Size, Trade, Technology and the Division of Labor

Nuno Limão[*], Yang Xu[†]

June 2021

Abstract

We model the implications of the classical ideas that larger markets allow for a finer division of labor and this division feeds back into larger market size. Market size affects specialization due to firm-level increasing returns to scale arising from fixed costs of adopting intermediate-intensive technologies. The impacts are magnified in general equilibrium by an endogenous multiplier—due to input-output linkages in a roundabout structure—and a selection effect due to heterogeneous fundamental productivity and entry costs.

Market size expansions imply (i) larger real income gains than under fixed specialization; (ii) an increase in the aggregate variable cost share for intermediates and a decrease for labor; (iii) increased concentration; (iv) increased average productivity for survivors; and (v) an increase in the intermediate trade share. We derive similar results for intermediate productivity improvements. The effects in (ii)-(v) are absent in a similar model with exogenous specialization.

In a calibration to U.S. manufacturing in 1987-2007 we isolate trade and intermediate productivity shocks and quantify their effects. Trade cost reductions increased effective market size by 7 log points (lp) and generated (i) a real income gain 1.4 times higher than under exogenous specialization; (ii) increases in the intermediate share in production and trade of 2 lp and a reduction in the labor share of value added of similar magnitude. Two counterfactuals highlight the importance of industrial and trade policy. First, a tax that induces firms to specialize increases real income; so the initial equilibrium is inefficient. Second, an increase in trade costs of 16 lp—similar to the recent trade war—reduces market size and real income substantially: almost half way to trade autarky.

[*]Department of Economics, University of Maryland & NBER. Email: limao@econ.umd.edu. I thank Kerem Cosar, Allan Drazen, Eunhee Lee, Felipe Saffie, David Riker and participants at the 2018 Lisbon Theory Conference; 2021 WAITS and AEA Conferences.

[†]School of Economics & Wang Yanan Institute for Studies in Economics (WISE), Xiamen University. Email: xuya0005@e.ntu.edu.sg. I thank Eunhee Lee, Felipe Saffie and other committee faculty for comments on the 2017 dissertation chapter this paper is based on.

1 Introduction

> "(...) *the division of labour depends upon the extent of the market, but the extent of the market also depends upon the division of labour. In this circumstance lies the possibility of economic progress.*" (Young, 1928, p. 539)

In "Increasing Returns and Economic Progress" Allyn Young builds on the core idea in Smith (1776) that labor specialization is limited by the size of the market by emphasizing a feedback effect. Young argues that labor specialization reduces production costs mainly due to a roundabout production structure where the sources of increasing returns go beyond the individual firm and require "industrial operations be seen as an interrelated whole" (p. 539). We formalize this idea using a model with input-output linkages where heterogeneous firms can increase their specialization by adopting intermediate intensive technologies thus reducing their own product's cost and those of other firms using it as an input. We show that increases in market size and intermediate productivity lead to higher aggregate income and firm concentration, and changes in factor shares in production consistent with recent evidence. International trade is a key determinant of market size and thus of these outcomes, as our reduced form evidence and subsequent quantification for U.S. manufacturing in 1987-2007 illustrates.

We model specialization as a form of directed technical change (Acemoglu, 2002): firms can invest in technologies with different intermediate-labor intensity. Firms initially draw a technology with constant returns to labor and heterogeneous productivity to manufacture a variety sold under monopolistic competition. After learning their productivity firms can invest in more specialized technologies that use intermediate inputs sourced from others, which implies a lower share of labor in their production. This is isomorphic to a setting where firms concentrate labor on tasks where it is most productive and outsource the rest. A firm specializes if the productivity advantage of using intermediates, henceforth the *specialization premium*, is sufficiently high relative to the adoption cost. We assume fixed costs of adoption so the model predicts that larger firms are more specialized and thus have lower labor share, which is consistent with the data (Autor et al., 2020).

The feedback mechanism that amplifies market size and technology shocks is the endogenous specialization premium and multiplier. This premium is common across firms and decreasing in the effective relative price of intermediates. An increase in market size lowers the relative price of intermediates and this effect is amplified due to the input-output linkages. The resulting aggregate input-output multiplier is central in other models but constant. We show this multiplier is determined by the aggregate cost share of intermediates and thus exogenous and independent of market size in models where firms use a common intermediate technology. In contrast, our model features an endogenous multiplier that is increasing in market size because of new adoption and sales re-allocation towards more specialized firms due to entry selection.

We derive a condition for a unique equilibrium and analytical comparative statics in a single sector economy with free entry. In equilibrium the initially most productive firms become more specialized; and increases in market size (via changes in labor endowment or trade liberalization) imply (i) larger real income gains relative to fixed specialization; (ii) an increase in the aggregate variable cost share for intermediates and a decrease for labor; (iii) increased firm selection; (iv) increased concentration in the profit and sales distributions; (v) increased average productivity over surviving firms and (vi) an increase in the intermediate trade share. A similar model with exogenous specialization does not generate any of the effects in (ii)-(vi), and is thus unable to explain either the recent increases in intermediate shares described above or the increases in concentration documented in Bajgar et al. (2019). Firms do not internalize the aggregate benefits of specialization so the market equilibrium is generally inefficient, which raises the value of policies that promote specialization—either directly (e.g. adoption incentives) or by expanding market size (e.g. trade policy).

We provide a calibration to U.S. manufacturing in 1987-2007 to unpack different shocks and mechanisms as well as perform counterfactual policy changes. In section 2 we provide motivating evidence that guides the calibration. We document an increase in the intermediate share in costs of almost 5 percentage points, which is highly correlated with trade openness and the relative labor/intermediate price. The intermediate/labor intensity within industries is increasing in market size—consistent with the modelled economies of scale in specialization.

The calibration yields a reduction in the marginal trade cost and an improvement in intermediate productivity; these changes allow us to match key targeted moments exactly, e.g. the increases in intermediate intensity and relative labor/intermediate price, and generate outcomes consistent with untargeted moments of the data, e.g. changes in real value added/worker and sales distribution. The improved intermediate productivity is equivalent to a lower marginal cost of using intermediates (both domestic and imported), hence it has a direct effect on the specialization premium and accounts for a substantial fraction of its increase.[1]

We then isolate and quantify the impact of trade costs. The marginal cost reduction in 1987-2007 implies an effective market size increase of over 7 log points (lp) and thus has significant effects including: (i) real income gains larger than under no specialization (4.6 times) or fixed specialization (1.4 times); (ii) an increase in the intermediate share in production and trade of 2 lp and a reduction in the labor share of value added of similar magnitude; and (iii) substantial increases in the fraction of firms specializing.

We conduct two counterfactual experiments to highlight the importance of industrial and trade policy. First, an industrial policy (e.g. a tax/subsidy) that induces all firms to specialize would increase real income; so the initial equilibrium is inefficient. That income increase is significant if the policy were implemented in 1987 but negligible in 2007 since by the latter period trade and technical change had induced sufficient specialization. Second, the impact in the 2007 economy of an increase in trade costs of 16 lp—similar to the recent trade war—reduces market size and real income substantially: almost half way to trade autarky.

We discuss related literature in section 1.1 and then present motivation evidence on changes in aggregate production specialization and imported intermediates in section 2. In sections 3, 4 and 5 we

[1] We can interpret this as a reduction in the marginal cost of outsourcing, which is central in models such as Grossman and Rossi-Hansberg (2008), and a source of production fragmentation towards domestic and foreign sources (Fort, 2016).

respectively develop the framework, characterize the equilibrium and derive size and technology comparative statics in a closed economy. In section 6 we extend the results to an open economy and use it as the basis for the quantification in section 7.

1.1 Related Literature

Our focus on intermediate adoption as a source of firm specialization, and the role of trade in spurring it, is motivated by several facts. First, intermediates are a large fraction of international trade (Johnson and Noguera, 2012, 2017); central to understanding its growth (Yi, 2003, 2010; Hummels et al., 2001); and imported intermediates increase firm productivity (Amiti and Konings, 2007; Halpern et al., 2015) and lower costs (De Loecker et al., 2016). Second, imported intermediates are negatively correlated with the labor share in value added in U.S. manufacturing (Elsby et al., 2013).[2] This labor share decline is present in other countries and its potential causes include a substitution towards capital (Karabarbounis and Neiman, 2014) or increasing profits due to higher concentration (Barkai, 2020; Autor et al., 2020). The labor share declines may also be caused by our specialization channel, which entails increases in the intermediate cost share in manufacturing that we document in section 2.

A growing literature on endogenous production networks provides interesting predictions about firm-level sourcing. Some of this research assumes a constant expenditure share of intermediates (e.g. Antras et al., 2017; Fieler et al., 2018; Oberfield, 2018) and thus can't explain increases in specialization. Other work is numerically consistent with specialization but not amenable to analytical solutions (e.g. Acemoglu and Azar, 2020; Dhyne et al., 2021). By using a roundabout structure and adoption decisions we provide a tractable framework that focuses on general equilibrium effects of market size and the implications for other outcomes such as documented increases in concentration in sales and declines in labor share (Autor et al., 2020).

There are three pervasive features of modern industries: productivity heterogeneity, input-output linkages and technical change. We show the interaction of all three is important in understanding the effect of globalization on the gains from trade and productivity; moreover we bridge models that study them separately. If the market is small enough then firms use only labor and the equilibrium in our model is similar to Melitz (2003). If the market is large enough then all firms adopt the most specialized technology in equilibrium, which has the constant intermediate share typically *assumed* in models with linkages (Ethier, 1982; Krugman and Venables, 1995; Eaton and Kortum, 2002). Income gains from international trade are larger in models with intermediate linkages (Costinot and Rodríguez-Clare, 2014; Caliendo and Parro, 2015); these gains are further magnified with multiple stages of production (Melitz and Redding, 2014) or endogenous choice of imported inputs (Ramanarayanan, 2020); and there is evidence that imported intermediates lower consumer prices (Blaum et al., 2018). The gains from trade due to intermediates are also larger in our model and magnified by endogenous adoption of both domestic

[2] They are also negatively correlated with relative demand for production labor in U.S. manufacturing in 1972-1990 (Feenstra and Hanson, 1999).

and imported intermediates so trade affects outsourcing not just offshoring.[3] Trade liberalization induces some firms to improve productivity (Lileeva and Trefler, 2010; Bustos, 2011; Bloom et al., 2015). In our model, liberalization increases labor productivity for the most productive firms but can lower it for the least productive (due to downgrading from competition)—so it is consistent with the increased productivity dispersion found in Berlingieri et al. (2017). Aggregate productivity (both labor and TFPQ) increases with liberalization due to selection and re-allocation effects.

Finally, our work relates to the literature on the implications of market size and scale economies for development emphasized by Young (1928). Rosenstein-Rodan (1943) argues that adopting increasing returns technologies may require coordinated investments across sectors in underdeveloped economies. Murphy et al. (1989) formalize this idea and show pecuniary externalities can create multiple Pareto ranked equilibria. Our model also features pecuniary externalities—via the love-of-variety production (Ethier, 1982)—and endogenous specialization can generate multiple equilibria with real income increasing in specialization. This potential under-adoption inefficiency implies a role for coordinating investments across firms within a sector. The calibration has a unique equilibrium; however it still reflects under-adoption that is significant in 1987 but negligible by 2007—by the latter period trade and technical change had induced sufficient specialization.

2 Motivation Evidence

We present evidence of increases in production specialization in manufacturing and its relationship with market size and trade. We focus on U.S. data for 1987-2007 here and provide related data for 67 countries in 1997-2007 in Appendix A.

2.1 Aggregate Facts

We use the NBER-CES industry aggregation of U.S. firm based measures of sales, input costs and prices covering 459 SIC-4 industries from 1987 to 2007.

A key summary statistic is the average intermediate share, $\bar{\alpha}^m$, which is defined as the sales weighted average of intermediate cost shares across firms (indexed by φ) in industry m:

$$\bar{\alpha}^m \equiv \frac{\Sigma_\varphi \alpha_\varphi^m Y_\varphi^m}{Y^m} = \frac{I^m}{C^m}.$$

Sales, Y_φ^m, are proportional to variable costs in our model so we construct $\bar{\alpha}^m$ as the industry intermediate expenditure, I^m, relative to all variable costs, C^m (the sum of non-energy materials, I^m, energy, labor and investment). We use the industry cost share in 1997 ($\bar{\alpha} = \Sigma_m \frac{C_{97}^m}{C_{97}} \bar{\alpha}^m$) to aggregate them, so variation over time reflects within industry changes, not composition. In Figure 1 we see an increase of almost 5

[3] Kee and Tang (2016) find globalization lowered Chinese materials prices and increased domestic sourcing by Chinese firms.

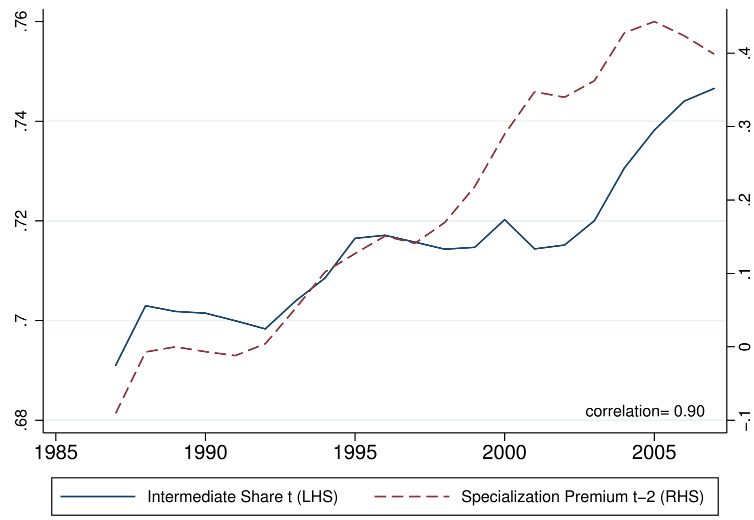


Figure 1: Intermediate cost share and specialization premium: 1987–2007

percentage points in $\bar{\alpha}$—half of it since China's WTO entry.[4]

In the model, firm adoption of intermediate-intensive technologies increases $\bar{\alpha}$; one of its determinants is the relative price of labor to intermediates, w_t^m/P_t^m. In Figure 1 we examine if this correlation is present by constructing a change in relative costs relative to 1987, $S_t^m \equiv \ln\left(\frac{w_t^m/P_t^m}{w_{87}^m/P_{87}^m}\right)$, and its cost weighted average, $\bar{S}_t \equiv \Sigma_m \frac{C_{97}^m}{C_{97}} S_t^m$. Unanticipated shocks to the premium that induce firms to invest at $t-n$ can take time to be reflected in increased intermediate intensity thus in Figure 1 we plot $\bar{S}_{t-2}$, and find it has a correlation of 0.9 with $\bar{\alpha}_t$. [5]

The annual growth trend of S_t^m between 1987 and 2007 is 2.5 lp; our subsequent calibration accounts for it in terms of increased productivity and trade cost reductions. Here we show this premium and the intermediate share are strongly correlated with imports. We measure import penetration in industry m by its import value relative to domestic absorption: $\lambda_t^m \equiv \frac{\text{Imports}_t^m}{dom_t^m}$, where $dom_t^m \equiv$Sales$_t^m$-Exports$_t^m$+Imports$_t^m$ and its aggregate counterpart at fixed weights is $\bar{\lambda}_t \equiv \Sigma_m \frac{dom_{97}^m}{dom_{97}} \lambda_t^m$. In the left panel of Figure 2 we see a 10 percentage point increase in $\bar{\lambda}_t$ between 1990 and 2007 (we start in 1990 due to industry data availability). The increase occurred in a period with several liberalization episodes. NAFTA was signed in 1993 and the WTO was established in 1995, and both lead to substantial reductions in both U.S. applied trade barriers and uncertainty (e.g. after China's WTO accession). The correlation of import penetration, $\bar{\lambda}_t$, is over 0.9 with respect to both the premium, $\bar{S}_t$, and intermediate share, $\bar{\alpha}_t^m$, which suggests imports affected the incentive to adopt intermediates.

[4]This measure assumes that existing capital captures fixed costs, which explains why the intermediate and labor shares are higher than when total costs are used. The calibration considers alternative measures of investment and capital use in variable costs.

[5]It is slightly lower if we use $\bar{S}_{t-1}$ and 0.77 for $\bar{S}_t$. Figure 1 raises the question of whether the relationship is mechanical or simply reflects a simple CES production—where relative input shares are proportional to their relative prices. If so then the correlation should be similarly large for other periods. However, we find it is much lower, only 0.1, between year 1967 and 1987. Below we also find differential impacts of S_t^m on relative industry shares consistent with our model.

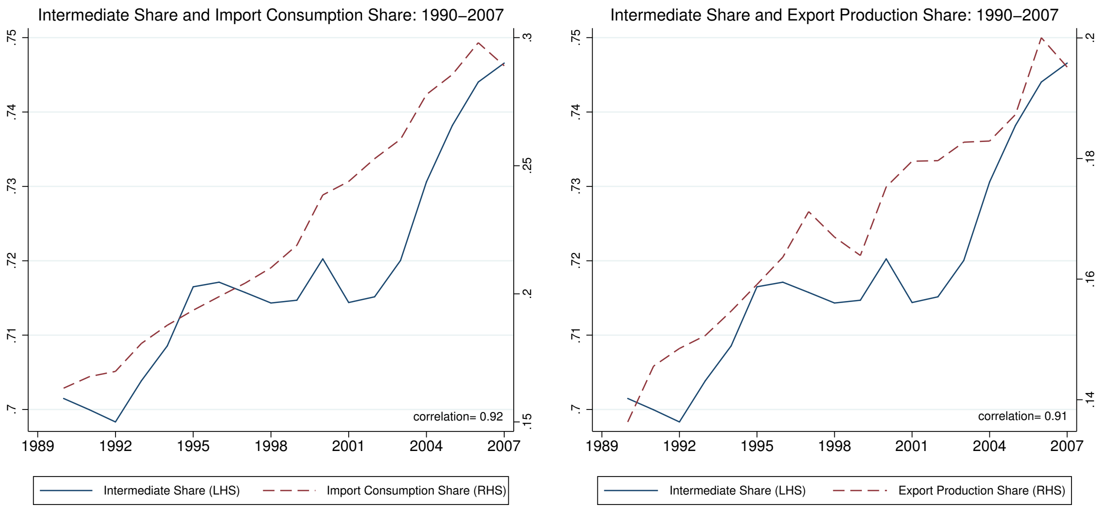


Figure 2: Import consumption share, export production share, and intermediate cost share

These trade agreements also reduced foreign barriers, which along with other trade cost reductions increased the effective market size for U.S. producers. This is reflected in the increased export share of production shown in the right panel of Figure 2.[6] There were increases in the wake of NAFTA and the WTO, followed by a stagnant share, and new increases after China's WTO accession. The correlation between this export intensity and the intermediate share is over 0.9.

