	(2)	(3)	(4)
$\mathcal{SI}$	0.0599***	0.0097***		
	(0.000)	(0.001)		
$\widetilde{\mathcal{SI}}$			0.1326***	0.0282***
			(0.000)	(0.001)
scope	-0.0292***	-0.0057***	-0.0013***	-0.0028***
	(0.000)	(0.000)	(0.000)	(0.001)
Observations	3,319,518	3,234,863	3,269,030	3,183,439
R-squared	0.808	0.914	0.812	0.913
Firm-Attribute-Module	N	Y	N	Y
Attribute-Module-Time-Age	Y	Y	Y	Y

Note: The table shows the results of estimating equation 28. The dependent variable is $\mathbb{D}_{cmit\tau}$, the diffusion of characteristic c, launched in module m, by firm i between periods t and $t+\tau$. The independent variable in column (1) and (2) is the scalability $\mathcal{SI}_{amit-1}$ of a given attribute a in module m, firm i and period $t-1$. Details on the construction of these variables can be found in Section 3.3. The independent variable in column (3) and (4) is the scalability $\widetilde{\mathcal{SI}}_{cmit-1\tau}$ of a given attribute a in firm i and module m in period $t-1$. All the specifications include a control for the total number of products sold by firm i in module m at time t.

positive relationship between diffusion and lagged scalability is robust to this alternative measure.

Overall, we find a positive relation between scalability and diffusion implying that shocks increasing the diffusion of knowledge within the firm also increase the diffusion of knowledge outside the firm to the entire industry.

6 Conclusion

The paper develops and validates a rich model of firm size, based on the idea of firms as composite of multiple 'units'. Central to the forces at work is the concept of *scalability* of the firm's expertise. The analysis delivers a simple, yet empirically relevant, insight: the effects of changes in the external environment can depend on a firm's scope and unit-level fundamentals.

There are many promising directions for future research. Our theoretical framework was kept intentionally simple and abstracts from many realistic elements (e.g. within-firm heterogeneity). We also abstracted from dynamics and stochastic fundamentals, both of which are no doubt essential to paint a complete picture of firm heterogeneity. Incorporating these elements and undertaking a full-fledged quantitative analysis is a natural, if ambitious, next step. Finally, exploring the aggregate implications of the link between scalability and diffusion (demonstrated in the previous section) is another interesting direction.

References

Acemoglu, Daron, David Autor, David Dorn, Gordon H Hanson, and Brendan Price, "Import competition and the great US employment sag of the 2000s," *Journal of Labor Economics*, 2016, *34* (S1), S141–S198.

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li, "A theory of falling growth and rising rents," 2019.

_ , Nicholas Bloom, Brian Lucking, Raffaella Sadun, and John Van Reenen, "Turbulence, Firm Decentralization, and Growth in Bad Times," *American Economic Journal: Applied Economics*, 2021, *13* (1), 133–69.

_ , Robin Burgess, Stephen J Redding, and Fabrizio Zilibotti, "The unequal effects of liberalization: Evidence from dismantling the License Raj in India," *American Economic Review*, 2008, *98* (4), 1397–1412.

Akcigit, Ufuk and William R Kerr, "Growth through heterogeneous innovations," *Journal of Political Economy*, 2018, *126* (4), 1374–1443.

Argente, David, Munseob Lee, and Sara Moreira, "The Life Cycle of Products: Evidence and Implications," *Mimeo*, 2020.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The fall of the labor share and the rise of superstar firms," *The Quarterly Journal of Economics*, 2020, *135* (2), 645–709.

Autor, David H., David Dorn, and Gordon H. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, October 2013, *103* (6).

Bai, Liang and Sebastian Stumpner, "Estimating US consumer gains from chinese imports," 2019.

Baldwin, John and Wulong Gu, "The impact of trade on plant scale, production-run length and diversification," in "producer dynamics: New Evidence from micro data," university of Chicago Press, 2009, pp. 557–592.

Barkai, Simcha, "Declining labor and capital shares," *The Journal of Finance*, 2020, *75* (5), 2421–2463.

Basker, Emek and Timothy Simcoe, "Upstream, Downstream: Diffusion and Impacts of the Universal Product Code," *NBER Working Paper No. 24040*, 2017.

Benkard, C Lanier, Ali Yurukoglu, and Anthony Lee Zhang, "Concentration in Product Markets," Technical Report, National Bureau of Economic Research 2021.

Bernard, Andrew B, Stephen J Redding, and Peter K Schott, "Multiproduct firms and trade liberalization," *The Quarterly journal of economics*, 2011, *126* (3), 1271–1318.

Bustos, Paula, "Trade liberalization, exports, and technology upgrading: Evidence on the impact of MERCOSUR on Argentinian firms," *American Economic Review*, 2011, *101* (1), 304–40.

Cao, Dan, Henry R Hyatt, Toshihiko Mukoyama, and Erick Sager, "Firm Growth through New Establishments," 2020.

Crane, Leland Dod and Ryan Decker, "Business Dynamics in the National Establishment Time Series (NETS)," 2019.

Crouzet, Nicolas and Neil R Mehrotra, "Small and large firms over the business cycle," *American Economic Review*, 2020, *110* (11), 3549–3601.

Davis, Steven J and John Haltiwanger, "Gross job creation, gross job destruction, and employment reallocation," *The Quarterly Journal of Economics*, 1992, *107* (3), 819–863.

Dhingra, Swati, "Trading away wide brands for cheap brands," *American Economic Review*, 2013, *103* (6), 2554–84.

Fort, Teresa C, John Haltiwanger, Ron S Jarmin, and Javier Miranda, "How firms respond to business cycles: The role of firm age and firm size," *IMF Economic Review*, 2013, *61* (3), 520–559.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J Klenow, "How destructive is innovation?," *Econometrica*, 2019, *87* (5), 1507–1541.

Garcia-Santana, Manuel and Josep Pijoan-Mas, "The reservation laws in India and the misallocation of production factors," *Journal of monetary economics*, 2014, *66*, 193–209.

Garicano, Luis, Claire Lelarge, and John Van Reenen, "Firm size distortions and the productivity distribution: Evidence from France," *American Economic Review*, 2016, *106* (11), 3439–79.

Gaulier, Guillaume and Soledad Zignago, "Baci: international trade database at the product-level (the 1994-2007 version)," 2010.

Gourio, François and Nicolas Roys, "Size-dependent regulations, firm size distribution, and reallocation," *Quantitative Economics*, 2014, *5* (2), 377–416.

Guner, Nezih, Gustavo Ventura, and Yi Xu, "Macroeconomic Implications of Size-Dependent Policies," *Review of Economic Dynamics*, 2008, *11* (4), 721–744.

Holmes, Thomas J, "The diffusion of Wal-Mart and economies of density," *Econometrica*, 2011, *79* (1), 253–302.

_ and John J Stevens, "An alternative theory of the plant size distribution, with geography and intra-and international trade," *Journal of Political Economy*, 2014, *122* (2), 369–421.

Hopenhayn, H.A., "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60* (5), 1127–1150.

Hottman, Colin J, Stephen J Redding, and David E Weinstein, "Quantifying the sources of firm heterogeneity," *Quarterly Journal of Economics*, 2016, *131* (3), 1291–1364.

Hsieh, Chang-Tai and Esteban Rossi-Hansberg, "The Industrial Revolution in Services," 2020.

Hsieh, C.T. and P.J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

Hyun, Jay and Ryan Kim, "Spillovers and Redistribution through Intra-Firm Networks: The Product Replacement Channel," *Available at SSRN 3480961*, 2019.

Klette, Tor Jakob and Samuel Kortum, "Innovating firms and aggregate innovation," *Journal of political economy*, 2004, *112* (5), 986–1018.

Lashkari, Danial, Arthur Bauer, and Jocelyn Boussard, "Information technology and returns to scale," *Available at SSRN 3458604*, 2018.

Lileeva, Alla and Daniel Trefler, "Improved Access to Foreign Markets Raises Plant-level Productivity...for Some Plants," *The Quarterly Journal of Economics*, 2010, *125* (3), 1051–1099.

Luttmer, Erzo GJ, "On the mechanics of firm growth," *The Review of Economic Studies*, 2011, *78* (3), 1042–1068.

Mariscal, Asier, "Firm Organization and Information Technology: Micro and Macro Implications," Technical Report 2018.

Mayer, Thierry, Marc J Melitz, and Gianmarco IP Ottaviano, "Product mix and firm productivity responses to trade competition," *Review of Economics and Statistics, forthcoming*, 2020.

Melitz, Marc, "The impact of trade on aggregate industry productivity and intra-industry reallocations," *Econometrica*, 2003, *71* (6), 1695–1725.

Mian, Atif and Amir Sufi, "House prices, home equity-based borrowing, and the US household leverage crisis," *American Economic Review*, 2011, *101* (5), 2132–56.

— and — , “What explains the 2007–2009 drop in employment?,” *Econometrica*, 2014, *82* (6), 2197–2223.

Moscarini, Giuseppe and Fabien Postel-Vinay, “The contribution of large and small employers to job creation in times of high and low unemployment,” *American Economic Review*, 2012, *102* (6), 2509–39.

Nocke, Volker and Stephen Yeaple, “Globalization and multiproduct firms,” *International Economic Review*, 2014, *55* (4), 993–1018.

Park, Ziho, “Trade Adjustment: Establishment-Level Evidence,” 2020.

Peters, Michael, “Heterogeneous markups, growth, and endogenous misallocation,” *Econometrica*, 2020, *88* (5), 2037–2073.

Restuccia, D. and R. Rogerson, “Policy Distortions and Aggregate Productivity with Heterogeneous Establishments,” *Review of Economic Dynamics*, 2008, *11* (4), 707–720.

Ridder, Maarten De, “Market power and innovation in the intangible economy,” 2019.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter, “Diverging trends in national and local concentration,” *NBER Macroeconomics Annual*, 2021, *35* (1), 115–150.

Rubinton, Hannah, “The Geography of Business Dynamism and Skill Biased Technical Change,” *FRB St. Louis Working Paper*, 2020, (2020-020).

Saiz, Albert, “The geographic determinants of housing supply,” *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Smith, Dominic and Sergio Ocampo, “The Evolution of US Retail Concentration,” 2020.

Stroebel, Johannes and Joseph Vavra, “House prices, local demand, and retail prices,” *Journal of Political Economy*, 2019, *127* (3), 1391–1436.

A Proofs and derivations

A.1 Derivation of the General Equilibrium Shifter, G_j

With wage w, firm i's unit cost of variety u is $\frac{q}{A_{ij}Z_{uij}}$. The firm is infinitesimal relative to the industry, so given demand $Q_{uij} = \left(\frac{P_{uij}}{P_{ij}}\right)^{-\epsilon}\left(\frac{P_{ij}}{P_j}\right)^{-\theta} Q_j$, the optimal price is $\frac{\theta}{\theta-1}\frac{q}{A_{ij}Z_{uij}}$. The price index summarizing i's output satisfies

$$P_{ij}^{1-\epsilon} = \int_0^{N_j} P_{uij}^{1-\epsilon} du = \int_0^{N_j} \left(\frac{\theta}{\theta-1}\frac{w}{A_{ij}Z_{uij}}\right)^{1-\epsilon} du.$$

Firm i's profit is thus

$$\begin{aligned}
\Pi_{ij}+ =& \int_0^{N_{ij}} \left(\frac{P_{uij}}{P_{ij}}\right)^{-\epsilon}\left(\frac{P_{ij}}{P_j}\right)^{-\theta} Q_j \left[P_{uij} - \frac{W}{A_{ij}Z_{uij}}\right] du - \mathcal{F}_j(N_{ij}) \\
=& \int_0^{N_{ij}} \left(\frac{P_{uij}}{P_{ij}}\right)^{-\epsilon}\left(\frac{P_{ij}}{P_j}\right)^{-\theta} Q_j \left[P_{uij} - \frac{\theta-1}{\theta}P_{uij}\right] du - \mathcal{F}_j(N_{ij}) \\
=& \frac{1}{\theta}Q_j P_j^{\theta} P_{ij}^{1-\theta} - \mathcal{F}_j(N_{ij}) \\
=& \frac{1}{\theta}Q_j P_j^{\theta} \left(\int_0^{N_j} \left(\frac{\theta}{\theta-1}\frac{A_{ij}Z_{uij}}{W}\right)^{\epsilon-1} du\right)^{\frac{\theta-1}{\epsilon-1}} - \mathcal{F}_j(N_{ij})
\end{aligned}$$

We thus have $G_j = \frac{1}{\theta}\left(\frac{\theta}{\theta-1}\right)^{\theta-1} Q_j P_j^{\theta} W^{1-\theta}$.

A.2 Heterogeneous Products

In this section, we explore a version of the model in which a firm's productivity varies exogenously across its products. Firm i in industry j chooses a set U_{ij} of products, and its productivity in producing product $u \in U_{ij}$ is $B_{uij}Z_{uij}$. Its profit is

$$\pi_{ij} = \int_{u\in U_{ij}} \left(P_{uij} - \frac{w}{B_{uij}Z_{uij}}\right) Q_{uij} du - \mathcal{F}_j\left(|U_{ij}|\right)$$

where the demand curve is $Q_{uij} \leq Q_j P_j^{\theta} P_{ij}^{\varepsilon-\theta} P_{uij}^{-\epsilon}$ and the price index is $P_{ij} = \left(\int_{u\in U_{ij}} P_{uij}^{1-\epsilon} du\right)^{\frac{1}{1-\epsilon}}$. The optimal price for product u is $P_{uij} = \frac{\theta}{\theta-1}\frac{w}{B_{uij}Z_{uij}}$, so the price index satisfies $P_{ij} = \frac{\theta}{\theta-1}\frac{w}{\left(\int_{u\in U_{ij}}(B_{uij}Z_{uij})^{\epsilon-1}\right)^{\frac{1}{\epsilon-1}}}$.

