Scalable Expertise: How Standardization Drives Scale and Scope*

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

Yale Northwestern Cornell NYU Stern

August 2025

Abstract

We develop and test a theory of firm size and scope centered on standardization: the extent to which a firm's products are similar. Heterogeneous firms choose the number, quality and standardization of their products. Standardizing makes expanding scope less costly, which then increases the return to standardization. Standardized firms have endogenously higher *marginal returns to scale*—their marginal costs rise less sharply with scale—implying greater responsiveness to demand shifts. We construct a novel measure of standardization using detailed product characteristics and show support for the theory's main predictions. We show that new product features from standardized firms diffuse more quickly.

Keywords: Standardization, returns to scale, firm size, multi-unit firm, concentration, diffusion
JEL: O3, D2, L25

*We thank Marios Angeletos, Lorenzo Caliendo, Ben Jones, Pete Klenow, Sam Kortum, Will Gamber, Andres Drenik, Giuseppe Moscarini, Esteban Rossi-Hansberg, Chad Syverson, as well as seminar and conference participants for useful comments and discussions. We would like to thank Lichen Zhang, Jiarui Cao and Yifan Shi for excellent research assistance. Researcher(s)' own analyses calculated (or derived) 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 those of the researcher(s) 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: david.argente@yale.edu. Address: Evans Hall, 165 Whitney Ave, New Haven, CT 06511, USA.

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

§Email: ezraoberfield@cornell.edu. Address: Uris Hall, Ithaca, NY 14853, USA

¶Email: vvenkate@stern.nyu.edu. Address: 44 W 4th Street, New York, NY 10012, USA

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 products or operate across a number of locations while others are more limited in scope. At the heart of these vast differences across firms are *returns to scale and scope*, which determine the costs of expanding operations. In his seminal work, Alfred Chandler (1990) emphasized that the hallmark of a modern economy is large enterprises that can take advantage of returns to scale and scope, and are large, in part, because they do so.

In this paper, we study, both theoretically and empirically, a fundamental source of returns to scale: standardization. We analyze the problem of a multi-unit firm in an otherwise standard model of heterogeneous, monopolistically competitive firms in the tradition of Hopenhayn (1992) and Melitz (2003). Firms choose scope (the number of units), standardization (the extent of similarity between these units) and quality (the value derived by the customer from each product). To make these concepts concrete, consider the product portfolio of a manufacturing firm. Each product is a bundle of features. Features can be *scalable* across products (firm-wide) or *local* (i.e. product-specific). A firm achieves greater standardization by investing more heavily in developing features shared by multiple products, i.e. in *scalable expertise*. For example, the furniture manufacturer IKEA designs its furniture so that many parts are identical across products, including interchangeable components like screws, connectors, and cam locks. These standardized elements simplify production and reduce costs as the same parts can be used in multiple product lines. These ideas are also relevant in the service sector, with units appropriately re-interpreted as branches or locations and features interpreted as processes. Consider Starbucks, which has developed streamlined procedures and standardized training programs for baristas that are used across all its locations.

Our analysis begins with a simple version of the model, in which firms differ only in an exogenous demand (or productivity) shifter and investment in expertise is limited by a firm-level capacity constraint. Later, we show how the main insights carry over to a more general setting with richer cost structures and heterogeneity.

Our first set of results shows that standardization and scope interact to amplify exogenous differences (or changes) in firm-level fundamentals. An increase in demand sets off a chain of mutually reinforcing actions. It directly leads to an expansion of scope, which raises the value of standardization. In turn, this reduces the need for product-specific investments, making adding products even more attractive.

Second, the theory predicts that the extent of amplification is heterogeneous. The effect of demand on firm size is increasing in the degree of standardization – more precisely, in the *scalable share* (the fraction of the firm's investment in expertise that is scalable). Before discussing the intuition, it is worth noting why the scalable share varies across firms. When scalable and local expertise are relatively substitutable in determining product quality – a natural assumption in our setting – incentives to invest in scalable expertise rise more than proportionately with increases in scope and so, the scalable share increases with demand.[1]

Since higher scalable shares imply greater responsiveness to demand, the positive correlation of size and scalable shares means (log) firm size is convex in (log) fundamentals (unlike the log-linear relationship in Melitz (2003)).[2] Moreover, a rise in industry demand leads to an increase in the relative size of large firms, or equivalently, in market concentration.

Third, this heterogeneity in standardization also has normative implications, in particular for the effects of taxes and subsidies. In our setting, markups distort production, making subsidies welfare-improving. Our theory implies that, at the *laissez-faire* equilibrium, the social return per dollar of subsidy is heterogeneous – specifically, it is increasing in scalability. That is, if resources available for subsidies are limited, rendering a full offset of markups infeasible, it is optimal to tilt subsidies towards scalable firms.

Underlying these results is how standardization mediates the relationship between a firm's marginal costs and its scale. Formally, *marginal returns to scale* (MRTS), defined as (minus) the elasticity of marginal costs with respect to output, is increasing in the scalable share. Note that, as pointed out by Menger (1954) and Shephard (1953), MRTS is distinct from *average returns to scale* (ARTS), which measures how a firm's *average* cost changes with output.[3] E.g., a production function with a fixed cost and a constant unit cost, as in, e.g., Dixit-Stiglitz, Krugman, or Romer has marginal cost that is invariant to size (MRTS = 0), but average cost declining with size. MRTS (along with the curvature of demand) is the key determinant of responses to change in demand.[4] When demand rises, a high MRTS firm—whose marginal costs rises more slowly with output—has a stronger incentive to increase its size. This intuition also explains why MRTS (and, therefore, standardization) is key to characterizing the marginal social return per dollar of subsidy.

Why does standardization affect MRTS? When products are relatively less standardized, expanding scope requires larger product-specific investments, steepening the firm's effective marginal cost curve. In contrast, scalable firms can increase the number of products they offer with a smaller marginal cost penalty. In this sense, these firms embody Chandler's insight: higher returns to scale (i.e. a flatter marginal cost curve) enables them to grow larger. For example, the standardized and interchangeable nature of its parts allows IKEA to offer a larger suite of products than a manufacturer whose products have less in common with each other.[5] Similarly, Starbucks can operate

[1] In the baseline model with only one source of heterogeneity, size and standardization are tightly linked. With additional sources of heterogeneity, size and scalable share need not be perfectly correlated.

[2] This mechanism can, under some conditions, become so powerful that size follows a power law distribution even with a bounded distribution of exogenous productivity/demand.

[3] ARTS is also equal to the ratio of marginal cost to average cost minus 1. When ARTS is higher (lower) than 0, the firm is said to exhibit decreasing (increasing) returns to scale.

[4] As Lashkari et al. (2024) show, ARTS along with the elasticity of demand—which governs the markup—determine the revenue-to-cost ratio of the firm.

[5] The connection between standardization and costs can take many forms. One incarnation, Eli Whitney's interchangeable parts or Henry Ford's assembly line standardize a process to make the production process more efficient,

more locations more efficiently thanks to its investments in standardization. Crucially, the scalable share captures the key forces driving these results: we show that it remains a sufficient statistic for MRTS and responsiveness to demand under fairly general conditions.

In Sections 3 and 4, we take our theory to the data. We propose a firm-level measure of standardization using detailed information on attributes of consumer goods products from the NielsenIQ dataset. Guided by the theory, our measure captures the extent to which the products in a firm's portfolio share common characteristics. We then test the model's predictions, about cross-sectional patterns as well as the more stringent ones about responsiveness to shocks.

First, we show that standardization is positively correlated with firm size and scope. Next, to test the model's prediction about responsiveness to shocks, we construct a demand shock based on the intensity of competition from Chinese imports, following Autor, Dorn and Hanson (2013). We find support for the central prediction of the theory: standardization is associated with a greater responsiveness (of size, scope, and standardization itself) to demand shocks. This finding can also be interpreted as validating our assumption that standardized and non-standardized characteristics are substitutes, in which case standardization is linked to higher marginal returns to scale. Third, we examine implications for market concentration. We show that, consistent with the theory, sectors where standardization covaries more strongly with firm size in the cross-section experience larger increases in concentration when demand rises. Finally, we return to the prediction that more standardized firms respond more to shocks using the National Establishment Time Series (NETS) data. While we do not have information to measure standardization directly for this set of firms, we find that firms with higher size (employment) or scope (number of establishments) respond more to changes in demand.[6]

While our baseline model is static, Section 5 presents empirical patterns relating standardization to firm and industry dynamics. First, we study standardization within innovation bursts. Berlingieri, De Ridder, Lashkari and Rigo (2024) show that such episodes play a central role in driving variation in firm-level growth, industry concentration and growth. We show that when a firm introduces a set of new products, if the new products are more standardized then they tend to comprise a larger innovation burst. Second, we demonstrate a connection between within-firm standardization and knowledge diffusion. New characteristics introduced by firms with a history of standardization are more likely to be adopted by other firms. This suggests that changes in incentives to standardize lead to ideas that can spill over to the whole industry. It also suggests underinvestment in standardization.

i.e. lower the unit cost. These examples are similar in spirit to the ones we pursue. In their simplest interpretation, they reflect purely returns to scale: they lower unit costs. An alternative version, closer to our model, is that interchangeability of parts lowers the cost of introducing new products.

[6] This is consistent with Chan, Hong, Hubmer, Ozkan and Salgado (2024), who find that returns to scale are systematically higher for larger firms.

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.[7] 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 – and endogenize – the degree of standardization across units and show how it shapes returns to scale and responses to fundamentals. Our interest in returns to scale is shared by a few recent papers. Engbom, Malmberg, Porzio, Rossi and Schoellman (2024) explore the role of different types of labor in generating heterogeneous returns to scale. Chan, Hong, Hubmer, Ozkan and Salgado (2024) estimate production functions with heterogeneous returns to scale and find that RTS is systematically higher for larger firms.

A growing literature examines how intangible capital of different types shapes firm-level outcomes. These include R&D (Hoberg and Phillips, 2025), knowledge inputs (Ding, 2023), software and ICT (Rubinton, 2020; Jiang and Rubinton, 2024; Lashkari, Bauer and Boussard, 2024), managerial inputs (Lucas, 1978; Akcigit, Alp and Peters, 2021; Chen, Habib and Zhu, 2023; Grobovšek, 2020) and headquarters services (Kleinman, 2022). The non-rival nature of these inputs leads to increasing returns to scale by enabling firms to spread fixed costs or knowledge across multiple units or products (Bilir and Morales, 2020; Crouzet et al., 2024). Relative to this literature, our paper makes several contributions. First, we highlight standardization as a distinct mechanism that also facilitates scalability. Second, we develop a measure of standardization in a multi-product setting. Third, we develop several subtle predictions of the theory and validate them empirically. Finally, 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 (2024). We highlight a different mechanism—investments in standardization lead to heterogeneous marginal returns to scale, which implies that a common increase in demand can, on its own, increase concentration.[8]

More broadly, our analysis pinpoints the importance of the elasticity of substitution between scalable and non-scalable investments. By mediating how scalable investments shape a firm's marginal returns to scale, it governs firms' responsiveness to shocks, and hence how common shocks affect concentration and which firms should be targeted with optimal industrial policy. This elasticity has received little attention, theoretical or empirical, in much of the existing work on intangible

[7]Scope is 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), Cao, Hyatt, Mukoyama and Sager (2020), Hsieh and Rossi-Hansberg (2020), and Rossi-Hansberg et al. (2021).

[8]A rise in concentration has been emphasized by, 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.

inputs.[9]

Our main empirical exercise, which relates standardization to the effects of the widely-studied "China shock," speaks to the literature following Autor, Dorn and Hanson (2013). Our finding that scalable firms were the most responsive echoes that of Holmes and Stevens (2014), who find that import competition in the furniture industry had a disproportionate impact on larger plants relative to smaller, more 'specialty' plants.[10]

2 Model

This section presents a theory in which heterogeneous firms choose the scope and size of their operations. The model makes a number of predictions about the cross-sectional distribution of firm size, responses to shocks, and concentration. In the subsequent 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):

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

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 implies the following demand function for product u of firm i:

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

where Q_j denotes sector-wide output while P_j and P_{ij} are the ideal price indices for the sector and

[9]Two exceptions are Koh and Raval (2025) and Jiang and Rubinton (2024). Koh and Raval (2025) use our framework to study shared and non-shared inputs among large firms. In the 1970s, the FTC surveyed roughly 500 large firms about their use of inputs across different production lines. The firms were asked to trace the portion of each input that was used by each line as well as the portion that was shared across lines (the two inputs with the largest portion shared across lines were management and capital). In line with our theory, they find that larger firms, as measured by size or scope, had a higher fraction of expenditures that were shared across lines. Further, using a production function estimation approach, they find that local and scalable inputs are indeed substitutes. In their study of software, Jiang and Rubinton (2024) estimate that, once a fixed cost has been paid, software is a complement of other inputs.

[10]It is also related to Aghion, Bloom, Lucking, Sadun and Van Reenen (2021), who find that firms with more centralized management – an example of more scalable (or firm-wide) investments – were more sensitive to the turbulence of the Great Recession, even controlling for size. Other papers studying the effects of demand shocks on firms include Mayer, Melitz and Ottaviano (2020), Hyun and Kim (2019), Park (2020), Lileeva and Trefler (2010), Bustos (2011), and Baldwin and Gu (2009).

the firm respectively:

$$Q_j = \left[\int (Q_{ij})^{1-\frac{1}{\theta}} \, di\right]^{\frac{\theta}{\theta-1}}, \qquad P_j = \left[\int P_{ij}^{1-\theta} di\right]^{\frac{1}{1-\theta}}, \qquad P_{ij}^{1-\epsilon} = \left[\int_0^{N_{ij}} (P_{uij})^{1-\epsilon} \, du\right]^{\frac{1}{1-\epsilon}}.$$

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 term expertise and describe in detail later. The firm is subject to a fixed operating cost 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:

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

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 increased profitability. Importantly, the expertise relevant to a particular product u is a combination of two types of knowledge – *scalable* (or firm-wide) or *local* (product-specific). The former, denoted x_i, is the same for all products of the firm, while the latter, y_{ui}, reflects knowledge unique to a product.

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

Investment in expertise is subject to costs $\mathcal{C}(x_i, \{y_{ui}\})$. In Section 2.2, we consider a simple specification – a capacity constraint – before turning to a more general formulation in Section 2.3.

2.1 Product Characteristics: A Micro-foundation For Expertise

In this subsection, we characterize the expertise function $Z(x, y)$ in a multi-product setting where expertise is embodied in product characteristics. While this micro-foundation is not crucial for our theoretical results, it will guide our empirical strategy for measuring standardization.

A firm has a portfolio of products, $u = 1, ..., N$. Each product has a (finite) set of attributes, indexed by $a = 1, ..., \mathcal{A}$. As an example, consider oral care products offered by Colgate-Palmolive. Attributes include "flavor", "size", "color", "packaging", or "ingredients". For every product-attribute pair (e.g. "flavor" of a particular toothpaste), the firm chooses a characteristic from an exogenously given set. The set contains 2 elements – one that is present in choice sets for all products (referred to as the "scalable" characteristic) and one specific to that product (or "local" characteristic).[11] For "flavor", a scalable characteristic could be "mint", used across a range of products, while a local characteristic is a more unique flavor used only for specific products.[12]

Each product has an effective quality (or appeal) that is a composite of the quality of each of its attributes. Formally, let b_{ua} denote the quality of attribute a for product n. Then, the effective quality of the product Z_u is given by

$$Z_u = \left(\sum_{a=1}^{A} b_{ua}^{\eta-1} \right)^{\frac{1}{\eta-1}} . \tag{2}$$

For attribute a of product u, a firm can choose the scalable characteristic, which has quality b_{ua}^x, or the local characteristic, which has quality b_{ua}^y, so that $b_{ua} = \max\{b_{ua}^y, b_{ua}^x\}$.