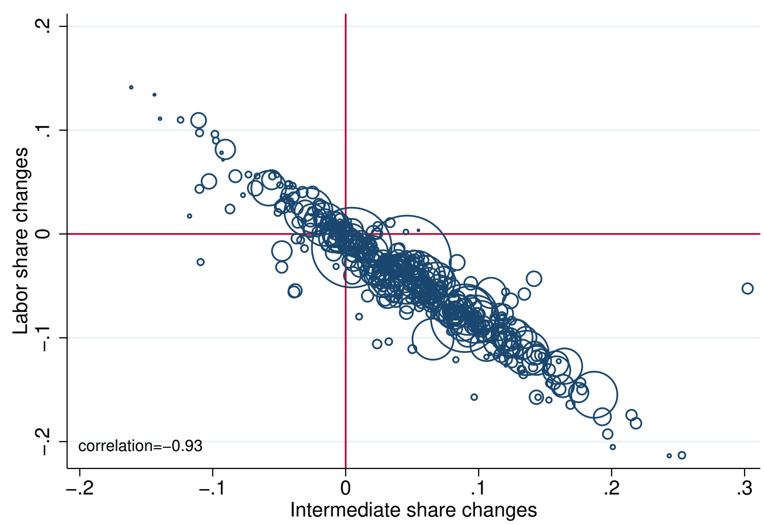


Figure 3: Intermediate and labor cost shares: 1987–2007 change

The model has one type of labor so we aggregate production and non-production pay to compute total labor expenditures and denote their share in industry cost by $\bar{\alpha}_w^m$. The corresponding aggregate

[6]This measure is calculated as $\frac{\text{Exports}_t^m}{Sales_t^m}$ for each industry and the corresponding aggregate is $\Sigma_m \frac{Sales_{97}^m}{Sales_{97}} \frac{\text{Exports}_t^m}{Sales_t^m}$.

share, $\bar{\alpha}_w = \Sigma_m \frac{C^m}{C} \bar{\alpha}^m_w$, fell from 0.23 in 1987 to 0.16 in 2007. The correlation between the labor and intermediate share, $\bar{\alpha}_w$ and $\bar{\alpha}$, is -0.97, which is consistent with the model's focus on the substitution between these inputs.[7]

2.2 Industry Panel Evidence

Disaggregated industry data provides additional evidence for the model's mechanisms and implications. First, changes in shares are also strongly negatively correlated within industries: the correlation is -0.93.[8] The relation is shown in Figure 3 where the change in the intermediate share, $\Delta\bar{\alpha}^m$, the x-axis, is positive for most industries and most where $\bar{\alpha}^m$ increased also experienced declines in $\bar{\alpha}^m_w$.[9] Importantly, there is considerable variation in the magnitude of $\Delta\bar{\alpha}^m$ across industries that we can exploit. There is also considerable variation within the specialization premium; the standard deviation of S^m_t over time is 0.13 for the typical industry (it ranges from 0.03 to .73).

In Table 1 we examine the impact of two potential determinants of these relative cost shares, $\ln \frac{\bar{\alpha}^m_t}{\bar{\alpha}^m_{w,t}}$, conditional on industry and time effects. We focus on the specialization premium and market size (measured by the value of shipments). We lag each of these to allow for time to build and also instrument each to minimize endogeneity concerns.[10] The first column shows that specialization increases with the premium, $\varepsilon_S \geq 0$, but only significantly so in concentrated industries, $\Delta\varepsilon_S > 0$.[11] Moreover, specialization is increasing in market size, $\varepsilon_Y > 0$.

$$\ln \frac{\bar{\alpha}^m_t}{\bar{\alpha}^m_{w,t}} = (\varepsilon_S + \Delta\varepsilon_S \times Conc^m)\, S^m_{t-1} + (\varepsilon_Y + \Delta\varepsilon_Y \times Conc^m) \ln Y^m_{t-1} + a_{t,m} + u_{tm}$$

This evidence rejects simple production structures and provides some support for the mechanisms in the model. First, the findings reject a standard production model (Cobb-Douglas or CES) where relative intermediate/labor shares are common across firms in an industry because in those cases the relative industry share would be independent of market size. The specialization bias from size suggests a composition effect: large firms are more specialized *and* their sales share is increasing in market size. This channel is present in our model and magnified by endogenous specialization. Table 1 is also consistent with endoge-

[7] In the model the labor share of *variable* costs is $1 - \bar{\alpha}^m$ so the two have a correlation of -1, but that need not be true in the data for two reasons. First, there are other inputs included in costs, such as energy and investment, so that $\bar{\alpha}$ and $\bar{\alpha}_w$ do not add up to unity, therefore $\bar{\alpha}$ can in principle have any correlation with $\bar{\alpha}_w$. Second, in the model labor is also used for fixed costs and to the extent that in practice they do use labor (in addition to investment expenditures) then they are reflected in $\bar{\alpha}_w$. This attenuates the correlation of $\bar{\alpha}$ and $\bar{\alpha}_w$ because the model predicts that adoption of intermediate technology increases the usage of fixed cost labor.

[8] We find a similar relationship with labor cost in value-added but with a lower correlation of -0.38.

[9] Industries where $\bar{\alpha}^m_w$ increased were fewer and smaller on average as illustrated by the smaller circles (proportional to the weights, C^m_{97}/C_{97}).

[10] Specifically, let M_2 denote a SIC-2 sector so for each $m \in M_2$ we construct an instrument for x^m based on the average of x in that period over all $m' \in M_2 \backslash m$. The first stage has high explanatory power and the instruments are not weak according to alternative F-statistics.

[11] Here we use the share of sales by top 20 but the result is robust to alternative concentration measures such as the HerfindahlHirschman Index for the 50 largest firms (HHI50). It also holds for industries with relatively higher average sales per firm.

Table 1: Annual intermediate/labor cost shares (log): 1987-2007, SIC-4 manufacturing

	IV	IV	IV
ln(Wage/Interm. Price) (lag)	0.004	0.083	0.057
	[0.059]	[0.050]	[0.050]
ln(Wage/Interm. Price) (lag)	0.296		0.247
× Top 20 share '87	[0.052]		[0.058]
ln(Size) (lag)		0.128	0.124
		[0.028]	[0.029]
ln(Size) (lag)			0.254
× Top 20 share '87			[0.071]
Observations	9,389	9,389	9,389

Notes: Robust standard errors in brakets, clustered at SIC-2 by year: the lowest level of variation of instruments. All specifications include year and SIC 4-digit fixed effects.

IV: Two-stage least square instrumenting specialization (or size) in SIC 4 by the value in the other industries of specialization (or size) sharing the same SIC-2 (and the respective interactions with the Top 20 variable when included).

Source: NBER CES Manufacturing database sic 1987 version and US Census of Manufacturing. Intermediate cost share $= \frac{\text{Non-energy material expenditure}}{\text{Total costs}}$, Total costs $=$ Intermediate $+$ Energy $+$ Investment $+$ Labor pay. $\ln\left(\frac{\text{Wage}}{\text{Intermediates price}}\right)$ calculated by SIC 4 and normalized by its respective 1987 value. Size: value of shipments. Top 20 Share 1987: Sales share of largest 20 firms in each SIC 4, demeaned.

nous specialization in specific industries. If all firms in an industry are too small to specialize then there is a negligible effect of the specialization premium provided production is Cobb-Douglas (as we will assume). In industries with high concentration it is more likely that at least some firms are sufficiently large to respond to increases in the premium (and/or market size) and specialize. The specialization premium elasticity is 0.07 higher for an industry with concentration 1 sd higher than the mean. In the calibration we find a similar relationship even though it is not targeted.

3 Environment and Firm Decisions

We model consumer preferences and firm technology in the absence of adoption similarly to Melitz (2003) and show how endogenous specialization affects aggregate real income, the labor share in production, profit and sales concentrations. We also show that endogenous specialization magnifies the importance of economies of scale and generates a feedback effect between productivity and size.

We develop the key results in a simple version of the model with a single industry in a closed economy. We then show that size increases are mostly isomorphic to trade cost reductions.

3.1 Preferences and Technology

There are L consumers with identical preferences over the set of differentiated varieties $\omega \in \Omega$ with constant elasticity of substitution $\sigma > 1$. This yields the standard demand function

$$q(\omega) = EP^{\sigma-1}p(\omega)^{-\sigma}, \tag{1}$$

where $p(\omega)$ is the variety price, E is consumer expenditure on all varieties, derived below, and P the associated price index given by $P = \left(\int_{\omega\in\Omega} p(\omega)^{1-\sigma} d\omega\right)^{1/(1-\sigma)}$.

After paying a sunk cost, each firm obtains a blueprint for a single variety ω identified by a core productivity φ drawn from a distribution $G(\varphi)$. Conditional on φ, the firm chooses between production technologies by paying f_i units of labor to access a CRS process with non-labor input cost shares $\alpha_i \in \{0, ..., \alpha_n\}$; where $1 - \alpha_n > 0$ is the minimum labor variable cost share constraint. We capture the gains from specialized inputs by modelling the non-labor inputs as a CES bundle of intermediates produced by other firms, similar to Ethier (1982). We assume a common CES aggregator in production and consumption so they have the same price index, P. The unit cost functions for alternative technologies are:

$$c_i(\varphi) = \frac{w^{1-\alpha_i}}{\varphi}\left(\frac{P}{\phi}\right)^{\alpha_i}, \quad i = 0, \ldots, n \geq 1; \tag{2}$$

where w is the wage rate and, when intermediates are adopted, we allow for a fixed productivity change parameter, ϕ, so P/ϕ is the effective unit price of intermediates.[12,13]

The degree of **firm specialization** is defined by the share α_i it chooses, which depends in part on the following specialization premium.

Definition 1. ***(Specialization Premium)***

The premium, s_I, from a technology with more specialized inputs $\delta_I \equiv \alpha_{i+I} - \alpha_i > 0$ is

$$s_I \equiv \frac{c_i(\varphi)}{c_{i+I}(\varphi)} = \left(\frac{w}{P/\phi}\right)^{\delta_I}. \tag{3}$$

The specialization premium is increasing in the relative effective cost of labor to intermediates, $\frac{w}{P/\phi}$, and independent of firm core productivity. Moreover, a firm only adopts a more specialized technology if $s_I > 1$, or equivalently if wages are sufficiently high, $w > P/\phi$. This condition is independent of the degree of specialization.

Our framework is isomorphic to one where the firm pays f_i and uses a fraction of labor to produce inputs in house using a linear technology and the remaining fraction, $1 - \alpha$, to assemble it. In appendix B, we map our baseline cost structure to one with a two-stage production process, where firms can pay f_i

[12] More precisely, the cost function is $c_i(\varphi) = \frac{w^{1-\alpha_i}P^{\alpha_i}}{\varphi\phi_i(1-\alpha_i)^{1-\alpha_i}\alpha_i^{\alpha_i}}$ and we assume $\phi_i = \frac{\phi^{\alpha_i}}{(1-\alpha_i)^{1-\alpha_i}\alpha_i^{\alpha_i}}$. Under this specification ϕ_i is first increasing than decreasing in α_i.

[13] Alternatively, the inverse of ϕ can be interpreted as the impediments to specialization in an iceberg fashion, such as communication costs between firms emphasized by Becker and Murphy (1992).

to replace some of their first stage input production with outsourced inputs. Doing so reallocates workers towards the more productive inputs while lowering the overall labor share.

3.2 Entry and Technology Decisions

The operating profit for firm φ using technology i is defined by $\pi_i(\varphi) = [p_i(\varphi) - c_i(\varphi)]\, q_i(\varphi)$. This firm faces the isoelastic demand from consumers and in a model without intermediates the profit maximizing price under monopolistic competition is given by the constant markup over marginal cost, $\tilde{p}_i(\varphi) = \frac{\sigma}{\sigma-1} c_i(\varphi)$. This is also the optimal price charged to firms in our setting because the input aggregator is CES with elasticity σ. Therefore the aggregate demand faced by firm φ is

$$\tilde{q}_i(\varphi) = XP^{\sigma-1}\left[p_i(\varphi)\right]^{-\sigma}, \tag{4}$$

where X is the total expenditure of consumers and firms, which is taken as given by the firm and derived below. Replacing $\tilde{p}$ and $\tilde{q}$ in π_i we obtain the maximized operating profit for any i

$$\tilde{\pi}_i(\varphi) = \tilde{\sigma} XP^{\sigma-1}\left[c_i(\varphi)\right]^{1-\sigma} \tag{5}$$

where $\tilde{\sigma} \equiv \frac{1}{\sigma-1}\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma}$. We define the equilibrium profit as a function of firm productivity after it chooses its technology, where the firm pays $f_i > 0$ to produce using technology i, or zero if it decides not to enter.

$$\tilde{\pi}(\varphi) \equiv \arg_i \max\left\{0, \tilde{\pi}_i(\varphi) - wf_i\right\} \tag{6}$$

Assumption 1. ***(Technology)***

1. *There is a continuum of firms with heterogeneous core productivity drawn from $G(\varphi)$ with support $[\varphi_l, \infty)$;*

2. *Specialization technology has constant share increments: $\Delta\alpha_{i+1} \equiv \alpha_{i+1} - \alpha_i = \delta$ thus $s_I = (s_1)^I$ and increasing costs: $\Delta f_{i+1}/f_i = \hat{f} > 1$.*

The equilibrium is characterized by entry and adoption thresholds, so for partial specialization we require heterogeneous firms; the continuum and unbounded productivity assumptions are made to simplify the analysis. The share and cost assumptions are used to show that specialization increases with core productivity.

The solution to the production and entry decisions using technology i is characterized by the zero profit cutoff productivity derived from setting $\tilde{\pi}_i(\bar{\varphi}_{i,e}) = wf_i$, which is

$$\left(\bar{\varphi}_{i,e}\right)^{\sigma-1} = \frac{wf_i}{\tilde{\sigma}X} \cdot P^{1-\sigma} \cdot \left(ws_1^{-i}\right)^{\sigma-1}. \tag{7}$$

All else equal entry is easier, and so the cutoff lower, if fixed costs can be spread over a larger market (the first term), competitor prices are higher (the second) and marginal costs are lower (the third). If entrants use only labor then $i = 0$ and the third term reflects only the wage, but if there is adoption then the marginal cost relative to the labor technology is reduced by the specialization factor s_1^{-i}, which is the only difference in the cutoff relative to a standard model without adoption.

A firm that is indifferent between technologies $i+1$ and i, has productivity $\bar{\varphi}_{i+1}$ defined by the equality between the operating profit increase of two adjacent technologies and the differential adoption cost, $\Delta\tilde{\pi}_{i+1}(\bar{\varphi}_{i+1}) = w\Delta f_{i+1}$, which yields:

$$(\bar{\varphi}_{i+1})^{\sigma-1} = S_{i+1}^{-1} \cdot (\bar{\varphi}_{i,e})^{\sigma-1} . \tag{8}$$

The expression is proportional to the entry cutoff expression if the entrant adopted i by a factor representing the relative gain in profits relative to fixed costs from switching technologies:

$$S_{i+1} \equiv \frac{\tilde{\pi}_{i+1}/\tilde{\pi}_i - 1}{\Delta f_{i+1}/f_i} = \frac{s_1^{\sigma-1} - 1}{\hat{f}} = S_1$$

Under the second part of technology assumption 1, S_{i+1} is the same for all i and thus the relative cutoff across any two technologies is simply $(\bar{\varphi}_{i+1}/\bar{\varphi}_i)^{\sigma-1} = \left(1+\hat{f}\right)/s_1^{\sigma-1}$. Therefore we can write the adoption cutoff for any technology $i+I$ as

$$\left(\frac{\bar{\varphi}_{i+I}}{\bar{\varphi}_{i,e}}\right)^{\sigma-1} = S_1^{-1} \cdot \left[\left(1+\hat{f}\right)/s_1^{\sigma-1}\right]^{I-1} . \tag{9}$$

Solving for the cutoff $i+I$ we obtain the function $\varphi_I \equiv \varphi\left(\bar{\varphi}_{i,e}, s_1, \hat{f}, I\right)$ for $I \geq 1$.

In Proposition 1 we characterize how firm specialization depends on the premium, fixed costs and productivity. We say that specialization is increasing in productivity if $\bar{\varphi}_{i+1} > \bar{\varphi}_i$ for all i and adopt the convention that an indifferent firm adopts the higher technology.

Proposition 1. ***(Heterogeneous specialization)***

Under technology assumption 1, specialization:

1. *occurs iff there is a premium:* $s_1 > 1$*;*
2. *is increasing in productivity,* φ*, iff* $s_1^{\sigma-1} < 1+\hat{f}$ *and heterogeneous across firms for* $(s_1)^{\sigma-1} \in (1, 1+\hat{f})$*;*
3. *is common at* $\alpha_n = \alpha$ *iff* $(s_1)^{\sigma-1} \geq 1+\hat{f}$.

The proof is in Appendix C.1. Part 1 shows it is necessary for $s_1 > 1$, so specialization lowers marginal costs enough to offset the fixed costs, otherwise no firms specialize; the condition is also sufficient for specialization because of the existence of sufficiently high productivity firms. Part 2 shows the condition for sorting, $\bar{\varphi}_{i+1} > \bar{\varphi}_i$, namely that the operating profit growth is no higher than that of fixed costs, $S_1 < 1 \Leftrightarrow s_1^{\sigma-1} - 1 < \hat{f}$. Moreover, when $S_1 < 1$ we obtain $\bar{\varphi}_{i+1} > \bar{\varphi}_{i,e}$ so the marginal producer does not choose the most specialized technology and thus there is heterogeneity in adoption. Part 3 shows that if $S_1 > 1$ then all firms gain from any incremental specialization and thus adopt the maximum possible.