Profit is therefore

$$
\begin{aligned}
\pi_{ij} &= \frac{(\theta-1)^{\theta-1}}{\theta^{\theta}} Q_j P_j^{\theta} w^{1-\theta} \left(\int_{u \in U_{ij}} (B_{uij} Z_{uij})^{\epsilon-1} \, du \right)^{\frac{\theta-1}{\epsilon-1}} - \mathcal{F}_j (N_{ij}) \\
&= G_j \left(\int_{u \in U_{ij}} (B_{uij} Z_{uij})^{\epsilon-1} \, du \right)^{\frac{\theta-1}{\epsilon-1}} - \mathcal{F}_j (N_{ij})
\end{aligned}
$$

We make two assumptions. First, expertise is the same across products, so that $Z_{uij} \equiv Z_{ij}$. Second, we assume that the distribution of productivities follows a power law. In particular, i can produce with productivity higher than B is $A_{ij}^{\tau} B^{-\tau}$. Then the cutoff associated with choosing N products is $A_{ij} N^{-1/\tau}$. Profit is

$$
\begin{aligned}
\pi_{ij} &= G_j Z_{ij}^{\theta-1} \left(\int_{u \in U_{ij}} B_{uij}^{\epsilon-1} du \right)^{\frac{\theta-1}{\epsilon-1}} - \mathcal{F}_j (N_{ij}) \\
&= G_j Z_{ij}^{\theta-1} \left(\int_{A_{ij} N_{ij}^{-1/\tau}}^{\infty} B^{\epsilon-1} A_{ij}^{\tau} \tau B^{-\tau-1} dB \right)^{\frac{\theta-1}{\epsilon-1}} - \mathcal{F}_j (N_{ij}) \\
&= G_j (A_{ij} Z_{ij})^{\theta-1} \left(\frac{1}{1-\frac{(\epsilon-1)}{\tau}} N_{ij}^{1-\frac{\epsilon-1}{\tau}} \right)^{\frac{\theta-1}{\epsilon-1}} - \mathcal{F}_j (N_{ij}) \\
&= \frac{G_j}{\left(1-\frac{(\epsilon-1)}{\tau}\right)^{\frac{\theta-1}{\epsilon-1}}} (A_{ij} Z_{ij})^{\theta-1} N_{ij}^{\left(1-\frac{\epsilon-1}{\tau}\right)\frac{\theta-1}{\epsilon-1}} - \mathcal{F}_j (N_{ij})
\end{aligned}
$$

A.3 Returns to Scale

In our multi-product setting, output refers to the composite good $Q_i = \left(\int_0^{N_i} Q_{iu}^{\frac{\epsilon-1}{\epsilon}} du \right)^{\frac{\epsilon}{\epsilon-1}}$. Noting the symmetry across products, the cost of producing Y_i is the solution to

$$
c_i(Q_i) = \min_{x_i, y_i, N_i, L_i, Q_{iu}} N_i W L_i + F N_i
$$

subject to the capacity constraint $x_i + N_i y_i \leq 1$, technology $A_i Z(x_i, y_i) L \geq Q_{iu}$, and within firm aggregation $Q_i \geq N_i^{\frac{\epsilon}{\epsilon-1}} Q_{iu}$. Substituting for L and Y_u yields

$$
c_i(Q_i) = \min_{x_i, y_i, N_i} \frac{W N_i Q_i}{N_i^{\frac{\epsilon}{\epsilon-1}} A_i Z(x_i, y_i)} + F N_i \qquad \text{subject to } x_i + N_i y_i \leq 1
$$

Finally, eliminating y, gives

$$
c_i(Q_i) = \min_{k_i, N_i} \frac{W N_i Q_i}{N_i^{\frac{\varepsilon}{\varepsilon-1}} A_i \frac{z(k_i)}{k_i + N_i}} + F N_i.
$$

Using the envelope theorem and $\epsilon = 2$, marginal cost is $c_i'(Q_i) = \frac{W}{N_i A_i \frac{z(k_i)}{k_i + N_i}}$. Marginal returns to scale can be found by differentiating once more:

$$\begin{aligned} MRTS_i = -\frac{Q_i c_i''(Q_i)}{c_i'(Q_i)} &= \frac{d \ln \frac{N_i z(k_i)}{k_i+N_i}}{d \ln Y_i} \\ &= \underbrace{\frac{\partial \ln \frac{N_i z(k_i)}{k_i+N_i}}{\partial \ln N_i}}_{=S_i} \frac{d \ln N_i}{d \ln Q_i} + \underbrace{\frac{\partial \ln \frac{N_i z(k_i)}{k_i+N_i}}{\partial \ln k_i}}_{=0} \frac{d \ln k_i}{d \ln Q_i} \end{aligned}$$

The envelope theorem implies that changes in k_i do not directly affect the marginal cost – we only need to consider changes in N_i. As before, let $\mathbb{K}(N_i)$ denote the relationship between scalability ratio and scope from (7). The optimality condition (8) yields the optimal scope as a function of desired output Q_i and scalability ratio, denoted $\widetilde{\mathbb{N}}_i(Q_i, k_i)$. Differentiating gives

$$\frac{d \ln N_i}{d \ln Q_i} = \underbrace{\frac{\partial \ln \widetilde{\mathbb{N}}_i}{\partial \ln Q_i}}_{=1/2} + \underbrace{\frac{\partial \ln \widetilde{\mathbb{N}}_i}{\partial \ln k_i}}_{=\frac{1-S_i}{2}} \underbrace{\frac{N \mathbb{K}'(N_i)}{\mathbb{K}(N_i)}}_{=\sigma} \frac{d \ln N_i}{d \ln Q_i} \tag{30}$$

$$= \frac{1}{2 - (1 - S_i)\,\sigma} \tag{31}$$

Together, these imply that the elasticity of marginal cost with respect to the firm's output is

$$MRTS_i = -\frac{Q_i c_i''(Q_i)}{c_i'(Q_i)} = \frac{S_i}{2 - (1 - S_i)\,\sigma} \tag{32}$$

A.4 Power Law Example

Lemma 1 $\lim_{k\to\infty} \frac{\log\left(\frac{z(k)}{k} - \frac{z'(k)^2}{Z_x(1,0)}\right)}{\log\left(\frac{z(k)}{k} - z'(k)\right)} = 1$

Proof 1 *As an intermediate step, we first show that* $\lim_{k\to\infty} \frac{\frac{d}{dk}\left(\frac{z}{k} - \frac{z'^2}{Z_x}\right)}{\frac{d}{dk}\left(\frac{z}{k} - z'\right)}$ *exists and has a finite, strictly positive magnitude. To do this, note first that the homogeneity of Z implies that* $\lim_{k\to\infty} z'(k) = \lim_{k\to\infty} Z_x(k, 1) = \lim_{k\to\infty} Z_x\left(1, \frac{1}{k}\right) = Z_x(1, 0)$. *Second,* $z'' = -\frac{z'(z-kz')}{kz}\frac{1}{\sigma}$. *With these we can derive expressions for the numerator and denominator*

$$\frac{d\left(\frac{z}{k} - \frac{z'^2}{Z_x}\right)}{ds} = \frac{z'}{k} - \frac{z}{k^2} - \frac{2z'}{Z_x} z'' = \frac{z'}{k} - \frac{z}{k^2} + \frac{2z'}{Z_x}\left(\frac{z'(z-kz')}{kz}\frac{1}{\sigma}\right) = \left\{\frac{z'}{Z_x}\frac{2}{\sigma}\frac{kz'}{z} - 1\right\}\frac{z - z'k}{k^2}$$

and

$$\frac{d}{dk}\left(\frac{z}{k} - z'\right) = \frac{z'}{k} - \frac{z}{k^2} - z'' = \frac{z'}{k} - \frac{z}{k^2} + \frac{z'(z-kz')}{kz}\frac{1}{\sigma} = \left\{\frac{1}{\sigma}\frac{kz'}{z} - 1\right\}\frac{z - z'k}{k^2}$$

Together, these imply that

$$\lim_{k\to\infty}\frac{\frac{d}{dk}\left(\frac{z}{k}-\frac{z'^2}{Z_x}\right)}{\frac{d}{dk}\left(\frac{z}{k}-z'\right)}=\lim_{k\to\infty}\frac{\frac{z'}{Z_x}\frac{2}{\sigma(k)}\frac{kz'}{z}-1}{\frac{1}{\sigma(s)}\frac{kz'}{z}-1}=\frac{\frac{2}{\bar\sigma}-1}{\frac{1}{\bar\sigma}-1}=\frac{2-\bar\sigma}{1-\bar\sigma}$$

where $\bar\sigma=\lim_{k\to\infty}\sigma(k)$. *Since that limit exists, we can now use L'Hopital's rule twice to get*

$$\begin{aligned}\lim_{k\to\infty}\frac{\log\left(\frac{z}{k}-\frac{z'^2}{Z_x}\right)}{\log\left(\frac{z}{k}-z'\right)}&=\lim_{k\to\infty}\frac{\frac{z}{k}-z'}{\frac{z}{k}-\frac{z'^2}{Z_x}}\frac{\frac{d}{ds}\left(\frac{z}{k}-\frac{z'^2}{Z_x}\right)}{\frac{d}{dk}\left(\frac{z}{k}-z'\right)}\\&=\lim_{k\to\infty}\frac{\frac{z}{k}-z'}{\frac{z}{k}-\frac{z'^2}{Z_x}}\lim_{s\to\infty}\frac{\frac{d}{dk}\left(\frac{z}{k}-\frac{z'^2}{Z_x}\right)}{\frac{d}{dk}\left(\frac{z}{k}-z'\right)}\\&=\lim_{k\to\infty}\frac{\frac{d}{dk}\left(\frac{z}{k}-z'\right)}{\frac{d}{dk}\left(\frac{z}{k}-\frac{z'^2}{Z_x}\right)}\lim_{k\to\infty}\frac{\frac{d}{dk}\left(\frac{z}{k}-\frac{z'^2}{Z_x}\right)}{\frac{d}{dk}\left(\frac{z}{k}-z'\right)}\\&=1\end{aligned}$$

Proposition 11 *Let the support of A is distributed according to the distribution function $H(A)$ with support $(\underline{A},\bar{A})$. Suppose further that $\lim_{A\nearrow\bar{A}}\frac{\log[1-H(A)]}{\log\left[1-A/\bar{A}\right]}=\kappa$. If $\bar{A}=A^*\equiv\frac{F}{GZ_x(1,0)}$, then the distribution of revenue follows a power law:*

$$\lim_{R\to\infty}\frac{\log\Pr\left(Revenue>R\right)}{\log R}=-\kappa(\bar\sigma-1)$$

where $\bar\sigma=\lim_{k\to\infty}\sigma(k)$. If $\bar{A}<A^*$, then the distribution of revenue is bounded. If $\bar{A}>A^*$, then there is a strictly positive fraction of firms that can earn infinite profit.

Proof 2 *Consider the problem*

$$\max_{k,N}GAN\frac{z(k)}{k+N}-FN$$

The first order conditions imply that scalability solves

$$\frac{GA}{F}\frac{kz'(k)^2}{z(k)}=1$$

and revenue is

$$Revenue=GA\left(z(k)-kz'(k)\right)$$

The optimal choice of scalability defines a strictly increasing function $\hat{A}(k)$ and its inverse, $\hat{k}(A)$ that satisfy $\frac{G\hat{A}(s)}{F}\frac{kz'(k)^2}{z(k)}=1$ and $\frac{GA}{F}\frac{\hat{k}(A)z'\left(\hat{k}(A)\right)^2}{z\left(\hat{k}(A)\right)}=1$ respectively. Using $\hat{k}$, we can express revenue *as a function of A:*

$$Revenue = E(A) \equiv GA\left(z\left(\hat{k}(A)\right) - \hat{k}(A)z'\left(\hat{k}(A)\right)\right)$$

If $\bar{A} < A^*$, then scalability for the largest firm will satisfy $\frac{A^*}{\bar{A}} = \frac{kz'(k)^2}{Z_x(1,0)z(k)}$. Since $\frac{kz'(k)^2}{z(k)Z_x(1,0)}$ is decreasing in k and $\lim_{k\to\infty} \frac{kz'(k)^2}{z(k)Z_x(1,0)} = 1$, the optimal choice of scalability, $\hat{k}(\bar{A})$, will be finite. Thus the largest firm's revenue would be finite, and the distribution of revenue would be bounded. If $\bar{A} > A^*$, revenue for all firms with $A \in (A^*, \bar{A})$ can attain infinite profit and infinite revenue by setting $y = 0$ and letting $N \to \infty$.

We now turn to the case of $\bar{A} = A^* \equiv \frac{F}{GZ_x(1,0)}$. We are interested in how the right tail of the distribution of revenue varies. This is:

$$\lim_{R\to\infty} \frac{\log \Pr\left(Revenue > R\right)}{\log R} = \lim_{R\to\infty} \frac{\log \Pr\left(E(A) > R\right)}{\log R} = \lim_{R\to\infty} \frac{\log \Pr\left(A > E^{-1}(R)\right)}{\log R}$$

We assumed a condition on distribution of productivity as it approaches the upper bound $\bar{A}$, and we can use that to get

$$\begin{aligned}\lim_{R\to\infty} \frac{\log \Pr\left(Revenue > R\right)}{\log R} &= \lim_{R\to\infty} \frac{\log \Pr\left(A > E^{-1}(R)\right)}{\log\left(1 - E^{-1}(R)/A^*\right)} \lim_{R\to\infty} \frac{\log\left(1 - E^{-1}(R)/A^*\right)}{\log R} \\ &= \kappa \lim_{R\to\infty} \frac{\log\left(1 - E^{-1}(R)/A^*\right)}{\log R}\end{aligned}$$

Since $E\left(\hat{A}(s)\right)$ approaches infinity as scalability k approaches infinity, we can express this as

$$\begin{aligned}\lim_{R\to\infty} \frac{\log \Pr\left(Revenue > R\right)}{\log R} &= \kappa \lim_{E\left(\hat{A}(k)\right)\to\infty} \frac{\log\left(1 - E^{-1}\left(E\left(\hat{A}(s)\right)\right)/A^*\right)}{\log E\left(\hat{A}(k)\right)} \\ &= \kappa \lim_{s\to\infty} \frac{\log\left(1 - E^{-1}\left(E\left(\hat{A}(k)\right)\right)/A^*\right)}{\log E\left(\hat{A}(k)\right)} \\ &= \kappa \lim_{s\to\infty} \frac{\log\left\{1 - \left[\frac{z(k)}{kz'(k)^2}\frac{F}{G}\right]/A^*\right\}}{\log\left\{\frac{z(k)}{kz'(k)^2} F\left[z(k) - kz'(k)\right]\right\}}\end{aligned}$$

Using $A^* \equiv \frac{F}{GZ_x(1,0)}$ and rearranging gives

$$\lim_{R\to\infty} \frac{\log \Pr\left(Revenue > R\right)}{\log R} = \kappa \lim_{s\to\infty} \frac{\log\left\{1 - \frac{z(k)}{kz'(k)^2} Z_x(1,0)\right\}}{\log\left\{F\frac{z(k)}{kz'(k)}\frac{z(k)-kz'(k)}{z'(k)}\right\}}$$

We can rearrange the limit into four terms. as follows

$$\lim_{R\to\infty}\frac{\log\Pr\left(Revenue>R\right)}{\log R}=\kappa\lim_{k\to\infty}\frac{\frac{\log\left(\frac{z}{k}-z'\right)}{-\log k}\frac{\log\left\{\frac{z'(k)^2}{Z_x(1,0)}-\frac{z(k)}{k}\right\}}{\log\left(\frac{z}{k}-z'\right)}-\frac{\log\frac{Z_x(1,0)}{z'(k)^2}}{\log k}}{\frac{\log\left\{F\frac{z(k)}{kz'(k)}\frac{z(k)-kz'(k)}{z'(k)}\right\}}{-\log k}}$$
$$=\kappa\frac{B_1B_2-B_3}{B_4}$$

Lemma 1 shows that $B_2\equiv\lim_{s\to\infty}\frac{\log\left\{\frac{z'(k)^2}{Z_x(1,0)}-\frac{z(k)}{k}\right\}}{\log\left(\frac{z}{k}-z'\right)}=1$. *Since* $\lim_{k\to\infty}z'(k)=Z_x(1,0)\in(0,\infty)$, $B_3\equiv\lim_{k\to\infty}\frac{\log\frac{Z_x(1,0)}{z'(k)^2}}{\log k}=0$. *Finally, noting that L'Hopital's rule and the definition of* σ *give* $\lim_{k\to\infty}\frac{\log\frac{z(k)-kz'(k)}{z'(k)}}{\log k}=\lim_{k\to\infty}\sigma(k)=\bar{\sigma}$, *the limits of* B_1 *and* B_4 *are, respectively,*

$$B_1\equiv\lim_{k\to\infty}\frac{\log\left(\frac{z}{k}-z'\right)}{-\log k}=\lim_{k\to\infty}\frac{\log\left(\frac{z-z'k}{z'}\right)-\log k+\log z'}{-\log k}=-\frac{1}{\bar{\sigma}}+1+0=\frac{\bar{\sigma}-1}{\bar{\sigma}}$$
$$B_4\equiv\lim_{k\to\infty}\frac{\log\left\{F\frac{z(k)}{kz'(k)}\frac{z(k)-kz'(k)}{z'(k)}\right\}}{-\log s}=\lim_{k\to\infty}\frac{\log\left\{F\frac{z(k)}{kz'(k)}\right\}}{-\log k}+\frac{\log\frac{z(k)-kz'(k)}{z'(k)}}{-\log k}=0-\frac{1}{\bar{\sigma}}=-\frac{1}{\bar{\sigma}}$$

Together, these imply that

$$\lim_{R\to\infty}\frac{\log\Pr\left(Revenue>R\right)}{\log R}=-\kappa\left(\bar{\sigma}-1\right)$$

A.5 Intermediate Model

In this section we study an intermediate version of the model. We later will show that more general model in the next section can be mapped exactly into this intermediate model.

Consider the problem of a firm with a fixed capacity of attention

$$\pi_i=\max_{x_i,y_i,N_i}GN_i^{\phi}\left(A_iZ(x_i,y_i)\right)^{\psi}-FN_i^{\omega}\qquad\text{subject to}\qquad x_i+N_iy_i\le 1$$

As in the simple model, we can express the firm's problem in terms of scalability and scope. Let $k_i\equiv\frac{x_i}{y_i}$ and $z(k)\equiv Z(k,1)$. Let $\sigma(k)\equiv-\frac{z'(z-kz')}{kzz''}$ be the elasticity of substitution between scalable and local knowledge, and let $S_i\equiv\frac{x_i}{x_i+N_iy_i}$ be the share of knowledge devoted to scalable expertise.