The characteristic-level qualities are random, but depend on the firm's investments in expertise. Given an investment I, the quality of the characteristic is drawn from a distribution with cumulative distribution function $e^{-(b/I)^{-\nu}}$, with $\nu > \max(0, \eta - 1)$. If x_a is the investment in the scalable characteristic for attribute a, then $\Pr(b_{ua}^x \leq b) = e^{-(b/x_a)^{-\nu}}$. Similarly, if y_{ua} is the investment in the product-specific characteristic for attribute a for product u, then $\Pr(b_{ua}^y \leq b) = e^{-(b/y_{na})^{-\nu}}$. The distribution of the product-attribute quality $b_{ua} = \max\{b_{ua}^x, b_{ua}^y\}$ is then given by $\Pr(b_{ua} \leq b) = e^{-(x_a^\nu + y_{ua}^\nu) b^{-\nu}}$.

Total investment across characteristics is $\sum_a (x_a + \sum_u y_{ua})$. Given the symmetry across attributes and products, the firm will choose the same investment in scalable characteristics across attributes (i.e. $x_a = x$) and the same investment in local characteristics across all products and attributes (i.e. $y_{ua} = y$). Thus, total investment is $\mathcal{A}(x + Ny)$. As the number of attributes $\mathcal{A}$ grows large, the law of large numbers implies that product quality converges to

$$\lim_{\mathcal{A} \to \infty} Z_u = \Gamma \left(1 - \frac{\eta - 1}{\nu} \right)^{\frac{1}{\eta-1}} (x^\nu + y^\nu)^{\frac{1}{\nu}} \quad \equiv \quad Z(x, y).$$

Under these conditions, a firm's investment in scalable expertise can be measured using information on product-level attributes. Define a firm's *scalability index*, $\mathcal{SI}$ as one minus the share of unique characteristics in a firm's product portfolio. The share of unique characteristics is the number of unique characteristics across all the products of the firm by the total number of possible

[11] The assumption that there is only one characteristic of each type in the set is only for simplicity in exposition and is not crucial; it is straightforward to allow for multiple scalable and product-specific characteristics.

[12] Other examples of scalable characteristics in Colgate-Palmolive's portfolio include antibacterial agents and fluoride while local characteristics include additives for whitening and formulation options (such as alcohol-free mouthwash). To take another example, in Sephora's private-label skincare lines, scalable characteristics include standardized packaging, such as airless pump bottles and glass jars, while local characteristics include specialized ingredients like hyaluronic acid and vitamin C.

characteristics that could appear. A high $\mathcal{SI}$ indicates that the firm's products share many characteristics; i.e., the number of unique characteristics is relatively small. As N becomes large, $\frac{\mathcal{SI}}{1-\mathcal{SI}}$ converges to $\left(\frac{x}{y}\right)^{\nu}$.[13] We leverage this insight in Section 4.1 to construct a measure of standardization at the firm level.

2.2 A Tractable 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 transparent fashion. These assumptions will be relaxed in the next subsection, which will show that the key results obtain in a more general setting as well.

First, we set $\theta = \epsilon = 2$, i.e. the within- and across- firm substitution elasticities are assumed to be equal to 2. Then, $\mathcal{F}(N_i) = FN_i$, i.e. the fixed operating cost is linear in the number of products. Then, the expression for profits in (1) simplifies to

$$\Pi_i = G \int_0^{N_i} A_i Z_{ui} du - FN_i. \tag{3}$$

Next, the cost of expertise is modeled as a capacity constraint:

$$x_i + \int_0^{N_i} y_{ui}\, du \leq 1. \tag{4}$$

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

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

$$Z_{ui} = Z(x_i, y_{ui}) = \left(x_i^{\frac{\sigma-1}{\sigma}} + y_{ui}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \qquad 0 < \sigma < 2\,.$$

The upper bound on σ ensures an interior solution to the firm's problem.

Note that the product characteristics formulation in Section 2.1 satisfies this assumption with $\sigma = \frac{1}{1-\nu}$. Since $\nu > 0$, this micro-foundation of expertise implies $\sigma > 1$, i.e. scalable and local expertise are gross substitutes. Accordingly, we view $\sigma > 1$ as the natural case, but, for completeness,

[13] For each attribute, the number of unique characteristics is the sum of the number of unique local characteristics and the number of unique scalable characteristics. The probability that a local characteristic is used is $\frac{y^{\nu}}{x^{\nu}+y^{\nu}}$. Since this likelihood is independent across products, as the number of products grows large, the proportion of unique local attributes converges to its expected value, $\frac{y^{\nu}}{x^{\nu}+y^{\nu}}$. For each attribute, the number of unique scalable characteristics is bounded between 0 and 1, so the proportion of product-attributes populated by scalable characteristics is bounded above by $1/N$, which converges to zero as N grows large. As a result,

$$\mathcal{SI} = 1 - \frac{\#\text{ unique char.}}{N \times \mathcal{A}} = 1 - \frac{\#\text{ unique local char.}}{N \times \mathcal{A}} - \frac{\#\text{ unique scalable char.}}{N \times \mathcal{A}} \to 1 - \frac{y^{\nu}}{x^{\nu}+y^{\nu}} - 0 = \frac{x^{\nu}}{x^{\nu}+y^{\nu}}$$

.

we present all results for general σ.

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

$$\max_{N_i, x_i, y_{ui}} \quad G \int_0^{N_i} A_i Z(x_i, y_{ui}) - FN_i \qquad \text{s.t.} \qquad x_i + \int_0^{N_i} y_{ui}\, du \leq 1. \tag{5}$$

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

$$\max_{N_i, x_i, y_i} \quad G \;\; A_i N_i Z(x_i, y_i) - FN_i \qquad \text{s.t.} \qquad x_i + N_i y_i \leq 1. \tag{6}$$

Since the capacity constraint will bind at the optimum, it can be rearranged as $y_i = \frac{1}{x_i/y_i + N_i}$. We can use this to express expertise in terms of scope and the scalability ratio, $k_i \equiv x_i/y_i$: $Z(x_i, y_i) = y_i z(k_i) = \frac{z(k_i)}{k_i + N_i}$. The problem thus reduces to a choice over scope and the scalability ratio:

$$\max_{N_i, k_i} \quad G \; A_i N_i \frac{z(k_i)}{k_i + N_i} - FN_i. \tag{7}$$

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

$$k_i: \qquad \frac{k_i z'(k_i)}{z(k_i)} = \frac{k_i}{k_i + N_i} \qquad \equiv \qquad S_i \tag{8}$$

$$N_i: \qquad GA_i \frac{z(k_i)}{k_i + N_i} - F = GA_i \frac{N_i}{k_i + N_i} \frac{z(k_i)}{k_i + N_i}\;, \tag{9}$$

Equation (8) equates the cost and benefits of making expertise more scalable (for a given scope N_i). The right hand side is the share of capacity devoted to scalable expertise, $\frac{k_i}{k_i+N_i} = \frac{x_i}{x_i+N_i y_i}$. We term this the *scalable share* and denote it by 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. This implies that when x and y are gross substitutes, the scalable share S_i and the scalability ratio k_i are positively related.

Equation (8) can also be rearranged as a relationship between the scalability ratio and scope:

$$\frac{Z_y}{Z_x} = \frac{z(k_i) - k_i z'(k_i)}{z'(k_i)} = N_i\;. \tag{10}$$

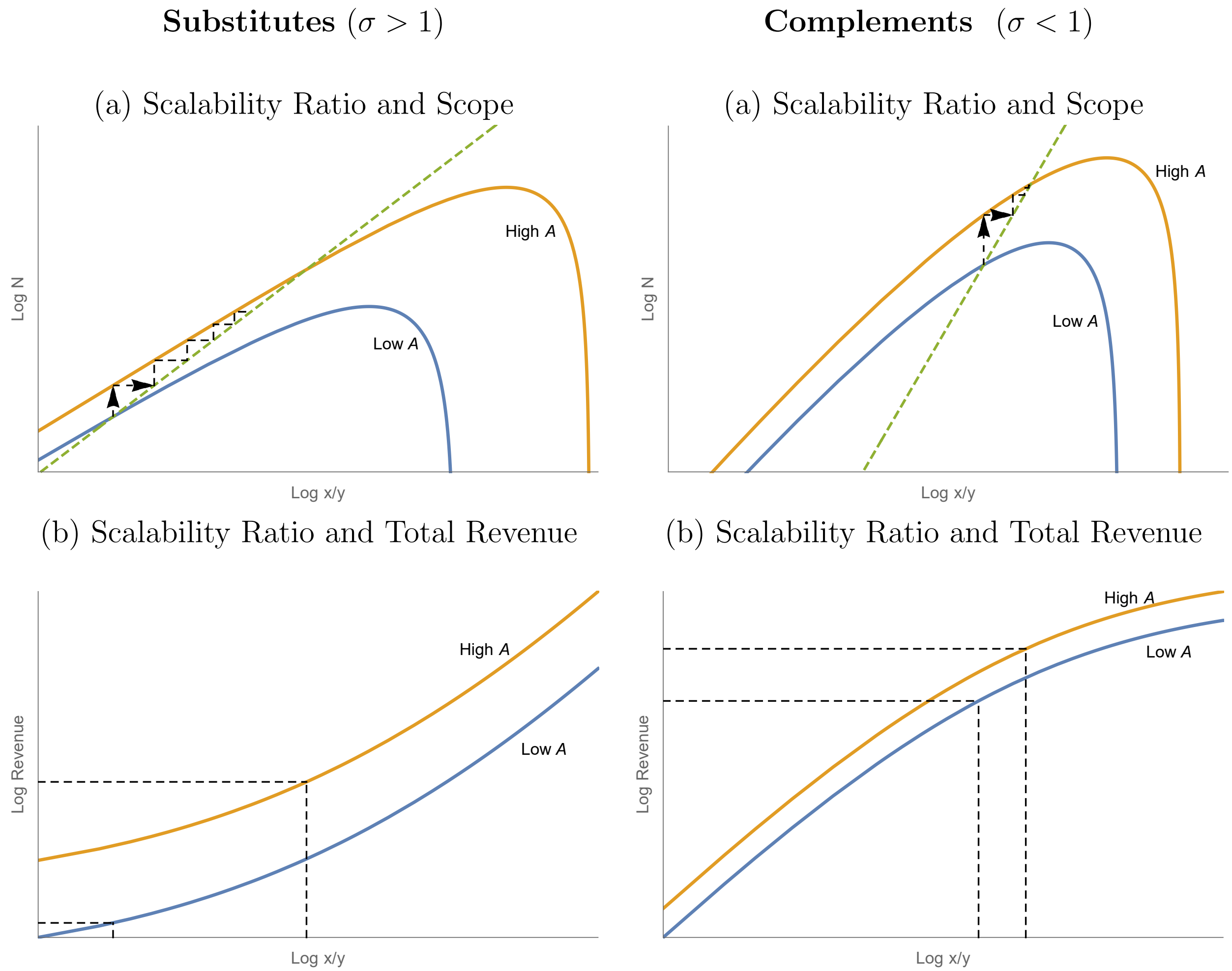
Thus, at the optimum, the firm equates the marginal benefit of increasing scalable expertise, NZ_x, to that of the non-scalable component $\frac{NZ_y}{N} = Z_y$. The marginal rate of substitution $\frac{Z_y}{Z_x}$ is increasing in the scalability ratio k_i, so this equation describes a positive relationship between scope and scalability ratio.

Let $\mathbb{K}(N)$ denote the optimal k_i as a function of scope. Note that $\frac{d\ln\mathbb{K}}{d\ln N} = \frac{d\ln k}{d\ln\frac{Z_y}{Z_x}} = \sigma$. When $\sigma > 1$, x_i/y_i rises more than one-for-one with N_i. Since $\frac{S_i}{1-S_i} = \frac{x_i}{N_i y_i}$, this implies that the scalable share rises as well. In other words, when $\sigma > 1$, the scalable share and the scalability ratio are positively linked. This insight will prove useful in interpreting our empirical results.

Equation (9) determines the optimal scope for a given k_i. Here, the fixed capacity leads to a tradeoff: the extra variety increases revenue but raises the firm's fixed cost and requires expertise. This equation can be rearranged to yield an expression for optimal scope as a function of k_i and A_i:

$$\mathbb{N}(k_i; A_i) = k_i\left[\left(\frac{GA_i}{F}\frac{z(k_i)}{k_i}\right)^{1/2} - 1\right]. \tag{11}$$

Figure 1: Scalability and Scope



Note: The dashed line in panel (a) plots (10), while the solid lines plot (11) for two values of A_i. The solid lines in (b) plot revenue $\mathbb{R}(k, \mathbb{N}(k; A_i); A_i)$, as a function of the scalability ratio for two 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 $\frac{GA_i}{F} = 20$.

The left panel of Figure 1 shows equations (10) and (11) on a log-log plot. The solid line displays (10), while the dashed ones show (11) for two levels of A_i.[14] The panel also depicts how the optimal choice of scalability and scope amplify exogenous differences in productivity (or equivalently, demand). Higher A_i shifts the dashed line upward (note that the solid line, equation (10), is independent of A_i). The arrows show successive rounds of adjustment. Higher productivity induces a larger scope, holding k_i fixed (the first vertical arrow). This in turn leads to an increase in k_i (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, when scalable expertise is very substitutable with the local kind, a high productivity firm's incentives to accumulate scalable expertise are very strong, leading to a corner solution. We assume that parameters are such that this case is precluded.

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

$$\frac{d\ln N}{d\ln A} = \frac{\partial \ln \mathbb{N}}{\partial \ln A} + \frac{\partial \ln \mathbb{N}}{\partial \ln k}\frac{d\ln \mathbb{K}}{d\ln N}\frac{d\ln N}{d\ln A} = \frac{\frac{\partial \ln \mathbb{N}}{\partial \ln A}}{1 - \frac{\partial \ln \mathbb{N}}{\partial \ln k}\frac{d\ln \mathbb{K}}{d\ln N}}. \tag{12}$$

The denominator in (12) captures the amplification: an increase in A_i sets off the sequence of mutually reinforcing changes we saw in the top panels of the figure. We can use (11) to derive[15]

$$\frac{\partial \ln \mathbb{N}}{\partial \ln k} = \frac{1}{2} \tag{13}$$

$$\frac{\partial \ln \mathbb{N}}{\partial \ln A} = \frac{1}{2}\frac{1}{1-S_i}. \tag{14}$$

Thus, firms who use a relatively larger share of capacity for scalable expertise (high S_i) adjust their scope by more in response to changes in A_i (holding k_i constant). Intuitively, increasing scope involves not only an additional fixed cost F but also investment in product-specific expertise. When such expertise is relatively small (S_i is high), increasing scope exerts less pressure on the capacity constraint, i.e. induces a smaller reduction in expertise. In this sense, scalable 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.

[14] Equation (11) can be hump-shaped, but the intersection will always be in the upward sloping part.

[15] To derive (13), re-arrange (9) 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 (8) yields

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

Finally, using $\frac{d\ln \mathbb{K}}{d\ln N} = \sigma$, equations (12)-(14) can be combined to yield:

$$\frac{d\ln N}{d\ln A} = \frac{\frac{\partial \ln \mathbb{N}}{\partial \ln A}}{1-\frac{\sigma}{2}} = \frac{1}{2-\sigma}\left(\frac{1}{1-S_i}\right) \;>\; 0\,. \tag{15}$$

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 for our applications, the elasticity of scope is increasing in k_i.