The proposition illustrates how the specialization premium spans three types of equilibria: no specialization ($S_1 \leq 0$), full specialization as in models where all firms must use a fixed share of intermediates ($S_1 \geq 1$), and partial specialization ($0 < S_1 < 1$), which is the novel range we focus on.

4 General Equilibrium

We aggregate the consumer and firm decisions and solve for the general equilibrium in a closed economy. We use the wage as the numeraire, $w = 1$, and define the equilibrium as follows.

Definition 2. ***(General Equilibrium)***

1. Consumers choose the quantity of each variety and labor supply to maximize their utility subject to their budget constraint taking prices as given.

2. Firms with productivity φ choose a production technology i to maximize profits taking aggregate prices (P), expenditure (X), costs (f_i), and the specialization premium (s_i) as given.

3. Goods and labor markets clear.

4. There is free entry by ex-ante identical potential producers to obtain a blueprint at cost f_E.

4.1 Endogenous Specialization Mechanisms and Number of Firms

Endogenous specialization features two novel mechanisms absent from standard models (e.g. Melitz, 2003, or variants without endogenous input adoption). First, an **endogenous multiplier effect**: an increase in market size generates adoption, which further reduces input prices via intermediates and amplifies the initial shock. Second, a **selection effect**: the marginal entrant depends on market size and thus economic scale impacts the equilibrium distribution of firm performance (size, labor demand, etc). We derive the multiplier as a function of productivity thresholds and the specialization premium, and then use free entry to pin down the entry threshold. After that we characterize the equilibria as a function of the specialization premium and price index. Subsequently we show that both effects amplify the impacts of economic scale on real income and generate novel predictions about the interaction of size and technology.

4.1.1 Aggregate Expenditure and Multiplier

We aggregate consumer and firm decisions to derive total expenditure, X, and the associated endogenous multiplier due to intermediate linkages.

Consumers do not value leisure and so supply their unit of labor at any wage. Thus aggregate labor income is L, as is consumer expenditure, E, because there are no aggregate profits under free entry.

A fraction α_i of a firm's variable costs is spent on intermediates, which under CES and monopolistic competition is simply a fraction $\frac{\sigma-1}{\sigma}\alpha_i$ of its sales. Thus total expenditure is

$$X = L + \frac{\sigma-1}{\sigma}\sum_{i=0}^{n}\alpha_i Y_i = L \cdot \underbrace{\left(1-\frac{\sigma-1}{\sigma}\bar{\alpha}\right)^{-1}}_{\text{Multiplier: } \bar{a}}, \tag{10}$$

where we used $E = L$ and Y_i represents aggregate sales of all firms using the i^{th} technology. The second equality uses the market clearing condition $X = Y \equiv \sum_{i=0}^{n} Y_i$. The multiplier, henceforth $\bar{a}$, is unity without specialization and larger otherwise. In standard models with intermediate linkages the technology is fixed and so is the multiplier. A contribution of this paper is deriving how this endogenous multiplier depends on the aggregate intermediate cost share and its determinants

$$\bar{\alpha} \equiv \sum_{i=0}^{n}\alpha_i \frac{Y_i}{Y}. \tag{11}$$

Below we provide an analytical solution for this share. For now we note that the CES structure implies that Y_i/Y is a function of the relative price index for each group of firms and thus depends on the cost functions and cutoff vector, $\{\bar{\varphi}_{\mathbf{i}\geq\mathbf{e}}\}$. Moreover, the technology assumptions allow us to write those relative prices as a function of the specialization premium s_1^i, and all cutoffs as a function of the entry one, $\bar{\varphi}_{i,e}$. Thus we denote the solution as a function of two endogenous variables, s_1 and $\bar{\varphi}_{i,e}$, and underlying parameters.

$$\tilde{\alpha} \equiv \bar{\alpha}\left(s_1, \hat{f}, I, \bar{\varphi}_{i,e}, G\right) = \sum_{i=0}^{n}\alpha_i \frac{\left(s_1^{i-e}\frac{\hat{\varphi}_i}{\hat{\varphi}}\right)^{\sigma-1}}{\sum_{i=0}^{n}\left(s_1^{i-e}\frac{\hat{\varphi}_i}{\hat{\varphi}}\right)^{\sigma-1}}. \tag{12}$$

In the next proposition we derive this expression. Here we note two points. First, conditional on the endogenous variables, it is independent of two parameters: size, L, and technology ϕ. Second, it depends on $\bar{\varphi}_{i,e}$ via the terms $\left(\frac{\hat{\varphi}_i}{\hat{\varphi}}\right)^{\sigma-1}$, which measure the average productivity of firms $\varphi \in [\bar{\varphi}_i, \bar{\varphi}_{i+1}]$ relative to all active firms and would represent their sales share in the absence of specialization. Thus the average intensity is a premium weighted average of the underlying sales share in a model without adoption. Thus we can also decompose the impacts of s_1 into an intensive margin (holding the cutoffs constant) and an extensive margin.

Proposition 2. ***(Heterogeneous specialization, multiplier and intermediate share)***

The average intermediate cost share in (12) and the associated multiplier $\bar{\alpha}$ are increasing in the specialization premium if and only if there is heterogeneous specialization. Moreover, the share elasticity can be decomposed as

$$\frac{d\ln\tilde{\alpha}}{d\ln s_1} = \underbrace{\frac{\sigma-1}{\delta\bar{\alpha}}\sum_{i=e}^{n}\lambda_i\left(\alpha_i-\bar{\alpha}\right)^2}_{\textit{intensive}\ \geq 0} + \underbrace{\frac{1}{\bar{\alpha}}\sum_{i=e}^{n}\lambda_i(\alpha_i-\bar{\alpha})\left(\sum_{j=e}^{n}\frac{\partial\ln Y_i}{\partial\ln\bar{\varphi}_j}\frac{d\ln\bar{\varphi}_j}{d\ln s_1}\right)}_{\textit{extensive}\ \geq 0}. \tag{13}$$

The proof is in appendix C.2, here we provide intuition. From (13) we see that under homogeneous specialization $\alpha_i = \bar{\alpha}$ for all i so there is no impact of s_1. The intensive impact is positive because more specialized firms are more exposed to changes in the premium; the intensive elasticity is also higher when there is more heterogeneity in specialization, as measured by the sales share weighted coefficient of variation of α_i. The extensive margin elasticity is also positive because technology upgrading of more specialized firms outweighs potential downgrading by those less specialized. We note that the latter result accounts for changes in selection, and thus requires us to derive the impact of s_1 on $\bar{\varphi}_{i,e}$, which we do in the next section.

The change in s_1 can be driven by technology, e.g. ϕ, or a parameter affecting P, e.g. size, L. If increases in ϕ (or L) increase s_1 (as we later derive) then the proposition above also shows that increases in ϕ (or L) increase the intermediate share because it only depends on these parameters via the premium.

4.1.2 Selection under Free Entry

The ex-ante identical potential firms obtain their core productivity by paying f_E units of labor. Thus under free entry the ex-ante period profit from this blueprint—the expectation of $\tilde{\pi}(\varphi)$ in (6) over the productivity distribution G must equal that cost:

$$\int_{\bar{\varphi}_e}^{\infty}\tilde{\pi}(\varphi)dG(\varphi) = f_E. \tag{14}$$

In the following proposition we show (14) can be rewritten similarly to models without adoption by using the cutoffs in (7) and (8), and defining F_i as Δf_i for $i > e$ and f_e otherwise:

$$\sum_{i=e}^{n}F_i\int_{\bar{\varphi}_i}^{\infty}\left[\left(\frac{\varphi}{\bar{\varphi}_i}\right)^{\sigma-1}-1\right]dG(\varphi) = f_E. \tag{15}$$

In an equilibrium with homogeneous specialization we have $n = e$ and obtain a standard expression where the marginal production cutoff, $\bar{\varphi}_e$, is determined by f_E/f_e and G. This includes models without intermediates (e.g. Melitz, 2003)—captured endogenously in our model when $s_1 < 1$—and those with an exogenous common share of intermediates for all firms—when s_1 is sufficiently large in our model (Proposition 1). In those two special cases of our model with homogeneous specialization, selection is independent of market size and this independence extends to any outcomes determined only by fixed costs and the survival probability, $1 - G(\bar{\varphi}_e)$, e.g. average operating profits, firm size and their distributions.

Under heterogeneous specialization, we still use (14) to solve for $\bar{\varphi}_e$. Replacing the cutoffs $\bar{\varphi}_I = \varphi\left(\bar{\varphi}_{i,e}, s_1, \hat{f}, I\right)$ from (9) in (15) we see that the equilibrium $\bar{\varphi}_e$ depends on the determinants of $\bar{\varphi}_I$, fixed costs, and G; thus we denote the resulting function by

$$\bar{\varphi}_e \equiv \varphi_e\left(s_1, \mathbf{f}, I, G\right). \tag{16}$$

The following proposition shows how this cutoff depends on the specialization premium, s_1. Moreover, conditional on s_1, selection depends on fixed costs ($\mathbf{f} = \{\hat{f}, f_E, f_e\}$ for adoption, operation, and entry, respectively), technology upgrades available (I), the underlying productivity distribution (G), and σ (omitted).

Proposition 3. ***(Heterogeneous specialization and selection)***

The entry cutoff $\bar{\varphi}_e$ in (16) is increasing in the specialization premium if and only if there is heterogeneous specialization. Moreover,

$$\frac{d\ln\bar{\varphi}_e}{d\ln s_1} = \frac{\bar{\alpha} - \alpha_e}{\delta} \geq 0. \tag{17}$$

We prove this in appendix C.3. Holding $\bar{\varphi}_e$ constant, an increase in s_1 lowers all adoption cutoffs and increases ex-ante expected profit above the entry cost. This induces additional entry and eventually a lower probability of survival due to higher competition that re-establishes the equality in (15). The elasticity of selection in (17) is zero if all firms adopt the same technology such that $\bar{\alpha} = \alpha_e$.

The change in s_1 can be driven by technology, e.g. ϕ, or a parameter affecting P, e.g. size, L. If increases in ϕ (or L) increase s_1 (as we later derive) then this proposition also shows that they increase selection because $\bar{\varphi}_e$ only depends on these parameters via the premium. This also further clarifies that the selection effect due to size in our model is only present under endogenous specialization.

4.1.3 Number of Firms

Using free entry and the fact that aggregate operating profits equal a fraction of total expenditure, X/σ, we can express the mass of entering firms as

$$M = \frac{X}{\sigma\left(f_E + \bar{F}\right)} = \frac{L \cdot \bar{a}}{\sigma\left(f_E + \bar{F}\right)}, \tag{18}$$

where we recall $\bar{a}$ is the multiplier and $\bar{F} \equiv \int_{\bar{\varphi}_e}^{\infty} f_{i(\varphi)} dG(\varphi)$ is the average expenditure of entering firms on fixed production costs. The second equality uses (10) and shows how the multiplier affects M, which will be an important channel for the impact of size and trade on real income.

Under homogeneous specialization we obtain a standard expression for M: log linear in size L and independent of the specialization premium. In such a setting, $\bar{\alpha}$ is fixed and $\bar{F}$ is determined by the entrant production cost $f_{e,i}$ required for $\bar{\alpha}$ and $\bar{\varphi}_e$, where the latter is independent of size as described before. So, in models without endogenous specialization M is determined by $(L, f_E, f_{e,i})$, and this is also the case in our model if s_1 is outside the heterogeneous specialization range (in which case the entrant uses either $i = 0$ or n). This implies that without heterogeneous specialization there is no amplification of market size on the mass of entrant or surviving firms even in the presence of intermediate linkages.

With heterogeneous specialization, size can affect M via changes in s_1. We denote the mass as a function of the premium and entry cutoffs by $\tilde{M} \equiv M\left(L, \bar{\alpha}\left(s_1, \bar{\varphi}_e\right), \bar{F}\left(s_1, \bar{\varphi}_e\right)\right)$. An increase in s_1 increases $\bar{\alpha}$ and thus M. Increases in s_1 also affect $\bar{F}$: holding $\bar{\varphi}_e$ fixed, it increases $\bar{F}$ because of increased adoption but this increase is partially offset by selection since higher s_1 reduces the fraction of entrants that produce.

The labor market will also clear and we can verify it yields the same condition for the mass of firms. For completeness we describe the labor income-expenditure equality showing its allocation here

$$L = \underbrace{\frac{\sigma - 1}{\sigma}(1 - \bar{\alpha})X}_{\text{variable labor}} + M\left(\bar{F} + f_E\right). \tag{19}$$

The first component is labor expenditure on variable costs; since intermediates account for $\bar{\alpha}\frac{\sigma-1}{\sigma}X$ and the process is constant returns, labor accounts for the remaining. The second represents the total fixed cost payments in production $M\bar{F}$ and entry costs given by Mf_E. Using the goods market condition (10) we obtain (18).

4.2 Equilibrium Specialization and Price

4.2.1 Characterization

The price index, P, determines real income and specialization, via s_1. Thus we write all variables in terms of these aggregates, show conditions for existence and uniqueness of the equilibrium and derive comparative statics.

To obtain the relationship between P and s_1 we substitute the goods market clearing condition (10) into the production entry expression (7), and obtain:

$$\tilde{P}\left(s_1\right) = \left(\frac{f_e}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}} \cdot \underbrace{\left[\bar{a}\left(s_1\right)\right]^{-\frac{1}{\sigma-1}}}_{\text{multiplier}} \cdot \underbrace{\left[\bar{\varphi}_e(s_1)\right]^{-1}}_{\text{selection}} \cdot (s_1)^{-e}, \tag{20}$$

which reflects how s_1 affects P via the endogenous multiplier, selection and the entrant technology. If we then substitute the definition $s_1 \equiv \left(\frac{\phi}{P}\right)^{\delta}$ we obtain the equilibrium P as the fixed point(s) ($\bar{P}$) that satisfy (20).

$$(\bar{P})^{1-\alpha_e} = \left(\frac{f_e}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}} \cdot \underbrace{\left[\bar{a}\left(s_1\left(\bar{P}\right)\right)\right]^{-\frac{1}{\sigma-1}}}_{\text{multiplier}} \cdot \underbrace{\left[\bar{\varphi}_e\left(s_1\left(\bar{P}\right)\right)\right]^{-1}}_{\text{selection}} \cdot \phi^{-\alpha_e} \tag{21}$$

In our model under heterogeneous specialization the marginal entrant optimally chooses the minimum level of specialization (shown in Lemma C.4 in the appendix), so $\alpha_e = \alpha_0$. If we take α_0 to be zero so $f_e = f_0$ then we can immediately see that the equilibrium price in our model under heterogeneous specialization is lower than in a model without specialization (where $\bar{P}_0 = \left(\frac{f_0}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}} \bar{\varphi}_0^{-1}$) and the difference across models is captured by the multiplier and selection terms in (21).

To provide some insight, we analyze the equilibrium by using $\tilde{P}(s_1)$ and the specialization schedule, $P_s(s_1)$, obtained from the definition 1 of the specialization premium:

$$P_s(s_1) = \phi s_1^{-\frac{1}{\delta}}. \tag{22}$$

In Figure 4 we plot the log of this and expression (20) against $\ln s_1$. Their intersection, E, represents the equilibrium and we depict a case when it is unique in the **heterogeneous specialization range**, defined by $\mathbf{s} : s_1 \in \left(1, \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}\right)$.

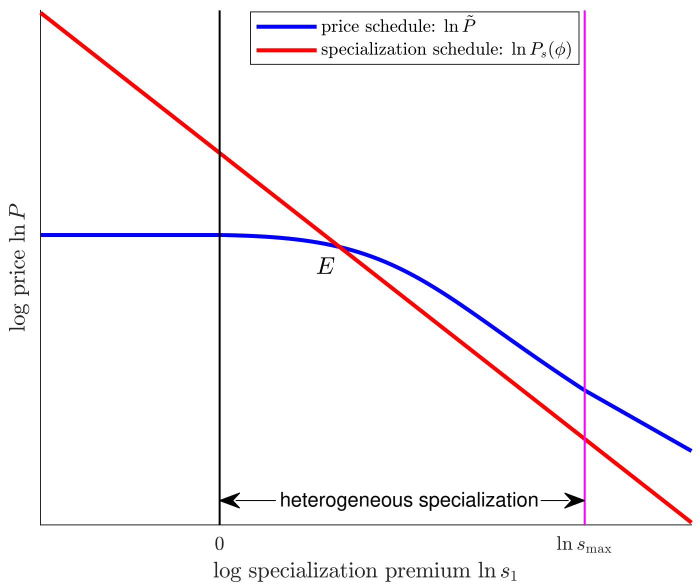


Figure 4: Existence and uniqueness of equilibrium

Proposition 4 establishes the existence of a unique equilibrium and characterizes it in terms of specialization, $\bar{\alpha}$, and heterogeneity, depending on the size of the economy and the productivity of intermediates, indexed by L_i and ϕ_i, respectively. The characterization is qualitatively similar using either $x = \{L, \phi\}$ so we refer to a change in either L or ϕ as a change in x and use $x \in [x_0, x_n]$ to denote either $L \in [L_0, L_n]$ at a given ϕ_i or vice versa.

Proposition 4. ***(Equilibrium)***

1. *In an economy $x = \{L, \phi\}$ with size L and intermediates productivity ϕ, there is an equilibrium with $\bar{\alpha}(x) \in [\alpha_0, \alpha_n]$;*

2. *The equilibrium is unique and has homogenous specialization given by*

   (a) $\bar{\alpha} = \alpha_0$ *if* $x < x_0$*;*

   (b) $\bar{\alpha} = \alpha_n$ *if* $x > x_n$*.*

3. *There exist equilibria with heterogeneous specialization and* $\bar{\alpha}^j(x) \in (\alpha_0, \alpha_n)$ *for some* $x \in [x_0, x_n]$.