The firm's problem can be expressed as

$$\pi_i=\max_{k_i,y_i,N_i}GN_i^{\phi}\left(A_iz(k_i)y_i\right)^{\psi}-FN_i^{\omega}\qquad\text{subject to}\qquad y_i\le\frac{1}{k_i+N_i}$$

Substituting in the constraint by eliminating y_i gives

$$\pi_i = \max_{k_i, N_i} GA_i^{\psi} N_i^{\phi} \left(\frac{z(k_i)}{k_i + N_i} \right)^{\psi} - FN_i^{\omega} \tag{33}$$

Abusing notation, we can express the firm's profit as

$$\pi_i = \max_N \pi(N; GA_i^{\psi})$$

where

$$\pi(N; GA_i^{\psi}) \equiv \max_k GA_i^{\psi} N^{\phi} \left(\frac{z(k)}{k + N} \right)^{\psi} - FN^{\omega}.$$

Given N, this is a strictly concave problem with an interior solution, so the first order condition $\frac{z'(k)}{z(k)} = \frac{1}{k+N}$ is necessary and sufficient to characterize the choice of k. This first order condition defines a function $\mathbb{K}(N)$. Note that $\frac{N\mathbb{K}'(N)}{\mathbb{K}(N)} = \sigma(\mathbb{K}(N))$.[30] We can thus express the firm's decision as

$$\pi_i = \max_N \pi(N; GA_i^{\psi}) = \max_N GA_i^{\psi} N^{\phi} \left(\frac{z(\mathbb{K}(N))}{\mathbb{K}(N) + N} \right)^{\psi} - FN^{\omega}$$

This is a unidimensional problem and we are interested in finding conditions under which there is a unique solution that is interior. To do this, we examine the first order condition

$$\phi GA_i^{\psi} N^{\phi-1} \left(\frac{z(\mathbb{K}(N))}{\mathbb{K}(N) + N} \right)^{\psi} - \psi \frac{1}{\mathbb{K}(N) + N} GA_i^{\psi} N^{\phi} \left(\frac{z(\mathbb{K}(N))}{\mathbb{K}(N) + N} \right)^{\psi} - \omega FN^{\omega-1} = 0$$

Define the function

$$\mathbb{H}(N) = \frac{GA_i^{\psi}}{F\omega} N^{\phi-\omega} \left(\frac{z(\mathbb{K}(N))}{\mathbb{K}(N) + N} \right)^{\psi} \left(\phi - \psi \frac{N}{\mathbb{K}(N) + N} \right)$$

The first results about the existence and uniqueness of a solution concern finding conditions under which $\lim_{N\to 0} \mathbb{H}(N) > 1$, $\lim_{N\to\infty} \mathbb{H}(N) < 1$, and such that $\mathbb{H}(N)$ is strictly decreasing for any N^* such that $\mathbb{H}(N^*) = 1$.

To find how a change in demand affects firm choices, we can differentiate the first order condition $\mathbb{H}(N) = 1$ with respect to G. This gives the following claim

Proposition 12 *Suppose that there is a unique optimum that is interior. Then*

$$\frac{d\ln N_i}{d\ln G} = \frac{1}{\omega - \phi + \psi(1 - S_i) + \frac{\psi(1-S_i)}{\phi - \psi(1-S_i)} S_i (1 - \sigma_i)} \tag{34}$$

where $S_i = \frac{x_i}{x_i + N_i y_i} = \frac{k_i}{k_i + N_i}$ *is firm i's scalable share of knowledge and* σ_i *is the local elasticity of*

[30] Differentiating $\frac{z'(k)}{z(k)} = \frac{1}{k+N}$ gives $\frac{d\ln k}{d\ln N}\left[\frac{kz''}{z'} - \frac{kz'}{z}\right] = -\frac{k}{k+N}\frac{d\ln k}{d\ln N} - \frac{N}{k+N}$. Using $\frac{k}{k+N} = \frac{kz'(k)}{z(k)}$, this can be rearranged as $\frac{d\ln k}{d\ln N} = -\left[\frac{z'}{kz''}\right]\frac{z-kz'}{z} = \sigma$.

substitution.

Proof 3 *Let* $\mathbb{N}\left(GA_i^{\psi}, k\right)$ *be the optimal choice of scope given productivity, demand, and the choice of scalability, which satisfies the first order condition*

$$\left(\phi - \psi\frac{\mathbb{N}\left(GA_i^{\psi}, k\right)}{k + \mathbb{N}\left(GA_i^{\psi}, k\right)}\right) GA_i^{\psi}\mathbb{N}\left(GA_i^{\psi}, k\right)^{\phi}\left(\frac{z(k)}{k + \mathbb{N}\left(GA_i^{\psi}, k\right)}\right)^{\psi} = \omega F\mathbb{N}\left(GA_i^{\psi}, k\right)^{\omega}$$

The optimal choice of N satisfies the fixed point problem of

$$N_i = \mathbb{N}\left(GA_i^{\psi}, \mathbb{K}(N_i)\right)$$

How does the choice of scope change with demand? Taking logs, differentiating, and rearranging gives (and using the notation $\mathbb{N}_G$ and $\mathbb{N}_k$ to denote partial derivatives with respect to the first and second arguments, respectively):

$$\begin{aligned}\frac{d\ln N_i}{d\ln G} &= \frac{G\mathbb{N}_G}{\mathbb{N}} + \frac{\mathbb{K}(N_i)\mathbb{N}_k}{\mathbb{N}}\frac{N_i\mathbb{K}'(N_i)}{\mathbb{K}(N_i)}\frac{d\ln N_i}{d\ln G} \\ &= \frac{\frac{G\mathbb{N}_G}{\mathbb{N}}}{1 - \sigma_i\frac{\mathbb{K}(N_i)\mathbb{N}_k}{\mathbb{N}}}\end{aligned}$$

These derivatives are:

$$\frac{-\psi\frac{\mathbb{N}}{k+\mathbb{N}}}{\phi - \psi\frac{\mathbb{N}}{k+\mathbb{N}}}\left(\frac{G\mathbb{N}_G}{\mathbb{N}} - \frac{\mathbb{N}}{k+\mathbb{N}}\frac{G\mathbb{N}_G}{\mathbb{N}}\right) + 1 + \phi\frac{G\mathbb{N}_G}{\mathbb{N}} - \psi\frac{\mathbb{N}}{k+\mathbb{N}}\frac{G\mathbb{N}_G}{\mathbb{N}} = \omega\frac{G\mathbb{N}_G}{\mathbb{N}}$$

$$\frac{-\psi\frac{\mathbb{N}}{k+\mathbb{N}}}{\phi - \psi\frac{\mathbb{N}}{k+\mathbb{N}}}\left(\frac{k\mathbb{N}_k}{\mathbb{N}} - \frac{k}{k+\mathbb{N}} - \frac{\mathbb{N}}{k+\mathbb{N}}\frac{k\mathbb{N}_k}{\mathbb{N}}\right) + \phi\frac{k\mathbb{N}_k}{\mathbb{N}} + \psi\left(\frac{kz'(k)}{z(k)} - \frac{k}{k+\mathbb{N}} - \frac{\mathbb{N}}{k+\mathbb{N}}\frac{k\mathbb{N}_k}{\mathbb{N}}\right) = \omega\frac{k\mathbb{N}_k}{\mathbb{N}}$$

Solving for $\frac{G\mathbb{N}_G}{\mathbb{N}}$ and $\frac{k\mathbb{N}_k}{\mathbb{N}}$, evaluating at the optimal N and k, and using $S = \frac{k}{k+N}$ gives

$$\frac{G\mathbb{N}_G}{\mathbb{N}} = \frac{1}{\omega - \phi + \psi(1-S) + \frac{\psi(1-S)}{\phi-\psi(1-S)}S}$$

$$\frac{k\mathbb{N}_k}{\mathbb{N}} = \frac{\frac{\psi(1-S)}{\phi-\psi(1-S)}S}{\omega - \phi + \psi(1-S) + \frac{\psi(1-S)}{\phi-\psi(1-S)}S}$$

Plugging these in and rearranging gives

$$\frac{d\ln N}{d\ln G} = \frac{\frac{1}{\omega-\phi+\psi(1-S)+\frac{\psi(1-S)}{\phi-\psi(1-S)}S}}{1-\sigma\frac{\frac{\psi(1-S)}{\phi-\psi(1-S)}S}{\omega-\phi+\psi(1-S)+\frac{\psi(1-S)}{\phi-\psi(1-S)}S}}$$
$$= \frac{1}{\omega-\phi+\psi(1-S)+\frac{\psi(1-S)}{\phi-\psi(1-S)}S(1-\sigma)}$$

The responses of other outcomes to a shift in demand can also summarized in terms of scalability. For other outcomes:

$$\frac{d\ln k}{d\ln G} = \frac{d\ln x/y}{d\ln G} = \sigma\frac{d\ln N}{d\ln G} \tag{35}$$

$$\frac{d\ln\frac{S}{1-S}}{d\ln G} = \frac{d\ln\frac{kz'(k)}{z(k)-kz'(k)}}{d\ln k}\frac{d\ln k}{d\ln G} = \left(1+\frac{kz''}{z'}-\frac{kz'-z''k}{z-kz'}\right)\sigma\frac{d\ln N}{d\ln G} = (\sigma-1)\frac{d\ln N}{d\ln G} \tag{36}$$

$$\frac{d\ln R}{d\ln G} = \frac{d\ln GA^{\psi}N^{\phi}\left(\frac{z(k)}{k+N}\right)^{\psi}}{d\ln G} = 1+[\phi-\psi(1-S)]\frac{d\ln N}{d\ln G} \tag{37}$$

$$\frac{d\ln\bar{R}}{d\ln G} = 1+[\phi-1-\psi(1-S)]\frac{d\ln N}{d\ln G} \tag{38}$$

A.5.1 Intermediate Model, Case 1: $\sigma < 1$

Assumption 4 *The parameters of the intermediate model satisfy*

(i) $\sigma < 1$

(ii) $\omega \geq \phi$

(iii) If $\omega = \phi$ *then* $A_i > \bar{A}$, *where* $\bar{A}$ *satisfies* $\frac{G\bar{A}_i^{\psi}Z_x(0,1)}{F} = 1$.

Proposition 13 *Suppose that Assumption 4 holds. Then there exists a unique solution, and it is interior.*

Proof 4 *First,* $\sigma < 1$ *implies that* $\frac{\mathbb{K}(N)}{N}$ *is strictly decreasing in* N*. To see this, note that* $\frac{d\ln(\mathbb{K}(N)/N)}{d\ln N} = \sigma(\mathbb{K}(N)) - 1 < 0$*. As a result* $\phi - \psi\frac{N}{\mathbb{K}(N)+N}$ *is strictly decreasing in* N *whenever it is positive. Second,* $\frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N}$ *is strictly decreasing in* N*: since* $\mathbb{K}(N)$ *maximizes* $\frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N}$*, the envelope theorem gives* $\frac{d}{dN}\left(\frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N}\right) = \frac{d}{dN}\max_k\left(\frac{z(k)}{k+N}\right) = -\frac{z(\mathbb{K}(N))}{(\mathbb{K}(N)+N)^2} < 0$*. Finally* $N^{\phi-\omega}$ *is weakly decreasing in* N*.* $\mathbb{H}(N)$ *is thus continuous and strictly decreasing whenever it is positive:*

$$\mathbb{H}(N) = \frac{GA_i^{\psi}}{F\omega}\underbrace{N^{\phi-\omega}}_{\text{weakly decreasing}}\underbrace{\left(\frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N}\right)^{\psi}}_{\text{strictly decreasing}}\underbrace{\left(\phi-\psi\frac{N}{\mathbb{K}(N)+N}\right)}_{\text{strictly decreasing}}$$

In other words, if $\phi \geq \psi$, then $\mathbb{H}(N)$ is strictly decreasing for all N. If $\phi < \psi$, letting $\bar{N}$ denote the unique positive solution to $\phi = \psi \frac{\bar{N}}{\mathbb{K}(\bar{N})+\bar{N}}$, $\mathbb{H}(N)$ is strictly decreasing on $\left[0, \bar{N}\right]$, $\mathbb{H}\left(\bar{N}\right) = 0$, $\mathbb{H}(N) < 0$ for $N > \bar{N}$.

We next show that $\lim_{N\to 0} \mathbb{H}(N) > 1$. $\sigma < 1$ implies $\lim_{N\to 0} \frac{\mathbb{K}(N)}{N} = \infty$ and $\lim_{x\to 0} Z_x(x,y) < \infty$, or equivalently $\lim_{k\to 0} \frac{z(k)}{k} = Z_x(0,1) < \infty$. The former implies $\lim_{N\to 0} \frac{N}{\mathbb{K}(N)+N} = 0$, and together, they imply $\lim_{N\to 0} \frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N} = \lim_{N\to 0} \frac{\mathbb{K}(N)}{\mathbb{K}(N)+N} \lim_{N\to 0} \frac{z(\mathbb{K}(N))}{\mathbb{K}(N)} = 1 \times \lim_{k\to 0} \frac{z(k)}{k} = Z_x(0,1)$. Together, these imply that

$$\lim_{N\to 0} \mathbb{H}(N) = \frac{GA_i^\psi}{F\omega} \left(\lim_{N\to 0} N^{\phi-\omega}\right) \underbrace{\left(\lim_{N\to 0} \frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N}\right)^\psi}_{=Z_x(0,1)} \underbrace{\lim_{N\to 0}\left(\phi - \psi \frac{N}{\mathbb{K}(N)+N}\right)}_{=\phi}$$

If $\omega > \phi$, then $\lim_{N\to 0} N^{\phi-\omega} = \infty$ and hence $\lim_{N\to 0} \mathbb{H}(N) = \infty$. If $\phi = \omega$ then $N^{\phi-\omega} = 1$ and $\lim_{N\to 0} \mathbb{H}(N) = \frac{GA_i^\psi}{F\omega} Z_x(0,1)^\psi \phi = \frac{GA_i^\psi Z_x(0,1)^\psi}{F} > 1$.

If $\psi > \phi$, then we are done: $\mathbb{H}(N)$ is continuous and strictly decreasing on $N \in \left[0, \bar{N}\right]$ with $\mathbb{H}(0) > 1$, $\mathbb{H}\left(\bar{N}\right) = 0$, and $\mathbb{H}(N) < 0$ for $N > \bar{N}$. Thus there is a unique N^* that satisfies $\mathbb{H}(N^*) = 1$, and $N^* \in \left(0, \bar{N}\right)$.

We complete the proof by showing that, in the case of $\phi \geq \psi$, $\lim_{n\to\infty} \mathbb{H}(N) = 0$. $\sigma < 1$ implies $\lim_{x\to\infty} Z_x(x,y) = 0$, or equivalently $\lim_{k\to\infty} z'(k) = 0$. Since $\mathbb{K}'(N) > 0$, $\frac{\mathbb{K}'(N)}{\mathbb{K}'(N)+1} \in (0,1)$. Together, these imply $\lim_{N\to\infty} \frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N} = \lim_{N\to\infty} \frac{z'(\mathbb{K}(N))\mathbb{K}'(N)}{\mathbb{K}'(N)+1} \leq \lim_{N\to\infty} z'(\mathbb{K}(N)) = \lim_{k\to\infty} z'(k) = 0$. In addition, $\lim_{N\to\infty} \frac{N}{\mathbb{K}(N)+N} = 1$ We thus have

$$\lim_{N\to\infty} \mathbb{H}(N) = \frac{GA_i^\psi}{F\omega} \underbrace{\lim_{N\to\infty} N^{\phi-\omega}}_{\leq 1} \underbrace{\lim_{N\to\infty}\left(\frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N}\right)^\psi}_{=0} \underbrace{\lim_{N\to\infty}\left(\phi - \psi\frac{N}{\mathbb{K}(N)+N}\right)}_{=\phi-\psi} = 0$$

Proposition 14 *Suppose Assumption 4 holds. Then a firm responds to higher demand by increasing size, scope, and scalability, but decreasing the scalable share of expertise: $\frac{d\ln N}{d\ln G} > 0$, $\frac{d\ln k}{d\ln G} = \frac{d\ln x/y}{d\ln G} > 0$, $\frac{d\ln \frac{S}{1-S}}{d\ln G} < 0$, $\frac{d\ln R}{d\ln G} > 0$. If $\omega \geq 1$, then the firm responds to an increase in demand by raising size per unit, $\frac{d\ln \bar{R}}{d\ln G} > 0$.*

Proof 5 *We begin with the response of scope. With $\sigma < 1$ and $\omega \geq \phi$, all terms in the denominator of (34) are weakly positive, as $\phi > \psi(1-S)$ at any interior solution. Further, if $S \in (0,1)$ then the denominator must be strictly positive. Therefore $\frac{d\ln N}{d\ln G} > 0$.*

$\frac{d\ln k}{d\ln G} = \frac{d\ln x/y}{d\ln G} > 0$, $\frac{d\ln \frac{S}{1-S}}{d\ln G} < 0$ and $\frac{d\ln R}{d\ln G} > 0$ follow directly from (35), (36), and (37) using $\sigma \in [0,1)$ and $\phi > \psi(1-S)$.