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

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

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

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

The first term is the direct effect, while the other two capture indirect effects (from induced changes in scope and scalability). The second term is 0 by an envelope argument (since k is chosen to maximize $\frac{z(k)}{k+N}$). 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 interestingly, the cost in expertise from increased scope is smaller when the share of scalable expertise is high, leading to a larger response of revenue:

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

Along with (15), this implies scalable firms experience a larger increase in size for a given change in productivity: [16]

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

We summarize these results in the following three propositions:

Proposition 1 *Firms with higher productivity have higher total revenue, scope, scalability ratio and (when $\sigma > 1$) scalable shares.*

The second proposition states a key result: the scalable share is a sufficient statistic for the firm's

[16]The impact of productivity on revenue *per unit* is more subtle. The direct effect is still equal to $\frac{\partial \ln \mathbb{R}}{\partial \ln A} = 1$, but the indirect effect stemming from the change in scope is negative. In this simple case, the strength of the indirect effect (and therefore, the sign of the total effect) depends on the degree of substitutability between scalable and local expertise. If $\sigma < 1$, the direct effect dominates so revenue per unit rises with productivity. The opposite happens if $\sigma > 1$. In the more general model below in which capacity can be increased at a cost, however, revenue per unit may rise even if $\sigma > 1$.

responsiveness to changes in demand or productivity. Note that the responsiveness to changes in demand (e.g. due to a shift in the equilibrium shifter G) and productivity (A_i) are identical, as these enter the firm's profit function in the same way.

Proposition 2 *Firms with a higher scalable share S_i exhibit higher elasticities of scalability ratio, scope, and size to common demand shocks G or idiosyncratic productivity A_i.*

The next proposition expresses the result in Proposition 2 in terms of the scalability ratio $\frac{x_i}{y_i}$. It forms the basis of the tests presented in Section 4, where we construct a measure of the scalability ratio at the firm-level and proxies for sectoral demand shocks. Recall that the scalable share and the scalability ratio are positively related when $\sigma > 1$. This in turn implies that firms with higher scalability ratios will be more responsive to changes in demand.

Proposition 3 *When $\sigma > 1$, firms with higher scalability ratios exhibit higher elasticities of scalability ratios, scope, and size to common demand shocks G or idiosyncratic productivity A_i.*

The Elasticity of Substitution and Responsiveness to Shocks

In this subsection, we provide a graphical illustration of how the elasticity of substitution σ determines on the responses to a common demand shock.

Figure 2 shows the response of key variables to a change in G. The three left panels show the case where $\sigma > 1$, the relevant one in our multi-product context, while the right panels show the same variables when $\sigma < 1$. In each panel, two sets of lines are depicted – they show the effect of a shift from G (red lines) to G' (black lines) for two levels of productivity, $\underline{A}$ and $\overline{A}$ where $\underline{A} < \overline{A}$.

The top panels show the effect 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 k, i.e. equation (8). The dashed lines show the combinations of k and S consistent with the optimal choice of scope,equation (9). The optimal choice of the firm is the intersection of the dashed and solid lines. These can respectively be rearranged as:

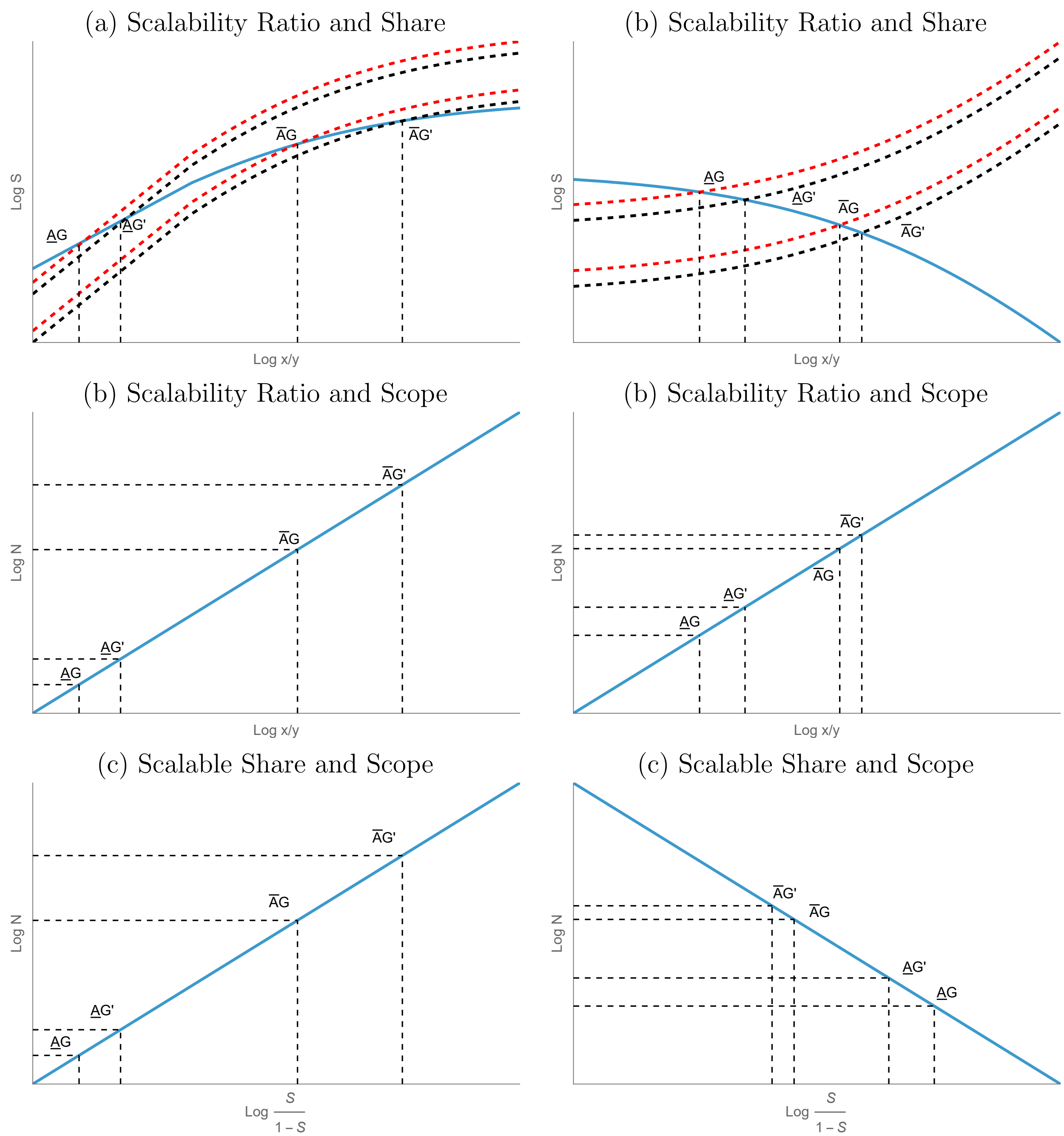
$$\ln S_i = \ln \frac{k_i z'(k_i)}{z(k_i)} \tag{16}$$

$$\ln S_i = -\frac{1}{2} \ln \left(\frac{z(k_i)}{k_i} A_i \right) - \frac{1}{2} \ln \left(\frac{G}{F} \right) . \tag{17}$$

The right hand side of (16) 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 substitutable (complementary), i.e. when σ is larger (smaller) than 1. The right hand side of (17) is increasing in k_i for all σ, so the dashed lines are increasing in both panels. Their curvature does depend on σ: they are concave (convex) when σ is larger (smaller) than 1.

Figure 2: Response to a Shock

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



Note: The dashed lines in (a) plot (17), while the solid line depicts (16). Panel (b) shows (10) in logs, while the relationship in (c) is derived from (8). In the left panels (Substitutes), we use $\sigma = 1.5$ and $\underline{A}/F = 0.2$ and $\bar{A}/F = 0.6$, with $G = 1$ and $G' = 1.3$. For the right panels (Complements), the corresponding values are $\sigma = 0.67$, $\underline{A}/F = 10$, $\bar{A}/F = 30$, and again $G = 1$ and $G' = 1.3$.

What happens when a common demand shock pushes G up to G'? Equation (16) remains the same. Since G enters (17) log-linearly, increases in $\ln G$ induce parallel, downward shifts in the dashed lines. Intuitively, higher demand implies a higher optimal scope and so, a lower scalable share for all k. In both the left and right panels, the higher scope increases the incentive to accumulate scalable expertise, so the new optimum features a higher scalability ratio. 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, so the scalability ratio rises more than proportionately with scope, resulting in a higher scalable share.

How does a firm's productivity A affect the response to a change in G? The answer depends on whether σ is larger or smaller than 1, which determines the relationship between productivity and the scalable share, S. If scalable and local expertise are substitutes (left panel), high productivity firms have higher scalable shares and so respond more. The opposite happens if they are complements (right panel). The middle and bottom panels show that this pattern extends to the scope: high productivity exhibit larger (smaller) changes in scope when the two forms of expertise are substitutes (complements).

Marginal Returns to Scale

The intuition for these results stems from the effect of scalability on a firm's marginal cost curve. Define *marginal returns to scale* ($MRTS$) 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. Note that $MRTS$ is distinct from *average returns to scale* ($ARTS$), which measures how a firm's *average* cost changes with output.[17] To see the distinction, consider, e.g. a production function with a fixed cost and a constant unit cost, as in, e.g., Dixit-Stiglitz, Krugman, or Romer. Marginal cost is invariant to size (so $MRTS = 0$), but average cost declines with size.

$MRTS$ (along with the curvature of demand) is the key determinant of responses to change in demand. To see this, consider a firm that faces iso-elastic demand, $Q_i = D_i P_i^{-\theta}$, 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 thus

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

[17] $ARTS$ is also equal to one minues the ratio of marginal cost to average cost. When $ARTS$ is higher (lower) than 0, the firm is said to exhibit increasing (decreasing) returns to scale.

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

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

Thus, the curvature of demand and $MRTS$ are sufficient to determine responses of output and revenue to a change in demand.[18]

In our multi-product setting, Q_i can be mapped to the firm-level composite good. Using the symmetry across products, the cost function of the firm 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 $Q_i \geq N_i^2 Q_{iu}$. Then, $MRTS_i$ is given by (see Appendix A.3 for details):

$$MRTS_i = -\frac{Q_i c_i''(Q_i)}{c_i'(Q_i)} = \frac{S_i}{2-(1-S_i)\sigma} = \frac{1}{\frac{2-\sigma}{S_i}+\sigma} \ , \tag{18}$$

which rises with S_i. Intuitively, a higher S_i implies that the firm can increase scope with a smaller 'cost' in terms of foregone expertise, which moderates the rise in marginal costs. This is stated formally in the following result:

Proposition 4 *Firms with higher scalable share have higher marginal returns to scale.*

MRTS and Policy

Consider an economy with a single sector and a representative household that supplies labor elastically and enjoys utility from the composite $Q = \left(\int_i Q_i^{\frac{\theta-1}{\theta}} di\right)^{\frac{\theta}{\theta-1}}$ according to $u(Q, L)$, where u is increasing in Q and decreasing in L. Apart from labor income, the household also owns all firms and receives their profits as dividends.

In the *laissez faire* equilibrium, each firm charges a markup of $\frac{\theta}{\theta-1}$ and is therefore too small, relative to the welfare-maximizing allocation. We define the marginal social return to subsidizing firm i to be the increase in social welfare from a small subsidy to firm i's revenue relative to the fiscal cost of the subsidy (funded with a lump sum tax). The following result links the marginal social return of subsidizing a firm to its MRTS.

Proposition 5 *At the laissez-faire equilibrium, the marginal social return to subsidizing firm i is proportional to $\frac{1}{\frac{1}{\theta}-MRTS_i}$, with a constant of proportionality that is common across firms.*

Together with (18), Proposition 5 shows that, with homogeneous demand elasticities, scalability

[18] $ARTS$, along with the markup, determines the revenue-to-cost ratio for the firm. See Lashkari et al. (2024).

is a sufficient statistic for the welfare impact of small subsidies.[19] Therefore, scalable firms offer the largest bang for the buck at the *laissez-faire* equilibrium. The intuition is similar to what we saw earlier: scalability increases the sensitivity of production to subsidy. This implies that, when resources/fiscal capacity is limited (rendering the full markup-correcting subsidy infeasible), it is optimal to direct/tilt subsidies towards scalable firms.

The Size Distribution

In this subsection, we show how scalability can aid our understanding of the enormous variation in firm size, a classic question in economics and one that has attracted renewed interest recently. As noted earlier, the endogenous response of scope and scalability amplifies the effects of productivity variation on firm size, inducing a convex relationship between total revenue and productivity (both in logs). Without these forces, our iso-elastic environment reduces to the standard Hopenhayn-Melitz framework with a log-linear relationship between total revenue and productivity. The following result states this formally.

Proposition 6 *Suppose* $\sigma > 1$*. Then, firm revenue 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 example 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 is the beta distribution with bounds $\underline{A}$ and $\bar{A}$ and parameters η and κ.[20] The next result states the conditions under which we obtain a power law distribution in size.

Proposition 7 *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:*

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

[19]The constant of proportionality in Proposition 5 approaches zero as labor supply becomes less elastic. In the limit, with perfectly inelastic labor, the equilibrium is efficient.

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

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[21], but it serves as a stark illustration of the potential of scalability to amplify productivity variation 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 when the largest firms in an industry are also the most scalable and therefore, respond most to the higher demand.

Proposition 8 *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.*

This follows directly from Propositions 1 (larger firms have higher shares of scalable expertise) and 2 (firms with higher scalable shares see larger increases in revenue in response to common shock). This result 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 $\sigma > 1$. An increase in G raises concentration of sales among a fixed set of firms.*

2.3 General Model

In this subsection, we relax assumptions on curvature and heterogeneity made in the previous subsection. The firm's problem is now assumed to be given by:

$$\max_{x_i, y_i, N_i} \qquad G N_i^{\phi} \left(A_i Z(x_i, y_i)\right)^{\psi} - F N_i^{\omega} - H \cdot \left[x_i^{\mu} + N_i y_i^{\mu}\right]^{\gamma} \ . \tag{19}$$

There are several differences relative to the version analyzed in the previous subsection: (i) elasticities of substitution between among products in a firm's portfolio can be different than among firm-level composite, i.e. $\phi = \frac{\theta-1}{\varepsilon-1} \neq 1$ [22] (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

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

[22] 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 (19).

(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: since $Z(x, y)$ is CRS, the elasticity is a function of the scalability ratio x/y. Define the transformed CRS function $\tilde{Z}(\tilde{x}, \tilde{y}) \equiv Z\left(\tilde{x}^{1/\mu}, \tilde{y}^{1/\mu}\right)^{\mu}$. Let $\tilde{\sigma}(\tilde{x}/\tilde{y})$ be the elasticity of substitution associated with $\tilde{Z}$. Appendix A.7 shows that 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 satisfy the following conditions:*

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

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

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

These restrictions are sufficient to guarantee a unique, interior solution to the firm's problem (see Appendix A.7): *(i)* ensures that curvature of the cost of scope and expertise is larger than that of the corresponding benefit *(ii)* ensures that the scope-scalability feedback is not explosive *(iii)* ensures that the feedback between reducing scope and increasing local expertise is not explosive.

As in the previous subsection, firms with higher productivity raise size, scope and scalability. Whether the share of scalable expertise, $S_i \equiv \frac{x_i^{\mu}}{x_i^{\mu}+N_i y_i^{\mu}}$, rises with size depends on the substituability between scalable and local expertise.

Proposition 9 *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:* $\frac{d\ln R_i}{d\ln G} > 0$, $\frac{d\ln N_i}{d\ln G} > 0$, *and* $\frac{d\ln x/y}{d\ln G} > 0$. *The scalable share* S_i *also rises if* $\sigma > 1$.

Thus, as in the simple version, heterogeneity in the 'level' of productivity endogenously generates heterogeneity in effective curvature. Moreover, even with all these rich interactions, responsiveness to shocks is summarized by a single variable, the scalable share of expertise.[23]

Proposition 10 *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.*

The following assumption restricts attention to the more 'natural' region of the parameter space, where the two forms of expertise are gross substitutes.

[23] As we will show in Section 2.4, this is robust to adding more heterogeneity (again, in levels).