4. *If* $\frac{d \ln \tilde{P}}{d \ln s_1} > -\frac{1}{\delta}$ *for all* s_1 *then* $\bar{\alpha}$ *is unique and non-decreasing in* x*.*

We provide the proof in Appendix C.4. The existence in part 1 follows from continuity of P_s and of $\tilde{P}$ in s_1 coupled with the facts that $P_s \in (0, \infty)$ and it is steeper than $\tilde{P}(s_1)$ at least for $s_1 \notin$ s, i.e. at the extremes. Thus, $\tilde{P}$ must cross P_s at least once from below as illustrated in Figure 4, which occurs whenever it is flatter, what we refer to as the **stability condition**

$$\frac{d \ln \tilde{P}(s_1)}{d \ln s_1} > \frac{d \ln P_s}{d \ln s_1} = -\frac{1}{\delta}. \tag{23}$$

Figure 4 assumes stability holds for all s_1. This need not be the case for the full range but does hold at least when $s_1 \notin$ s. To see this note from (20) that $\tilde{P}$ depends on s_1 via the entrant technology choice, the multiplier and selection effect. But for $s_1 \notin$ s the last two effects are absent (since either none or all specialize) so $\frac{d \ln \tilde{P}(s_1)}{d \ln s_1} = -e$, the number of adoption steps of the marginal producer, which is at most n. If the minimum technology is $\alpha_0 = 0$ then $\ln \tilde{P}(s_1)$ is initially flat, as in Figure 4, and when all adopt, it is $-n = -\frac{\alpha_n}{\delta} > -\frac{1}{\delta}$ since we assume $\alpha_n < 1$ (labor is necessary).

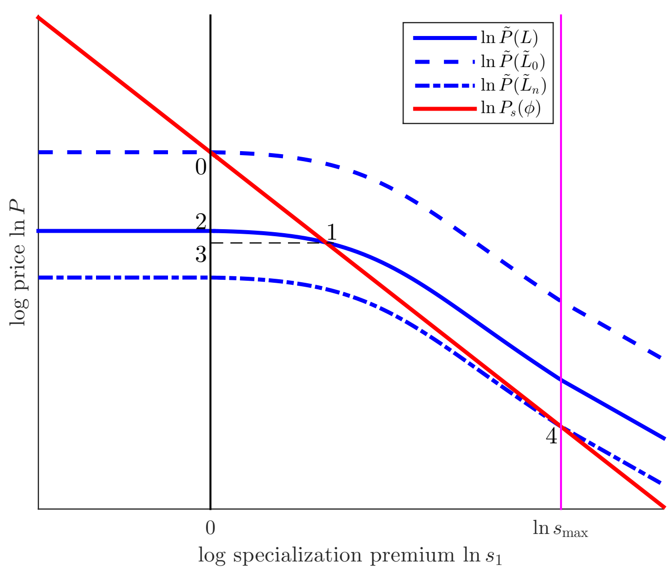


Figure 5: Specialization under unique equilibrium

Part 2a shows that in an economy that is sufficiently small or unproductive all firms use the minimum intermediate share, α_0. Part 2b establishes that in large/productive economies all firms adopt the maximum, α_n. It also establishes the critical x at which this occurs. In Figure 5 we illustrate the critical values for L. Note that increases in L shift only $\ln \tilde{P}$ and do so proportionally without affecting the slope. Thus starting at L_1 we can find a lower $\tilde{L}_0$ such that it just supports the minimum heterogeneous specialization equilibrium defined by $\tilde{P}(1, \tilde{L}_0) = P_s(1)$ in the figure, thus any further reductions in L do not affect $\bar{\alpha}$. In this case $\tilde{L}_0 = L_0$. A similar argument applies to part 2b and in Figure 5 we see $\tilde{L}_n = L_n$. The same logic applies to increases in ϕ; these shift $\ln P_s$ up without affecting the slope. We can find critical values of ϕ to support equilibria along $\tilde{P}$ with different s_1.

Part 4 uses the insights above to show that when (23) holds globally the model explains how increases in x monotonically increase specialization. We use Figure 5 to illustrate it for market size. We start at point 0 and an increase in L shifts $\tilde{P}$ down to point 1, further increases eventually imply maximum adoption at point 4. A simple corollary is that our model subsumes as special cases the standard Melitz model ($x < x_0$ and $\alpha_0 = 0$) and a version of it with fixed share of intermediates ($x > x_n$). Given part 4, which assumes (23) holds globally, it is clear that in those conditions there exist heterogeneous specialization equilibria, as stated in part 3.[14]

5 Qualitative Implications

We provide additional qualitative implications of the model. First, we show there is a potential for inefficiency of the market equilibrium in small/unproductive economies due to under-adoption. Second, we derive equilibrium elasticities of various model outcomes with respect to $x = \{L, \phi\}$. Furthermore, we decompose those elasticities into the novel endogenous specialization terms consisting of selection and multiplier effects.

While a government may not literally be able to increase the number of workers or technology within its borders, it may increase the relevant market size via trade with others. In section 6 we show that the size increases have identical impacts to trade liberalization for nearly all outcomes.

5.1 Market Inefficiency

The market equilibrium may not be efficient if some firms remain unspecialized. The resulting possibility of efficiency-enhancing policies stands in contrast to Melitz (2003), which Dhingra and Morrow (2019) show is socially optimal.[15] Our model yields the Melitz (2003) market equilibrium if the minimum spe-

[14] Heterogeneous specialization equilibria also exist if (23) does not hold globally; in that case there are multiple equilibria. We can show at least one is stable and non-decreasing in x.

[15] The equilibrium coincides with the solution of a planner that maximizes consumer utility (since there are no aggregate profits) subject to the labor resource constraint and linear technology by choosing the mass of entrants, the minimum productivity firm, and a production quantity function that is continuously differentiable in productivity, $\{M, \bar{\varphi}_e, q(\varphi)\}$ respectively. They show this occurs in the one sector economy with CES (similar to our framework) because the quantity-variety trade off exactly offset: firms produce less than the optimal quantity given their monopoly power but this is offset by the extra entry of

cialization possible is $\alpha_0 = 0$ and the economy is sufficiently small/unproductive (part 2a of proposition 4). If that is the unique equilibrium then it is also socially efficient under the Dhingra-Morrow planner, which restricts the technology to use only labor. If firms can choose different technologies and use intermediates, then the planner problem would still involve maximizing utility subject to the resource and technology constraints. However, now the choice variables should also include the thresholds for adoption and quantity functions for firms in each group. This would allow the planner to replicate any market equilibrium in our model.

Instead of a full planner solution we simply ask if a government can improve on some market equilibrium, in which case we will say the latter is inefficient. This is the case even if the government has a single instrument at its disposal: a tax on fixed costs of operating with any technology other than the most specialized. Moreover, we restrict this tax to be prohibitive so that in equilibrium all firms adopt the most specialized technology. Any improvement from this prohibitive tax rate on the market equilibrium provides a lower bound relative to an optimal tax rate (or planner solution).

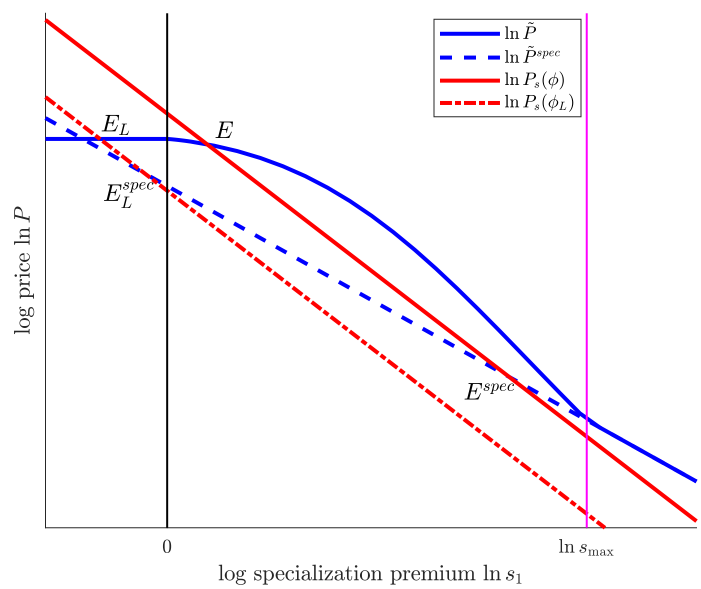


Figure 6: Inefficiency of heterogeneous specialization equilibrium

Utility is increasing in real wages, $1/P$, and a prohibitive tax rate generates no revenue. Therefore we need only determine when P is lower under the specialization tax. This requires deriving the equilibrium price, $\bar{P}^{spec}$, which is equal to what we obtain in an economy where only the most specialized technology is available. In Appendix C.5 we show that $\tilde{P}^{spec}(s_1)$ is loglinear in s_1 as depicted in Figure 6 and the equilibrium is at its intersection with $P_s(s_1)$—the latter is still defined by (22) as in the full model. Note that the price schedules $\tilde{P}^{spec}(s_1)$ and $\tilde{P}(s_1)$ coincide if $s_1 \geq s_{\max}$ and the equilibrium price is the same if $x > x_n$. The reason is simple: in the full model with a unique equilibrium and large economy all firms adopt the most specialized technology so removing the option of adopting other technologies is

monopolistic firms that ignore their impact on the profits of competitors.

irrelevant. If $x < x_n$ and the heterogeneous specialization equilibrium is at E then the equilibrium with a single technology is E^{spec}, with lower prices and thus higher utility. In Appendix C.5 we show that there is always some x close enough to x_n such that $\tilde{P}^{spec}$ lies below $\tilde{P}$ since the latter is steeper when s_1 increases to $s_{\max}$—the effect of the premium has an additional impact on P in that case from the firms adopting.

Alternatively, there are economies sufficiently small or unproductive where forcing adoption is not efficient so $\bar{P}^{spec} > \bar{P}$.[16] However, as long as the price schedules intersect left of $s_1 = 1$ (as depicted in Figure 6) then a market equilibrium such as E_L without specialization is inefficient.

In the quantitative section we show the equilibrium reflects this inefficiency and that it is not simply due to the input-output linkage but rather the endogenous specialization. Moreover, we find it is stronger in the initial period when there is less specialization.

5.2 Size and Technology Elasticities

We derive the impacts of economic size and technology on aggregate and firm outcomes. Moreover, we contrast them with alternative models with a fixed specialization premium. To do so we start with an equilibrium outcome function $o\left(x, s_1\left(x\right)\right)$ and decompose the impact of a parameter x (e.g. L or ϕ) as follows

$$\begin{array}{ccccc} \frac{d\ln o(x,s_1(x))}{d\ln x} & = & \frac{d\ln o}{d\ln x}\Big|_{s_1} & + & \frac{d\ln o}{d\ln s_1}\cdot\frac{d\ln s_1}{d\ln x} \\ \epsilon_o^x & = & \underset{\text{fixed } s_1}{\bar{\epsilon}_o^x} & + & \underset{\text{endogenous } s_1}{\epsilon_o^s\cdot\frac{d\ln s_1}{d\ln x}} \end{array} \tag{24}$$

The overall elasticity of outcome o, denoted by ϵ_o^x, reflects the direct elasticity holding s fixed, denoted by $\bar{\epsilon}_o^x$ and the novel endogenous response in the model, ϵ_o^s, from selection and multiplier effects. We will show how $\bar{\epsilon}_o^x$ can also be interpreted as the elasticity in certain standard models with homogeneous specialization where either no firm adopts or *all* do so but at a common level $\bar{\alpha}$, which will be a convenient form to compare the additional impacts of endogenous specialization. Moreover, we also illustrate how to relate the effect of endogeneous specialization to the multiplier effect, $\frac{d\ln\bar{\alpha}}{d\ln s_1}$, and show that this multiplier elasticity is sufficient to compute those size and technology elasticities in the quantitative section. We focus on locally stable equilibria, so $\frac{d\ln s}{d\ln x} > 0$, as shown in the proof of proposition 5.

5.2.1 Income Gains

The elasticities ϵ_W^x of aggregate real income are simply the inverse of those for the price index (since $W \equiv 1/P$ given no aggregate profits) and we focus on size and technology shocks: $x = \{L, \phi\}$.

Proposition 5. ***(Real income elasticities)***

[16] Graphically this would occur if the equilibrium s_1 under endogenous specialization were left of the intersection of $\tilde{P}^{spec}$ and $\tilde{P}$.

Real income $W = 1/P$ increases with economic size (L) and intermediates technology (ϕ) around stable equilibria and the respective elasticities are

$$\epsilon_W^L = \bar{\epsilon}_W^L \left(1 - \delta\epsilon_W^s\right)^{-1} > 0 \quad ; \quad \epsilon_W^\phi = \left(\bar{\epsilon}_W^\phi + \delta\epsilon_W^s\right)\left(1 - \delta\epsilon_W^s\right)^{-1} > 0. \tag{25}$$

In an equilibrium with heterogeneous specialization ($x \in (x_0, x_n)$ so $s_1 \in \mathbf{s}$) the specialization premium elasticity reflects multiplier and selection effects.

$$\delta\epsilon_W^s\left(\bar{\alpha}\right) = \underbrace{\frac{\bar{\alpha}\delta}{\sigma - \bar{\alpha}\left(\sigma - 1\right)}}_{\textit{Multiplier}} \frac{d\ln\bar{\alpha}}{d\ln s_1} + \underbrace{\bar{\alpha}}_{\textit{Selection}} \geq \bar{\alpha}. \tag{26}$$

Thus, these income elasticities exceed those of alternative models with homogeneous specialization fixed at any level $\alpha_e \leq \bar{\alpha}$

$$\epsilon_W^L|_{s_1\in\mathbf{s}} = \left[\left(\sigma - 1\right)\left(1 - \delta\epsilon_W^s\left(\bar{\alpha}\right)\right)\right]^{-1} > \left[\left(\sigma - 1\right)\left(1 - \alpha_e\right)|_{\alpha_e\leq\bar{\alpha}}\right]^{-1} = \epsilon_W^L|_{\alpha_e\leq\bar{\alpha},\bar{s}} \tag{27}$$

$$\epsilon_W^\phi|_{s_1\in\mathbf{s}} = \left(1 - \delta\epsilon_W^s\left(\bar{\alpha}\right)\right)^{-1} - 1 > \left[\left(1 - \alpha_e\right)|_{\alpha_e\leq\bar{\alpha}}\right]^{-1} - 1 = \epsilon_W^\phi|_{\alpha_e\leq\bar{\alpha},\bar{s}}. \tag{28}$$

We derive the expressions in Appendix C.6, here we provide the insight for L, which extends to ϕ. Consider an initial economy with $L < L_0$ such that, according to proposition 4 there is no specialization. We can see directly that increases in L shift (20) down by $\bar{\epsilon}_W^L = 1/\left(\sigma - 1\right)$ (if $\alpha_0 = 0$) and thus raise real income. That elasticity is constant until $L = L_0$, the equilibrium point 0 in Figure 5, where some firms are indifferent about adopting. Further increases in L continue to shift the price schedule down by $\bar{\epsilon}_W^L$ and would increase real income from point 0 to 2 if s_1 was fixed (and thus $\bar{\alpha} = 0$) but by more in our model as s_1 falls to the final equilibrium at point 3. The magnification is captured by $\left(1 - \delta\epsilon_W^s\right)^{-1}$, which captures both a selection and a multiplier effect as shown by the expression for $\delta\epsilon_W^s$. The elasticity is also larger when compared to an alternative model where all firms use the same intermediate technology with a share no higher than $\bar{\alpha}$, the average in our model. In the quantitative section we show how to construct this alternative fixed premium economy such that it has the same initial price index as our model. The elasticity of welfare with respect to size in this fixed premium economy is obtained by using (21) directly holding s fixed and is equal to $\bar{\epsilon}_W^L|_{\alpha_e=\bar{\alpha}(s_1)}$. We can see that this is smaller than $\epsilon_W^L|_{s_1\in\mathbf{s}}$ for any $\alpha_e \leq \bar{\alpha}$, when that economy uses exactly $\bar{\alpha}$ then the difference between them is due to the multiplier effect. This holds whenever $\delta\epsilon_W^s\left(\bar{\alpha}\right) \geq \bar{\alpha}$, which requires average specialization to be increasing in s_1, $\frac{d\ln\bar{\alpha}}{d\ln s_1} \geq 0$, as shown in proposition 2.

5.2.2 Aggregate Cost Shares and Labor Productivity

We first derive the elasticity of alternative aggregate variable cost shares with respect to size and technology. In addition to the intermediate share $\bar{\alpha}$ we also examine the variable labor cost share in production, $l_{sc} = 1 - \bar{\alpha}$. The model implies that the correlation between these cost shares is -1, which is consistent with what we find in the U.S. manufacturing between 1987–2007. We also consider an alternative labor share common in the literature, that in value added, l_{sv}, which for variable cost is defined as

$$l_{sv} = \frac{\frac{\sigma-1}{\sigma}(1-\bar{\alpha})Y}{Y - \frac{\sigma-1}{\sigma}\bar{\alpha}Y} = \frac{(\sigma-1)(1-\bar{\alpha})}{1+(\sigma-1)(1-\bar{\alpha})}. \tag{29}$$

The numerator is the aggregate labor share in variable costs. The denominator represents the gross value-added, calculated as total sales minus intermediates expenditure.[17]

Using the decomposition in (24) we write the size elasticities for these outcomes as

$$\epsilon^L_{\bar{\alpha}} = \frac{d\ln\bar{\alpha}}{d\ln s_1}\cdot\delta\epsilon^L_W \;\; ; \;\; \epsilon^L_{l_{sc}} = -\frac{\bar{\alpha}}{1-\bar{\alpha}}\cdot\epsilon^L_{\bar{\alpha}} \;\; ; \;\; \epsilon^L_{l_{sv}} = -\frac{\bar{\alpha}}{1-\bar{\alpha}}\cdot\frac{\epsilon^L_{\bar{\alpha}}}{1+(\sigma-1)(1-\bar{\alpha})}. \tag{30}$$

Size affects variable production cost shares only via the endogenous specialization premium elasticity, ϵ^s_o. To see this note that $\bar{\epsilon}^L_o = 0$ since conditional on $\bar{\alpha}$ all shares are independent of L and the same is true for $\bar{\alpha}$ conditional on s_1 (from $\tilde{\alpha}$ in (12) and $\bar{\varphi}_e$ in (16)). We summarize these results below and derive them in Appendix C.7:

Proposition 6. ***(Aggregate variable cost share elasticities)*** *The elasticities of aggregate variable cost shares with respect to size are given by (30); non-zero iff there is heterogeneous specialization and positive for intermediates, $\epsilon^L_{\bar{\alpha}} > 0$, negative for both labor, $\epsilon^L_{l_{sc}} < 0$, and for labor value added in production, $\epsilon^L_{l_{sv}} < 0$.*

The intermediate share elasticity $\epsilon^L_{\bar{\alpha}}$ depends only on the effect of size on s_1, which is positive in a stable equilibrium (proposition 5) and the effect of specialization premium on intermediate cost share, $\frac{d\ln\bar{\alpha}}{d\ln s_1} > 0$ iff there is heterogeneous specialization (proposition 2). The labor results follow directly since they can be written as functions of $\bar{\alpha}$, so their elasticity is inversely proportional to $\epsilon^L_{\bar{\alpha}}$ as shown in (30).