For size per unit, (38) and (34) give

$$\frac{d\ln\bar{R}}{d\ln G} = 1 + \frac{[\phi - 1 - \psi(1-S)]}{\omega - \phi + \psi(1-S) + \frac{\psi(1-S)}{\phi-\psi(1-S)}S(1-\sigma)}$$
$$= \frac{\omega + \frac{\psi(1-S)}{\phi-\psi(1-S)}S(1-\sigma) - 1}{\omega - \phi + \psi(1-S) + \frac{\psi(1-S)}{\phi-\psi(1-S)}S\,(1-\sigma)}$$

which is positive for $S \in (0,1)$ if $\omega \geq 1$.

Proposition 15 *Suppose Assumption 4 holds. Suppose further that, if $\phi > \psi$, that $\omega \geq 1-\sigma$. For firms with higher scalable share of knowledge, size is more sensitive to a change in demand, i.e., $\frac{d}{dS}\frac{d\ln R}{d\ln G} > 0$.*

Proof 6 *The response of size to demand is*

$$\frac{d\ln R}{d\ln G} = 1 + \frac{[\phi - \psi(1-S)]}{\omega - \phi + \psi(1-S) + \frac{\psi(1-S)}{\phi-\psi(1-S)}S(1-\sigma)}$$
$$= 1 + \frac{[\phi - \psi(1-S)]}{\omega - (\phi - \psi(1-S)) + \left(\frac{\phi}{\phi-\psi(1-S)} - 1\right)S(1-\sigma)}$$

Differentiating and letting $Denom \equiv \omega - (\phi - \psi(1-S)) + \left(\frac{\phi}{\phi-\psi(1-S)} - 1\right)S(1-\sigma)$ gives

$$\begin{aligned}
\frac{d}{dS}\frac{d\ln R}{d\ln G} =& \frac{\psi}{Denom} + \frac{\phi-\psi(1-S)}{Denom^2}\left\{\psi + \frac{\phi}{[\phi-\psi(1-S)]^2}S(1-\sigma)\psi - \left(\frac{\phi}{\phi-\psi(1-S)} - 1\right)\left[(1-\sigma) - S\frac{d\sigma}{dS}\right.\right. \\
=& \frac{\psi}{Denom^2}\left\{Denom + [\phi-\psi(1-S)] + \frac{\phi}{\phi-\psi(1-S)}S(1-\sigma) - (1-S)\left[(1-\sigma) - S\frac{d\sigma}{dS}\right]\right\} \\
=& \frac{\psi}{Denom^2}\left\{\omega + \left(2\frac{\phi}{\phi-\psi(1-S)} - 1\right)S(1-\sigma) - (1-S)\left[(1-\sigma) - S\frac{d\sigma}{dS}\right]\right\} \\
=& \frac{\psi}{Denom^2}\left\{\omega + \left(2\frac{\phi S}{\phi-\psi(1-S)} - 1\right)(1-\sigma) + (1-S)S\frac{d\sigma}{dS}\right\}
\end{aligned}$$

Since S is inversely related to scalability, $\frac{d\sigma}{dS}$ is weakly positive as long as σ is non-increasing in scalability. Consider first $\phi \leq \psi$. Then $\frac{\phi S}{\phi-\psi(1-S)} \geq 1$, so that the term in brackets is positive. Consider next $\phi > \psi$. Then $\frac{\phi S}{\phi-\psi(1-S)} \in [0,1]$ and the term in brackets is positive as long as $\omega > 1-\sigma$.

Proposition 16 *Suppose Assumption 4 holds. Suppose also that $\phi \leq \psi$ and that σ is non-increasing in scalability. Then, in response to the same increase in demand, firms with higher* *scalable share of expertise raise scope and scalability by more, i.e.,*

$$\frac{d}{dS}\left(\frac{d\ln N}{d\ln G}\right)>0, \frac{d}{dS}\left(\frac{d\ln x/y}{d\ln G}\right)>0\ .$$

Proof 7 *The expression for the scope elasticity from Proposition 12 can be rearranged as*

$$\frac{d\ln N_i}{d\ln G}=\frac{1}{\omega-\phi+\psi(1-S_i)+(1-S_i)\frac{\psi S_i}{\phi-\psi(1-S_i)}(1-\sigma_i)} \tag{39}$$

Note that $\psi \geq \phi$ implies that $\frac{\psi S_i}{\phi-\psi(1-S_i)}$ is weakly decreasing in S_i. Further, if σ is non-increasing in scalability, then $(1-\sigma)$ is non-increasing in the scalable share of expertise (because S is decreasing in x/y). Together, these imply that the denominator of 39 is strictly decreasing in S, and hence $\frac{d\ln N_i}{d\ln G}$ is strictly increasing in S.

Since σ is weakly increasing in S and $\sigma \in [0,1)$, $\frac{d\ln x/y}{d\ln G}=\sigma\frac{d\ln N}{d\ln G}$ is increasing in S.

A.5.2 Intermediate Model, Case 2: $\sigma > 1$

Proposition 17 *If $\sigma \in (1, 1+\omega)$, $\omega \geq \phi \geq \psi$, and, if $\omega=\phi=\psi$ then $\frac{GA_i^\psi}{F}Z_x(1,0)^\psi<1$. Then there exists a unique solution, and it is interior.*

Proof 8 *Consider first $N \to 0$. Since $\lim_{N\to 0}\frac{N}{\mathbb{K}(N)+N}=1$. Suppose first that $\phi>\psi$. Then $\lim_{N\to 0}\left(\phi-\psi\frac{N}{\mathbb{K}(N)+N}\right)=\phi-\psi$. Also, since $Z\left(0,1\right)>0$, $\lim_{N\to 0}\frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N}=\lim_{n\to 0}\frac{Z(0,1)}{\mathbb{K}(N)+N}=\infty$. We thus have*

$$\lim_{N\to 0}\mathbb{H}(N)=\frac{GA_i^\psi}{F\omega}\underbrace{\lim_{N\to 0}N^{\phi-\omega}}_{\geq 1}\underbrace{\lim_{N\to 0}\left(\frac{z\left(\mathbb{K}(N)\right)}{\mathbb{K}(N)+N}\right)^\psi}_{=\infty}\underbrace{\lim_{N\to 0}\left(\phi-\psi\frac{N}{\mathbb{K}(N)+N}\right)}_{=\phi-\psi>0}=\infty$$

Next suppose $\phi=\psi$. Since $\sigma<1+\omega$ implies that $\lim_{N\to 0}\frac{\mathbb{K}(N)}{N^{1+\omega}}=\infty$. To see this, note that since $\sigma\left(0\right)<1+\omega$, and σ is continuous, so that there is a $\bar{k}$ and a $\tilde{\sigma}\in\left(\sigma\left(0\right),1+\omega\right)$ such that $\sigma\left(k\right)\leq\tilde{\sigma}$ for all $k\in\left[0,\bar{k}\right]$. Since $\frac{N\mathbb{K}'(N)}{\mathbb{K}(N)}=\sigma\left(\mathbb{K}(N)\right)\leq\tilde{\sigma}$, which implies that for $k<\bar{k}$, $\mathbb{K}(N)\geq\bar{k}\mathbb{K}^{-1}\left(\bar{k}\right)^{-\tilde{\sigma}}N^{\tilde{\sigma}}$, so $\lim_{N\to 0}\frac{\mathbb{K}(N)}{N^{1+\omega}}\geq\lim_{N\to 0}\frac{\bar{k}\mathbb{K}^{-1}\left(\bar{k}\right)^{-\sigma}N^{\tilde{\sigma}}}{N^{1+\omega}}=\infty$. Using this, we have

$$\begin{aligned}\lim_{N\to 0}\mathbb{H}(N) &= \frac{GA_i^\psi}{F\omega}\phi Z\left(0,1\right)^\psi\lim_{N\to 0}N^{\phi-\omega}\left(\frac{1}{\mathbb{K}(N)+N}\right)^\phi\frac{\mathbb{K}\left(N\right)}{\mathbb{K}(N)+N}\\ &= \frac{GA_i^\psi}{F\omega}\phi Z\left(0,1\right)^\psi\lim_{N\to 0}\frac{\mathbb{K}(N)}{N^{1+\omega}}\\ &= \infty\end{aligned}$$

Consider next $N\to\infty$. Since $\lim_{N\to\infty}\frac{N}{\mathbb{K}(N)+N}=0$, $\lim_{N\to\infty}\phi-\psi\frac{N}{\mathbb{K}(N)+N}=\phi$. In addition, $\lim_{N\to\infty}z'\left(\mathbb{K}(N)\right)=\lim_{k\to\infty}z'(k)=Z_x(1,0)>0$. We also have $\lim_{N\to\infty}\frac{\mathbb{K}'(N)+1}{\mathbb{K}'(N)}=1+$

$\lim_{N\to\infty} \frac{1}{\mathbb{K}'(N)} = 1+\lim_{N\to\infty} \frac{\frac{N}{\mathbb{K}(N)}}{\frac{N\mathbb{K}'(N)}{\mathbb{K}(N)}} = 1$. *Together, these imply* $\lim_{N\to\infty} \frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N} = \lim_{N\to\infty} \frac{z'(\mathbb{K}(N))\mathbb{K}'(N)}{\mathbb{K}'(N)+1} = Z_x(1,0)$.

$$\lim_{N\to\infty} \mathbb{H}(N) = \frac{GA_i^\psi}{F\omega} \lim_{N\to\infty} N^{\phi-\omega} \underbrace{\lim_{N\to\infty} \left(\frac{z\left(\mathbb{K}(N)\right)}{\mathbb{K}(N)+N}\right)^\psi}_{=Z_x(1,0)^\psi} \underbrace{\lim_{N\to\infty} \left(\phi - \psi \frac{N}{\mathbb{K}(N)+N}\right)}_{=\phi}$$
$$= \frac{GA_i^\psi}{F\omega} \phi Z_x(1,0)^\psi \lim_{N\to\infty} N^{\phi-\omega}$$

If $\phi < \omega$, *then* $\lim_{N\to\infty} \mathbb{H}\left(N\right) = 0$. *If* $\phi = \omega$, $\lim_{N\to\infty} \mathbb{H}(N) < 1$ *only if* $\frac{GA_i^\psi}{F} Z_x(1,0)^\psi < 1$.

Finally, we find conditions under which $\mathbb{H}(N)$ *is decreasing*

$$\mathbb{H}(N) = \frac{GA_i^\psi}{F\omega} N^{\phi-\omega} \left(\frac{z\left(\mathbb{K}(N)\right)}{\mathbb{K}(N)+N}\right)^\psi \left(\phi - \psi \frac{N}{\mathbb{K}(N)+N}\right)$$

Consider any N *such that* $S \in (0,1)$. *Taking logs and differentiating gives*

$$\frac{d\ln \mathbb{H}(N)}{d\ln N} = \phi - \omega + \psi \frac{d\ln \left(\frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N}\right)}{d\ln N} + \frac{-\psi \frac{N}{\mathbb{K}(N)+N}}{\phi - \psi \frac{N}{\mathbb{K}(N)+N}} \frac{d\ln \frac{N}{\mathbb{K}(N)+N}}{d\ln N}$$
$$= \phi - \omega + \psi \left[\frac{\mathbb{K}(N) z'\left(\mathbb{K}(N)\right)}{z\left(\mathbb{K}\left(N\right)\right)} \frac{N\mathbb{K}'(N)}{\mathbb{K}(N)} - \frac{\mathbb{K}(N)\frac{d\ln \mathbb{K}(N)}{d\ln N} + N}{\mathbb{K}(N)+N}\right] + \frac{-\psi \frac{N}{\mathbb{K}(N)+N}}{\phi - \psi \frac{N}{\mathbb{K}(N)+N}} \left(1 - \frac{\mathbb{K}(N)\frac{d\ln \mathbb{K}(N)}{d\ln N} +}{\mathbb{K}(N)+N}\right.$$
$$= \phi - \omega + \psi \left[S\sigma - \left(S\sigma + (1-S)\right)\right] + \frac{-\psi(1-S)}{\phi - \psi\left(1-S\right)} S(1-\sigma)$$
$$= \phi - \omega - \psi(1-S) - \frac{\psi(1-S)}{\phi - \psi(1-S)} S(1-\sigma)$$
$$= \frac{\psi}{\phi - \psi(1-S)} \left\{\left[\frac{\phi-\omega}{\psi} - (1-S)\right] \left[\phi - \psi\left(1-S\right)\right] - (1-S)S(1-\sigma)\right\}$$

Using $\phi \geq \psi$ *and* $\phi - \omega \leq 0$, *we can derive an upper bound by substituting* ϕ *for* ψ *inside the curly brackets:*

$$\frac{d\ln \mathbb{H}(N)}{d\ln N} \leq \frac{\psi}{\phi - \psi(1-S)} \left\{\left[\frac{\phi-\omega}{\phi} - (1-S)\right] \left[\phi - \phi(1-S)\right] - \left(1-S\right) S (1-\sigma)\right\}$$
$$= \frac{\psi}{\phi - \psi(1-S)} \left\{\left[\frac{\phi-\omega}{\phi} - (1-S)\right] \phi S - (1-S) S\left(1-\sigma\right)\right\}$$
$$= \frac{\psi}{\phi - \psi(1-S)} S\left\{\left[\phi - \omega - (1-S)\phi\right] - (1-S)(1-\sigma)\right\}$$
$$= \frac{\psi S}{\phi - \psi(1-S)} \left\{\left(\phi - \omega\right) S - \left(\omega + (1-\sigma)\right)\left(1-S\right)\right\}$$
$$< 0$$

Thus for any N such that $S \in (0,1)$ (i.e., and $N \in (0,\infty)$) $\frac{d\ln \mathbb{H}(N)}{d\ln N} < 0$. This means that there is a unique N^* such that $\mathbb{H}(N^*) = 1$, and it is interior and the global optimum.

Assumption 5 *Assume $\sigma \in (1, 1+\omega)$ and $\omega > \phi \geq \psi$.*

Proposition 18 *Suppose that Assumption 5 holds. Firms respond to increase in demand by increasing size, scope, scalability, and the scalable share of knowledge: $\frac{d\ln R}{d\ln G} > 0$, $\frac{d\ln N}{d\ln G} > 0$, $\frac{d\ln x/y}{d\ln G} > 0$, $\frac{d\ln \frac{S}{1-S}}{d\ln G} > 0$. If ω is sufficiently large ($\omega > \sigma$), then firms respond to an increase in demand by raising size per unit, i.e., $\frac{d\ln \bar{R}}{d\ln G} > 0$.*

Proof 9 *We first show that scope increases with demand. To do this, we rearrange (34) as*

$$\frac{d\ln N}{d\ln G} = \frac{1}{\frac{\psi}{\phi-\psi(1-S)}\left\{\left[\frac{\omega-\phi}{\psi} + (1-S)\right][\phi - \psi(1-S)] + (1-S)S(1-\sigma)\right\}}$$

To show that the term in the curly brackets is positive for $S \in (0,1)$, we can use $\phi \geq \psi$ to get

$$\begin{aligned}\left[\frac{\omega-\phi}{\psi} + (1-S)\right][\phi - \psi(1-S)] &\geq \left[\frac{\omega-\phi}{\phi} + (1-S)\right][\phi - \phi(1-S)] \\ &= \left[\frac{\omega-\phi}{\phi} + (1-S)\right]\phi S \\ &= (\omega-\phi)S^2 + (1-S)S\omega \\ &> (1-S)S(\sigma-1)\end{aligned}$$

where the last line uses $\omega \geq \phi$ and $\omega > \sigma$.