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 11 *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:*

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

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 11 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 a change in* G*, firms with higher size or higher scope have larger changes in size, scope, scalability ratio, and scalable share:*

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

2.4 Additional Heterogeneity

In this subsection, we add more dimensions of heterogeneity. 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 Section 2.2-2.3, now firms also vary in the operating cost, F_i and in the cost of expertise, both through the overall cost shifter H_i, as well as in type-specific parameters a_i^x and a_i^y.

This richer heterogeneity breaks the tight links between size, scope, and scalability in the cross-section. Nevertheless, the elasticities of size, scope and scalability with respect to demand are exactly the same as in Section 2.3. 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 10 and 11 continue to hold.

We next turn to the implications for concentration of rising industry demand. Recall that, with only 1 dimension of heterogeneity, a common rise in demand increased the market share of the largest firms (when $\sigma > 1$). Now, with additional dimensions of heterogeneity, scalability is not necessarily tightly linked to size, but is still the key determinant of responsiveness to demand. Therefore, the effect of rising industry demand on concentration depends on how scalability covaries with size: loosely speaking, the more positive this covariance, the larger is the rise in concentration to a positive demand shock.

Formally, we first define a partial ordering of the scalability-size covariance: scalability is said to covary 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 corresponding distribution 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. Our notion of a rise in concentration is the same as in Proposition 8: an increase in any metric 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.

Proposition 12 *An increase in G raises 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 now test the key predictions of the model using detailed product- and establishment-level data. Our primary application focuses on multi-product firms: we propose a firm-level measure of scalability, guided by Section 2.1. For the analysis on multi-establishment firms, we do not have an analogous measure of scalabaility, so we test the model's predictions leveraging the size-scalability correlation implied by the theory.

This section describes the data sources and the construction of demand shocks. In the next section, we describe the construction of the scalability measure and present the empirical validation of the model's predictions.

3.1 Multi-Product Firms

We use comprehensive data on firms and products in the consumer packaged goods (CPG) industry from 2006 to 2015, collected by the NielsenIQ Retail Scanner Data (RMS) and provided by the Kilts Data Center at the University of Chicago Booth School of Business. The dataset is based on point-of-sale records from grocery, drug, and general-merchandise stores. The CPG industry accounts for approximately 14% of total goods consumption in the U.S. The NielsenIQ RMS data covers about 40% of industry sales and includes nearly the full universe of firms and products in the sector.

The RMS dataset contains over one million distinct products identified by barcodes, enabling us to track sales over time at a highly granular level. Barcodes are organized hierarchically. Each barcode is assigned to one of 1,070 product modules (e.g., lamps, flashlights, Christmas lights, razor blades, shave cream, shaving accessories), which in turn belong to one of 104 broader product groups (e.g., light bulbs, shaving needs). Following Hottman, Redding and Weinstein (2016) and Argente, Lee and Moreira (2020), we define sectors as product groups, a more granular level than the 4-digit Standard Industrial Classification (SIC).

We link products to firms using information from GS1 US, the official source of barcode assignments.[24] For each firm-sector-year observation, we compute total sales (*size*) and the number of distinct products (*scope*). Because each barcode corresponds to a unique combination of product attributes, barcode counts provide a close mapping to scope in the model. Finally, the dataset also provides rich information on each firm's product portfolio, including detailed attributes (e.g., size, packaging, formula, flavor) and observable characteristics (e.g., variants such as blue or red), which we use to measure standardization.

3.2 Sectoral Demand Shocks

Our theory makes predictions about how firms respond to common shocks. To test these predictions, we use a widely used sectoral demand shock: Chinese import competition. Specifically, we use changes in the import penetration ratio to construct sector-level trade exposures, closely following Autor, Dorn and Hanson (2013) and Acemoglu, Autor, Dorn, Hanson and Price (2016). Data on trade between China and the U.S. is from UN Comtrade, NBER-CES, and UNIDO. Our baseline measure of trade exposure for sector j is defined as

$$\Delta IP_{j,06-15} = \frac{M_{j,15} - M_{j,06}}{Y_{j,06} + M_{j,06} - E_{j,06}} \times 100 \tag{20}$$

where $M_{j,t}$ denotes imports from China into the US in year t, $E_{j,t}$ denotes exports, and $Y_{j,t}$ denotes industry shipments. We use 2006 as the baseline year because it is the earliest period for which we have simultaneous trade, product, and establishment data. Intuitively, $\Delta IP_{j,06-15}$ captures sector-level increases in import competition from China. To address endogeneity concerns, we also use an instrumental variable based on Chinese imports in other high-income countries. The instrument exploits the fact that high-income economies are similarly exposed to Chinese supply shocks but are unaffected by U.S.-specific demand shocks.[25]

[24] Appendix B provides detailed information on the NielsenIQ RMS and GS1 US datasets.

[25] See Appendix B.3 for more details on the construction of the instrument.

We interpret rising import penetration as a decline in the residual demand faced by U.S. producers. Accordingly, we define a sector's demand shock as the additive inverse of import penetration: $\Delta G_j = -\Delta IP_{j,06-15}$. For the product-level data, we use the concordance developed by Bai and Stumpner (2019) to map NielsenIQ product groups to trade data and construct shocks at the product group level. For the establishment data, we use a standard mapping from SIC codes to trade sectors to assign shocks at the 4-digit SIC level.[26]

4 Mapping Theory to Data

In this section, we describe our measure of scalability and use it to test the following key predictions of the theory: (i) in the cross-section, firms with greater size or scope exhibit higher scalability (Proposition 1). (ii) in response to increase in sectoral demand, firms with higher scalability expand size, scope, and scalability by more (Proposition 11); (iii) market concentration should rise more in sectors where scalability covaries more strongly with firm size (Proposition 12); and (iv) firms with greater size or scope should exhibit larger adjustments along both the size and scope margins (Corollary 2). We group these predictions into two sets. The first three leverage our scalability measure and are tested using multi-product firm data. The fourth prediction relies only on firm size and scope and is tested using both multi-product and multi-establishment data.

4.1 Measuring Scalability

Our measure of scalability captures the extent to which characteristics are shared across products. Specifically, we construct the scalability index, $\mathcal{SI}$, as introduced in Section 2.1, using detailed information on product attributes. In the NielsenIQ data, each product is associated with a number of attributes, such as package, size, flavor, formula.[27] These attributes can take different values, which we term "characteristics". For example, the product module razor blades has 5 attributes: form, consumer type, scent, skin condition, and generic. The attribute "form" can take the following characteristics: "adjustable", "assorted", "injector", "moving", "pivoting", etc.

The scalability index for firm i in module m in sector j at time t is defined as the fraction of common attributes across its product portfolio:

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

where Unique_{imjt} denotes the number of distinct characteristics in the product portfolio of firm i

[26] Figure B.3 shows the variation in Chinese import penetration across sectors. There is substantial heterogeneity both within the consumer product industry and across manufacturing more broadly. As expected, sectors producing semi-durable and durable goods were more affected by Chinese import competition and therefore experienced more negative demand shocks than sectors focused on food products.

[27] This information is available for approximately 61% of all barcodes in the data. We use a total of 20 distinct attributes, with each product module containing between 4 and 8 active attributes.

in module m, sector j at time t. This is divided by the total number of attribute cells to be filled: module-level scope Scope_{imjt} (the number of products) times the number of attributes per product in that module, $\text{NumAttributes}_{mjt}$. Thus, $\mathcal{SI}_{imjt}$ captures the share of common characteristics across a firm's product portfolio. If no characteristic is shared across products, the scalability index equals zero. For example, a single-product firm will have as many characteristics as attributes, resulting in a scalability index of zero. In contrast, when characteristics are almost always shared, the index approaches one.

We aggregate this object (using revenue weights) to obtain the scalability index at the firm-sector-year level, $\mathcal{SI}_{ijt}$. The following transformation maps $\mathcal{SI}$ to the scalability ratio:

$$\ln\left(\frac{x_{ijt}}{y_{ijt}}\right) \quad \propto \quad \ln\left(\frac{\mathcal{SI}_{ijt}}{1-\mathcal{SI}_{ijt}}\right) . \tag{21}$$

One potential concern is that coarseness in NielsenIQ's tabulation of product characteristics could lead to a mechanical bias in our measurement of scalability ratios.[28] To partly address this, we adjust the observed index with a bootstrap procedure: for each firm-module-sector-year, we compute a counterfactual scalability index by assigning N_{imjt} randomly selected products from that module to the firm. This is aggregated to the firm level and then transformed as in (21) obtain a counterfactual scalability ratio for a given firm-sector-year. Our adjusted scalability measure is obtained by subtracting this counterfactual (log) scalability ratio from the observed one. The adjusted measure can be interpreted as capturing excess scalability with respect to a random portfolio of products of the same size. Additional details on the construction, including examples and robustness checks, are provided in Appendix C.

4.2 Scalability in the Cross-section

We begin by testing the model's predictions about the cross-sectional pattern of scalability: in particular, as Proposition 1 predicts, we assess whether the scalability ratio is positively related to size and scope.

$$\begin{aligned} \ln\left(\frac{x_{ijt}}{y_{ijt}}\right) &= \alpha_R + \beta_R \ \text{size}_{ijt} + \lambda_i^R + \Gamma_{jt}^R + \varepsilon_{ijt}^R \\ \ln\left(\frac{x_{ijt}}{y_{ijt}}\right) &= \alpha_N + \beta_N \ \text{scope}_{ijt} + \lambda_i^N + \Gamma_{jt}^N + \varepsilon_{ijt}^N \end{aligned} \tag{22}$$

[28]There are two potential issues. First, within an attribute, the set of characteristics may be partitioned into coarse groupings, so that products with characteristics that are different may be reported as sharing the same characteristic. For example, the attribute 'style' for flashlights is only described as "waterproof," "with batteries," or "without batteries." In this case, our raw scalability measure could overstate the degree of standardization. Second, fully describing a product may require describing a very large set of attributes, while the data contains information about a more limited subset. The severity of these issues is likely to be heterogeneous across modules and attributes.

where $\frac{x_{ijt}}{y_{ijt}}$ is the scalability ratio as defined in the previous subsection, and size_{ij} and scope_{ij} are the log of total revenue and scope, respectively, and $\{\lambda_i^R, \lambda_i^N\}$ and $\{\Gamma_{jt}^R, \Gamma_{jt}^N\}$ are firm- and sector-by-year fixed effects respectively.[29]

The results, in Table 1, show that $\beta_R > 0$ and $\beta_N > 0$, consistent with the model's prediction that firms with larger size or scope have more scalable knowledge. Columns (2) and (4) show that these results are robust to controlling for firm fixed effects.[30]

Table 1: Cross-Sectional Relationship: Scalability, Scope and Size

	(1) $\ln\frac{x}{y}$	(2) $\ln\frac{x}{y}$	(3) $\ln\frac{x}{y}$	(4) $\ln\frac{x}{y}$
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

The table reports the results from estimating (22). The dependent variable is the log of scalability, as defined in equation (22), and the independent variables are the logs of firm size (revenue) and scope. All variables are standardized relative to the mean of their sector and time period (year). Standard errors are robust. Scalability is adjusted relative to its bootstrapped counterpart.

4.3 Scalability and Response to Shocks

Next, we examine how scalability mediates a firm's response to a common sectoral shock. Proposition 11 predicts that, when the two forms of expertise are relatively substitutable, firms with higher scalability experience larger changes in size, scope, and scalability in response to the same shock. We test this prediction using the following specifications:

$$\begin{aligned}\Delta\ \text{size}_{ij} &= \alpha_{RS}\ \Delta G_j + \beta_{RS}\ \left(\Delta G_j \times \ln\frac{x_{ij}}{y_{ij}}\right) + \gamma_{RS}\mathbf{X}_{ij} + \psi_j^{RS} + \epsilon_{ij}^{RS} \\ \Delta\ \text{scope}_{ij} &= \alpha_{NS}\ \Delta G_j + \beta_{NS}\ \left(\Delta G_j \times \ln\frac{x_{ij}}{y_{ij}}\right) + \gamma_{NS}\mathbf{X}_{ij} + \psi_j^{NS} + \epsilon_{ij}^{NS} \\ \Delta \ln\left(\frac{x_{ij}}{y_{ij}}\right) &= \alpha_{SS}\ \Delta G_j + \beta_{SS}\ \left(\Delta G_j \times \ln\frac{x_{ij}}{y_{ij}}\right) + \gamma_{SS}\mathbf{X}_{ij} + \psi_j^{SS} + \epsilon_{ij}^{SS} \end{aligned} \tag{23}$$

[29]With a slight abuse of notation, when we write $\ln(x_{ijt}/y_{ijt})$, we are referring to the log of the ratio between the observed scalability ratio and its bootstrapped counterpart.

[30]Table D1 in the Appendix shows that these results are robust across different categories in the NielsenIQ data and to disaggregating sectors further.

where dependent variables are the log changes in total revenue, number of products, and scalability ratio, respectively, between 2006 and 2015. The variable G_j is the China import penetration shock for sector j.[31] We control for sector fixed effects to account for heterogeneity in sector-specific responses and include firm-level controls $\mathbf{X}_{ij}$: log scalability, log size, log size squared, log scope, and log scope squared. All controls are measured prior to the shock to address potential systematic associations between size and growth, as well as mean reversion effects. The coefficients of interest are β_{RS}, β_{NS}, and β_{SS}, which capture how firms with different levels of scalability respond to the same shock. The theory predicts all three should be positive.

Columns (1)-(3) in Table 2 show that firms with higher scalability ratio change their size, scope and scalability more in response to the shock. There are two ways to interpret this result. On the one hand, if one takes as given that scalable and local knowledge are substitutes (as we argue is natural in this setting), then this is a test of Proposition 11. Alternatively, one could accept the model but be agnostic about σ, and view Columns (1)-(3) as evidence that scalable and local expertise are substitutes.[32]

One might worry that scalability may simply be standing in for size (or scope), which may be associated with a firm's responsiveness to shocks for reasons outside the model. This assertion is difficult to assess, as (i) our theory predicts that the same differences in fundamentals drive changes in scalability, scope, and size jointly, and (ii) scope and size are arguably measured more precisely than scalability. Together, these imply that a regression that gauges responsiveness using scalability as well as size or scope would likely load the explanatory power on size or scope. Remarkably, when we estimate the same specifications, controlling for the interactions of the shock with scope in columns (4)-(6) or with size in columns and (7)-(9), the effect of scalability survives (though in some cases, we lose statistical significance). These results show that scalability continues to be a significant predictor of firm responsiveness, indicating that it captures an independent dimension of firm capabilities not accounted for by size alone. Overall, the results provide strong support for the central prediction of the theory – more scalable firms are more sensitive to demand.

4.4 Market Size and Concentration

Proposition 12 states that a given increase in G will raise the concentration of firm size more in sectors where scalability is associated more strongly with size. To identify such sectors, we calculate, for each sector j, the covariance between scalability and size in 2006 ($W_{j,2006}$), focusing on firms

[31]Note that the shock, ΔG_j, is sector-specific. Therefore, when we estimate these specifications with sector fixed effects, the coefficient α on the shock cannot be identified.

[32]In other contexts, expenditure on scalable and non-scalable investments can be measured directly, so that the relationship between size and the share of scalable investments can be used as suggestive evidence of substitutability. Take, for example, the use of managerial input, a form of scalable (or firm-wide) expertise. Akcigit, Alp and Peters (2021) and Chen, Habib and Zhu (2023) find that larger firms systematically hire more managers while Grobovšek (2020) documents that the share of managerial compensation rises with size, consistent with relatively high substitutability.