The same points and proposition apply to the technology parameter ϕ, the only difference is that we replace $\delta\epsilon^L_W = \frac{d\ln s_1}{d\ln L}$ with $\frac{d\ln s_1}{d\ln\phi}$ in (30).

The model also generates novel predictions for two measures of worker productivity. Similarly to the cost shares we focus on workers used in assembly, i.e. the variable component, and denote their aggregate quantity by L_v. We then define aggregate real productivity per variable worker in terms of output, $\bar{\varphi}_Q \equiv Y/PL_v$, and real value-added: $\bar{\varphi}_{VA} \equiv Y\left(1-\frac{\sigma-1}{\sigma}\bar{\alpha}\right)/PL_v$.

Proposition 7. ***(Aggregate variable labor productivity)*** *The elasticities of aggregate variable worker productivity with respect to size are*

$$\epsilon^L_{\bar{\varphi}_Q} = \epsilon^L_W - \epsilon^L_{l_{sc}} > 0 \;\; ; \;\; \epsilon^L_{\bar{\varphi}_{VA}} = \epsilon^L_W - \epsilon^L_{l_{sv}} > 0. \tag{31}$$

[17]This is the measure Karabarbounis and Neiman (2014) employ to document the global decline in labor share; whereas they emphasize substitution towards capital, our explanation relies on adoption of less labor intensive technologies and a re-allocation of production towards the more productive less labor intensive firms.

These are positive and exceed those of alternative models with homogeneous specialization fixed at any level $\alpha_e \leq \bar{\alpha}$.

This is a corollary of propositions 5 and 6, as we show in Appendix C.8. First, both real output and value added increase as the price index falls and thus reflect ϵ^L_W, which we show is positive in proposition 5 and larger in our model under heterogeneous specialization. Second, Y/L_v is inversely proportional to the share l_{sc}, and $\bar{\varphi}_{VA} = 1/Pl_{sv}$ so there is an additional effect as those shares fall as shown in proposition 6.

5.2.3 Selection and Firm Entry

We show that size increases selection into production and thus lowers the fraction of entrants that actually produce. Moreover, there are productivity distributions such that the size effect from endogenous specialization increases the mass of entrants but is dominated by the selection effect, which results in a reduction of active firms (and thus variety).

We define selection into production as the fraction of firms that invested to enter and remain active, $M_a/M = 1 - G(\bar{\varphi}_e)$, and summarize the effect of size in the following proposition.

Proposition 8. ***(Selection, entrants, and producer mass)***

1. *The size elasticities for selection, entrants, and active firms are respectively*

$$\epsilon^L_{M_a/M} = \epsilon^s_{M_a/M} \cdot \delta\epsilon^L_W \ \ ; \ \ \epsilon^L_M = 1 + \epsilon^s_M \cdot \frac{d\ln s_1}{d\ln L} \ \ ; \ \ \epsilon^L_{M_a} = 1 + \left(\epsilon^s_M + \epsilon^s_{M_a/M}\right) \cdot \delta\epsilon^L_W; \tag{32}$$

2. *The size elasticity components from specialization for selection and entrants are respectively*

$$\epsilon^s_{M_a/M} = -\frac{\bar{\varphi}_e g(\bar{\varphi}_e)}{1 - G(\bar{\varphi}_e)} \frac{\bar{\alpha} - \alpha_e}{\delta} \ \ ; \ \ \epsilon^s_M = (\bar{a} - 1)\frac{d\ln\bar{\alpha}}{d\ln s_1} - \frac{\bar{F}}{f_E + \bar{F}}\frac{d\ln\bar{F}}{d\ln s_1}; \tag{33}$$

3. *Larger economies have more selection into production if and only if there is heterogeneous specialization: $\epsilon^L_{M_a/M} < 0$ iff $s \in \mathbf{s}$;*

4. *There are distributions $g_k(\cdot)$ s.t. size increases the mass of firms via specialization, $\epsilon^s_M > 0$, and simultaneously reduces active firms, $\epsilon^s_M + \epsilon^s_{M_a/M} < 0$, for an economy with sufficiently large L and α_n.*

We prove this in Appendix C.9. Holding s_1 constant, the mass of entrants and active firms is directly proportional to size—the effect in standard models. The novel elasticity components due to specialization are in (33). The mass of active firms falls relative to entrants, so selection increases if there is heterogeneous specialization, otherwise there is no effect since $\bar{\alpha} = \alpha_e$. The specialization increases the mass of entrants due to the multiplier effect, $\frac{d\ln\bar{\alpha}}{d\ln s_1}$>0, but that may be partly offset if the increased premium also increases the average production fixed cost over all entrants, i.e. if $\frac{d\ln\bar{F}}{d\ln s_1} > 0$. In part 4 we show that the mass of entrants unambiguously increases with size due to specialization under certain productivity distributions, namely any $g_k(\varphi)$ with constant elasticity in φ, e.g. if G is Pareto. Under these distributions we obtain $\frac{d\ln\bar{F}}{d\ln s_1} = 0$: some firms incur increased fixed cost expenditures to upgrade, however this is offset by the reduction in the fraction of producers (and thus lower production fixed costs per entrant). For such distributions we can also show that for economies that are sufficiently large, the selection effect offsets the mass effect and so $\epsilon^s_M + \epsilon^s_{M_a/M} < 0$, which implies that specialization can simultaneously increase overall entry and reduce active firms. This existence result is present in our quantification so it is useful to understand how it arises in the theory.

We obtain similar results if we focus on technology improvements, ϕ.[18]

In sum, endogenous specialization introduces a novel set of predictions for the impacts of size and technology that are absent in standard models due to the presence of multiplier effect.

5.3 Profit, Sales and Productivity Distributions

The model has implications for the role of size on profits net of fixed costs and sales distributions. We show these distributions are invariant with respect to size in standard models without changes in specialization. However, in our model increases in size generate a mean preserving spread in profits, and so higher concentration. The same is true for sales under the most standard type of productivity distribution, Pareto. We conclude by discussing the size effects on alternative measures of productivity.

5.3.1 Profit

Profits net of production costs are given by $\tilde{\pi}(\varphi, L)$ in (6) and the free entry condition in (14) implies its mean is constant across any equilibria with the same entry cost, f_E. Thus we can focus on measuring its concentration.

We characterize the distributional impacts of size shocks using the **profit CDF:**

$$\Phi(x, L) \equiv \Pr(\tilde{\pi}(\varphi, L) \leq x);\, x \in [0, \infty); \tag{34}$$

and the **profit cumulative share** among all entrants of firms with productivity at least $\bar{\varphi}$:

$$\Pi(\bar{\varphi}, L) \equiv \frac{\int_{\bar{\varphi}}^{\infty} \tilde{\pi}(\varphi, L) dG(\varphi)}{\int_{\varphi_{\min}}^{\infty} \tilde{\pi}(\varphi, L) dG(\varphi)}. \tag{35}$$

We say that after a change in size from L to L' **profit concentration increases if** either $\Phi(x, L)$ second-order stochastic dominates (SSD) $\Phi(x, L')$ or $\Pi(\bar{\varphi}, L) \leq \Pi(\bar{\varphi}, L')$ for all $\bar{\varphi}$ (with some inequality).

[18] Simply adjust (32) to use $\frac{d\ln s_1}{d\ln\phi}$ instead of $\frac{d\ln s_1}{d\ln L}$ and remove the direct effect since $\bar{\epsilon}^{\phi}_M = \bar{\epsilon}^{\phi}_{M_a} = 0$.

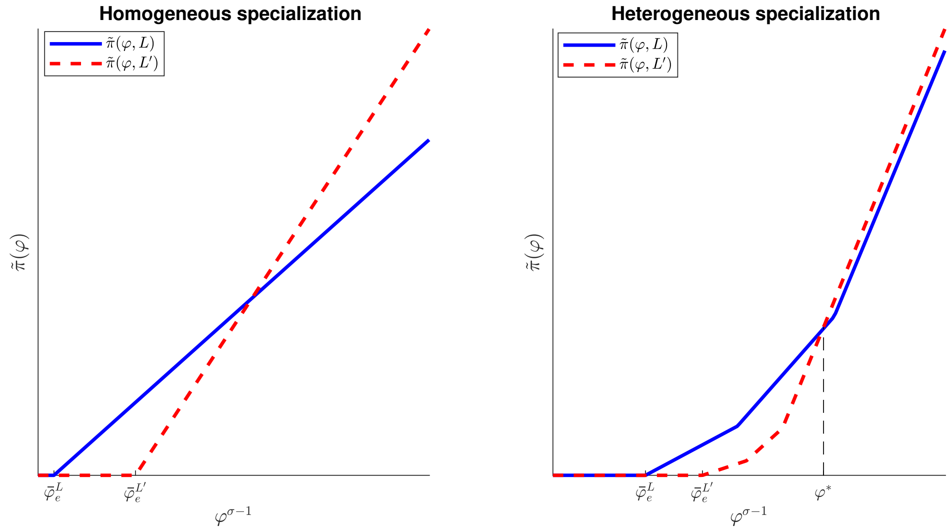


Figure 7: Size and profit distributions ($L' > L$)

Proposition 9. ***(Profit distribution)*** *An increase in size implies an increase in profit concentration iff it changes specialization. The resulting profit distribution is a mean-preserving spread of the original with higher cumulative shares for all $\bar{\varphi}$.*

We provide a formal proof in Appendix C.10. A change in specialization is necessary after the shock because otherwise technology is unchanged and so is selection (shown in section 4.1.2). Thus, if the economy remains either unspecialized or fully specialized then the marginal entrant and $\tilde{\pi}(\cdot)$ are the same and so are $\Pi(\cdot)$ and $\Phi(\cdot)$. What if the increase in size were large enough to move the economy from no specialization to full specialization? This special case illustrates the basic intuition: the shock increases selection so $\tilde{\pi}(\varphi, L') \leq \tilde{\pi}(\varphi, L)$ for $\varphi \leq \varphi^*$; but the expected profit is still f_E (due to free entry) so $\tilde{\pi}(\varphi, L') \geq \tilde{\pi}(\varphi, L)$ above some φ^*. Since $\tilde{\pi}(\cdot)$ is continuous and the only difference across firms (within each equilibrium) is their productivity we can show the intersection φ^* is unique and the new $\tilde{\pi}(\cdot)$ intersects the original only once from below as depicted in the left panel of Figure 7. In sum, a smaller share of entrants produces, and the ones that adopt have sales and operating profit rise faster in φ than non-adopters. This implies a higher share of cumulative profits, $\Pi(\bar{\varphi}, L') \geq \Pi(\bar{\varphi}, L)$, and also that the original distribution $\Phi(x, L)$ SSD $\Phi(x, L')$. Since the mean profits are equal across equilibria the two conditions are equivalent in this setting.[19]

A similar intuition applies to a shock under heterogeneous specialization. The shock leads to selection and we show $\tilde{\pi}(\varphi, L')$ is continuous and crosses $\tilde{\pi}(\varphi, L)$ once from below and thus implies higher concentration. The right panel of Figure 7 illustrates a case with two specialization technologies and the proof extends it to arbitrary number of technologies. A key part of the insight is that initially the most productive firms already have the technology most intensive in intermediates and thus benefit the most from reductions in their price, P.

[19] Atkinson (1970) shows that $\Phi(x, L')$ being a mean-preserving spread of $\Phi(x, L)$ is equivalent to $\Phi(x, L)$ Lorenz-dominating, which in our setting requires $\Pi(\bar{\varphi}, L) \leq \Pi(\bar{\varphi}, L')$ for any $\bar{\varphi}$.

5.3.2 Sales

We establish the analogue of proposition C.10 for firm sales. To isolate the effect of size on concentration we provide a sufficient condition that implies constant average sales per entrant, Y/M. Using (18) we have $Y/M = \sigma\left(f_E + \bar{F}\right)$ where $\bar{F} \equiv \int_{\varphi_e}^{\infty} f_{i(\varphi)} dG(\varphi)$ so average sales are constant if the change in size leaves $\bar{F}$ unchanged, which holds in the following two cases. First, if there is no change in specialization and thus no changes in selection or adoption, in which case the full sales distribution remains unchanged. Second, if there is a change in specialization but the change in entry fixed costs (selection) are exactly offset by the change in upgrading costs. In part 4 of proposition 8 we provide the following sufficient **productivity condition for constant mean sales** under heterogeneous specialization:

$$\frac{d\gamma(\bar{\varphi})}{d\bar{\varphi}} = 0 \text{ for all } \bar{\varphi} > \varphi_{\min} \text{ and } z \in [1, \sigma];\ \gamma(\bar{\varphi}, z) \equiv -\frac{d \ln \int_{\bar{\varphi}}^{\infty} \varphi^{z-1} dG(\varphi)}{d \ln \bar{\varphi}}. \tag{36}$$

The condition is satisfied by the untruncated Pareto where $\gamma = k - (z-1)$.[20]

Similarly to profit, we define the **sales cumulative share** among all entrants of firms with productivity at least $\bar{\varphi}$ as:

$$\lambda\left(\bar{\varphi}, L\right) \equiv \frac{\int_{\bar{\varphi}}^{\infty} \tilde{y}(\varphi, L) dG(\varphi)}{\int_{\varphi_{\min}}^{\infty} \tilde{y}(\varphi, L) dG\left(\varphi\right)}. \tag{37}$$

Sales concentration increases if either the original distribution SSD the new one or $\lambda\left(\bar{\varphi}, L\right) \leq \lambda\left(\bar{\varphi}, L'\right)$ for all $\bar{\varphi}$ (with some inequality).

The shares of operating profits are also given by (37) so the proposition below applies to these as well. Operating profits differ from overall profits due to endogenous fixed costs so we need to extend the results in Proposition 9 to account for this.

Proposition 10. ***(Sales distribution)***

An increase in size implies an increase in sales concentration iff it changes specialization and $\frac{d\gamma(\bar{\varphi})}{d\bar{\varphi}} = 0$*. The resulting sales distribution is a mean-preserving spread of the original with higher cumulative shares for all* $\bar{\varphi}$*.*

We provide the proof in Appendix C.11. If $i = e$ for all i and a shock to L does not change specialization then the sales distribution remains unchanged. With homogeneous specialization $\bar{F}$ is constant and so are mean sales; moreover the cumulative shares are also unchanged: $\lambda\left(\bar{\varphi}, L\right) = \int_{\bar{\varphi}}^{\infty} \varphi^{\sigma-1} dG(\varphi) / \int_{\bar{\varphi}_e}^{\infty} \varphi^{\sigma-1} dG(\varphi)$ and $\bar{\varphi}_e$ is unchanged as shown in section 4.1.2. This share expression also shows that if the shock changes specialization from none to full then the increased selection

[20] Recall that firm sales are proportional to $\varphi^{\sigma-1}$ so we can interpret $\gamma(\bar{\varphi}, \sigma) > 0$ as the elasticity of cumulative sales with respect to $\bar{\varphi}$ (of firms with productivity above $\bar{\varphi}$ that share a given technology). These elasticities weight the impacts of changes in cutoffs as s_1 increases. Melitz and Redding (2015) define $\gamma\left(z = \sigma - 1\right)$ as the "hazard function for the distribution of log fim size in a market" and show that it determines if the partial trade elasticity is constant and is thus important in determining the gains from trade.

leads to higher $\lambda(\bar{\varphi}, L)$ if $\frac{d\gamma(\bar{\varphi})}{d\bar{\varphi}} = 0$. A similar insight applies when we move from no specialization to some or within heterogeneous specialization equilibria. The proposition applies to the Pareto distribution used in the calibration.[21]

5.3.3 Productivity

We focus on a baseline firm measure of TFPQ given by the inverse of unit cost net of inputs, which is simply the fundamental productivity parameter[22]:

$$C_i^{-1} \equiv \left(\frac{c_i}{w^{1-\alpha_i} (P/\phi)^{\alpha_i}} \right)^{-1} = \varphi. \tag{38}$$

We focus on surviving firms and define $\mu(\cdot)$ **as the survivor density of TFPQ:** $\mu(\varphi, L) = g(\varphi) / \left(1 - G\left(\bar{\varphi}_e^L\right)\right)$ for $\varphi \geq \bar{\varphi}_e^L$ and zero otherwise, where $\bar{\varphi}_e^L$ is the production cutoff under L.

Proposition 11. ***(Survivor TFPQ distribution)***

An increase in size changes the TFPQ distribution iff it changes specialization and then implies $\mu(\varphi, L')$ FOSD $\mu(\varphi, L)$.

Changes in specialization are necessary because otherwise there is no selection, as previously discussed. Large enough shocks from no specialization to full specialization will generate selection and thus a higher mean productivity of surviving firms. Selection is also the only channel operating under heterogeneous specialization since L only affects $\mu(\varphi, L)$ via $\bar{\varphi}_e$.

In appendix C.12 we discuss how alternative weighted averages of TFPQ are affected by size due to selection and re-allocation.