The positive responses of revenue, scalability, and scalable share directly from $\frac{d\ln N}{d\ln G} > 0$ and (37), (35), (36).

For size per unit, (38) and (34) give

$$\begin{aligned}\frac{d\ln \bar{R}}{d\ln G} &= 1 + \frac{[\phi - 1 - \psi(1-S)]}{\omega - \phi + \psi(1-S) + \frac{\psi(1-S)}{\phi-\psi(1-S)}S(1-\sigma)} \\ &= \frac{\omega + \frac{\psi(1-S)}{\phi-\psi(1-S)}S(1-\sigma) - 1}{\omega - \phi + \psi(1-S) + \frac{\psi(1-S)}{\phi-\psi(1-S)}S(1-\sigma)} \\ &= \left[\omega + \frac{\psi(1-S)}{\phi-\psi(1-S)}S(1-\sigma) - 1\right]\frac{d\ln N}{d\ln G}\end{aligned}$$

The term in brackets is positive because $\frac{\psi(1-S)}{\phi-\psi(1-S)}S \leq \frac{\phi(1-S)}{\phi-\phi(1-S)}S = 1-S \leq 1$, so that $\frac{\psi(1-S)}{\phi-\psi(1-S)}S(1-\sigma) \geq (1-\sigma)$. Thus if $\omega > \sigma$, then the term in brackets is strictly positive. This along with $\frac{d\ln N}{d\ln G} > 0$ implies that $\frac{d\ln \bar{R}}{d\ln G} > 0$.

Proposition 19 *Suppose that assumption 5 holds and that σ is non-decreasing in scalability. Then firms with higher scalability have a higher sensitivity of size to demand, i.e., $\frac{d}{dS}\left(\frac{d\ln R}{d\ln G}\right) > 0$.*

Proof 10 *The response of size to demand is*

$$\begin{aligned}\frac{d\ln R}{d\ln G} &= 1 + \frac{[\phi - \psi(1-S)]}{\omega - \phi + \psi(1-S) + \frac{\psi(1-S)}{\phi-\psi(1-S)}S(1-\sigma)} \\ &= 1 + \frac{[\phi - \psi(1-S)]}{\omega - (\phi - \psi(1-S)) + \left(\frac{\phi}{\phi-\psi(1-S)} - 1\right)S(1-\sigma)}\end{aligned}$$

Differentiating and letting $Denom \equiv \omega - (\phi - \psi(1-S)) + \left(\frac{\phi}{\phi-\psi(1-S)} - 1\right)S(1-\sigma)$ gives

$$\begin{aligned}\frac{d}{dS}\frac{d\ln R}{d\ln G} =& \frac{\psi}{Denom} + \frac{\phi-\psi(1-S)}{Denom^2}\left\{\psi + \frac{\phi}{[\phi-\psi(1-S)]^2}S(1-\sigma)\psi - \left(\frac{\phi}{\phi-\psi(1-S)} - 1\right)\left[(1-\sigma) - S\frac{d\sigma}{dS}\right]\right\} \\ =& \frac{\psi}{Denom^2}\left\{Denom + [\phi - \psi(1-S)] + \frac{\phi}{\phi-\psi(1-S)}S(1-\sigma) - (1-S)\left[(1-\sigma) - S\frac{d\sigma}{dS}\right]\right\} \\ =& \frac{\psi}{Denom^2}\left\{\omega + \left(2\frac{\phi}{\phi-\psi(1-S)} - 1\right)S(1-\sigma) - (1-S)\left[(1-\sigma) - S\frac{d\sigma}{dS}\right]\right\} \\ =& \frac{\psi}{Denom^2}\left\{\omega + \left(2\frac{\phi S}{\phi-\psi(1-S)} - 1\right)(1-\sigma) + (1-S)S\frac{d\sigma}{dS}\right\} \\ & \frac{\psi}{Denom^2}\left\{\omega + 1 - \sigma + 2\left(1 - \frac{\phi S}{\phi-\psi(1-S)}\right)(\sigma - 1) + (1-S)S\frac{d\sigma}{dS}\right\}\end{aligned}$$

The term in brackets is positive because $\omega + 1 - \sigma > 0$ and $\phi \geq \psi$ which implies $\frac{\phi S}{\phi-\psi(1-S)} \in [0,1]$.

Proposition 20 *Suppose that Assumption 5 holds. Suppose also that $\sigma < 1 + \phi$ and that σ is non-decreasing in k. Then firms with higher scalability respond to an increase in demand by raising scope, scalability, and the scalable share of expertise. i.e.,*

$$\frac{d}{dS}\left(\frac{d\ln N}{d\ln G}\right) > 0, \frac{d}{dS}\left(\frac{d\ln k}{d\ln G}\right) > 0, \frac{d}{dS}\left(\frac{d\ln\frac{S}{1-S}}{d\ln G}\right) > 0.$$

Proof 11 *We first show that if $\sigma \leq 1 + \phi$ then $\frac{d\ln N}{d\ln G}$ is increasing in S. Recall from (34) as that we can express the response to scope as*

$$\frac{d\ln N}{d\ln G} = \frac{1}{\omega - \phi + \psi(1-S)\left\{1 - \frac{\phi S}{\phi-\psi(1-S)}\frac{\sigma-1}{\phi}\right\}}$$

Note that $\frac{\phi S}{\phi-\psi(1-S)} \in [0,1]$ and is increasing in S. Since $(1-S)$ is decreasing in S and $1-\frac{\phi S}{\phi-\psi(1-S)}\frac{1-\sigma}{\phi}$ is decreasing in S (because $\frac{\sigma-1}{\phi} \leq 1$ and σ is weakly increasing in S), so that the denominator is decreasing in S. As a result, $\frac{d\ln N}{d\ln G}$ is increasing in S.

Since σ is weakly increasing in S and $\sigma > 1$, $\frac{d\ln k}{d\ln G} = \sigma\frac{d\ln N}{d\ln G}$ and $\frac{d\ln\frac{S}{1-S}}{d\ln G} = (\sigma-1)\frac{d\ln N}{d\ln G}$ are both increasing in S.

A.6 General Model

The firm's profit is

$$\pi_i = \max_{N,x,y,E} GN^{\phi}\left[A_i Z\left(x,y\right)\right]^{\psi} - FN^{\omega} - HE^{\gamma} \text{ subject to } x^{\mu} + Ny^{\mu} \leq E$$

Let $z\left(k\right) = Z\left(k^{1/\mu},1\right)^{\mu}$, where $k = \left(\frac{x}{y}\right)^{\mu}$. This can be expressed as

$$\pi_i = \max_{N,k,y,E} GN^{\phi}A_i^{\psi}z(k)^{\psi/\mu}y^{\psi} - FN^{\omega} - HE^{\gamma} \text{ subject to } (k+N)y^{\mu} \leq E$$

Eliminating y gives

$$\begin{aligned}\pi_i &= \max_{N,k,E} GN^{\phi}A_i^{\psi}\left[z(k)\frac{E}{k+N}\right]^{\psi/\mu} - FN^{\omega} - HE^{\gamma}\\ &= \max_{N,k,E} GA_i^{\psi}N^{\phi}\left(\frac{z(k)}{k+N}\right)^{\psi/\mu}E^{\psi/\mu} - FN^{\omega} - HE^{\gamma}\end{aligned}$$

The optimal choice of E satisfies

$$GA_i^{\psi}N^{\phi}\left(\frac{z(k)}{k+N}\right)^{\psi/\mu}\frac{\psi}{\mu}E^{\psi/\mu-1} = \gamma HE^{\gamma-1}$$

or $E = \left[GA_i^{\psi}N^{\phi}\left(\frac{z(k)}{k+N}\right)^{\psi/\mu}\frac{\psi}{\gamma\mu}\frac{1}{H}\right]^{\frac{1}{\gamma-\psi/\mu}}$. Plugging this into the expression for profit yields

$$\pi_i = \max_{N,k}\left\{1-\frac{\psi}{\gamma\mu}\right\}\left[\frac{\psi}{\gamma\mu}\frac{1}{H}\right]^{\frac{\psi/\mu}{\gamma-\psi/\mu}}\left[GA_i^{\psi}N^{\phi}\left(\frac{z(k)}{k+N}\right)^{\psi/\mu}\right]^{\frac{1}{1-\frac{\psi}{\gamma\mu}}} - FN^{\omega}$$

Define the following variables

$$\begin{aligned}\tilde{\phi} &= \phi\frac{1}{1-\frac{\psi}{\gamma\mu}}\\ \tilde{\psi} &= \frac{\frac{\psi}{\mu}}{1-\frac{\psi}{\gamma\mu}}\\ \tilde{A}_i &= \frac{A_i^{\mu}}{H^{1/\gamma}}\\ \tilde{G} &= \left\{1-\frac{\psi}{\gamma\mu}\right\}\left[\frac{\psi}{\gamma\mu}\right]^{\frac{\tilde{\psi}}{\gamma}}G^{\frac{1}{1-\frac{\psi}{\gamma\mu}}}\end{aligned}$$

Substituting these into the expression for firm i's profit gives

$$\pi_i = \max_{N,k} \tilde{G}\tilde{A}_i^{\tilde{\psi}} N^{\tilde{\phi}} \left(\frac{z(k)}{k+N}\right)^{\tilde{\psi}} - FN^{\omega}$$

Note that the maximization problem is the same as (33) of the intermediate model, so the same analysis applies. Further, as in the medium version of the problem, the scalable share of knowledge is

$$S \equiv \frac{x^{\mu}}{x^{\mu} + Ny^{\mu}} = \frac{k}{k+N}$$

Finally, $\tilde{G} = \left\{1 - \frac{\psi}{\gamma\mu}\right\} \left[\frac{\psi}{\gamma\mu}\right]^{\frac{\tilde{\psi}}{\gamma}} G^{\frac{1}{1-\frac{\psi}{\gamma\mu}}}$ and $x/y = k^{1/\mu}$ imply

$$\frac{d\ln R}{d\ln G} = \left(1 - \frac{\psi}{\mu\gamma}\right) \frac{d\ln R}{d\ln \tilde{G}} > 0 \tag{40}$$

$$\frac{d\ln N}{d\ln G} = \left(1 - \frac{\psi}{\mu\gamma}\right) \frac{d\ln N}{d\ln \tilde{G}} > 0 \tag{41}$$

$$\frac{d\ln x/y}{d\ln G} = \frac{1}{\mu}\left(1 - \frac{\psi}{\mu\gamma}\right) \frac{d\ln k}{d\ln \tilde{G}} > 0 \tag{42}$$

$$\frac{d\ln S}{d\ln G} = \left(1 - \frac{\psi}{\mu\gamma}\right) \frac{d\ln S}{d\ln \tilde{G}} > 0 \tag{43}$$

$$\frac{d\ln \bar{R}}{d\ln G} = \left(1 - \frac{\psi}{\mu\gamma}\right) \frac{d\ln \bar{R}}{d\ln \tilde{G}} > 0 \tag{44}$$

Lemma 2 *Let $\sigma(x/y)$ be the elasticity of substitution of $Z(x,y)$ and let $\tilde{\sigma}(\tilde{x}/\tilde{y})$ be the elasticity of substitution of the production function $\tilde{Z}(\tilde{x},\tilde{y}) \equiv Z\left(\tilde{x}^{1/\mu}, \tilde{y}^{1/\mu}\right)^{\mu}$. Then σ and $\tilde{\sigma}$ are related by $\frac{\tilde{\sigma}(k)-1}{\tilde{\sigma}(k)} = \frac{1}{\mu}\frac{\sigma\left(k^{1/\mu}\right)-1}{\sigma\left(k^{1/\mu}\right)}$.*

Proof 12 *Note that $\sigma = \frac{Z_x Z_y}{Z Z_{xy}}$, and $\tilde{\sigma} = \frac{\tilde{Z}_{\tilde{x}}\tilde{Z}_{\tilde{y}}}{\tilde{Z}\tilde{Z}_{\tilde{x}\tilde{y}}}$. Since both Z and $\tilde{Z}$ exhibit constant returns to scale, their derivatives are linked by*

$$\tilde{Z}(k,1) = Z\left(k^{1/\mu}, 1\right)^{\mu}$$

The first derivatives are linked by

$$\tilde{Z}_{\tilde{x}} = Z^{\mu-1} Z_x k^{1/\mu - 1}$$

or, letting $\alpha \equiv \frac{k\tilde{Z}_{\tilde{x}}(k,1)}{\tilde{Z}(k,1)}$,

$$\alpha \equiv \frac{k\tilde{Z}_{\tilde{x}}}{\tilde{Z}} = \frac{k\left(Z^{\mu-1}Z_x k^{1/\mu-1}\right)}{Z^{\mu}} = \frac{k^{1/\mu}Z_x}{Z}$$

Constant returns to scale implies

$$\tilde{Z}_{\tilde{y}}(k,1) = \tilde{Z} - k\tilde{Z}_x = Z^{\mu} - Z^{\mu-1}Z_x k^{1/\mu}$$

Differentiating once more with respect to k gives an expression for the cross-derivative:

$$\tilde{Z}_{\tilde{x}\tilde{y}}(k,1) = Z^{\mu-1}Z_x k^{1/\mu-1} - \left(1-\frac{1}{\mu}\right) Z^{\mu-2}Z_x^2 k^{1/\mu}k^{1/\mu-1} - \frac{1}{\mu}Z^{\mu-1}Z_{xy}k^{1/\mu}k^{1/\mu-1} - \frac{1}{\mu}Z^{\mu-1}Z_x k^{1/\mu-1}$$

Using $\alpha = \frac{k^{1/\mu}Z_x}{Z} = 1-\frac{Z_y}{Z}$, the fact that constant returns to scale of Z implies that $k^{1/\mu}Z_{xx}\left(k^{1/\mu},1\right) = -Z_{xy}\left(k^{1/\mu},1\right)$, and $\sigma \equiv \frac{Z_x Z_y}{ZZ_{xy}}$, we can rearrange this as

$$\begin{aligned}
\frac{k\tilde{Z}_{\tilde{x}\tilde{y}}(k,1)}{\tilde{Z}(k,1)} &= \frac{Z_x k^{1/\mu}}{Z} - \left(1-\frac{1}{\mu}\right)\left(\frac{Z_x k^{1/\mu}}{Z}\right)^2 - \frac{1}{\mu}\frac{k^{1/\mu}k^{1/\mu}Z_{xx}}{Z} - \frac{1}{\mu}\frac{Z_x k^{1/\mu}}{Z} \\
&= \alpha - \left(1-\frac{1}{\mu}\right)\alpha^2 + \frac{1}{\mu}\alpha(1-\alpha)\frac{ZZ_{xy}}{Z_x Z_y} - \frac{1}{\mu}\alpha \\
&= \alpha(1-\alpha)\left[\left(1-\frac{1}{\mu}\right) + \frac{1}{\mu}\frac{1}{\sigma}\right]
\end{aligned}$$

Finally, the $\tilde{\sigma}$ can be expressed as

$$\begin{aligned}
\frac{1}{\tilde{\sigma}} &= \frac{\tilde{Z}\tilde{Z}_{\tilde{x}\tilde{y}}}{\tilde{Z}_{\tilde{x}}\tilde{Z}_{\tilde{y}}} = \frac{1}{\frac{k\tilde{Z}_{\tilde{x}}}{\tilde{Z}}}\frac{1}{\frac{\tilde{Z}_{\tilde{y}}}{\tilde{Z}}}\frac{k\tilde{Z}_{\tilde{x}\tilde{y}}}{\tilde{Z}} = \frac{1}{\alpha(1-\alpha)}\alpha(1-\alpha)\left[\left(1-\frac{1}{\mu}\right) + \frac{1}{\mu}\frac{1}{\sigma}\right] \\
&= \left(1-\frac{1}{\mu}\right) + \frac{1}{\mu}\frac{1}{\sigma}
\end{aligned}$$

Note that this can be rearranged as

$$\frac{\tilde{\sigma}-1}{\tilde{\sigma}} = \frac{1}{\mu}\frac{\sigma-1}{\sigma}$$

Note that $\tilde{\sigma}$ is between 1 and σ, and approaches 1 as μ grows large.