Table 2: **Response to Shocks: Scalability**

	(1) Δ size	(2) Δ scope	(3) $\Delta \ln\left(\frac{x}{y}\right)$	(4) Δ size	(5) Δ scope	(6) $\Delta \ln\left(\frac{x}{y}\right)$	(7) Δ size	(8) Δ scope	(9) $\Delta \ln\left(\frac{x}{y}\right)$
$\Delta G \times \ln\left(\frac{x}{y}\right)$	0.0262**	0.0141**	0.0119***	0.0165	0.0090	0.0116***	0.0162	0.0138**	0.0116***
	(0.012)	(0.006)	(0.003)	(0.013)	(0.006)	(0.003)	(0.013)	(0.006)	(0.003)
$\Delta G \times$ scope				0.0414***	0.0220***	0.0014			
				(0.011)	(0.005)	(0.003)			
$\Delta G \times$ size							0.0965***	0.0028	0.0030
							(0.016)	(0.007)	(0.003)
Obs.	14,186	14,186	13,467	14,186	14,186	13,467	14,186	14,186	13,467
R-squared	0.167	0.121	0.257	0.168	0.123	0.257	0.170	0.121	0.257
Sector	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: The table reports the results from estimating (23) for the period 2006–2015. The dependent variable in Columns (1), (4), and (7) is the change in the (log of) size (revenue) of firm i in sector j. In Columns (2), (5), and (8), the change in scope, and in Columns (3), (6), and (9), the change in scalability. The key independent variable is the China import penetration shock from 2006-2015 interacted with the firm's baseline level of scalability, scope, or size in 2006. Standard errors are robust. All specifications include as controls the shock, sector fixed effects and controls for scalability, firm size, size squared, scope, and scope squared.

active throughout the 2006–2015 period.[33] We then estimate the following specification:

$$\Delta Y_j = \alpha \; \Delta G_j + \beta \; \left(\Delta G_j \times \hat{W}_j\right) + \gamma \hat{W}_j + \epsilon_j \tag{24}$$

where ΔY_{jt} is the change in concentration in sector j between 2006 and 2015. As before, ΔG_j is the sector-level Chinese import penetration shock. The coefficient of interest is β. According to the theory, $\beta > 0$, i.e. concentration is more sensitive to demand when scalability covariance more positively with size.

Table 3 shows the results for various measures of concentration. Column (1) uses the Herfindahl index, column (2) the concentration ratio —the share of revenue held by the top five firms in the sector— and column (3) the additive inverse of the number of firms accounting for 80% of market share. For all three measures, we find that $\beta > 0$, showing that the covariance between scalability and size is informative about how industry concentration responds to demand.

4.5 Size and Response to Shocks

Corollary 2 states that, in response to the same change in sectoral demand G, firms with greater initial size or scope exhibit larger changes in both size and scope. This prediction can be tested even in the absence of a direct measure of scalability. This allows us to extend our analysis beyong

[33]Specifically, we estimate the regression $\ln\left(\frac{x_{ij}}{y_{ij}}\right) = \lambda_j + W_j \text{ scope}_{ij} + \varepsilon_{ij}$ 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. The estimated coefficient on size is positive in over 80% of sectors.

Table 3: Market Size and Concentration

	(1)	(2)	(3)
	Δ HHI	Δ Share Top 5	-Δ ln # firms 80% share
$\Delta G \times \hat{W}_{2006}$	0.0073***	0.0088***	0.0422***
	(0.002)	(0.002)	(0.008)
Observations	112	112	110
R-squared	0.065	0.080	0.099

Note: The table reports the results of estimating (24). The dependent variable in Columns (1) is the change in the Herfindahl index in sector j from 2006-2015. In Column (2), the change in the market share held by the top five firms in the sector and in Column (3), the log change in the number of firms accounting for 80% of total market share, multiplied by minus one. The variable $\hat{W}_{j,2006}$ is the covariance between scalability and size in 2006, estimated separately for each sector using a reduced-form approach and considering only firms with positive sales in both 2006 and 2015. Standard errors are robust. All columns use the China import penetration shock. The data are drawn from the NielsenIQ dataset.

the multi-product setting: in particular, we can perform this test with establishment-level data on revenues and employment.

Our data source is the National Establishment Time Series (NETS) from Dun & Bradstreet, which covers almost the entire universe of firms and sectors in the US.[34] It provides annual information on employment and sales for "lines of business" at specific locations, which we refer to as establishments. Each establishment is assigned a Data Universal Numbering System identifier, allowing us to track its employment and sales over time. For each establishment, we observe its location, industry classification, and parent company; we use the parent company identifier to define firms. For our baseline analysis, we classify establishments at the 4-digit SIC level.[35] We construct measures of firm *size* (total employment) and *scope* (number of establishments), using employment as the preferred size metric due to its higher accuracy in NETS.[36]

We run the following specification on both the multi-product and multi-establishment data:

$$\begin{aligned}
\Delta\text{size}_{ij} &= \alpha_{RR}\ \Delta G_j + \beta_{RR}\ (\Delta G_j \times \text{size}_{ij}) + \gamma_{RR}\mathbf{X}_{ij} + \psi_{RRj} + \epsilon_{RRij} \\
\Delta\text{scope}_{ij} &= \alpha_{NR}\ \Delta G_j + \beta_{NR}\ (\Delta G_j \times \text{size}_{ij}) + \gamma_{NR}\mathbf{X}_{ij} + \psi_{NRj} + \epsilon_{NRij} \\
\Delta\text{size}_{ij} &= \alpha_{RN}\ \Delta G_j + \beta_{RN}\ (\Delta G_j \times \text{scope}_{ij}) + \gamma_{RN}\mathbf{X}_{ij} + \psi_{RNj} + \epsilon_{RNij} \\
\Delta\text{scope}_{ij} &= \alpha_{NN}\ \Delta G_j + \beta_{NN}\ (\Delta G_j \times \text{scope}_{ij}) + \gamma_{NN}\mathbf{X}_{ij} + \psi_{NNj} + \epsilon_{NNij}.
\end{aligned} \tag{25}$$

[34] The data is provided by Walls & Associates. Appendix B provides detailed information.

[35] We apply sample restrictions and robustness exercises – following Crane and Decker (2019), who compare NETS to administrative data sources – to ensure that our results are representative.

[36] Appendix D.2 shows that results are similar when using revenue. The number of establishments aligns with measures of firm expansion along intensive and extensive margins, as in Cao, Hyatt, Mukoyama and Sager (2020).

where Δsize_{ij} and Δscope_{ij} represent the log changes in firm size and scope, respectively, between 2006 and 2015. As before, the demand shock, ΔG_j, is the sector-level change in Chinese import penetration. The controls $\mathbf{X}_{ij}$ include pre-shock values: log size, log size squared, log scope, and log scope squared. We also include sector fixed effects. To facilitate interpretation and comparability, we standardize the size and scope variables within each sector. The coefficients of interest – β_{RR}, β_{NR}, β_{RN}, and β_{NN} – capture whether larger firms (in terms of size or scope) exhibit systematically stronger responses to shocks. The model predicts that all four should be positive.

Table 4 presents our baseline estimates. Columns (1)–(4) use multi-product data from NielsenIQ, while Columns (5)–(8) use multi-establishment data from NETS. Each column reports the coefficient on the interaction between the shock and either firm size or scope. It shows significant positive estimates across all cases, i.e. firms with greater initial size (or scope) exhibit larger changes in both size and scope in response to a common demand shock.[37]

The magnitudes are economically meaningful: in response to 1 standard deviation decline in Chinese import penetration, a firm that is 1 standard deviation larger than the sector average increases size by 4.4 percentage points and scope by 2.4 percentage points more than the sector average. Notably, in the NETS data, the estimated response of size is almost twice as large as that for scope, implying that larger firms not only expand more overall but also on a per-unit basis.

Appendix D contains a number of robustness checks. The first examines sensitivity to the definition of growth and the sample of firms used. Our baseline specification measures log changes in size and scope among surviving firms, capturing the intensive margin but abstracting from differences in entry and exit dynamics. To address this, we re-estimate our regressions using the bounded growth rates of Davis and Haltiwanger (1992), which incorporate firm exit. As shown in Tables D3 and D5, our main findings remain qualitatively unchanged. We also explore non-parametric versions of equation 25, estimating the effect of shocks across different quantiles of the size and scope distribution. Figure D1 shows that responses to the China shock are increasing in firm size and scope. For example, the largest firms (top 1%) exhibit significantly stronger responses than firms below the median, consistent with the monotonic patterns reported in Table 4.

5 Dynamic Patterns

In this section, we present additional empirical results that link scalability to firm dynamics and aggregate growth. These go beyond the implications for the simple theory presented in this paper, but serve to highlight the broader relevance of standardization. Specifically, we show that

[37]Heterogeneous responses to demand shocks is related to, but differs from the literature studying the cyclicality of large vs. small firms, e.g., Moscarini and Postel-Vinay (2012), Fort, Haltiwanger, Jarmin and Miranda (2013), and Crouzet and Mehrotra (2020). These papers focus on whether employment of large firms exhibits a stronger correlation (usually, unconditional) with aggregate economic activity 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.

Table 4: Response to Shocks: Size and Scope

	(1) Δ size	(2) Δ scope	(3) Δ size	(4) Δ scope	(5) Δ size	(6) Δ scope	(7) Δ size	(8) Δ scope
ΔG × size	0.017***	0.001			0.044**	0.024***		
	(0.005)	(0.004)			(0.022)	(0.009)		
ΔG × scope			0.014***	0.017***			0.117***	0.057***
			(0.005)	(0.003)			(0.042)	(0.021)
Obs.	17,138	17,138	17,138	17,138	321,115	321,115	321,115	321,115
R-squared	0.163	0.150	0.163	0.152	0.047	0.023	0.043	0.242
Sector	Y	Y	Y	Y	Y	Y	Y	Y
Data	NielsenIQ: Multi-product				NETS: Multi-establishment			

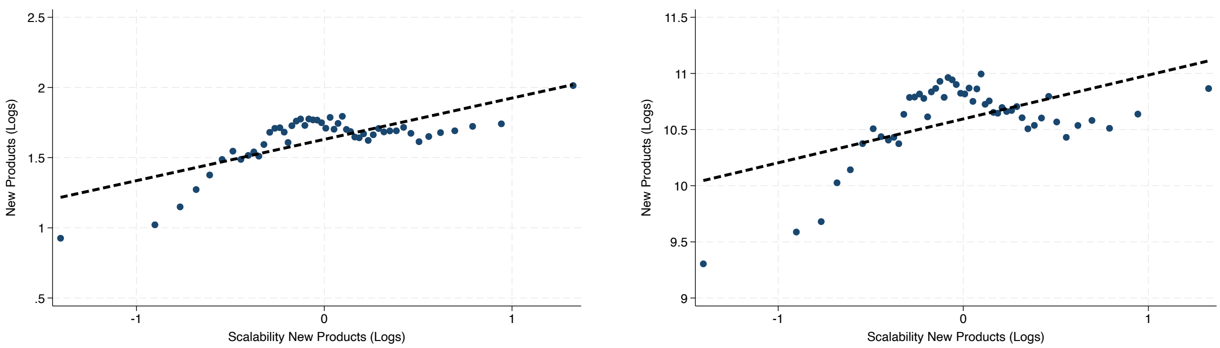
Note:The table reports the results of estimating (25). The dependent variable is either the log change in total employment of firm i in sector j from 2006 to 2015, or the change in the number of products or plants. The table reports β_{RR} (Columns 1 and 5), β_{NR} (Columns 2 and 6), β_{RN} (Columns 3 and 7), and β_{NN} (Columns 4 and 8). Columns (1)–(4) use NielsenIQ data, while Columns (5)–(8) use NETS data. All regressions use the China import penetration shock from 2006 to 2015. Specifications include robust standard errors, sector fixed effects, and firm-level controls: scalability, log size, log size squared, log scope, and log scope squared.

standardization is associated with innovation "bursts" and innovations that diffuse more widely.

5.1 Scalability and Innovation Bursts

Berlingieri, De Ridder, Lashkari and Rigo (2024) have shown the importance of innovation bursts – the rapid addition of new products to a firm's portfolio – for understanding variation in firm-level growth, industry concentration and growth. We show that when a firm introduces a set of new products, if the newly introduced products are more standardized then they are more likely to comprise an innovation burst. Panel (a) of Figure 3 plots the relationship between the scalability of newly introduced products (adjusted using the bootstrap procedure referred to earlier) and the number of new products launched, controlling for sector-period fixed effects. It shows a strong positive relationship. Panel (b) shows a similar positive association of scalability with the total revenue generated by the new products. Together, the panels suggest that standardization plays an important role in episodes of extreme firm growth.

Figure 3: Standardization and Innovation



(a) New products (b) Revenue New Products

Note: Panel (a) shows a binscatter plot of the relationship between the number of new products and the scalability of those products (i.e., $\frac{x}{y}$) (in logs). Panel (b) shows a binscatter plot of the relationship between the revenue from new products (in logs) and their scalability (in logs). The scalability index is adjusted relative to the alternative bootstrapped index. Both figures control for period-sector fixed effects.

5.2 Scalability and Knowledge Diffusion

Finally, we show that scalability has implications for diffusion of knowledge in the economy. The underlying hypothesis is that scalability — applicability of knowledge within a firm — also makes such knowledge more useful to other firms in the industry. This captures the intuitive notion that the process of standardization involves practices which likely also facilitate the transfer of knowledge beyond firm boundaries. An alternative, complementary, mechanism is that scalable expertise reflects innovations that are inherently more attractive or relevant for other firms to adopt.

To test this hypothesis, we construct a novel measure of diffusion based on attribute-level product data from NielsenIQ. Specifically, we track how widely a new product characteristic spreads across firms after it is first introduced. For each characteristic c introduced by firm i in product module m at time t, we count the number of products introduced by other firms that adopt the same characteristic within a future time window $[t, \tau]$. This provides a *ex-post* measure of how broadly a given feature diffuses within the market after its introduction. Formally, for each characteristic c introduced by firm i, we define:

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

The numerator counts the number of times characteristic c, first introduced by firm i in module m, appears in products launched by other firms over a horizon τ and the denominator, the total number of products introduced by those other firms in the same time window. By construction, $\mathbb{D}_{c,m,i,t,\tau}$ ranges from 0 to 1.

We now examine how the diffusion of a new characteristic is related to the scalability of the firm that introduced it. Given our interest in characteristic-level outcomes (rather than firm-level), we run the following regression at that more granular level:

$$\mathbb{D}_{cimt\tau} = \alpha + \beta\, \mathcal{SI}_{aimt-1} + \gamma\, scope_{imt} + \lambda_{amt\tau} + \theta_{aim} + \epsilon_{cimt\tau} \tag{26}$$

where c, a, i, m and t index the characteristic, attribute, firm, product module, and time period, respectively. The coefficient of interest in β, the coefficient on $\mathcal{SI}_{aimt-1}$, the Scalability Index, measured at the attribute-module-firm level in the prior period. The lagging is done to mitigate reverse causality or simultaneity concerns. We control for firm scope and a rich set of fixed-effects to address unobserved heterogeneity. Our most saturated specification includes both attribute $\times$ module $\times$ time $\times$ age fixed effects ($\lambda_{amt\tau}$) and firm $\times$ attribute $\times$ module fixed effects (θ_{aim}). The results, shown in columns (1) and (2) of Table E1, indicate a robust and positive association between scalability and diffusion.[38]

Table 5: Scalability and Knowledge Diffusion

Diffusion	(1)	(2)
$\mathcal{SI}$	0.0599***	0.0097***
	(0.000)	(0.001)
scope	-0.0292***	-0.0057***
	(0.000)	(0.000)
Observations	3,319,518	3,234,863
R-squared	0.808	0.914
Firm-Attribute-Module	N	Y
Attribute-Module-Time-Age	Y	Y

Note: The table shows the results of estimating (26). The dependent variable, $\mathbb{D}_{cmit\tau}$, measures the diffusion of characteristic c in module m, introduced by firm i between periods t and $t+\tau$. The key independent variable is the scalability index $\mathcal{SI}_{aimt-1}$, measured for attribute a in firm i and module m in period $t-1$. All specifications control for the total number of products sold by firm i in module m at time t.