6 International Trade

We allow for international trade between symmetric countries and show the size comparative statics are isomorphic to moving from autarky to free trade; or to partial trade liberalization. This extends the applicability of the results in an important direction because trade policy is a government lever to change market size. Moreover, the extension provides a direct connection between trade costs and specialization, which we highlight in the motivation facts and exploit in the calibration.

We also provide two additional insights about the effect of liberalization. With endogenous specialization we show that liberalization generates (i) higher income gains and (ii) increases the share of trade in intermediates (in contrast to standard models where it is constant). These occur only for economies

[21] When $\frac{d\gamma(\bar{\varphi})}{d\bar{\varphi}} > 0$ there are competing effects and we can't show that the size shock leads to higher cumulative shares for *all* $\bar{\varphi}$.

[22] The result applies to "true" quantity TFP, which requires a measure of the effective intermediate price index. In practice, we measure aggregate expenditure in intermediates, which implies that shocks increasing variety would typically have an additional impact on measured productivity.

with size or intermediates technology in the heterogeneous specialization range, and thus imply different impacts of trade liberalization across countries (or industries) and over time.

6.1 Market Size Equivalence

We show size and trade shocks have nearly isomorphic results in the absence of fixed export costs for most outcomes.

6.1.1 Free Trade Areas

We first illustrate the basic equivalence between liberalization and market size with a simple example. Consider an initial equilibrium where point 0 in Figure 5 represents a closed economy of size L. In a fully integrated economy of size $N \times L$ the new equilibrium can be represented by point 1, which results from shifting $\tilde{P}$ down as the result of increasing size by a factor N; this is also the equilibrium if we only allow for free trade in goods between N symmetric countries.[23]

In this setting (and without fixed exporting costs), an advalorem export cost factor $\tau \geq 1$ enters $\tilde{P}$ similarly to L, as we show explicitly below. So an identical reduction in τ between N symmetric countries from infinity (autarky) to $\tau = 1$ (free trade) can also be represented by Figure 5. If we interpret N as a measure of a continuum of countries then we can derive elasticities with respect to κ: the fraction of countries with free trade. The income elasticity with respect to κ is equal to ϵ_W^L given in proposition 5. Moreover, the effects of κ are similar to those of L in propositions 6, 7, 9, and 10; since in them L works only through the price index and specialization.[24]

In proposition 5 we show that size expansions have larger income effects for countries in the endogenous specialization range relative to those below it. Combining this and the isomorphism implies that expansions of free trade areas can have heterogeneous and threshold effects, which are larger if the expansion of the area is sufficiently large (or between countries with sufficiently advanced intermediate technology ϕ). This provides one possible rationale for mega-regional deals.

6.1.2 Trade Liberalization

The liberalization-size isomorphism described above extends to partially integrated economies: those with iceberg export costs, $\tau \geq 1$. In this setting the price index is still given by (21) but with $\tilde{L} \equiv L \times l^*$, where l^* is the **trade market size factor**, which we derive below as

$$l^* \equiv 1 + (N-1)\,\tau^{1-\sigma} \geq 1. \tag{39}$$

[23] Free trade and identical preferences imply a firm sets the same price across markets but faces higher aggregate expenditure, NX, and competition (lower P) relative to autarky. Firms in each country with a given productivity make the same specialization decisions. Similar prices, technologies, and size imply that in the free trade equilibrium, which replicates the fully integrated economy outcome, the income-expenditure constraints hold for individual countries.

[24] The main difference is that L has a direct effect on the mass of each country's firms, reflected in the first term of $\epsilon_M^L = 1 + \epsilon_M^s \cdot \frac{d \ln s_1}{d \ln L}$, whereas κ does not.

This factor is increasing in the number of trading partners and decreasing in export costs and spans the two special cases in the previous section: free trade case, $\tau = 1$, and $l^* = N$, and autarky, $\tau = \infty$ and $l^* = 1$.

To derive l^*, note that total operating profits from domestic and foreign sales are now

$$\begin{aligned} \tilde{\pi}_i(\varphi) &= \tilde{\sigma} X P^{\sigma-1} \left[c_i(\varphi)\right]^{1-\sigma} + \tilde{\sigma}\,(N-1)\, X^*(P^*)^{\sigma-1} \left[\tau c_i(\varphi)\right]^{1-\sigma} \\ &= \tilde{\sigma}\left[l^* X\right] P^{\sigma-1} \left[c_i(\varphi)\right]^{1-\sigma}, \end{aligned} \tag{40}$$

where the export profit in the first line reflects the $N-1$ additional markets where sales incur an extra marginal cost. The second equality reflects the symmetry, which implies identical aggregate expenditures, $X = X^*$, at home (no asterisk) and each foreign market, (asterisk). In the absence of fixed export costs all firms export, so all firms have the same profit expression. Moreover, the profit expression is identical to what we had under autarky except that it reflects "effective expenditure" $\tilde{X} \equiv l^* X$. Thus the expressions for productivity thresholds are unchanged but now use $\tilde{X}$.

The expression for each country's total expenditure is still the first equality in (10); total sales of a given country's firms, Y, must equal world expenditure on them. In a symmetric equilibrium this implies we still obtain $Y = X$ and (10) holds: $X = L\bar{a}$. The free entry expression does not reflect expenditure so it remains unchanged.

In sum, the only difference in the price index is that the entry cutoff in (7) reflects $\tilde{X}$ so in (21) we use $\tilde{L} \equiv L \times l^*$, implying that P has the same elasticity with respect to L and l^*. The following proposition summarizes the equivalence between increases in L and **trade liberalization**—i.e., increases in l^* due to a reduction in τ or an increase in N.

Proposition 12. ***(Impacts of trade liberalization)***

A trade liberalization increases effective market size, $\tilde{L} = L \times l^*$, and has the same impacts as increased size in propositions 5 through 11; except that $\epsilon_M^{l^*} = \epsilon_M^L - 1$ and $\epsilon_{M_a}^{l^*} = \epsilon_{M_a}^L - 1$.

The proof is straightforward. First, the trade liberalization parameters, τ or N, only affect s_1 via P, which we already showed has equal elasticity for both L and l^*. Second, with one exception, the effects of size increases in propositions 5 through 11 work through the specialization elasticities and don't depend directly on L or l^* (or its determinants); so $\epsilon_o^s \cdot \frac{d\ln s_1}{d\ln l^*} = \epsilon_o^s \cdot \frac{d\ln s_1}{d\ln L}$ implies these shocks have similar impacts. The exception is the mass of firms, which depends directly on a country's L but not on l^*, hence $\epsilon_M^{l^*} = \epsilon_M^L - 1$, with a similar modification for active firms.

We summarize some of the key results from proposition 12 here. Under endogenous specialization, trade liberalization implies (i) larger real income gains than alternative models with homogeneous specialization (proposition 5); (ii) an increase in the aggregate variable cost share for intermediates and a decrease for labor (proposition 6); (iii) increased selection (proposition 8); (iv) increased concentration in profit (proposition 9) and sales (proposition 10) and (v) a FOSD of surviving firm productivity distributions (proposition 11). All effects in (ii), (iii), (iv), (v) are absent in similar trade models with exogenous specialization and no fixed costs of exporting.

6.2 Intermediate Trade Share

Estimation and calibration often makes use of trade shares and costs. Thus we note two results related to these. First, using proposition 12 and (39) we can obtain elasticities in terms of τ, e.g. for real income we obtain

$$\epsilon_W^{\tau} = \epsilon_W^{l^*} \cdot \frac{d \ln l^*}{d \ln \tau} = \epsilon_W^{L} \cdot \left[- (\sigma - 1) \left(1 - 1/l^*\right) \right]. \tag{41}$$

Second, the trade share of intermediates increases with trade liberalization. Since the consumption and intermediates bundle are the same, the intermediate share of total trade, v, is simply the intermediate share of expenditures:

$$v = \frac{\left[(\sigma - 1)/\sigma \right] \bar{\alpha} Y}{Y} = \frac{\sigma - 1}{\sigma} \bar{\alpha}. \tag{42}$$

The elasticity of this share with respect to l^* is simply $\epsilon_{\bar{\alpha}}^{l^*}$, which is positive under endogenous specialization as shown in proposition 12—so trade liberalization increases v. Therefore, reductions in export costs, τ, common to final and intermediates can explain increases in v only under endogenous specialization.

7 Quantitative Implications

We have three objectives in the following simple calibration to U.S. manufacturing in 1987-2007. First, to assess if the model is consistent with first order changes in the data untargeted in the calibration. Second, to illustrate the relative importance of the selection and specialization mechanisms, and the extent of market inefficiency. Third, to quantify the effects of large shocks, such as moving from the 2007 equilibrium to autarky or a trade war on real income, intermediate trade, firm technology, and profit concentration and contrast to alternative models.

7.1 Calibration: Assumptions and Identification

We focus on a symmetric setting where the U.S. accounts for one fourth of world income ($N = 4$) and use it to calibrate the parameters in Table 2. We assume three potential intermediate technologies ($\alpha_i = 0, \alpha/2, \alpha$) so $\delta = \alpha/2$. Productivity follows a Pareto: $G(\varphi) = 1 - \left(\frac{\varphi_{\min}}{\varphi}\right)^k$, and technology parameters are constant as are all fixed costs, $\mathbf{f}$.

The baseline elasticity of substitution is $\sigma = 5$.[25] The following parameters are common to other models of heterogeneous firms and so we use the external standard values (L) and normalizations ($\varphi_{\min}, f_0, f_E$) listed in Table 2.[26] The most productive firms fully specialize in our model so their sales

[25] Melitz and Redding (2015) use a value of 4 and Costinot and Rodríguez-Clare (2014) use a value of 6.

[26] Our model is not neutral with respect to economic size but our calibration strategy allows us to find the "effective size" of the economy, which is a function of f_E, $\varphi_{\min}$, f_0, and ϕ. See the online appendix of the working paper version Limão and Xu (2021) for details.

Table 2: Calibration Parameters

Parameter	**Value**	**Source/data moment**
		Internal
Variable export cost	$\tau_0 = 2.28$, $\tau_T = 1.98$	Export intensity: 0.10, 0.163 (CM)
Maximum intermediate cost share	$\alpha = 0.746$	Average intermediate cost shares: $\bar{\alpha}_{t=0,T} = 0.699, 0.743$
Adoption fixed cost factor	$\hat{f} = 11.37$	Initial sales share of top $20V$ largest firms: 0.645 (CM)
TFP parameter	$\phi_0 = 0.209, \phi_T = 0.229$	Log growth in relative factor price: 0.383 (NBER-CES)
		External
Number of Countries	$N = 4$	Average US share of world income: 1/4
Potential Intermediate Technologies	$\alpha_i = 0, \alpha/2, \alpha$	N/A
Elasticity of substitution	$\sigma = 5$	
Productivity dispersion	$k = 5.67$	$\frac{k}{\sigma-1} = 1.42$ in Melitz and Redding (2015)
Minimum productivity *	$\varphi_{\min} = 1$	” ”
Fixed cost of entry and production *	$f_E = f_0 = 1$	” ”
Labor *	$L = L_0$	Initial Manufacturing employment: 17,718 (CM)

CM: Census of Manufacturing data for 1987 and 2007 (export intensity) and 1987 (sales share and employment). NBER CES database data moments are computed as three-year geometric averages in both the initial (1987–1989) and final period (2005–2007). * Normalizations, which affect the magnitude of ϕ but not other calibrated parameters or outcomes.

distribution is Pareto and thus we choose its parameter $k/(\sigma-1)=1.42$, which is consistent with sales distribution evidence.[27]

The calibration of variable export costs τ_t is also standard: it directly matches average firm exports as a share of sales, $(N-1)\tau_t^{1-\sigma}/\left(1+(N-1)\tau_t^{1-\sigma}\right)$, which we observe in the data in 1987 and 2007.

The remaining four parameters, α, $\hat{f}$, ϕ_0, and ϕ_T, are internally calibrated when solving for the equilibria. We sketch the data moments and two-loop algorithm here and provide details in Appendix D.1. The inner loop uses initial guesses for $\hat{f}$, α, and the specialization premia, $s_{1,t}$. Using these along with τ_t, the pre-assigned parameters and the expression in (12) we choose the unobserved maximum intensities α_t to match the observed average share in each period ($\bar{\alpha}_{t=0,T}$ =0.699, 0.743).[28]

The free entry condition is used to calculate the entry cutoff, $\bar{\varphi}_{e,t}$, and the zero cutoff profit condition yields P_t. From the specialization schedule (22), we obtain ϕ_t, and calculate the model-implied changes in relative factor prices $\Delta\ln\left(\frac{w}{P}\right)$.[29] If maximum intensity is constant and equal to the initial guess, $\alpha_{t=0,T}=\alpha$, and $\Delta\ln\left(\frac{w}{P}\right)=0.383$ (the change in relative intermediates costs to labor in the data) then the algorithm moves to the outer loop, otherwise it iterate over alternative values of $\hat{f}$ and $s_{1,t}$. The outer loop computes the top firms' sales share in the initial equilibrium and stops if it is equal to the one in the data in 1987 (0.645); otherwise it continues to iterate over α.[30]

7.2 Equilibrium, Parameters and External Validity

We illustrate the calibrated equilibria, discuss the resulting parameters and provide some external validity.

Figure 8 shows the calibrated equilibrium in the initial, E_0, and final period, E_T. The shift in the specialization and price schedules reflect the ϕ and τ shocks respectively. Each equilibrium is unique for the respective set of parameters and entails heterogeneous specialization. The change in specialization premium is 17.6 lp.

The model is exactly identified and replicates all the targeted data moments listed in the third column of Table 2. Thus we briefly discuss external validity by comparing (i) the internally calibrated parameters to other estimates and (ii) model predictions to key untargeted data moments.

7.2.1 Trade growth

The calibration implies a variable export cost factor in 2007 that is about 14 log points lower than 1987. This seems reasonable given the bilateral and multilateral liberalization in that 20-year period and the reductions in transport and information costs.

[27]For example, Kondo et al. (2018) estimate it to be 1.49 for US manufacturing in 1982 and 1.75 in 1992 using establishment data and around 1.15 using firms.

[28]We use non-energy material expenditure as a share of all costs: materials, labor and investment. In the appendix we examine robustness to alternative capital expenditure measures.

[29]From the specialization schedule (22), $\Delta\ln\left(\frac{w}{P}\right)=\frac{\Delta\ln s_{1,t}}{\alpha_0}-\Delta\ln\phi_t$.

[30]We measure the sales share as the fraction of total manufacturing sales done by the top 20 firms in each industry (equivalently an industry sales weighted measure of top 20 firm shares).

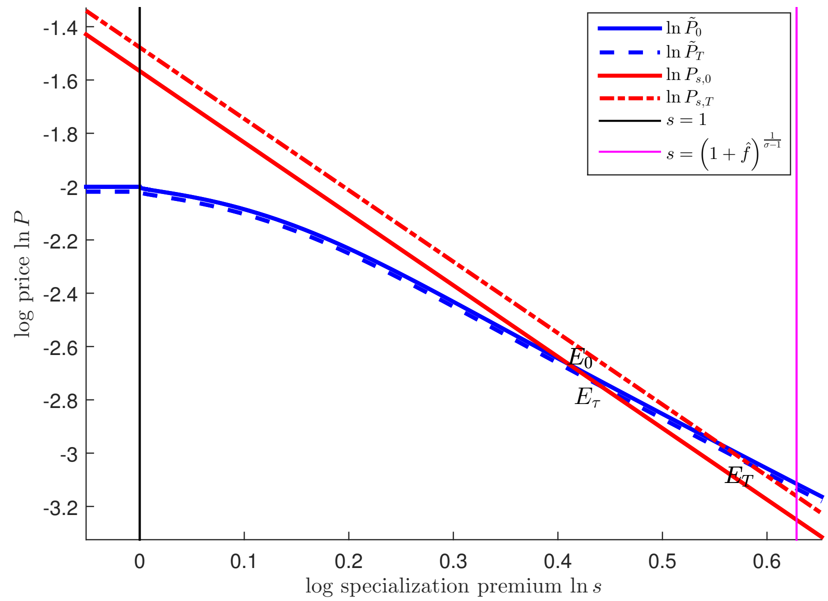


Figure 8: Calibrated equilibria

Given some stark simplifications of the model we ask if the calibrated model can explain key untargeted aggregate trade statistics. Specifically, the model assumes that all firms export and countries are symmetric. This implies that the growth of export intensity is common across firms and so equal to the aggregate growth of (i) exports to sales ratio (all firms export) and (ii) import penetration (country symmetry). The average of these two untargeted aggregate growth rates in the data is 2.6 log points/annum—very close to the 2.4 export intensity growth using the targeted firm level moments.[31]

7.2.2 Input shares and adoption

The calibrated model implies a change in the aggregate labor share of variable cost of -4.4 percentage points over the full period, which is similar to the corresponding untargeted change in the data: -4.0.[32]

The calibration replicates the average intermediate shares for the initial and final period, $\bar{\alpha}_{t=0,T}$; it implies a maximum intensity of 0.746, which as we would expect is above each of these targeted averages but also above the similarly computed $\bar{\alpha}_t$ for any other year in between. It also allows us to compute the fraction of adopters of each technology, $n = 0, 1, 2$. We obtain an increase from 0.24 to 0.77 in the share of adopters that prefer at least an intensity of 0.37 (and thus adopt either $n = 1, 2$).

[31] The aggregate growth rates employ NBER-CES data combined with the import and export data from the motivation section from year 1990 to 2007, hence the comparison of average yearly growth rates. These growth rates are computed within industry and aggregated with fixed weights, which is similar to how the intermediate shares are computed.

[32] We target the intermediate share in the data but not the labor share. While reductions in the latter mirror any intermediate share increase in the model, that is not necessarily the case in the data where total costs also reflect energy and investment.