A.6.1 General Model, Substitutes: $\sigma > 1$

Let $\tilde{\sigma}$ satisfy $\frac{\tilde{\sigma}-1}{\tilde{\sigma}} = \frac{1}{\mu}\frac{\sigma-1}{\sigma}$. Note that $\tilde{\sigma}$

Assumption 6 *The parameters satisfy*

- $1 < \tilde{\sigma} < \omega$
- $1 - \frac{\phi}{\omega} - \frac{\psi}{\mu\gamma} \geq 0$
- $\phi \geq \frac{\psi}{\mu}$
- *If* $1 - \frac{\phi}{\omega} - \frac{\psi}{\mu\gamma} = 0$ *and* $\phi = \frac{\psi}{\mu}$, *then* $\left(1-\frac{\psi}{\gamma\mu}\right)^{1-\frac{\psi}{\gamma\mu}}\left[\frac{\psi}{\gamma\mu}\right]^{\frac{\psi}{\gamma\mu}}\frac{GA_i^{\psi}Z_x(1,0)^{\frac{\psi}{\mu}}}{F^{1-\frac{\psi}{\gamma\mu}}H^{\frac{\psi}{\gamma\mu}}} < 1$

Proposition 21 *Under Assumption 6, there is a unique solution, and it is interior.*

Proof 13 *From Proposition 17, there is a unique solution that is interior if if* $1 < \tilde{\sigma} < \omega$ *and* $\omega \geq \tilde{\phi} \geq \tilde{\psi}$, *and, if* $\omega = \tilde{\phi} = \tilde{\psi}$, *that* $\frac{\tilde{G}\tilde{A}_i^{\tilde{\psi}}}{F} Z_x(1,0)^{\tilde{\psi}} < 1$. *These conditions are equivalent to Assumption 5.*

Proposition 22 *Suppose that Assumption 6 holds. Firms respond to an increase in demand by increasing size, scope, scalability, and the scalable share of knowledge:* $\frac{d\ln R}{d\ln G} > 0$, $\frac{d\ln N}{d\ln G} > 0$, $\frac{d\ln x/y}{d\ln G} > 0$, *and* $\frac{d\ln\frac{S}{1-S}}{d\ln G} > 0$. *If* ω *is sufficiently large, then* $\omega > \tilde{\sigma}$, *then firms respond to an increase in demand by increasing size per unit,* $\frac{d\ln\bar{R}}{d\ln G} > 0$.

Proof 14 *This follows from Proposition 18 and (40), (41), (42), (43), and (44).*

Proposition 23 *Suppose that Assumption 6 holds and that* σ *is non-decreasing in scalability. Then firms with higher scalability have a higher sensitivity of size to demand, i.e.,* $\frac{d}{dS}\left(\frac{d\ln R}{d\ln G}\right) > 0$.

Proof 15 *This follows from Proposition 19 and (40).*

Proposition 24 *Suppose that Assumption 6 holds. Suppose further that* $\tilde{\sigma} \leq 1 + \frac{\phi}{1-\frac{\psi}{\mu\gamma}}$ *and that* σ *is non-decreasing in scalability. Then firms with higher scalability respond to an increase in demand by raising scope, scalability, and the scalable share of knowledge by more, i.e.,*

$$\frac{d}{dS}\left(\frac{d\ln N}{d\ln G}\right) > 0,\ \frac{d}{dS}\left(\frac{d\ln x/y}{d\ln G}\right) > 0,\ \frac{d}{dS}\left(\frac{d\ln S}{d\ln G}\right) > 0$$

Proof 16 *This follows from Proposition 20 and (41), (42), and (43).*

A.6.2 General Model, Complements: $\sigma < 1$.

Again, let $\tilde{\sigma}$ satisfy $\frac{\tilde{\sigma}-1}{\tilde{\sigma}} = \frac{1}{\mu}\frac{\sigma-1}{\sigma}$. Note that $\tilde{\sigma} \in (\sigma, 1)$, with $\tilde{\sigma}(\sigma,\mu)$ approaching 1 as μ grows large.

Assumption 7 *The parameters satisfy*

- $\tilde{\sigma} < 1$
- $1 - \frac{\phi}{\omega} - \frac{\psi}{\mu\gamma} \geq 0$
- *If* $1 - \frac{\phi}{\omega} - \frac{\psi}{\mu\gamma} = 0$, *then* $\left\{1 - \frac{\psi}{\gamma\mu}\right\}^{1-\frac{\psi}{\gamma\mu}} \left[\frac{\psi}{\gamma\mu}\right]^{\frac{\psi}{\gamma\mu}} \frac{GA_i^{\psi} Z_x(0,1)^{\frac{\psi}{\mu}}}{F^{1-\frac{\psi}{\gamma\mu}} H^{\frac{\psi}{\gamma\mu}}} > 1$

Proposition 25 *Under Assumption 7, there is a unique solution, and it is interior.*

Proof 17 *From proposition 13, there is a unique solution that is interior if if* $\tilde{\sigma} < 1$ *and* $\omega \geq \tilde{\phi}$, *and, if* $\omega = \tilde{\phi}$, *that* $\frac{\tilde{G}\tilde{A}_i^{\tilde{\psi}}}{F_i} Z_x(1,0)^{\tilde{\psi}} > 1$. *These conditions are equivalent to Assumption 4.*

Proposition 26 *Suppose that Assumption 7 holds. Firms respond to an increase in demand by increasing size, scope, and scalability, and decreasing the scalable share of knowledge:* $\frac{d \ln R}{d \ln G} > 0$, $\frac{d \ln N}{d \ln G} > 0$, $\frac{d \ln x/y}{d \ln G} > 0$, *and* $\frac{d \ln \frac{S}{1-S}}{d \ln G} > 0$. *If* $\omega > 1$ *then firms respond to an increase in demand by increasing size per unit,* $\frac{d \ln \bar{R}}{d \ln G} > 0$.

Proof 18 *This follows from proposition 18 and (40), (41), (42), (43), and (44).*

Proposition 27 *Suppose that Assumption 7 holds and that* σ *is non-increasing in scalability. Suppose further that, if* $\phi > \frac{\psi}{\mu}$, *that* $\omega \geq 1 - \tilde{\sigma}$. *Then firms with higher scalability have a higher sensitivity of size to demand, i.e.,* $\frac{d}{dS}\left(\frac{d \ln R}{d \ln G}\right) > 0$.

Proof 19 *This follows from Proposition 15 and (40).*

Proposition 28 *Suppose that Assumption 7 holds. Suppose further that* $\phi \leq \frac{\psi}{\mu}$ *and that* σ *is non-increasing in scalability. Then firms with higher scalability respond to an increase in demand by raising scope and scalability by more, i.e.,*

$$\frac{d}{dS}\left(\frac{d \ln N}{d \ln G}\right) > 0, \ \frac{d}{dS}\left(\frac{d \ln x/y}{d \ln G}\right) > 0$$

Proof 20 *This follows from Proposition 20 and (41) and (42).*

A.7 Richer Heterogeneity

The firm's profit is

$$\pi_i = \max_{N,x,y} GN^{\phi}\left[A_i Z(x,y)\right]^{\psi} - F_i N^{\omega} - H_i\left[\left(\frac{x}{a_i^x}\right)^{\mu} + N\left(\frac{y}{a_i^y}\right)^{\mu}\right]^{\gamma}$$

We make the substitutions

$$\tilde{N} = \left(\frac{a_i^x}{a_i^y}\right)^{\mu} N$$
$$k = (x/y)^{\mu}$$
$$z(k) = Z(k^{1/\mu}, 1)^{\mu}$$

$$\pi_i = \max_{\tilde{N},k,y} GA_i^{\psi}\left(\frac{a_i^y}{a_i^x}\right)^{\mu\phi}\tilde{N}^{\phi} z(k)^{\psi/\mu} y^{\psi} - F_i\left(\frac{a_i^y}{a_i^x}\right)^{\mu\omega}\tilde{N}^{\omega} - \frac{H_i}{(a_i^x)^{\mu\gamma}}\left[k + \tilde{N}\right]^{\gamma} y^{\mu\gamma}$$

Solving for y, factoring out $F_i\left(\frac{a_i^y}{a_i^x}\right)^{\mu\omega}$, and making the substitutions

$$\tilde{\phi} = \frac{\phi}{1 - \frac{\psi}{\mu\gamma}}$$
$$\tilde{\psi} = \frac{\psi/\mu}{1 - \frac{\psi}{\mu\gamma}}$$
$$\tilde{G} = \left(1 - \frac{\psi}{\gamma\mu}\right)\left[\frac{\psi}{\gamma\mu}\right]^{\frac{\tilde{\psi}}{\gamma}} G^{\frac{1}{1-\frac{\psi}{\gamma\mu}}}$$
$$\tilde{A}_i = \frac{\left[(a_i^x)^{1+\frac{\omega-\tilde{\phi}}{\tilde{\psi}}}(a_i^y)^{-\frac{\omega-\tilde{\phi}}{\tilde{\psi}}} A_i\right]^{\mu}}{F_i^{1/\tilde{\psi}} H_i^{1/\gamma}}$$

yields

$$\pi_i = F_i\left(\frac{a_i^y}{a_i^x}\right)^{\mu\omega}\left\{\max_{\tilde{N},k} \tilde{G}\tilde{A}_i^{\tilde{\psi}}\tilde{N}^{\tilde{\phi}} z(k)^{\tilde{\psi}} - \tilde{N}^{\omega}\right\}$$

Let R_i, N_i, x_i/y_i, and S_i be firm i's size, scope, scalability, and scalable share. Let $\tilde{R}_i \equiv F_i\left(\frac{a_i^y}{a_i^x}\right)^{\mu\omega} R_i$ ans $\tilde{S}_i = \frac{k_i}{k_i+\tilde{N}_i}$. Since the problem inside the brackets is identical A.6, we can follow the exact derivations in Appendix A.6 to get that

$$\frac{d\ln\tilde{N}_i}{d\ln\tilde{G}} = \frac{1}{\omega - \tilde{\phi} + \tilde{\psi}(1-\tilde{S}_i) + \frac{\tilde{\psi}(1-\tilde{S}_i)}{\tilde{\phi}-\tilde{\psi}(1-\tilde{S}_i)}\tilde{S}_i(1-\tilde{\sigma}_i)}$$
$$\frac{d\ln x/y}{d\ln\tilde{G}} = \tilde{\sigma}\frac{d\ln\tilde{N}}{d\ln\tilde{G}}$$
$$\frac{d\ln\frac{\tilde{S}}{1-\tilde{S}}}{d\ln\tilde{G}} = (\tilde{\sigma}-1)\frac{d\ln\tilde{N}}{d\ln\tilde{G}}$$
$$\frac{d\ln\tilde{R}}{d\ln\tilde{G}} = 1 + \left[\tilde{\phi} - \tilde{\psi}(1-\tilde{S}_i)\right]\frac{d\ln\tilde{N}}{d\ln\tilde{G}}$$

Finally, note that $S_i = \frac{\left(\frac{x}{a_i^x}\right)^{\mu}}{\left(\frac{x}{a_i^x}\right)^{\mu} + N\left(\frac{y}{a_i^y}\right)^{\mu}} = \frac{k}{k+\tilde{N}} = \tilde{S}_i$, and that $\tilde{N}_i$ and $\tilde{R}_i$ are proportional to N_i and R_i with constants of proportionality that do not change with demand. As a result, the elasticities of size, scope, size per unit, scalability and scalable share with respect to demand are the same as in section A.6, and can be expressed in terms of only the scalable share S and parameters that are common across firms.

Since $S_i = \tilde{S}_i = \frac{k_i z'(k_i)}{z(k_i)} = \frac{\left(\frac{x_i}{y_i}\right)^{\mu} z'\left(\left(\frac{x_i}{y_i}\right)^{\mu}\right)}{z\left(\left(\frac{x_i}{y_i}\right)^{\mu}\right)}$, these elasticities can be expressed in terms of scalability and parameters that are common across firms.

Lastly, under Assumptions 2 and 3, the elasticities of size, scope, scalability, and scalable share are increasing in the scalable share, or equivalently, increasing in scope.

B Data Appendix

B.1 Datasets

B.1.1 Nielsen Data

The product data comes from the Nielsen Retail Measurement Services (RMS), generated by point-of-sale systems in retail stores, covering the period 2006-2015. Each individual store reports weekly sales volume and quantities sold of every barcode that had any sales volume during that week. A barcode is a 12-digit Universal Product Code (UPC) consisting of 12 numerical digits that is uniquely assigned to each specific good available in stores. UPCs were created to allow retail outlets to determine prices and inventory accurately and improve transactions along the supply chain distribution (Basker and Simcoe, 2017).

The main advantage of the RMS data set is its size and coverage. Overall, the RMS data consists of sales across retail establishments worth approximately $2 trillion, representing 53% of all sales in grocery stores, 55% in drug stores, 32% in mass merchandisers, 2% in convenience stores, and 1% in liquor stores. Another key distinctive feature of this database is that the collection points include more than 40,000 distinct stores from around 90 retail chains, across 371 metropolitan statistical areas (MSAs) and 2,500 counties. As a result, the data provide good coverage of the universe of products and firms in the CPG sector. In comparison to other scanner data sets collected at the store level, Nielsen RMS covers a much wider range of products and stores. In comparison to scanner data sets collected at the household level, Nielsen RMS also has a wider range of products because it reflects the universe of all transactions for the categories it covers, as opposed to the purchases made by a sample of households.

The original data consist of more than one million distinct products identified by barcodes, organized into a hierarchical structure. Each barcode is classified into one of the 1070 product modules, that are organized into 104 product groups, that are then grouped into 10 major departments. The ten major departments are: Health and Beauty Aids, General Merchandise, Dry Grocery (e.g., baby food, canned vegetables), Frozen Foods, Dairy, Deli, Packaged Meat, Fresh Produce, Non-Food Grocery, and Alcohol. For example, a 31-ounce bag of Tide Pods has UPC 037000930389, is produced by Procter & Gamble, and belongs to the product module "Detergent-Packaged" in product group "Detergent," which belongs to the "Non-Food Grocery" department. The product group "Detergent" includes several product modules, including automatic dishwasher compounds, detergents heavy duty liquid, detergents light duty, detergents packaged, dishwasher rinsing aids, and packaged soap.