These findings have a number of implications. One, they highlight the disproportionate role of large (more precisely, scalable) firms in driving aggregate growth. Two, changes in the environment that cause firms to change the scalability of their investments (such as the China shock studied in the preceding sections) have effects that go beyond the focal firms. Finally, on the normative front, these patterns to point to an externality: to the extent that firms do not internalize the value generated by diffusion, they will under-invest in scalability (relative to the social optimum). As a result, subsidizing scalable expertise becomes socially desirable.[39]

[38] In Appendix E, we construct an alternative measure of a firm's tendency to introduce scalable characteristics, using *ex-post* information about the firm's use of that characteristic in products that are introduced in the future. The strong positive relationship between diffusion and lagged scalability remains robust when we use this alternative measure.

[39] This echoes the notion that, in the presence of knowledge spillovers, there can be underinvestment in ideas, as

6 Conclusion

The paper develops and tests a model of multi-unit firms, centered on standardization. The analysis delivers a simple, empirically relevant, insight: when scalable and non-scalable investments are substitutes, standardization increases a firm's marginal returns to scale, making it more sensitive to changes in demand.

There are many promising directions for future research. Our theoretical framework was kept intentionally simple and abstracts from many realistic elements, such as richer heterogeneity across products. We also abstracted from dynamics and stochastic fundamentals, both of which are no doubt essential to paint a complete picture of firm heterogeneity and growth. Incorporating these elements and undertaking a full-fledged quantitative analysis is a natural and ambitious next step. Finally, integrating standardization into workhorse models of aggregate growth—particularly the link between standardization and diffusion—promises new positive and normative implications.

emphasized recently by Crouzet et al. (2024).

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.

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

_ , Harun Alp, and Michael Peters, "Lack of selection and limits to delegation: firm dynamics in developing countries," *American Economic Review*, 2021, *111* (1), 231–275.

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.

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

Berlingieri, Giuseppe, Maarten De Ridder, Danial Lashkari, and Davide Rigo, "Growth through Innovation Bursts," 2024.

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.

Bilir, L Kamran and Eduardo Morales, "Innovation in the global firm," *Journal of Political Economy*, 2020, *128* (4), 1566–1625.

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.

Chan, Mons, Guangbin Hong, Joachim Hubmer, Serdar Ozkan, and Sergio Salgado, "Scalable vs. Productive Technologies," Technical Report 2024.

Chandeler, Alfred D., *Scale and Scope: The Dynamics of Industrial Capitalism*, Harvard University Press, 1990.

Chen, Chaoran, Ashique Habib, and Xiaodong Zhu, "Finance, managerial inputs, and misallocation," *American Economic Review: Insights*, 2023, *5* (3), 409–426.

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.

_ , Janice Eberly, Andrea Eisfeldt, and Dimitris Papanikolaou, "Intangible Capital, Firm Scope, and Growth," Technical Report, Working Paper 2024.

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.

Ding, Xiang, "Industry linkages from joint production," Technical Report 2023.

Engbom, Niklas, Hannes Malmberg, Tommaso Porzio, Federico Rossi, and Todd Schoellman, "Economic Development According to Chandler," Technical Report, Technical report, mimeo, New York University 2024.

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.

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

Grobovšek, Jan, "Managerial delegation, law enforcement, and aggregate productivity," *The Review of Economic Studies*, 2020, *87* (5), 2256–2289.

Hoberg, Gerard and Gordon M Phillips, "Scope, Scale, and Concentration: The 21st-Century Firm," *The Journal of Finance*, 2025, *80* (1), 415–466.

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.

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

Jiang, Xian and Hannah Rubinton, "The Adoption of Non-Rival Inputs and Firm Scope," Technical Report, CESifo Working Paper 2024.

Kleinman, Benny, "Wage inequality and the spatial expansion of firms," Technical Report, Tech. rep., mimeo 2022.

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

Koh, Paul and Devesh Raval, "Economies of Scope from Shared Inputs," Technical Report 2025.

Lashkari, Danial, Arthur Bauer, and Jocelyn Boussard, "Information technology and returns to scale," *American Economic Review*, 2024, *114* (6), 1769–1815.

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.

Lucas, Robert E, "On the size distribution of business firms," *The Bell Journal of Economics*, 1978, pp. 508–523.

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.

Menger, Karl, "The logic of the laws of return: a study in meta-economics," *Economic activity analysis*, 1954, pp. 419–481.

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.

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).

Shephard, Ronald William, *Theory of cost and production functions*, Princeton University Press, 1953.

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 $\frac{w}{A_{ij}Z_{uij}}$. 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 $\frac{\theta}{\theta-1}\frac{W}{A_{ij}Z_{uij}}$. 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 Marginal 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 Q_i is the solution to

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

subject to the capacity constraint $x_i + N_i y_i \leq 1$, technology $A_i Z(x_i, y_i) L_{iu} \geq Q_{iu}$, and within firm aggregation $Q_i \geq N_i^{\frac{\epsilon}{\epsilon-1}} Q_{iu}$ and $L_i \geq \int_0^{N_i} L_{iu} du$. Eliminating L_i, L_{iu}, and Q_{iu} 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_i and using $k_i = x_i / y_i$ 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:

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

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 (8). The optimality condition (9) 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

$$\frac{d\ln N_i}{d\ln Q_i} = \underbrace{\frac{\partial \ln \widetilde{\mathbb{N}}_i}{\partial \ln Q_i}}_{=1/2} + \underbrace{\frac{\partial \ln \widetilde{\mathbb{N}}_i}{\partial \ln k_i}}_{=\frac{1-S_i}{2}}\underbrace{\frac{N\mathbb{K}'(N_i)}{\mathbb{K}(N_i)}}_{=\sigma}\frac{d\ln N_i}{d\ln Q_i} \tag{27}$$

$$= \frac{1}{2-(1-S_i)\,\sigma} \tag{28}$$

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

$$MRTS_i = -\frac{Q_i c_i''(Q_i)}{c_i'(Q_i)} = \frac{S_i}{2-(1-S_i)\,\sigma} \tag{29}$$

A.4 Marginal Returns to Scale and Policy

Consider a household maximizes with preferences $U = u(Q, L)$, where $Q \equiv \left(\int Q_i^{\frac{\theta-1}{\theta}} di\right)^{\frac{\theta}{\theta-1}}$ that faces the budget constraint $\int P_i Q_i di \leq WL + \Pi - T$ where T is a lump sum tax to pay for subsidies. Defining the price index $P_i \equiv \left(\int P_i^{1-\theta} di\right)^{\frac{1}{1-\theta}}$, the household's consumption of good i satisfies

$$P_i = P\left(\frac{Q_i}{Q}\right)^{-\frac{1}{\theta}}$$

and its choice of labor satisfies

$$\frac{P}{W} = -\frac{u_Q}{u_L}.$$

Market clearing for labor requires that $L = \int c_i(Q_i)\,di$.

Firm i maximizes profit subject to revenue subsidy $\mathfrak{s}_i$ which is paid for with a lump sum tax. Since firm i's revenue is $P_i Q_i = PQ^{1/\theta} Q_i^{1-1/\theta}$, its optimizes by choosing a quantity Q_i to maximize

$$\pi_i = \max_{Q_i}\,(1+\mathfrak{s}_i)\,PQ^{1/\theta}Q_i^{1-1/\theta} - Wc_i(Q_i)$$

The optimal quantity satisfies the FOC

$$\frac{\theta-1}{\theta}\left(1+\mathfrak{s}_i\right)\frac{PQ^{1/\theta}}{W}Q_i^{-1/\theta}=c_i'\left(Q_i\right)$$

which also implies $P_i=\frac{\theta}{\theta-1}\frac{Wc_i'(Q_i)}{1+\mathfrak{s}_i}$. This along with $\frac{P}{W}=-\frac{u_Q}{u_L}$ implies

$$\frac{\theta-1}{\theta}\left(1+\mathfrak{s}_i\right)\frac{u_Q Q^{1/\theta}}{-u_L}Q_i^{-1/\theta}=c_i'\left(Q_i\right) \tag{30}$$

Equation (30) along with the market clearing condition for labor, $L=\int c_i\left(Q_i\right)di$, are sufficient to characterize the allocation.

We next characterize how a subsidy to firm i affects each firm's quantity, $Q_{\tilde{\imath}}$

$$\begin{aligned}
\frac{d\log\frac{u_Q Q^{1/\theta}}{-u_L}}{d\mathfrak{s}_i}-\frac{1}{\theta}\frac{d\log Q_{\tilde{\imath}}}{d\mathfrak{s}_i} &= \frac{Q_{\tilde{\imath}}c''\left(Q_{\tilde{\imath}}\right)}{c_{\tilde{\imath}}'\left(Q_{\tilde{\imath}}\right)}\frac{d\log Q_{\tilde{\imath}}}{d\mathfrak{s}_i} \quad \text{for } \tilde{\imath}\neq i\\
\frac{1}{1+\mathfrak{s}_i}+\frac{d\log\frac{u_Q Q^{1/\theta}}{-u_L}}{d\mathfrak{s}_i}-\frac{1}{\theta}\frac{d\log Q_i}{d\mathfrak{s}_i} &= \frac{Q_i c''\left(Q_i\right)}{c_i'\left(Q_i\right)}\frac{d\log Q_i}{d\mathfrak{s}_i} \quad \text{for } \tilde{\imath}\neq i
\end{aligned}$$

Evaluating this derivative at $\mathfrak{s}_i=0$ and defining ρ_i to satisfy $\frac{1}{\rho_i}\equiv\frac{1}{\theta}+\frac{Q_ic_i''(Q_i)}{c_i'(Q_i)}=\frac{1}{\theta}-MRTS_i$. These can be expressed as

$$\frac{d\log\frac{u_Q Q^{1/\theta}}{-u_L}}{d\mathfrak{s}_i} = \frac{1}{\rho_{\tilde{\imath}}}\frac{d\log Q_{\tilde{\imath}}}{d\mathfrak{s}_i} \quad \text{for } \tilde{\imath}\neq i \tag{31}$$

$$-1+\frac{d\log\frac{u_Q Q^{1/\theta}}{-u_L}}{d\mathfrak{s}_i} = \frac{1}{\rho_i}\frac{d\log Q_i}{d\mathfrak{s}_i} \quad \text{for } \tilde{\imath}\neq i \tag{32}$$

Next, we characterize how $\frac{u_Q Q^{1/\theta}}{-u_L}$ responds to a subsidy to firm i. Since $Q^{\frac{\theta-1}{\theta}}\equiv\int Q_i^{\frac{\theta-1}{\theta}}di$,

$$d\log Q=\int\omega_{\tilde{\imath}}d\log Q_{\tilde{\imath}}d\tilde{\imath}$$

where $\omega_i\equiv\left(\frac{Q_i}{Q}\right)^{\frac{\theta-1}{\theta}}=\frac{P_iQ_i}{PQ}$. Note also that differentiating the market clearing condition and using each firm's choice of price gives

$$\begin{aligned}
d\log L &= \frac{\int Q_ic_i'\left(Q_i\right)d\log Q_{\tilde{\imath}}d\tilde{\imath}}{\int c_i\left(Q_i\right)d\tilde{\imath}}=\frac{\theta-1}{\theta}\frac{PQ}{WL}\int\omega_i d\log Q_{\tilde{\imath}}d\tilde{\imath}\\
&= \frac{\theta-1}{\theta}\frac{PQ}{WL}d\log Q
\end{aligned}$$

Together, these imply

$$\begin{aligned} d\log\frac{u_Q Q^{1/\theta}}{-u_L} &= \left(\frac{Qu_{QQ}}{u_Q}d\log Q+\frac{Lu_{QL}}{u_Q}d\log L-\frac{Qu_{LQ}}{u_L}d\log Q-\frac{Lu_{LL}}{u_L}d\log L+\frac{1}{\theta}d\log Q\right)\\ &= \left[\frac{Qu_{QQ}}{u_Q}-\frac{Qu_{LQ}}{u_L}+\frac{1}{\theta}+\frac{\theta-1}{\theta}\frac{PQ}{wL}\left(\frac{Lu_{QL}}{u_Q}-\frac{Lu_{LL}}{u_L}\right)\right]d\log Q \end{aligned}$$

Letting $B\equiv\left[\frac{Qu_{QQ}}{u_Q}-\frac{Qu_{LQ}}{u_L}+\frac{1}{\theta}+\frac{\theta-1}{\theta}\frac{PQ}{WL}\left(\frac{Lu_{QL}}{u_Q}-\frac{Lu_{LL}}{u_L}\right)\right]$, this last equation can be expressed as $d\log\frac{u_Q Q^{1/\theta}}{-u_L}=Bd\log Q$.

Using (31) and (32), the change in Q an be expressed as

$$\begin{aligned} \frac{d\log Q}{d\mathfrak{s}_i} &= \int\omega_{\tilde{\imath}}\frac{d\log Q_i}{d\mathfrak{s}_i}d\tilde{\imath}\\ &= \int\omega_{\tilde{\imath}}\rho_{\tilde{\imath}}\frac{d\log\frac{u_Q Q^{1/\theta}}{-u_L}}{d\mathfrak{s}_i}d\tilde{\imath}+\omega_i\rho_i di\\ &= \int\omega_{\tilde{\imath}}\rho_{\tilde{\imath}}B\frac{d\log Q}{d\mathfrak{s}_i}d\tilde{\imath}+\omega_i\rho_i di\\ &= \frac{\omega_i\rho_i di}{1-B\int\omega_{\tilde{\imath}}\rho_{\tilde{\imath}}d\tilde{\imath}} \end{aligned}$$

The change in welfare is

$$\begin{aligned} \frac{dU}{d\mathfrak{s}_i} &= u_Q Q\frac{d\log Q}{d\mathfrak{s}_i}+u_L L\frac{d\log L}{d\mathfrak{s}_i}\\ &= u_Q Q\frac{d\log Q}{d\mathfrak{s}_i}+u_L L\frac{\theta-1}{\theta}\frac{PQ}{WL}\frac{d\log Q}{d\mathfrak{s}_i} \end{aligned}$$

Using $\frac{W}{P}=-\frac{u_L}{u_Q}$, this can be simplified to

$$\begin{aligned} \frac{dU}{d\mathfrak{s}_i} &= \frac{1}{\theta}u_Q Q\frac{d\log Q}{d\mathfrak{s}_i}\\ &= \frac{\frac{1}{\theta}u_Q Q}{1-B\int\omega_{\tilde{\imath}}\rho_{\tilde{\imath}}d\tilde{\imath}}\omega_i\rho_i di \end{aligned}$$

The lump sum tax to pay for the subsidy (in units of final output) is $\frac{T}{P}=\frac{\int\mathfrak{s}_{\tilde{\imath}}P_{\tilde{\imath}}Q_{\tilde{\imath}}d\tilde{\imath}}{P}=\int\mathfrak{s}_{\tilde{\imath}}Q^{\frac{1}{\theta}}Q_{\tilde{\imath}}^{1-\frac{1}{\theta}}d\tilde{\imath}$. At the Laissez Faire equilibrium, the marginal change in tax revenue from a small subsidy to i is

$$\frac{d\left(T/P\right)}{d\mathfrak{s}_i}=Q^{\frac{1}{\theta}}Q_i^{1-\frac{1}{\theta}}di=Q\omega_i di$$

As a result, the marginal change in welfare relative to the marginal cost of the subsidy is

$$\frac{\frac{dU}{d\mathfrak{s}_i}}{\frac{d(T/P)}{d\mathfrak{s}_i}} = \frac{\frac{\frac{1}{\theta}u'(Q)Q}{1-B\int \omega_{\tilde{\imath}}\rho_{\tilde{\imath}}d\tilde{\imath}}\omega_i\rho_i di}{Q\omega_i di} = \frac{\frac{1}{\theta}u'(Q)}{1-B\int \omega_{\tilde{\imath}}\rho_{\tilde{\imath}}d\tilde{\imath}}\rho_i$$

A.5 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. 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

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

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 13 *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:*

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

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. 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

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

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:

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

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:

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

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

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

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

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

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.6 Intermediate Model

In this section, we study an intermediate version of the model. We later will show that the 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

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

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 $S_i \equiv \frac{x_i}{x_i + N_i y_i}$ be the share of knowledge devoted to scalable expertise.