7.2.3 Productivity Growth

The calibration yields an intermediate productivity improvement of $\Delta \ln \phi = 9$ lp. There is no direct data counterpart so we assess its reasonableness indirectly by comparing its implications (jointly with the rest of the parameters) for a measure of productivity it affects. In section 5.2.2 we defined the value added TFP and we can show that the change is $\Delta \ln \bar{\varphi}_{VA} = \Delta \ln (w/Pl_{sv}) = \Delta \ln \left(\frac{w}{P}\right) - \Delta \ln \left[\frac{(\sigma-1)(1-\bar{\alpha})}{1+(\sigma-1)(1-\bar{\alpha})}\right] = 45.8$ lp in the quantitative model, about 2.3 lp /annum from 1987 to 2007. In the data we compute it as real value added per production-worker hour and its change over the period is 3.65 lp/annum.[33] So the model explains almost 2/3 of this untargeted moment.[34]

7.2.4 Sales Distribution Change

The calibration targets the share of top 20 firm sales in 1987 but not its change relative to 2007. In Table 3 we show that the model-implied change is over two thirds of the observed when we consider changes in the top 20 and almost all observed in the top 50.[35]

Table 3: Changes in sales share: untargeted data vs. calibrated model (pp)

Moments (Change share, pp)	Data	Model (τ and ϕ shock)	Model (τ shock)
Top $8V$ firms	3.98	1.82	0.66
Top $20V$ firms	3.45	2.38	0.87
Top $50V$ firms	2.95	3.12	1.14

Notes: Comparison of untargeted changes in alternative sales concentration ratios in the data (1987-2007) with the model-implied changes due to the calibrated trade and productivity shocks. The V denotes the number of industries of US manufacturing industries with more than 100 firms used in the calculation. The aggregation of the concentration ratios across industries in each year t uses each industry's sales share. For the model-implied change, we compute the productivity cutoff that matches the fraction of firms corresponding to the top firm sales share in year 1987, and see how much changes in the sales share are generated at the same cutoff in the final equilibrium.

7.2.5 Model Versus Data Regressions: Specialization and Concentration

In Table 1 we found specialization increases with an industry's relative input price w_t^m/P_t^m, significantly so for more concentrated industries. To examine if the calibrated model can generate this finding we must bridge between the model and the multiple industries in the regression. The model could be extended to incorporate multiple industries if we allowed linkages only within industries (and Cobb-Douglas

[33] The within industry log change (1987-2007) weighted by 1997 industry total costs, divided by 20.

[34] One reason the calibration can match this is that it generates a large enough change in the specialization premium, about 18 lp, when we allow for trade and ϕ shocks, with the latter accounting for a large fraction of that growth. The premium would have increased by almost 15 lp in a counterfactual with only the shock to ϕ.

[35] That fraction is smaller when we consider the top 8, suggesting superstar firms may be specializing even more (the model could capture this by allowing more than 3 intermediate intensities).

aggregation across them on the consumption side). By doing so we could calibrate similarly to the single industry approach; however this would exactly replicate the intermediate shares, specialization premia and top sales share by industry in the data and thus the empirical relationship in Table 1. Therefore we follow an alternative approach. We start from the aggregate calibration, apply an exogenous shock to generate variation in the initial conditions and then apply a trade shock to it. We then use the model-generated outcomes to run a regression similar to Table 1.

We describe the basic approach here (for details see Appendix D.3). We take 500 draws of shocks (approximately the number of industries in the regression) from a lognormal to apply to the calibrated adoption cost, $\hat{f}$, and generate new economies that vary only in this dimension and thus in initial sales concentration. We re-calculate the initial equilibrium variables, denoted $\tilde{x}_0^m$. We then draw a trade shock (from a lognormal consistent with the mean and standard deviation in the data) and apply to each initial equilibrium. We compute the resulting final period outcomes holding all other parameters fixed. Finally, we compute the change in outcomes, $\Delta_t \tilde{x}^m$, and run a version of specification in column 1 of Table 1 in changes using the 500 model draws. We report the average OLS coefficients over 100 repetitions of this procedure, all of which are significant at the 1% level.

$$\Delta_t \ln\left(\frac{\widetilde{\alpha_t^m}}{\alpha_{w,t}^m}\right) = \left(0.40 + 1.67 \times \widetilde{Conc}^m\right) \Delta_t \ln\left(\frac{\widetilde{w_t^m}}{P_t^m}\right) + 0.015 + \Delta_t u_{tm} \tag{43}$$

We find a positive differential elasticity for industries with higher initial concentration. This is also the case in the data where the differential elasticity is 0.067 for industries with concentration 1 standard deviation above the mean in the data. The analogous differential using the model is 0.13 (the standard deviation of $\widetilde{Conc}^m$ is 0.08). In Appendix D.3 we argue that the data estimate is attenuated by measurement error of the concentration variable (top 20 share) relative to its model counterpart (share of sales by a fixed percentile of firms).

7.3 Equilibrium Trade Effects and Role of Endogenous Specialization

We now focus on the reduction in trade cost between 1987 and 2007 and examine its impacts on the initial equilibrium of the calibrated economy while holding intermediate technology (ϕ) fixed. The objectives are to illustrate the quantitative relevance of the model in studying trade shocks and contrast it with alternative models with fixed or no specialization.

7.3.1 Income, Cost Shares and Productivity

The first three rows of Table 4 illustrate the relative impacts of technology and trade shocks in explaining the observed changes between 1987 and 2007. We observe that shocks to ϕ generate most of the changes but the trade shock is still quantitatively relevant.

The third row of Table 4 shows the effects of the 14 lp decrease in trade costs. It increases real income

Table 4: Changes in real income, cost shares, and productivity (lp)

Model	Real income (W)	Intermediate share ($\bar{\alpha}$)	Labor VA share (l_{sv})	Labor VA productivity ($\bar{\varphi}_{VA}$)
Both shocks	37.9	6.10	-7.48	44.0
Technology shock	30.3	5.68	-6.89	35.9
Trade cost shock	8.34	1.96	-2.16	10.5
Fixed Specialization	6.03	0	0	6.03
No Specialization	1.81	N.A.	N.A.	1.81

Notes: The first three rows apply the calibrated changes in trade costs (τ) and technology (ϕ) jointly or separately to the initial equilibrium under endogenous specialization. The fourth and fifth present the outcomes with only the trade cost shock under fixed specialization and no specialization, respectively.

by over 8 lp. The aggregate intermediate share in variable cost increases by close to 2 lp; this is also the growth in the share of trade in intermediates (see (42)).

The share of labor in variable cost, l_{sc}, mirrors that of the intermediate share, so it falls by 1.96 lp , whereas the labor share in value added, l_{sv}, falls by 2.16. As we show in proposition 6, the growth in real productivity per variable production worker is given by the difference of real productivity (or income in first column) and that of the respective labor share (l_{sv} in the third), which yields the 10.5 increase in labor VA productivity.

The model provides a simple decomposition of the gains from liberalization. Using the equilibrium price expression in (20) that determines real income changes, we obtain

$$\begin{aligned} \Delta \ln W = & \underbrace{\frac{\Delta \ln \tilde{L}}{\sigma - 1}}_{\text{No Specialization}} + \underbrace{\frac{\Delta \ln \bar{a}}{\sigma - 1}}_{\text{Multiplier}} + \underbrace{\Delta \ln \bar{\varphi}_e}_{\text{Selection}} . \\ 8.34 = & \quad 1.81 \quad + \quad 0.63 \quad + \quad 5.9 \end{aligned} \tag{44}$$

Using the relationship in proposition 12 the 14 lp decrease in τ is equivalent to a size increase of $\Delta \ln \tilde{L} = 7.3$ lp. The no specialization component divides it by $\sigma - 1$. This yields a modest income increase of 1.81 lp, which is common to a variety of trade models.[36] The endogenous multiplier effect adds 0.63 lp and the selection effect is about 3.3 times larger than the basic no specialization effect. Overall, the income effect is 4.6 times higher with endogenous specialization than without any.

Existing models with fixed intermediate technology are also known to generate larger trade impacts. However, this is not all that our model captures. Using (21) we obtain a gain under fixed specialization of $\Delta \ln W^{FS} = \frac{1}{1-\alpha_e} \frac{\Delta \ln \tilde{L}}{\sigma - 1} = 6.0$, so it is about 3.3 times larger than without specialization—reflecting the fixed input multiplier $\frac{1}{1-\alpha_e}$ evaluated at the initial $\bar{\alpha}$ in the data. The endogenous specialization gain

[36] Specifically any other model that falls in the Arkolakis et al. (2012) class where trade elasticity and trade share growth are sufficient statistics to compute trade gains, these are captured by $\sigma - 1$ and $\Delta \ln \tilde{L}$ respectively in our setting without export selection and no specialization.

is still almost 1.4 times higher than what we obtain in the comparable fixed specialization calibration.[37]

In addition to the quantitative importance of endogenous specialization, the table also highlights the fact that similar models without that feature can't explain any changes in input shares, and thus, any changes to labor productivity in those models reflect only overall productivity.

7.3.2 Endogenous Multiplier and Size Elasticity

We use the model to compute the size elasticities for key outcomes, as derived in section 5.2. The resulting first order effects closely approximate the full model impacts.

The income and factor share size elasticities depend on the multiplier elasticity, $\delta \frac{d \ln \bar{\alpha}}{d \ln s_1}$, 0.27 in the calibrated initial equilibrium. Using this value and the formulas in (25) and (30), we obtain the elasticities listed in each column of Table 5.

Table 5: Size elasticity of real income, cost shares, and productivity

Model	Real income (ϵ^L_W)	Intermediate share ($\epsilon^L_{\bar{\alpha}}$)	Labor VA share ($\epsilon^L_{l_{sv}}$)	Labor VA productivity ($\epsilon^L_{\bar{\varphi}_{VA}}$)
Endogenous Specialization	1.16	0.32	-0.33	1.50
Fixed Specialization	0.83	0	0	0.83
No Specialization	0.25	N.A.	N.A.	0.25

Notes: Size elasticity of real income, cost shares, and productivity in the initial equilibrium.

In the second and third row we provide the elasticities for alternative models and note two points. First, we verify these are smaller in absolute value, as shown in section 5.2. Second, in the fixed or no specialization models the size elasticity is constant so when multiplied by the relevant size shock they reproduce the outcomes in the last rows of Table 4. For example, the 14 lp reduction in trade costs is equivalent to a size shock of 7.3 lp and thus a real income effect of around 6 lp using the fixed specialization elasticity. Under endogenous specialization the elasticity is variable but the same shock implies an income effect of 8.5 lp—very close to the 8.3 in the full model. This approximation is also close to the full effects for the factor shares and productivity.

In sum, armed with a multiplier elasticity, a value of σ, and an initial average intermediate share we can obtain close approximations of various impacts of size (or trade) shocks that, conditional on these sufficient statistics, are independent of productivity distribution and technology parameters. This suggests a high value to obtaining empirical estimates of the multiplier elasticity across countries as an alternative approach to quantifying these effects.

[37] In the working paper version Limão and Xu (2021) we show how we can calibrate this fixed specialization model to yield the same initial equilibrium welfare and all other endogenous variables while having all firms with the same intermediate intensity $\alpha_e = \bar{\alpha}$ as in our model but different TFP technology ϕ_i.

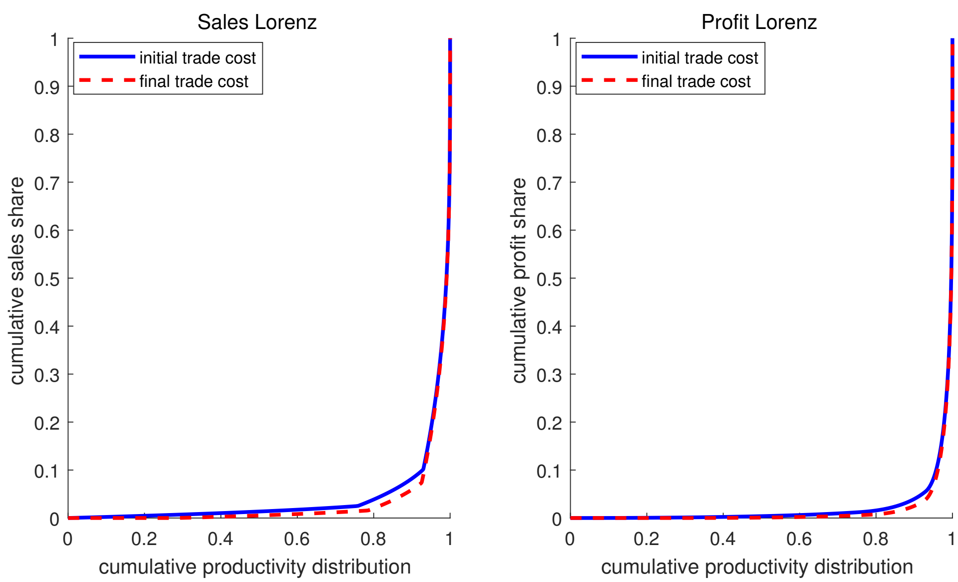
7.3.3 Selection, Entry and Adoption

The counterfactual of 14 lp reductions in trade costs also confirms and quantifies two of the insights in proposition 8. First, there is a selection effect with a 33.4 lp reduction in the fraction of active firms. Second, there is a 2.5 lp increase in entry, as predicted under Pareto. Moreover, in this calibration the selection is sufficiently large that the mass of active firms falls by 30.8 lp. The fraction of active firms adopting intermediates at $\alpha/2$ and α intensity increase by 13.3 and 39.2 lp, respectively.

As noted in section 5.2.3 a similar change in trade costs has zero effect on all these outcomes in the fixed and no specialization models.

7.3.4 Profits and Sales

Figure 9: Sales and profit Lorenz



Notes: Productivity distributions are truncated pareto with minimum productivity set to the initial level of the entry cutoff. Source: Authors' calculations.

Finally, we examine the calibration impact of trade costs on profit and sales distributions derived in propositions 9 and 10. Figure 9 shows the Lorenz curves. The dashed curves represent the equilibrium with lower trade costs and are below those of the initial equilibrium. Thus, we confirm that the increase in market size causes a MPS of sales and profits under a Pareto. Recall from the theory that this shock has no effect on distributions if there is no change in specialization.

7.4 Robustness

We check robustness to three alternative calibrations. First, we lower the elasticity of substitution from $\sigma = 5$ to 4. Second, instead of pre-assigning the Pareto shape parameter, k, we calibrate it by utilizing an additional moment: top 4 firms sales share. Lastly, we use capital expenditures from the EU KLEMS database to measure intermediates cost share. The conclusion is that the basic quantitative implications of the baseline calibration are robust to each alternative, and we provide them in detail in the working paper version Limão and Xu (2021).

7.5 Policy Experiments

We now examine counterfactual policy changes that can affect market size and efficiency.

7.5.1 Trade Taxes and Market Size

Since 2007 there has been increased trade protectionism, most recently characterized by Trump's trade war. In Table 6 we illustrate its implications for real income and intermediates if it applied to all U.S. partners and they retaliated symmetrically. We also compare it to going to autarky; which provides a benchmark number for gains from trade commonly used in the literature. All of the shocks are applied to the 2007 equilibrium.

Table 6: Real income and intermediates share under trade war and autarky

Policy scenario	Trade war ($\Delta \ln \tau$ =16 lp)	Autarky ($\Delta \ln \tau = \infty$)
Market size ($\tilde{L}$)	-8.02	-17.8
Intermediate share ($\bar{\alpha}$)	-0.46	-1.28
Real income: End. specialization (W)	-8.40	-18.8
Real income: No specialization (W^{ns})	-2.00	-4.45

Autarky entails a market size reduction of about 18 lp, which lowers the intermediate share by 1.3 lp. It reduces real income by almost 19 lp—over 4 times the amount it would in a setting without specialization.

For comparison with the rest of the quantification we model the trade war as a symmetric trade cost increase of 16 lp.[38] This magnitude reflects the increase in the U.S. average tariff rate factor on Chinese imports from about 1.03 to 1.21 between early 2018 and the end of 2019.[39] The implied market size reduction is almost half that of autarky and therefore so is the income reduction under no specialization. Under endogenous specialization the intermediate share falls by 0.5 lp and real income by 8 lp.

[38] We abstract from the fact that the tariff has revenue and its elasticity is slightly different from a pure export cost.

[39] This also uses the fact that US tariffs are applied to the cost at the border so the growth in the overall trade cost factor is $\ln \frac{1.21}{1.03}$, independent of the initial value of other costs.

7.5.2 Specialization Taxes and Market Inefficiency

In the qualitative section we note that a market inefficiency caused by a specialization externality is possible. We now show it is present in the initial calibrated equilibrium and can be eliminated by a simple government policy that yields a large income gain if specialization is as low as in 1987.

We consider a proportional tax on the operational cost of the less specialized technologies,

$$f_0' = (1 + \text{tax})f_0, \quad f_1' = (1 + \text{tax})f_0(f_a)^{\delta}, \quad f_2' = f_0(f_a)^{2\delta},$$

which increases the incentive to adopt the most specialized one.

A particularly simple policy is a prohibitive tax rate. This policy ensures that all firms optimally fully specialize and thus no tax revenues are collected.[40] We compute the required rate for this behavior in the 1987 equilibrium to be 118%. Since the policy is not optimally chosen there is no a priori reason that it must increase real income as shown in section 5.1. However, we find that it does generate a 5.2 lp gain—indicating a sizeable specialization externality. This differs from the standard input-output externality, which is captured in the fixed specialization framework. In fact, in the latter case the calibration shows that a similar prohibitive tax generates a 3.1 lp income *loss*.[41] This divergence in outcomes further highlights the importance of understanding endogenous specialization and the role for government policies to minimize the associated market inefficiencies.

We perform a similar experiment starting in 2007 and find it has a negligible income impact, 0.5 lp. This suggests that trade and productivity improvements induced sufficient specialization to eliminate the under-adoption inefficiency.

8 Conclusion

We provide a tractable framework to analyze the implications of the classical ideas that larger markets allow for a finer division of labor and this division feeds back into larger market size. By focusing on the adoption of intermediates by firms with heterogeneous productivity we capture key features of modern economies and provide new insights on the impacts of market size and technology on the structure of production, firm concentration and income gains. International trade is a key determinant of size and thus of these outcomes, as the quantification illustrates.

Market size affects specialization due to firm-level increasing returns to scale arising from adopting intermediate-intensive technologies. The impacts are magnified in general equilibrium by an endogenous multiplier and a selection effect. We show analytically that increases in market size or a trade liberaliza-

[40] To determine the optimal rate we can distribute any tax revenue (TR) to consumers so the goods' market clearing condition would become $X = \frac{L+TR}{1-\frac{\sigma-1}{\sigma}\bar{\alpha}}$.