Nielsen RMS data does not include information on manufacturing firms. However, products can be linked with firms using information obtained from the GS1 US Data Hub. GS1 is the company that issues to producers the prefix associates with barcodes. In order to issue a UPC, firms must first obtain a GS1 company prefix. The prefix is a five- to ten-digit number that identifies firms in

Table A.I: Summary Statistics of Firms by Censoring

	All	By Censoring Type			
		Complete	Right	Left	Right&Left
Total # of firms	22,938	4,425	4,726	6,107	7,680
Duration (quarters)					
average	23	11	17	16	40
less than 4	16	35	18	20	0
less than 16	44	81	58	61	0
above 28	43	3.9	18	18	100
Sales (quarterly, $1,000)					
mean	1,183	8.4	24	111	3,425
25th percentile	.6	.1	.1	1.3	8.9
median	6	.5	1.1	6.8	57
75th percentile	52	3.3	7.7	36	366
95th percentile	1,177	32	87	350	7,387
Products (quarterly)					
mean	12	2.1	3.2	5.3	27
25th percentile	1	1	1	1.3	2.7
median	2.8	1	1.8	3	6.7
75th percentile	6.6	2.5	3.5	5.5	18
95th percentile	37	5.8	10	16	98
Sectors (quarterly)					
mean	1.7	1.1	1.3	1.4	2.4
25th percentile	1	1	1	1	1
median	1	1	1	1	1.4
75th percentile	1.7	1	1.1	1.5	2.5
95th percentile	4	2	2.3	3	6.6

Notes: The table presents summary statistics of firms included in the baseline pooled sample for the period 2006q1-2015q4. Firms that are already active in 2006q1 and 2006q2 are left-censored, and firms with sales in 2015q3 and 2015q4 are right-censored. Firms that enter and exit in the period under analysis are classified as "Complete", firms for which we can determine entry but not exit are classified as "Right", firms for which we do not observe entry but we observe exit are classified as "Left", and firms for which both entry and exit cannot be determined are both right and left-censored ("Right&Left"). For each of these categories, we report the total number of observations and statistics on duration, sales, number of firms, and number of sectors. Under duration, we report the average duration and the share of observations with duration below 4, 16 and above 28 quarters. The duration refers to the number of quarters for which we observe the firms. The statistics for sales are computed by determining the average quarterly sales (in thousands of dollars), deflated by the Consumer Price Index for All Urban Consumers. The table presents the average and distribution statistics of the variables total sales, number of products and number of sectors. Sectors refers to the number of different product groups as classified by Nielsen.

their products' UPCs. The GS1 data include the name and address of the firm associated with each prefix, which allows us to append a firm name to the UPCs included in the Nielsen-RMS data. A "firm" in the database is defined based on the entity that purchased the barcodes from GS1, which is typically the manufacturer, such as Procter & Gamble. Table A.I presents the characteristics of firms by type of censoring. Among the approximately 23,000 firms covered in the sample, we can measure the age of about 9,000, and the remaining 14,000 are firms that were born before 2006.

B.1.2 NETS Data

The National Establishment Time-Series (NETS) Database is produced by Walls & Associates teamed up with Dun and Bradstreet to convert their archival establishment data into a time-series database of establishment information. We use the dataset that comprises annual observations on specific lines of business at unique locations over the period 1990–2017 operating in the U.S.

The NETS data allow us to observe sales and employment of each line of business identifier. For each identifier, we can track its sales and employment over time at the 8-digit level of Standard Industrial Classification (SIC) and at specific latitudes and longitudes. In addition, we can assign each line of business to a headquarters using firm identifiers. The NETS firm concept is based on a common headquarters establishment.

Crane and Decker (2019) explores the representativeness of NETS in the cross section, comparing the data to the U.S. Census Bureau's datasets, reporting that the static distributions of NETS data can be made reasonably comparable to official sources, on average, in terms of establishment size, industry, and geography cells, with some limitations when only covering small firms. The main limitations of the NETS data regards annual establishment growth and firm lifecycle patterns. Our focus is on cross sectional differences on long-term growth rates, and thus less likely to be strongly affected by measurement errors.

NETS data preparation – Our analysis uses changes in size and scope in the periods 2001-2006 and 2006-2011 for the housing shock, and 2006-2015 for the China Shock. Our baseline sample, includes all firm-sector observations for whom we can measure the shocks. We measure at the firm-sector-year and firm-sector-location-year, the total size (measures as total number of employees or total sales), and the number of business lines (scope). Sectors are mapped into broader 4-digit SIC codes, and locations are mapped into MSAs.

We completed several exercises to ensure robustness of the patterns documented in the paper. Crane and Decker (2019) makes several suggestions on how to use the NETS data in order to minimize measurement errors. Following their suggestions, we created samples that exclude small firm-sector observations (less than five employees) and excludes single-establishment firms (only one location).

We also explored the role of alternative measures of firm definition. Our baseline exercise uses an iterative procedure to determine the ultimate headquarters. The original data lists for each business line a direct headquarters, but some headquarters list some other headquarters, creating indirect headquarters. We link all establishments that are related through headquarters, either directly or indirectly, under a single firm identifier by "rolling up" firm identifiers, capturing the ultimate headquarter. For robustness, we use a procedure that further associates business lines to ensure that we capture mergers and acquisitions (Crane and Decker, 2019).

Table A.I presents the characteristics of firm-sector observations of the baseline and robustness samples underlying the results provided in the paper.

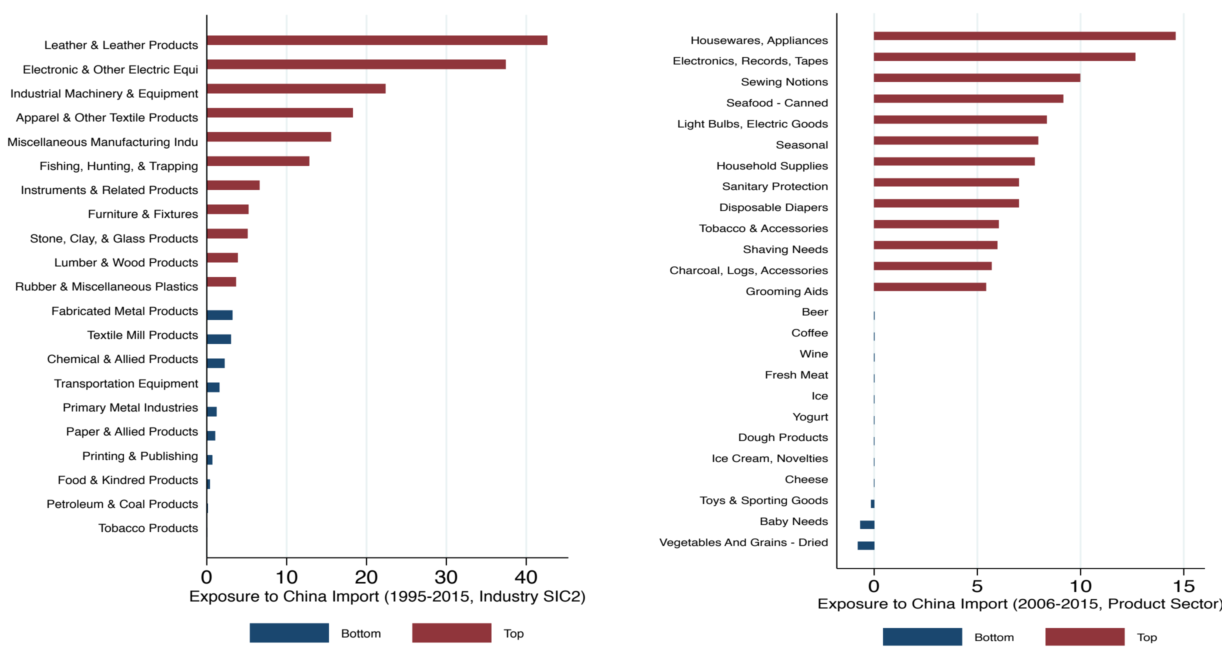
B.2 Sectoral Shocks

B.2.1 China Import Penetration Shock

In this section we describe the data sources used to update the China Shock measure by Autor et al. (2013). The U.S. value of shipments ($Y_{j,06}$) at the 4-digit 1987 SIC industry level from NBER-CES.[31] We obtain gross output ($YO_{j,06}$) at the 4-digit ISIC rev.3 industry level for several European countries from UNIDO. We pick the five largest European economies: Germany, France, UK, Italy and Spain following Bai and Stumpner (2019). These countries have the largest coverage at the 4-digit ISIC rev.3 industry level in the UNIDO data set. We use the trade flows ($M_{j,06}$,$E_{j,06}$) for both U.S. and European countries from UN Comtrade at the HS 6-digit level, obtained from CEPII-BACI (Gaulier and Zignago (2010)).[32] This data set is slightly different from the trade flows data used in Acemoglu et al. (2016) since their data on trade flows is directly from UN Comtrade. The data provided by CEPII-BACI is a harmonized version of UN Comtrade that reconciles the declarations of the exporter and the importer. This harmonization procedure extends considerably the number of countries (150 countries in CEPII_BACI) for which trade data are available, as compared to the original data set. Lastly, we use the PCE deflator provided by the BEA-NIPA for the US and the PCE deflator provided by Eurostat for the five European countries we consider.

Figure A.1: China import penetration 2006–2015 by sector

(a) NETS (b) Nielsen



Note: The figure shows the average value of the baseline measure of China import penetration 2006–2015 $\Delta IP^1_{j,06-15}$ and the values of a selective group of sectors: the top and bottom 20 sectors.

[31] http://www.nber.org/nberces

[32] http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=1

B.2.2 House Price Shock

Table A.II: Instrumental Variables Regression - First Stage

	(1)	(2)	(3)	(4)
Housing Supply Elasticity	-6.709***	-5.243***	7.260***	4.378***
	(0.838)	(0.825)	(0.980)	(0.947)
Observations	228	223	228	224
R-squared	0.221	0.371	0.195	0.418
Controls	N	Y	N	Y
Period	2001-2006	2001-2006	2006-2011	2006-2011

Note: Table shows results from the first-stage instrumental variable regression in equation 22. The unit of observation is an MSA, the dependent variable is house price growth over 2001-2006 in columns 1-2, and house price growth over 2007-2011 in columns 3-4. For the Saiz Elasticity Measure, higher values signal an MSA with more elastic housing supply. The control variables include the change in the number of retail establishments, the change in all establishments, the change in the construction share of employment, the change in the retail share of employment, and the change in the share of employment in the non-tradable sector.

C Scalability

C.1 Illustrative Example

$$\mathcal{SI}_{mijt} \equiv 1 - \frac{\text{Unique}_{mijt}}{\text{Scope}_{mijt} \times \text{NumAttributes}_{mjt}}$$

An illustrative example: Firm A has the following set of products

Table C1: Measuring Scalability: An example (Lamps, incandescent)

Firm	Product	Attribute	
		Style	Use
General Electric	1	Clear	Nite Fixture
General Electric	2	Halogen	Appliance
General Electric	3	Clear	Bath & Vanity
General Electric	4	Clear	Ceiling Fan
General Electric	5	Frost	Chandelier

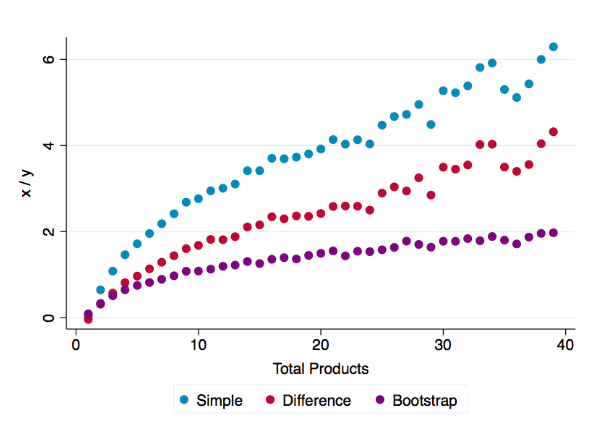
- Unique Characteristics: clear, halogen, nite fixture, etc.
- Attributes: style, use, etc.

$$\begin{aligned}
\text{Style:} \quad & S_{Style,GE} \equiv 1 - \frac{\text{Unique Characteristics}}{\text{Scope}} = 1 - \frac{3}{5} = 0.4 \\
\text{Use:} \quad & S_{Use,GE} \equiv 1 - \frac{\text{Unique Characteristics}}{\text{Scope}} = 1 - \frac{5}{5} = 0 \\
\text{Firm level:} \quad & S_{GE} \equiv 1 - \frac{\text{Unique Characteristics}}{\text{Scope} \times \text{Number of Attributes}} = 1 - \frac{5+3}{5+5} = 0.2
\end{aligned}$$

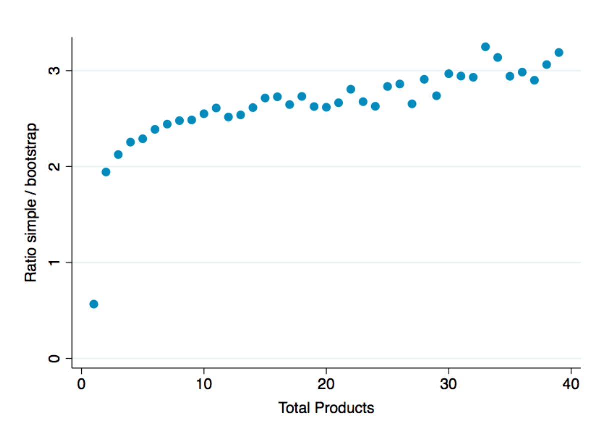
C.2 Bootstrap

A potential concern is that our scalability index mechanically increases as firms introduce more products. In order to minimize this concern, we construct an alternative index where we randomize products, within a sector, and assign them to firms of different sizes. Then, we compute an alternative (bootstrapped) scalability index that we use as reference. This index captures all the mechanical relationship between the scalability index and the total number of products. Thus, our baseline scalability index is throughout the paper used always relative to this alternative index.

Figure C1: Scalability - Alternative (Bootstraped) Version



(a) $\frac{x}{y}$



(b) $\frac{x}{y}$ Bootstrap

Note: Panel (a) shows $\frac{x}{y}$ as a function of the total number of products of the firm. The blue dots show the estimates using the original measure. The purple dots are the estimates of $\frac{x}{y}$ when the sample of products is randomized within modules and across firms. The red dots are the difference between the original measure and the bootstrapped version. Panel (b) shows the ratio of the original measure of $\frac{x}{y}$ and the bootstrapped version. $\frac{x}{y}$ is computed using the entire portfolio of product of the firms using data from 2006 to 2015.

Panel (a) in Figure C1 shows our measure of $\frac{x}{y}$ for both the original index and the bootstrapped version. The purple dots show that part of the positive relationship between $\frac{x}{y}$ and the total number of products comes from the fact that, as firms grow, they are more likely to have products sharing common attributes. Nonetheless, our measure captures a size-dependent relationship that goes beyond that established by chance. The red dots show the difference between the original measure and the bootstrapped version. As shown in the graph, the difference is also increasing with size. Panel (b) shows the ratio between the original and the bootstrapped version. It shows that the ratio increases as firms add products to the portfolio indicating that larger firms replicate characteristics across their products.

D Additional Results

D.1 Heterogeneous response to shocks with size and scope measures using establishment data

D.1.1 China import penetration shock

Table D2: Heterogeneous response with China import penetration shock for continuing firms (log change): using revenue as proxy for size

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_j \times$ size$_{ij}$	0.127***	0.022**		
	(0.033)	(0.011)		
$\Delta G_j \times$ scope$_{ij}$			0.119**	0.059***
			(0.051)	(0.021)
Observations	321,518	321,518	321,108	321,115
R-squared	0.096	0.019	0.030	0.242
Within R-Squared	.082	.015	.016	.239
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data and size measures based on revenue. The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j (revenue) and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. All the specifications include sector effects and robust standard errors.

Table D3: Heterogeneous response with China import penetration shock for continuing and exiting firms (alternative growth rates)

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_j \times$ size$_{ij}$	0.115***	0.123***		
	(0.016)	(0.015)		
$\Delta G_j \times$ scope$_{ij}$			0.051***	0.046***
			(0.013)	(0.011)
Observations	711,264	711,264	710,993	710,993
R-squared	0.035	0.043	0.027	0.031
Within R-Squared	.008	.013	0	0
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data. The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. The change in the dependent variables are calculated as in Davis and Haltiwanger (1992), i.e. $2(y_t - y_{t-1})/(y_t + y_{t-1})$. All the specifications include sector effects and robust standard errors.

Table D4: Heterogeneous response with China import penetration shock for continuing firms (log change): alternative firm definitions

Employment, firm definition 1

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij} \times$ size$_{ij}$	0.065***	0.015		
	(0.020)	(0.010)		
$\Delta G_{ij} \times$ scope$_{ij}$			0.091***	0.039**
			(0.029)	(0.016)
Observations	334,274	334,274	334,044	334,044
R-squared	0.048	0.018	0.028	0.136
Within R-Squared	.041	.013	.021	.132
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Employment, firm definition 2

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{j} \times$ size$_{ij}$	0.039*	0.013		
	(0.023)	(0.011)		
$\Delta G_{j} \times$ scope$_{ij}$			0.099***	0.046**
			(0.032)	(0.018)
Observations	326,049	326,049	325,806	325,806
R-squared	0.045	0.015	0.029	0.134
Within R-Squared	.038	.011	.022	.131
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data. The difference relative to the baseline results is that we use alternative definitions of firm: (1) uses information on headquarters in 2015, (2) uses the time-varying definition of headquarters. The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. All the specifications include sector effects and robust standard errors.