The firm's problem can be expressed as

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

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 $\frac{N\mathbb{K}'(N)}{\mathbb{K}(N)} = \sigma(\mathbb{K}(N))$.[40] 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 14 *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 $S_i = \frac{x_i}{x_i+N_iy_i} = \frac{k_i}{k_i+N_i}$ *is firm i's scalable share of knowledge and* σ_i *is the local elasticity of substitution.*

Proof. 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

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

[40]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 $\frac{d\ln k}{d\ln N} = -\left[\frac{z'}{kz''}\right]\frac{z-kz'}{z} = \sigma$.

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):

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

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:

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

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

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

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

A.6.1 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 15 *Suppose that Assumption 4 holds. Then there exists a unique solution, and it is interior.*

Proof. First, $\sigma<1$ implies that $\frac{\mathbb{K}(N)}{N}$ is strictly decreasing in N. To see this, note that $\frac{d\ln(\mathbb{K}(N)/N)}{d\ln N}=\sigma\left(\mathbb{K}(N)\right)-1<0$. 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:

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

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 $\phi=\psi\frac{\bar{N}}{\mathbb{K}\left(\bar{N}\right)+\bar{N}}$, $\mathbb{H}(N)$ 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 $\lim_{x\to 0}Z_x\left(x,y\right)<\infty$, or equivalently $\lim_{k\to 0}\frac{z(k)}{k}=Z_x\left(0,1\right)<\infty$. The former implies $\lim_{N\to 0}\frac{N}{\mathbb{K}(N)+N}=0$, and together, they imply $\lim_{N\to 0}\frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N}=\lim_{N\to 0}\frac{\mathbb{K}(N)}{\mathbb{K}(N)+N}\lim_{N\to 0}\frac{z(\mathbb{K}(N))}{\mathbb{K}(N)}=1\times\lim_{k\to 0}\frac{z(k)}{k}=Z_x\left(0,1\right)$. Together, these imply that

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

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 $\mathbb{H}\left(N^*\right)=1$, 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 $\lim_{N\to\infty}\frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N}=\lim_{N\to\infty}\frac{z'(\mathbb{K}(N))\mathbb{K}'(N)}{\mathbb{K}'(N)+1}\leq\lim_{N\to\infty}z'\left(\mathbb{K}(N)\right)=\lim_{k\to\infty}z'(k)=0$. In addition, $\lim_{N\to\infty}\frac{N}{\mathbb{K}(N)+N}=1$ We thus have

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

■

Proposition 16 *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. 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

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

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

Proposition 17 *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. 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

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

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 18 *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. The expression for the scope elasticity from Proposition 14 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.6.2 Intermediate Model, Case 2: $\sigma > 1$

Proposition 19 *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. 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 $Z(0,1) > 0$, $\lim_{N\to 0} \frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N} = \lim_{n\to 0} \frac{Z(0,1)}{\mathbb{K}(N)+N} = \infty$. We thus have

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

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 $\sigma(0) < 1+\omega$, and σ is continuous, so that there is a $\bar{k}$ and a $\tilde{\sigma} \in (\sigma(0), 1+\omega)$ such that $\sigma(k) \leq \tilde{\sigma}$ for all $k \in [0, \bar{k}]$. Since $\frac{N\mathbb{K}'(N)}{\mathbb{K}(N)} = \sigma(\mathbb{K}(N)) \leq \tilde{\sigma}$, 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

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

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, $\lim_{N\to\infty} z'(\mathbb{K}(N)) = \lim_{k\to\infty} z'(k) = Z_x(1,0) > 0$. 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 $\lim_{N\to\infty} \frac{z(\mathbb{K}(N))}{\mathbb{K}(N)+N} = \lim_{N\to\infty} \frac{z'(\mathbb{K}(N))\mathbb{K}'(N)}{\mathbb{K}'(N)+1} = Z_x(1,0)$.

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

$$= \frac{GA_i^{\psi}}{F\omega} \phi Z_x(1,0)^{\psi} \lim_{N\to\infty} N^{\phi-\omega}$$

If $\phi < \omega$, then $\lim_{N\to\infty} \mathbb{H}(N) = 0$. 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

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

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

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

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

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

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 20 *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. 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 21 *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. 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

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

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 22 *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. 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.7 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. 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 $\sigma \equiv \frac{Z_x Z_y}{Z Z_{xy}}$, 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.7.1 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 23 *Under Assumption 6, there is a unique solution, and it is interior.*

Proof. From Proposition 19, 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 24 *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. This follows from Proposition 20 and (40), (41), (42), (43), and (44). ■

Proposition 25 *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. This follows from Proposition 21 and (40). ■

Proposition 26 *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. This follows from Proposition 22 and (41), (42), and (43). ■

A.7.2 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 27 *Under Assumption 7, there is a unique solution, and it is interior.*

Proof. From proposition 15, 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 28 *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. This follows from proposition 20 and (40), (41), (42), (43), and (44). ■

Proposition 29 *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. This follows from Proposition 17 and (40). ■

Proposition 30 *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. This follows from Proposition 22 and (41) and (42). ■

A.8 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.7, we can follow the exact derivations in Appendix A.7 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.7, 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 NielsenIQ

The product data come from the NielsenIQ Retail Measurement Services (RMS), which track weekly sales volumes and quantities sold at the barcode (UPC) level using point-of-sale systems in retail stores, covering the period 2006–2015. Each UPC is a 12-digit Universal Product Code uniquely assigned to a specific good.

The main advantage of the RMS dataset is its size and coverage. It captures approximately $2 trillion in sales, representing 53% of grocery store sales, 55% of drug store sales, 32% of mass merchandiser sales, and smaller shares of convenience and liquor store sales. Data are collected from over 40,000 stores across 90 retail chains, covering 371 metropolitan statistical areas (MSAs) and 2,500 counties. In comparison to household-level scanner datasets, NielsenIQ RMS covers a wider range of products because it captures the universe of all transactions within the categories it covers, rather than the purchases made by a sample of households.

The original data consist of more than one million distinct products identified by barcodes and organized into a hierarchical structure. Each barcode is classified into one of 1,070 product modules, which are grouped into 104 product groups, and further aggregated into 10 major departments. These 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 is classified in the "Detergent-Packaged" module, within the "Detergent" product group, which falls under the "Non-Food Grocery" department. The "Detergent" product group includes several modules, such as automatic dishwasher compounds, heavy-duty liquid detergents, light-duty detergents, packaged detergents, dishwasher rinsing aids, and packaged soap.

NielsenIQ RMS data do not include direct information on manufacturing firms. However, products can be linked to firms using data from the GS1 US Data Hub. GS1 is the organization that issues barcode prefixes to producers. To obtain a UPC, a firm must first acquire a GS1 company prefix—a five- to ten-digit number that identifies the firm within its product UPCs. The GS1 database includes the name and address of the firm associated with each prefix, allowing us to match UPCs in the NielsenIQ RMS data to firm identities. A "firm" in this context is defined as the entity that purchased the prefix from GS1, which is typically the manufacturer—for example, Procter & Gamble.

Table A.I presents firm characteristics by type of censoring. Of the approximately 23,000 firms in the sample, the age of about 9,000 can be measured, while the remaining 14,000 were established before 2006.

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 for firms included in the baseline pooled sample covering the period 2006Q1–2015Q4. Firms that are already active in 2006Q1 or 2006Q2 are classified as left-censored, while firms with sales in 2015Q3 or 2015Q4 are classified as right-censored. Firms that both enter and exit within the sample period are labeled as "Complete." Firms for which we observe entry but not exit are labeled as "Right," and those for which we observe exit but not entry are labeled as "Left." Firms for which neither entry nor exit can be determined are classified as both right- and left-censored ("Right&Left"). For each group, the table reports the total number of observations, as well as summary statistics on firm duration, sales, number of firms, and number of sectors. Under "Duration," we report the average number of quarters a firm is observed, along with the share of firms observed for fewer than 4 quarters, between 4 and 16 quarters, and more than 28 quarters. Sales statistics refer to average quarterly sales (in thousands of dollars), deflated using the Consumer Price Index for All Urban Consumers. The table also reports the average and distribution statistics for total sales, number of products, and number of sectors. "Sectors" refers to the number of distinct product groups as classified by NielsenIQ..

B.2 NETS Data

The National Establishment Time Series (NETS) database is produced by Walls & Associates in collaboration with Dun & Bradstreet, converting archival establishment data into a longitudinal panel of establishment-level information. We use the version of the dataset that covers annual observations on specific lines of business at unique locations across the U.S. for the period 1990–2017. The NETS data allow us to observe sales and employment for each line of business identifier. For each identifier, we can track outcomes over time at the 8-digit Standard Industrial Classification (SIC) level and specific geographic coordinates (latitude and longitude). Additionally, each line of business can be linked to its headquarters using firm identifiers, with firms defined based on a common headquarters establishment.

Crane and Decker (2019) assesses the representativeness of NETS in the cross section by comparing it to U.S. Census Bureau data. They find that the static distributions of NETS are generally comparable to official sources across establishment size, industry, and geographic cells, although limitations arise particularly in the coverage of small firms. The main weaknesses of the NETS data pertain to annual establishment growth and firm lifecycle patterns. Since our analysis focuses on cross-sectional differences in long-term growth rates, it is less likely to be affected by these measurement limitations.

Our analysis focuses on changes in firm size and scope during the period 2006–2015 in response to the China shock. The baseline sample includes all firm-sector observations for which we can measure the shock. At both the firm-sector-year and firm-sector-location-year levels, we compute total size—measured as either total employment or total sales—and the number of distinct business lines (scope). Sectors are mapped to 4-digit SIC codes, and locations are mapped to metropolitan statistical areas (MSAs).

We conduct several robustness exercises to ensure the reliability of the patterns documented in the paper. Crane and Decker (2019) provides guidance on minimizing measurement error in the NETS data. Following their recommendations, we construct alternative samples that exclude small firm-sector observations (fewer than five employees) and single-establishment firms (those operating in only one location). We also examine the sensitivity of our results to alternative definitions of the firm. In our baseline, we use an iterative procedure to identify the ultimate headquarters. Although the original data link each business line to a direct headquarters, some of these headquarters are themselves subsidiaries of other headquarters. We resolve this by "rolling up" all directly and indirectly connected entities into a single firm identifier, capturing the ultimate corporate parent. As a further robustness check, we implement a procedure that associates business lines in ways that account for mergers and acquisitions, following the approach in Crane and Decker (2019).

B.3 Sectoral Shocks: China Import Penetration Shock

In this section, we describe the data sources used to update the China Shock measure developed by Autor et al. (2013). U.S. shipments ($Y_{j,06}$) at the 4-digit 1987 SIC industry level are obtained from the NBER-CES database.[41] Gross output ($YO_{j,06}$) at the 4-digit ISIC Rev. 3 industry level for several European countries comes from UNIDO. Following Bai and Stumpner (2019), we focus on the five largest European economies—Germany, France, the United Kingdom, Italy, and Spain—which also have the most comprehensive coverage at the 4-digit ISIC Rev. 3 level in the UNIDO dataset. Trade flows ($M_{j,06}$,$E_{j,06}$) for both the U.S. and European countries are sourced from UN Comtrade at the HS 6-digit level, using the CEPII-BACI database (Gaulier and Zignago, 2010).[42] While this differs slightly from the trade flow data used in Acemoglu et al. (2016), which come directly from

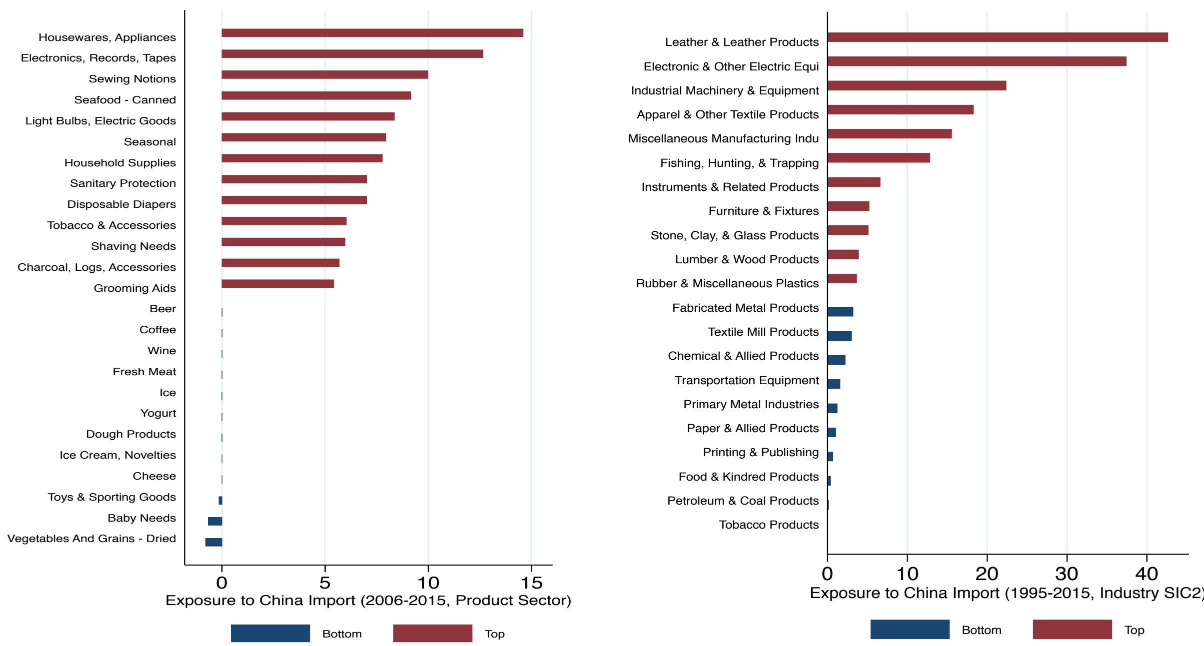
[41] http://www.nber.org/nberces

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

UN Comtrade, the CEPII-BACI dataset offers a harmonized version that reconciles discrepancies between exporter and importer reports. This harmonization substantially expands the number of countries with available trade data (up to 150 countries), relative to the original dataset. Lastly, we use the PCE deflator from BEA-NIPA for the U.S. and from Eurostat for the five European countries in our analysis.

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

(a) NielsenIQ (b) NETS



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

C Scalability

C.1 Illustrative Example

We measure firm-level scalability using the following index:

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

where Unique_{imjt} denotes the number of distinct characteristics used across all products sold by firm i in module m, sector j, and time t. Scope_{imjt} is the number of distinct products, and $\text{NumAttributes}_{mjt}$ is the number of active attributes in the module. We illustrate this with a simplified example based on a firm selling incandescent lamps.