Table D5: Heterogeneous response with China import penetration shock for continuing firms (log change): excluding micro-firms

	Firms with at least 2 plants			
	(1)	(2)	(3)	(4)
VARIABLES				
$\Delta G_{ij} \times$ size$_{ij}$	0.428*	0.130		
	(0.221)	(0.110)		
$\Delta G_{ij} \times$ scope$_{ij}$			0.244	0.068
			(0.258)	(0.132)
Observations	12,564	12,564	12,460	12,460
R-squared	0.085	0.050	0.094	0.142
Within R-Squared	.046	.013	.057	.11
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data. The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. All the specifications include sector effects and robust standard errors.

D.1.2 Housing price 2001-2006 shock

Table D6: Heterogeneous response with housing price 2001-2006 shock for continuing firms (log change): using revenue as proxy for size

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_j\times$ size$_{ij}$	0.029***	0.003***		
	(0.002)	(0.000)		
$\Delta G_j\times$ scope$_{ij}$			0.002	0.000
			(0.002)	(0.002)
Observations	5,931,743	5,931,816	5,931,743	5,933,769
R-squared	0.065	0.005	0.045	0.071
Within R-Squared	.021	.004	0	.069
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data and size measures based on revenue. The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j (revenue) and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. All the specifications include sector effects and robust standard errors.

Table D7: Heterogeneous response with housing price 2001-2006 shock for continuing and exiting firms (alternative growth rates)

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij}\times$ size$_{ij}$	0.007***	0.003***		
	(0.001)	(0.000)		
$\Delta G_{ij}\times$ scope$_{ij}$			0.003**	-0.003**
			(0.001)	(0.001)
Observations	4,140,905	4,140,905	4,140,905	4,140,905
R-squared	0.044	0.006	0.008	0.066
Within R-Squared	.038	.004	.002	.064
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data. The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. The change in the dependent variables are calculated as in Davis and Haltiwanger (1992), i.e. $2(y_t - y_{t-1})/(y_t + y_{t-1})$. All the specifications include sector effects and robust standard errors.

Table D8: Heterogeneous response with housing price 2001-2006 shock for continuing firms (log change): alternative definitions of firm

Employment, alternative firm definition 1

	(1)	(2)	(3)	(4)
VARIABLES				
$\Delta G_{ij} \times$ size$_{ij}$	0.021***	0.002***		
	(0.002)	(0.001)		
$\Delta G_{ij} \times$ scope$_{ij}$			0.001	-0.003
			(0.003)	(0.002)
Observations	5,899,725	5,899,725	5,899,725	5,899,725
R-squared	0.052	0.005	0.010	0.081
Within R-Squared	.047	.004	.004	.081
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Employment, alternative firm definition 2

	(1)	(2)	(3)	(4)
VARIABLES				
$\Delta G_{ij} \times$ size$_{ij}$	0.012***	0.004***		
	(0.002)	(0.001)		
$\Delta G_{ij} \times$ scope$_{ij}$			0.003	-0.002
			(0.002)	(0.002)
Observations	4,140,905	4,140,905	4,140,905	4,140,905
R-squared	0.049	0.005	0.008	0.064
Within R-Squared	.044	.004	.002	.063
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data. The difference relative to the baseline results is that we use alternative definitions of firm: (1) uses Crane and Decker (2019) methodology (2) uses the time-varying definition of headquarters. The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. All the specifications include sector effects and robust standard errors.

Table D9: Heterogeneous response with housing price 2001-2006 shock for continuing firms (log change): alternative sample of sectors

Employment, All sectors

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij} \times$ size$_{ij}$	0.019***	0.002***		
	(0.002)	(0.000)		
$\Delta G_{ij} \times$ scope$_{ij}$			0.004*	-0.001
			(0.002)	(0.001)
Observations	6,430,800	6,430,800	6,430,604	6,430,604
R-squared	0.052	0.006	0.008	0.072
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Employment, Non-tradables sectors

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij} \times$ size$_{ij}$	0.006**	0.005***		
	(0.003)	(0.001)		
$\Delta G_{ij} \times$ scope$_{ij}$			0.009**	0.006**
			(0.004)	(0.003)
Observations	1,745,126	1,745,126	1,745,126	1,745,126
R-squared	0.044	0.009	0.008	0.082
Within R-Squared	.041	.008	.004	.081
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data. Non-tradables are defined as in Mian and Sufi (2011). The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. All the specifications include sector effects and robust standard errors.

Table D10: Heterogeneous response with housing price 2001-2006 shock for continuing firms (log change): excluding micro-firms

Firms with at least 2 plants

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij} \times$ size$_{ij}$	0.022**	0.013**		
	(0.010)	(0.005)		
$\Delta G_{ij} \times$ scope$_{ij}$			0.005	0.008***
			(0.004)	(0.003)
Observations	200,924	200,924	200,924	200,924
R-squared	0.030	0.026	0.014	0.026
Within R-Squared	.018	.001	.001	.001
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Firms with at least 5 employees

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij} \times$ size$_{ij}$	0.031***	0.006***		
	(0.005)	(0.001)		
$\Delta G_{ij} \times$ scope$_{ij}$			-0.000	0.001
			(0.002)	(0.002)
Observations	1,727,617	1,727,617	1,727,617	1,727,617
R-squared	0.041	0.011	0.015	0.074
Within R-Squared	.028	.007	.002	.07
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data. The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. All the specifications include sector effects and robust standard errors.

D.1.3 Housing price 2006-2011 shock

Table D11: Heterogeneous response with housing price 2006-2011 shock for continuing firms (log change): using revenue as proxy for size

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_j \times$ size$_{ij}$	0.000	0.000		
	(0.002)	(0.000)		
$\Delta G_j \times$ scope$_{ij}$			0.006**	0.004**
			(0.003)	(0.002)
Observations	6,967,942	6,967,965	6,967,942	6,968,641
R-squared	0.095	0.009	0.028	0.091
Within R-Squared	.071	.007	.003	.089
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y
Firm x Sector	N	N	N	N
Sample	1	1	1	1

Note: The table shows the results of estimating equations 24 with the NETS data and size measures based on revenue. The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j (revenue) and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. All the specifications include sector effects and robust standard errors.

Table D12: Heterogeneous response with housing price 2006-2011 shock for continuing and exiting firms (alternative growth rates)

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij} \times$ size$_{ij}$	0.017***	0.016***		
	(0.002)	(0.002)		
$\Delta G_{ij} \times$ scope$_{ij}$			0.006***	0.007***
			(0.002)	(0.002)
Observations	8,022,454	8,022,454	8,022,374	8,022,374
R-squared	0.059	0.065	0.057	0.060
Within R-Squared	.003	.006	.001	.001
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data. The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. The change in the dependent variables are calculated as in Davis and Haltiwanger (1992), i.e. $2(y_t - y_{t-1})/(y_t + y_{t-1})$. All the specifications include sector effects and robust standard errors.

Table D13: Heterogeneous response with housing price 2006-2011 shock for continuing firms (log change): alternative definitions of firm

Employment, alternative firm definition 1

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij}\times$ size$_{ij}$	0.002*	-0.000		
	(0.001)	(0.001)		
$\Delta G_{ij}\times$ scope$_{ij}$			0.012***	0.008***
			(0.003)	(0.002)
Observations	6,947,499	6,947,499	6,947,499	6,947,499
R-squared	0.035	0.013	0.019	0.132
Within R-Squared	.028	.011	.012	.13
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Employment, alternative firm definition 2

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij}\times$ size$_{ij}$	0.003*	-0.003***		
	(0.001)	(0.001)		
$\Delta G_{ij}\times$ scope$_{ij}$			0.007***	0.006***
			(0.002)	(0.002)
Observations	5,043,577	5,043,577	5,043,577	5,043,577
R-squared	0.036	0.016	0.017	0.115
Within R-Squared	.028	.014	.009	.112
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data. The difference relative to the baseline results is that we use alternative definitions of firm: (1) uses Crane and Decker (2019) methodology (2) uses the time-varying definition of headquarters. The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. All the specifications include sector effects and robust standard errors.

Table D14: Heterogeneous response with housing price 2006-2011 shock for continuing firms (log change): alternative sample of sectors

Employment, All sectors

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij}\times$ size$_{ij}$	0.003***	0.001		
	(0.001)	(0.000)		
$\Delta G_{ij}\times$ scope$_{ij}$			0.004*	0.003**
			(0.002)	(0.002)
Observations	7,507,106	7,507,106	7,506,895	7,506,895
R-squared	0.036	0.012	0.016	0.095
Within R-Squared	.029	.01	.008	.092
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Employment, Non-tradables sectors

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij}\times$ size$_{ij}$	0.010***	-0.002*		
	(0.002)	(0.001)		
$\Delta G_{ij}\times$ scope$_{ij}$			0.002	-0.005
			(0.004)	(0.003)
Observations	1,990,063	1,990,063	1,990,063	1,990,063
R-squared	0.031	0.015	0.017	0.102
Within R-Squared	.024	.014	.01	.101
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data. Non-tradables are defined as in Mian and Sufi (2011). The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. All the specifications include sector effects and robust standard errors.

Table D15: Heterogeneous response with housing price 2006-2011 shock for continuing firms (log change): excluding micro-firms

Firms with at least 2 plants

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij} \times$ size$_{ij}$	0.038**	0.014		
	(0.019)	(0.012)		
$\Delta G_{ij} \times$ scope$_{ij}$			0.137***	0.096***
			(0.028)	(0.020)
Observations	146,601	146,601	146,601	146,601
R-squared	0.029	0.034	0.043	0.081
Within R-Squared	.017	.009	.031	.058
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Firms with at least 5 employees

VARIABLES	(1)	(2)	(3)	(4)
$\Delta G_{ij} \times$ size$_{ij}$	-0.005	0.000		
	(0.004)	(0.001)		
$\Delta G_{ij} \times$ scope$_{ij}$			0.006***	0.004**
			(0.002)	(0.002)
Observations	1,733,121	1,733,121	1,733,121	1,733,121
R-squared	0.016	0.019	0.015	0.098
Within R-Squared	.013	.013	.012	.092
Dep. Var.	Δ size	Δ scope	Δ size	Δ scope
Sector	Y	Y	Y	Y

Note: The table shows the results of estimating equations 24 with the NETS data. The dependent variable in columns (1) and (3) is the log change in the total size of firm i in sector j and in columns (2) and (4) is the log change in the number of establishments of firm i in sector j. All the specifications include sector effects and robust standard errors.

D.2 Heterogeneous response to shocks with size and scope measures using product data

Table D16: Heterogeneous Response to Demand Shocks - Extensive and Intensive Margins (Alternative Growth Rates)

	(1) Δ size	(2) Δ scope	(3) Δ size	(4) Δ scope	(5) Δ size	(6) Δ scope	(7) Δ size	(8) Δ scope
$\Delta g\times$ size	0.027***	0.000	0.112*	0.012				
	(0.005)	(0.004)	(0.063)	(0.040)				
$\Delta g\times$ scope					0.015***	0.011***	0.024	0.098***
					(0.005)	(0.003)	(0.062)	(0.037)
Obs.	24,146	25,917	28,137	30,001	24,146	25,917	28,137	30,001
R-squared	0.125	0.041	0.075	0.029	0.068	0.069	0.055	0.032
Sector	Y	Y	Y	Y	Y	Y	Y	Y
Sample	China	China	Housing	Housing	China	China	Housing	Housing
Period	2006-2015	2006-2015	2006-2011	2006-2011	2006-2015	2006-2015	2006-2011	2006-2011

Note: The table reports the results of estimating equations 24. The dependent variable in Columns (1)-(4) is the log change in the total sales of firm i in sector j in the period τ ($\Delta size_{ij,\tau}$) or the change in the number of products of firm i in sector j in period τ ($\Delta scope_{ij,\tau}$). The change in the dependent variables are calculated as in Davis and Haltiwanger (1992), i.e. $2(y_t - y_{t-1})/(y_t + y_{t-1})$. The reported coefficient is the effect of changes in exposure to demand shocks by firm size using the Nielsen data. Column 1, 2, 5, and 6 use the China import penetration shock. Columns 3, 4, 7 and 8 use as demand shock the housing price shock from 2001-2006. All the specifications include sector effects.

D.3 Relationship between scalability, size and scope

Table D17: Cross-Sectional Relationship: Scalability, Scope and Size (Sectors)

	(1)	(2)	(3)	(4)	(5)	(6)
size	0.214***		0.307***		0.352***	
	(0.021)		(0.037)		(0.014)	
scope		0.204***		0.281***		0.350***
		(0.022)		(0.031)		(0.011)
Observations	143,140	143,140	71,354	71,351	364,592	364,577
R-squared	0.368	0.376	0.446	0.465	0.297	0.331
Firm	Y	Y	Y	Y	Y	Y
Period-Sector	Y	Y	Y	Y	Y	Y
Sample	Food	Food	Non-Food	Non-Food	Module	Module

Note: The table shows the results estimating equation 25 in the Nielsen data. The dependent variable is (the log of) scalability from (25) and the independent variables are (the logs of) size (revenue) and scope. All the variables are standardize relative to the mean of the sector and time period (year). Scalability index is adjusted relative to the alternative (bootstrapped) index. Columns (1)-(2) include departments: dry grocery, frozen food, dairy, deli, packaged meat, and fresh produce . Columns (3)-(4) include departments health & beauty care, non-food grocery, and general merchandise. Column (5) defines a sector at the module-level.

D.4 Heterogeneous Response to Symmetric Shocks with Scalability Measures

Table D18: Response of Scalability to Shocks - Housing Price Shock

	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta g \times \frac{x}{y}$	-0.0305	0.0275	-0.0116	-0.0700	0.0003	-0.0176
	(0.143)	(0.040)	(0.039)	(0.148)	(0.041)	(0.040)
$\Delta g \times$ scope				0.1445	0.1915***	0.0245
				(0.176)	(0.055)	(0.027)
Observations	19,226	19,344	15,663	19,226	19,344	15,663
R-squared	0.078	0.035	0.234	0.085	0.051	0.236
Sector	Y	Y	Y	Y	Y	Y
Shock	Housing	Housing	Housing	Housing	Housing	Housing
Period	2006-2011	2006-2011	2006-2011	2006-2011	2006-2011	2006-2011

Note: The table shows the results of estimating equation 26. The dependent variable in Columns (1) and (4) is the change in the (log of) size (revenue) of firm i in sector j. The dependent variable in Columns (2) and (5) is the change in scope and in Columns (3) and (6) is the change in scalability of firm i in sector j. The independent variable is the demand shock interacted with the baseline level of scalability in 2006 or the same shock interacted with the levels of scope in 2006. In all specifications we use the housing shock from 2006-2011. All specifications include sector fixed effects.

Table D19: Response of Scalability to Shocks (Alternative Growth Rates)

	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta g \times \frac{x}{y}$	0.0094	0.0143***	0.0108***	0.0040	0.0074	0.0106***
	(0.007)	(0.005)	(0.002)	(0.007)	(0.005)	(0.003)
$\Delta g \times$ scope				0.0152**	0.0186***	0.0020
				(0.006)	(0.004)	(0.002)
Observations	14,186	14,186	13,488	14,186	14,186	13,488
R-squared	0.049	0.047	0.245	0.059	0.098	0.246
Sector	Y	Y	Y	Y	Y	Y
Shock	China	China	China	China	China	China
Period	2006-2015	2006-2015	2006-2015	2006-2015	2006-2015	2006-2015

Note: The table shows the results of estimating equation 26. The dependent variable in Columns (1) and (4) is the change in the (log of) size (revenue) of firm i in sector j. The dependent variable in Columns (2) and (5) is the change in scope and in Columns (3) and (6) is the change in scalability of firm i in sector j. The change in the dependent variables are calculated as in Davis and Haltiwanger (1992), i.e. $2(y_t - y_{t-1})/(y_t + y_{t-1})$. The independent variable is the demand shock interacted with the baseline level of scalability in year 2006 or the same shock interacted with the levels of scope in 2006. In all specifications we use the China import penetration shock from 2006-2015. All specifications include sector fixed effects.