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

Firm A sells 5 products, described along two attributes: style and use. Within these attributes, the firm uses several distinct characteristics:

- Unique characteristics: e.g., clear, halogen, nite fixture, etc.
- Attributes: style, use

We compute attribute-level and firm-level scalability as follows:

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

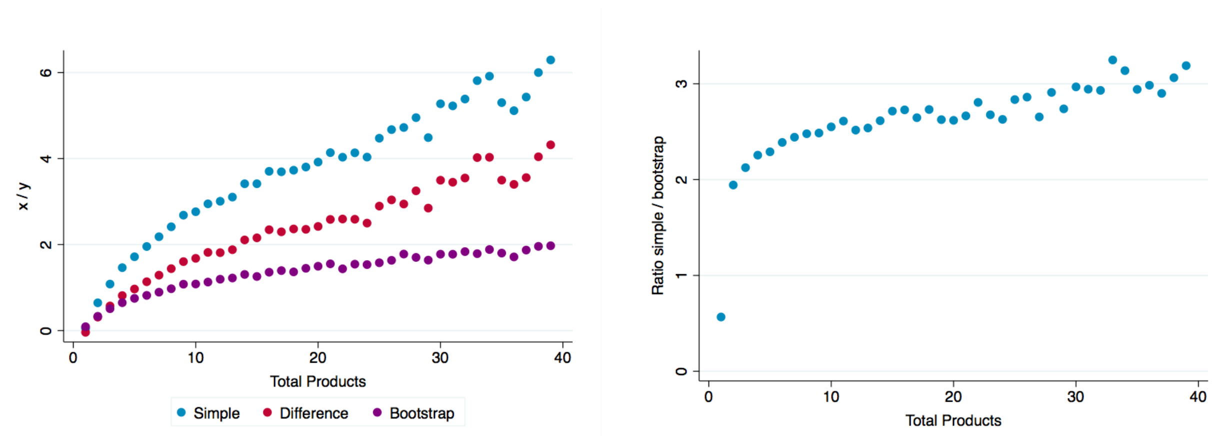
This simple example shows how the scalability index captures the extent to which firms rely on a common set of characteristics to define their products. Higher values of the index indicate more standardized product design, reflecting greater scalable expertise. In our data, the average scalability index is 0.47, with a standard deviation of 0.3.

C.2 Bootstrap

A potential concern is that our scalability index may mechanically increase as firms introduce more products. To address this, we construct an alternative index by randomly assigning products—within each sector—to firms of different sizes. We then compute a bootstrapped version of the scalability index, which serves as a reference. This alternative index captures the mechanical relationship between scalability and the number of products. Throughout the paper, our baseline scalability measure is always interpreted relative to this bootstrapped benchmark.

Panel (a) of Figure C1 displays the original index ($\frac{x}{y}$) alongside the bootstrapped version. The purple dots indicate that part of the positive relationship between $\frac{x}{y}$ and firm scope arises mechanically: as firms grow, they are more likely to have products that share attributes. However, our measure captures a size-dependent relationship that goes beyond what would occur by chance. The red dots represent the difference between the original and bootstrapped indices, which also increases with firm size. Panel (b) shows the ratio of the original to the bootstrapped index. Recall that we use the log of the ratio between the observed scalability ratio and its bootstrapped counterpart in all our specifications. This ratio increases as firms add products to their portfolio, indicating that larger firms replicate specific characteristics more frequently across products.

Figure C1: Scalability - Alternative (Bootstrapped) 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 sold by the firm. The blue dots represent estimates using the original measure. The purple dots correspond to estimates of $\frac{x}{y}$ when the sample of products is randomized within modules and across firms. The red dots represent the difference between the original measure and the bootstrapped version. Panel (b) displays the ratio of the original $\frac{x}{y}$ measure to the bootstrapped version. All values of $\frac{x}{y}$ are computed using the full product portfolio of each firm over the period 2006 to 2015.

D Robustness Checks and Supplementary Results

D.1 Multi-Product Data

Table D1: 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 from estimating equation 22 using NielsenIQ data. The dependent variable is the log of scalability, and the independent variables are the logs of size (revenue) and scope. All variables are standardized relative to the mean within sector and year. The scalability index is adjusted relative to the alternative (bootstrapped) version. Columns (1)–(2) include the following departments: Dry Grocery, Frozen Food, Dairy, Deli, Packaged Meat, and Fresh Produce. Columns (3)–(4) include Health & Beauty Care, Non-Food Grocery, and General Merchandise. Column (5) defines sectors at the module level.

Table D2: Response to Shocks: Scalability (Alternative Growth Rates)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Δ size	Δ scope	$\Delta \ln\left(\frac{x}{y}\right)$	Δ size	Δ scope	$\Delta \ln\left(\frac{x}{y}\right)$	Δ size	Δ scope	$\Delta \ln\left(\frac{x}{y}\right)$
$\Delta G \times \ln\left(\frac{x}{y}\right)$	0.0133**	0.0105**	0.0108***	0.0087	0.0066	0.0105***	0.0092	0.0102**	0.0106***
	(0.006)	(0.004)	(0.002)	(0.007)	(0.004)	(0.003)	(0.006)	(0.004)	(0.002)
$\Delta G \times$ scope				0.0195***	0.0168***	0.0017			
				(0.006)	(0.004)	(0.002)			
$\Delta G \times$ size							0.0398***	0.0028	0.0028
							(0.008)	(0.005)	(0.003)
Observations	14,186	14,186	13,488	14,186	14,186	13,488	14,186	14,186	13,488
R-squared	0.144	0.122	0.247	0.145	0.124	0.247	0.146	0.122	0.247
Sector	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: The table reports the results from estimating equation (23) for the period 2006–2015. The dependent variable in Columns (1) and (4) is the change in the (log of) size (revenue) of firm i in sector j; in Columns (2) and (5), the change in scope; and in Columns (3) and (6), the change in scalability. Changes in the dependent variables are computed following Davis and Haltiwanger (1992), as $2(y_t - y_{t-1})/(y_t + y_{t-1})$. The key independent variable is the China import penetration shock from 2006–2015, interacted with the firm's baseline level of scalability or scope in 2006. All specifications include sector fixed effects and controls for firm size, size squared, scope, and scope squared.

Table D3: Response to Shocks: Size and Scope (Alternative Growth Rates)

	(1) Δ size	(2) Δ scope	(3) Δ size	(4) Δ scope
ΔG × size	0.017***	0.001		
	(0.005)	(0.004)		
ΔG × scope			0.014***	0.017***
			(0.005)	(0.003)
Observations	17,138	17,138	17,138	17,138
R-squared	0.163	0.150	0.163	0.152
Sector	Y	Y	Y	Y

Note: The table reports the results of estimating (25). The dependent variable is either the log change in total employment of firm i in sector j from 2006 to 2015, or the change in the number of products or plants. 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 table reports β_{RR} (Columns 1), β_{NR} (Columns 2), β_{RN} (Columns 3), and β_{NN} (Columns 4). All regressions use the NielsenIQ data and the China import penetration shock from 2006 to 2015. Specifications include robust standard errors, sector fixed effects, and firm-level controls: scalability, log size, log size squared, log scope, and log scope squared.

D.2 Multi-Establishment Data

Table D4: Response to Shocks: Size and Scope (Revenue as Size)

	(1) Δ size	(2) Δ scope	(3) Δ size	(4) Δ scope
ΔG× size	0.127***	0.022**		
	(0.033)	(0.011)		
ΔG× scope			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
Sector	Y	Y	Y	Y

Note: The table reports the results of estimating (25). The dependent variable is either the log change in revenue of firm i in sector j from 2006 to 2015, or the change in the number of plants. The table reports β_{RR} (Column 1), β_{NR} (Column 2), β_{RN} (Column 3), and β_{NN} (Column 4). All regressions use the NETS data and the China import penetration shock from 2006 to 2015. All the specifications include sector effects and robust standard errors.

Table D5: Response to Shocks: Size and Scope (Alternative Growth Rates)

	(1) Δ size	(2) Δ scope	(3) Δ size	(4) Δ scope
$\Delta G\times$ size	0.115***	0.123***		
	(0.016)	(0.015)		
$\Delta G\times$ scope			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
Sector	Y	Y	Y	Y

Note: The table reports the results of estimating equation (25) using NETS data. The dependent variable is the log change in revenue (Columns 1 and 3) or the log change in the number of establishments (Columns 2 and 4) for firm i in sector j. Changes are computed as in Davis and Haltiwanger (1992), i.e., $2(y_t - y_{t-1})/(y_t + y_{t-1})$. All regressions include sector fixed effects and robust standard errors.

Table D6: Response to Shocks: Size and Scope (Alternative Firm Definitions)

Employment, firm definition 1

	(1) Δ size	(2) Δ scope	(3) Δ size	(4) Δ scope
$\Delta G\times$ size	0.065***	0.015		
	(0.020)	(0.010)		
$\Delta G\times$ scope			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
Sector	Y	Y	Y	Y

Employment, firm definition 2

	(1) Δ size	(2) Δ scope	(3) Δ size	(4) Δ scope
$\Delta G\times$ size	0.039*	0.013		
	(0.023)	(0.011)		
$\Delta G\times$ scope			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
Sector	Y	Y	Y	Y

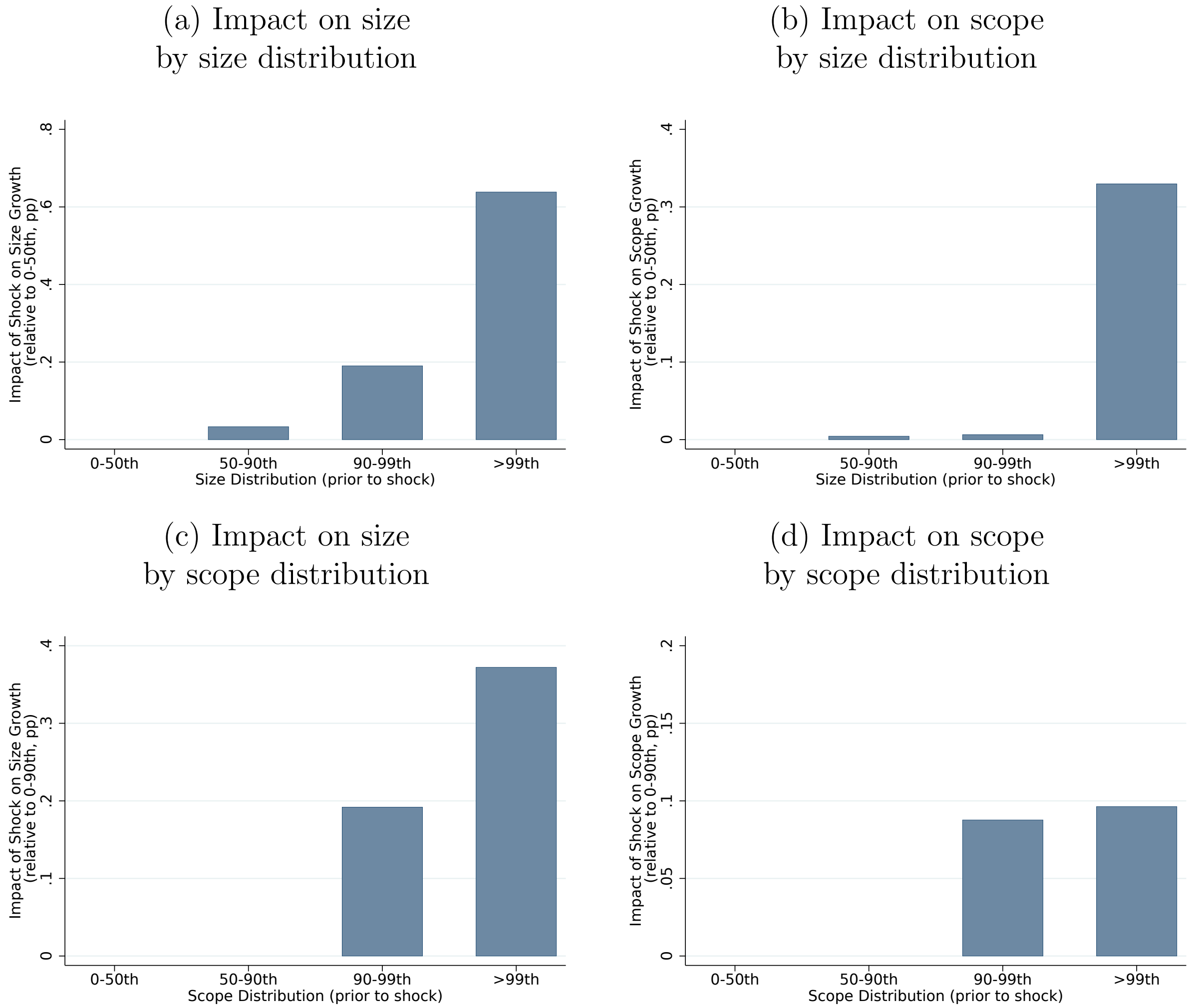
Note: The table reports the results of estimating equation (25) using NETS data with alternative firm definitions. We use alternative definitions of firm: (1) uses information on headquarters in 2015, (2) uses the time-varying definition of headquarters. The dependent variable is the log change in revenue (Columns 1 and 3) or the log change in the number of establishments (Columns 2 and 4) for firm i in sector j. All regressions include sector fixed effects and robust standard errors.

Table D7: Response to Shocks: Size and Scope (No Sector FE)

	(1)	(2)	(3)	(4)
	Δ size	Δ scope	Δ size	Δ scope
ΔG	0.290***	0.045*	0.290***	0.045*
	(0.098)	(0.027)	(0.098)	(0.027)
$\Delta G \times$ size	0.045**	0.025***		
	(0.022)	(0.009)		
$\Delta G \times$ scope			0.118***	0.058***
			(0.042)	(0.021)
Observations	321,114	321,114	321,114	321,114
R-squared	0.066	0.248	0.066	0.248
Sector	N	N	N	N

Note: The table shows the results of estimating equations 25 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. Specifications include robust standard errors.

Figure D1: Response to Shocks: Size and Scope (Non-Parametric Specification)



Note: The figure shows the impact of shocks across different size/scope quartiles. We plot the estimated coefficients $\hat{\beta}_k$ from the following 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 dummy variables for four quantile groups: below the median, 50–90th percentile, 90–99th percentile, and top 1%. For the scope regressions, we combine the first two groups into one dummy, since over 90% of firms operate only one establishment. The dependent variable $\Delta \mathrm{Y}_{ij}$ refers to changes in log size or log scope. Panel (a) shows the impact of shocks across size quantiles, corresponding to the interpretation of β_{RR}. Panels (b), (c), and (d) present analogous results for β_{NR}, β_{NR}, and β_{NN}, respectively. All estimates are based on establishment-level data and use the China import penetration shock as the source of variation.

E Scalability and Knowledge Diffusion: Alternative Measure of Scalabiltiy

Our diffusion strategy also suggests an alternative, forward-looking definition of scalability: namely, by treating scalability as the internal diffusion of a characteristic within a firm. This allows us to construct an *ex-post* measure that captures how widely a newly introduced characteristic is reused by the originating firm. Formally, for a characteristic c introduced by firm i in module m at time t, we define the alternative scalability index 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 then re-estimate our main regression using this alternative measure. Specifically, we estimate:

$$\mathbb{D}_{cmit\tau} = \alpha + \beta\, \widetilde{\mathcal{SI}}_{aimt-1\tau} + \gamma\, scope_{imt} + \lambda_{amt\tau} + \theta_{aim} + \epsilon_{aimt\tau} \tag{45}$$

where, as before, the scalability measure $\widetilde{\mathcal{SI}}$ s aggregated to the attribute level and lagged to mitigate endogeneity concerns. We also control for the total number of products sold by firm i in module m at t, along with a rich set of fixed effects. The results, presented in columns (3) and (4) of Table E1, confirm that the strong positive relationship between scalability and diffusion is robust to this alternative, forward-looking definition of scalability.

Table E1: Scalability and Knowledge Diffusion: Alternative Measure

Diffusion	(1)	(2)
$\widetilde{\mathcal{SI}}$	0.1326***	0.0282***
	(0.000)	(0.001)
scope	-0.0013***	-0.0028***
	(0.000)	(0.001)
Observations	3,269,030	3,183,439
R-squared	0.812	0.913
Firm-Attribute-Module	N	Y
Attribute-Module-Time-Age	Y	Y

Note: The table shows the results of estimating equation 26. The dependent variable is $\mathbb{D}_{cmit\tau}$, which measures the diffusion of characteristic c, introduced in module m by firm i between periods t and $t+\tau$. The independent variable in both columns is the forward-looking scalability measure $\widetilde{\mathcal{SI}}_{cmit-1\tau}$, defined as the share of future products by firm i that reuse characteristic c over the same window. All specifications control for the total number of products sold by firm i in module m at time t.