[41] Under fixed specialization all three technologies have the same intermediates share but different unit costs, so we implement the tax on the fixed operational costs of two technologies with higher unit costs as in the baseline, as we detail in the working paper.

tion imply (i) larger real income gains than alternative models with fixed specialization; (ii) an increase in the aggregate variable cost share for intermediates and a decrease for labor; (iii) increased firm selection; (iv) increased concentration in the profit and sales distributions; (v) a FOSD shift of TFPQ for surviving firms and (vi) an increase in the intermediate trade share. The effects in (ii)-(vi) are absent in similar models with exogenous specialization.

In the calibration we illustrate key analytical results; quantify the importance of trade and technology shocks in the U.S. in 1987-2007 and the role of selection and multiplier effects; we also examine counterfactual policy changes. The calibration yields reductions in trade costs and improvement in intermediate productivity in this period that allow us to match key targeted moments, and is consistent with untargeted moments of the data. A substantial fraction of the increase in specialization premium is due to the exogenous intermediate productivity. But the trade cost reduction still corresponds to an effective market size increase of over 7 lp and thus has significant effects including: (i) real income gains larger than without specialization (4.6 times) or fixed specialization (1.4 times); (ii) an increase in the intermediate share in production and trade of 2 lp and a reduction in the labor share of value added of similar magnitude—none of which possible in alternative models; (iii) substantial increases in the fraction of firms specializing and selection into production.

Two counterfactual experiments highlight the importance of trade and industrial policy. First, a tax that induces firms to specialize would increase real income; so the initial equilibrium is inefficient. Second, the impact in the 2007 economy of an increase in trade costs of 16 lp—similar to the recent trade war—reduces market size and real income substantially: almost half way to trade autarky.

The model nests the special cases of no specialization or homogeneous specialization technology; it also shows that increases in size or technology can explain how an economy can develop by endogenous specialization. Future research can extend this simple framework along interesting directions, e.g. adding capital owners to study redistribution; allowing variable markups so increases in concentration become a source of increasing market power; modelling multiple sectors and export selection to better match the data.

References

Acemoglu, D. (2002). Directed technical change. *Review of Economic Studies*, 69(4):781–809.

Acemoglu, D. and Azar, P. D. (2020). Endogenous production networks. *Econometrica*, 88(1):33–82.

Amiti, M. and Konings, J. (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia. *American Economic Review*, 97(5):1611–1638.

Antras, P., Fort, T. C., and Tintelnot, F. (2017). The margins of global sourcing: Theory and evidence from US firms. *American Economic Review*, 107(9):2514–2564.

Arkolakis, C., Costinot, A., and Rodríguez-Clare, A. (2012). New trade models, same old gains? *American Economic Review*, 102(1):94–130.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2(3):244–263.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *Quarterly Journal of Economics*, 135(2):645–709.

Bajgar, M., Berlingieri, G., Calligaris, S., Criscuolo, C., and Timmis, J. (2019). Industry concentration in Europe and North America. *OECD Productivity Working Papers*.

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

Becker, G. S. and Murphy, K. M. (1992). The division of labor, coordination costs, and knowledge. *Quarterly Journal of Economics*, 107(4):1137–1160.

Berlingieri, G., Blanchenay, P., and Criscuolo, C. (2017). The great divergence(s). (39).

Bernard, A. B., Jensen, J. B., and Lawrence, R. Z. (1995). Exporters, jobs, and wages in US manufacturing: 1976-1987. *Brookings Papers on Economic Activity. Microeconomics*, 1995:67–119.

Blaum, J., Lelarge, C., and Peters, M. (2018). The gains from input trade with heterogeneous importers. *American Economic Journal: Macroeconomics*, 10(4):77–127.

Bloom, N., Draca, M., and Van Reenen, J. (2015). Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity. *The Review of Economic Studies*, 83(1):87–117.

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of MERCOSUR on Argentinian firms. *American Economic Review*, 101(1):304–340.

Caliendo, L. and Parro, F. (2015). Estimates of the trade and welfare effects of NAFTA. *Review of Economic Studies*, 82(1):1–44.

Costinot, A. and Rodríguez-Clare, A. (2014). Trade theory with numbers: Quantifying the consequences of globalization. In *Handbook of International Economics*, volume 4, pages 197–261. Elsevier.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, 84(2):445–510.

Dhingra, S. and Morrow, J. (2019). Monopolistic competition and optimum product diversity under firm heterogeneity. *Journal of Political Economy*, 127(1):196–232.

Dhyne, E., Kikkawa, A. K., Mogstad, M., and Tintelnot, F. (2021). Trade and domestic production networks. *The Review of Economic Studies*, 88(2):643–668.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–1779.

Elsby, M. W., Hobijn, B., and Şahin, A. (2013). The decline of the US labor share. *Brookings Papers on Economic Activity*, 2013(2):1–63.

Ethier, W. J. (1982). National and international returns to scale in the modern theory of international trade. *American Economic Review*, 72(3):389–405.

Feenstra, R. C. and Hanson, G. H. (1999). The impact of outsourcing and high-technology capital on wages: estimates for the United States, 1979–1990. *Quarterly Journal of Economics*, 114(3):907–940.

Fieler, A. C., Eslava, M., and Xu, D. Y. (2018). Trade, quality upgrading, and input linkages: Theory and evidence from Colombia. *American Economic Review*, 108(1):109–146.

Fort, T. C. (2016). Technology and Production Fragmentation: Domestic versus Foreign Sourcing. *The Review of Economic Studies*, 84(2):650–687.

Grossman, G. M. and Rossi-Hansberg, E. (2008). Trading tasks: A simple theory of offshoring. *American Economic Review*, 98(5):1978–97.

Halpern, L., Koren, M., and Szeidl, A. (2015). Imported inputs and productivity. *American Economic Review*, 105(12):3660–3703.

Hummels, D., Ishii, J., and Yi, K.-M. (2001). The nature and growth of vertical specialization in world trade. *Journal of international Economics*, 54(1):75–96.

Johnson, R. C. and Noguera, G. (2012). Accounting for intermediates: Production sharing and trade in value added. *Journal of International Economics*, 86(2):224–236.

Johnson, R. C. and Noguera, G. (2017). A portrait of trade in value-added over four decades. *Review of Economics and Statistics*, 99(5):896–911.

Karabarbounis, L. and Neiman, B. (2014). The global decline of the labor share. *Quarterly Journal of Economics*, 129(1):61–103.

Kee, H. L. and Tang, H. (2016). Domestic value added in exports: Theory and firm evidence from china. *American Economic Review*, 106(6):1402–36.

Kondo, I. O., Lewis, L. T., and Stella, A. (2018). On the US firm and establishment size distributions. *Mimeo*.

Krugman, P. and Venables, A. J. (1995). Globalization and the Inequality of Nations*. *The Quarterly Journal of Economics*, 110(4):857–880.

Lileeva, A. and Trefler, D. (2010). Improved access to foreign markets raises plant-level productivity... for some plants. *Quarterly Journal of Economics*, 125(3):1051–1099.

Limão, N. and Xu, Y. (2021). Size, trade, technology and the division of labor. *NBER Working Paper No. 28969*.

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

Melitz, M. J. and Redding, S. J. (2014). Missing gains from trade? *American Economic Review*, 104(5):317–21.

Melitz, M. J. and Redding, S. J. (2015). New trade models, new welfare implications. *American Economic Review*, 105(3):1105–1146.

Murphy, K. M., Shleifer, A., and Vishny, R. W. (1989). Industrialization and the big push. *Journal of political Economy*, 97(5):1003–1026.

Oberfield, E. (2018). A theory of input-output architecture. *Econometrica*, 86(2):559–589.

Ramanarayanan, A. (2020). Imported inputs and the gains from trade. *Journal of International Economics*, 122:103260.

Rosenstein-Rodan, P. N. (1943). Problems of industrialisation of eastern and south-eastern europe. *Economic Journal*, 53(210/211):202–211.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*, volume 1. Strahan and Cadell, London, UK.

Yi, K.-M. (2003). Can vertical specialization explain the growth of world trade? *Journal of Political Economy*, 111(1):52–102.

Yi, K.-M. (2010). Can multistage production explain the home bias in trade? *American Economic Review*, 100(1):364–393.

Young, A. A. (1928). Increasing returns and economic progress. *Economic Journal*, 38(152):527–542.

A Cross Country Evidence on Production Specialization: 1997-2007

In the following, we use data from the Global Trade Analysis Project (GTAP) to provide some evidence on production specialization between 1997 and 2007. We define production specialization in a country c as any increase in the intermediates' share in production costs, i.e. $\Delta\alpha_c > 0$.[42] We denote the aggregate change over a set of countries as

$$\Delta\alpha = \sum_c \left[w_{c,07}\alpha_{c,07} - w_{c,97}\alpha_{c,97}\right], \tag{A.1}$$

where the $w_{c,t}$ represent production cost shares used to weight across countries. For the 67 countries available in the GTAP data we obtain an increase of 6.3 percentage points between 1997-2007 using manufacturing industries. We then decompose it into within and between country specialization as follows

$$\Delta\alpha = \underbrace{\sum_c \bar{w}_c \left[\alpha_{c,07} - \alpha_{c,97}\right]}_{\text{within}} + \underbrace{\sum_c \bar{\alpha}_c \left[w_{c,07} - w_{c,97}\right]}_{\text{between}} = 3.7 + 2.6,$$

where $\bar{w}_c = (w_{c,07} + w_{c,97})/2$, $\bar{\alpha}_c = (\alpha_{c,07} + \alpha_{c97})/2$ denote the average weights. The within change accounts for nearly 60% and the remaining reflects production re-allocation towards countries with higher intermediate shares. The simple average of $\Delta\alpha_c$ across countries is only 0.4 percentage points. Thus larger countries had larger increases in the intermediate share—suggesting a positive correlation between size and production specialization. International trade is one channel in our model that expands market size and access to inputs leading to increased specialization. Increases in specialization in this data are strongly correlated with imported intermediates. To see this we run a panel regression of $\Delta\alpha_c^m$ (the share change within country and each of the 9 GTAP manufacturing industries, indexed by m) on Δm_c^m (the change of the imported fraction of intermediates at the same level) and obtain:

$$\Delta\alpha_c^m = \underset{(0.04)}{0.49} \cdot \Delta m_c^m + a_c + a_m.$$

This evidence rejects a constant intermediate share production function, and suggests that share is increasing as imported intermediates are adopted.[43]

[42] Intermediates share is defined as expenditure on intermediates over total production cost, which in GTAP includes production costs on labor, capital, land, and intermediates.

[43] To address the concern that some of the correlation is mechanical we also instrument Δm_i^m with changes in tariffs and find a similar result.

B Two Stage Production

Our framework can be interpreted as a reduced form representation of alternative models of firm specialization. Here we provide one where increases in market size or trade integration induce within firm specialization in the sense of shifting its workers towards production of inputs (intermediates or tasks) in which they are more productive. The production has two stages. In the first stage, labor produces firm-specific inputs with the following technology

$$x_t = z_t l_t \; ; \; t \in \Omega^0 \tag{B.1}$$

where z_t is the productivity of each of the l_t units of labor and, Ω^0 denotes the set of n firm-specific inputs that an unspecialized firm must produce. In the second stage, the unspecialized firm produces the final good by aggregating labor and all n inputs with a constant returns to scale production:

$$y(\varphi) = \varphi l^{1-nd} \prod_{t \in \Omega^0} (x_t)^d, \tag{B.2}$$

where $nd \leq 1$. Firms are homogeneous in the production of first stage intermediates, but have heterogeneous assembly productivity in the second stage, captured by φ. We continue to assume a single type of worker with wage w. Normalizing the units of the inputs such that $\prod_{t \in \Omega^0} (z_t)^d = 1$ we see that the unit cost of production for the unspecialized firm is w/φ, as in the baseline model. Now allow a fixed cost of $f_0(f_a)^{rd}$ units of labor to replace $r \leq n$ of the firm-specific inputs with the market bundle. This bundle is similar to the baseline model: a CES aggregate of stage two goods in the economy with price equal to P. Denoting the set of intermediates the firm remains specialized in by Ω^r, we can write the unit cost of the firm as

$$c_r(\varphi) = \frac{w^{1-rd} P^{rd}}{\varphi \prod_{t \in \Omega^r} (z_t)^d}, \quad r = \{0, \ldots, n\} \tag{B.3}$$

Thus the unit cost structure is the same as in the baseline model with the following interpretations. The cost share of intermediates in the baseline model, α_r, is interpreted as the cost share of first stage inputs no longer performed internally, rd. The "productivity boost" parameter $\phi^{\alpha_r} = \prod_{t \in \Omega^r} (z_t)^d$, captures the productivity change from re-allocating workers in the first stage. We now summarize this and show how firms re-allocate labor focusing on the case where firms can choose a single arbitrary r. Without loss of generality we re-index inputs to be increasing in productivity, $z_1 < \ldots < z_n$.

Proposition B.1. *The baseline model is equivalent to a production process where firms can adopt a technology allowing them to outsource some first stage inputs. Firms that are sufficiently productive in assembly, $\varphi \geq \bar{\varphi}$, specialize in inputs $t \in \Omega^r = \{t \in \Omega^0 | t > r\}$ resulting in*

1. *a re-allocation of labor away from the r least productive inputs and increase in productivity: $\prod_{t \in \Omega^r} (z_t)^d > \prod_{t \in \Omega^0} (z_t)^d$;*

2. *a cost share reduction of labor in production equal to rd;*

3. *a specialization premium: $s_r \equiv c_0(\varphi)/c_r(\varphi) = \left(\prod_{t \in \Omega^r} (z_t)^d\right) \left(\frac{w}{P}\right)^{rd}$.*

C Proof of Propositions

C.1 Proof of proposition 1

1. Under costly specialization $f_I > f_0$ so firms specialize iff

$$\tilde{\pi}_I > \tilde{\pi}_0 \;\Leftrightarrow\; c_I < c_0 \;\Leftrightarrow\; s_I > 1 \Leftrightarrow s_1 > 1,$$

   where the first equivalence uses the definition of profits, the second one the definition of s_I, and the third the constant increments assumption, which implies that $s_I = (s_1)^I$ for any $I \geq 1$.

2. Using (9) we have that $\bar{\varphi}_{i+1} > \bar{\varphi}_i$ for all i iff

$$\frac{f_{i+1}}{f_i} \cdot \left(s_1^{-1}\right)^{\sigma-1} > 1 \;\Leftrightarrow\; (s_1)^{\sigma-1} < \frac{f_{i+1}}{f_i} \;\Leftrightarrow\; (s_1)^{\sigma-1} \leq \frac{f_{i+1} - f_i}{f_i} + 1 \;\Leftrightarrow\; s_1^{\sigma-1} < \hat{f} + 1$$

   If $s_1 > 1$ then at least some firms specialize (part 1), so for heterogeneous specialization we further require that the marginal producer to have lower profits under the most specialized available technology, n, relative to whatever other technology $i < n$ the entrant choses. The condition is the same as that for sorting:

$$\begin{aligned}
\tilde{\pi}_n\left(\bar{\varphi}_{i,e}\right) - \tilde{\pi}_i\left(\bar{\varphi}_{i,e}\right) &< w\left(f_n - f_i\right) \\
\frac{\tilde{\pi}_n\left(\bar{\varphi}_{i,e}\right)}{\tilde{\pi}_i\left(\bar{\varphi}_{i,e}\right)} - 1 &< \left(\frac{f_n}{f_i} - 1\right)\left(\frac{w f_i}{\tilde{\pi}_i\left(\bar{\varphi}_{i,e}\right)}\right) \\
\frac{\tilde{\pi}_n\left(\bar{\varphi}_{i,e}\right)}{\tilde{\pi}_i\left(\bar{\varphi}_{i,e}\right)} &< \frac{f_n}{f_i} \\
\left(\frac{c_n\left(\bar{\varphi}_{i,e}\right)}{c_i\left(\bar{\varphi}_{i,e}\right)}\right)^{1-\sigma} &< \left(\frac{f_n}{f_{n-1}}\right) \ldots \left(\frac{f_{i+1}}{f_i}\right) \\
\left(s_{n-i}\right)^{\sigma-1} &< \left(1 + \hat{f}\right)^{n-i} \\
\left(\frac{s_1^{\sigma-1}}{1 + \hat{f}}\right)^{n-i} &< 1
\end{aligned} \tag{C.1}$$

   The second line divides both sides by $\tilde{\pi}_i\left(\bar{\varphi}_{i,e}\right)$ and re-arranges the terms, the third uses the marginal entry condition, $w f_i = \tilde{\pi}_i\left(\bar{\varphi}_{i,e}\right)$, the fourth uses the profit expressions, the fifth uses the definition of s_I and the technology assumption for f_i, the last one uses the constant share increment assumption and re-arranges. solving for s_1 yields the condition in the proposition.

3. From part 2 we see the entrant is indifferent or prefers n to any i iff the inequality is reversed in equation (C.1) and thus iff $s_1^{\sigma-1} \geq 1 + \hat{f}$ where under equality we use the convention that n is adopted. □

C.2 Proof of Proposition 2

C.2.1 $\tilde{\alpha}$ in (12)

To obtain $\tilde{\alpha}$ we first derive an expression for $\frac{Y_i}{Y}$ and replace in $\bar{\alpha}$ at expression (11). The aggregate sales of technology i firms is

$$\begin{aligned} Y_i &= M \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \sigma \tilde{\pi}_i(\varphi) dG(\varphi) \\ &= M\sigma\tilde{\pi}_e(\bar{\varphi}_e) \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \frac{\tilde{\pi}_i(\varphi)}{\tilde{\pi}_e(\bar{\varphi}_e)} dG(\varphi) \\ &= M\sigma f_e \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \left(\frac{c_i(\varphi)}{c_e(\bar{\varphi}_e)} \right)^{1-\sigma} dG(\varphi) \end{aligned}$$

where the first line uses the definition $Y_i \equiv M \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} y_i(\varphi) dG(\varphi)$ and $y_i = \sigma\tilde{\pi}_i$ under monopolistic competition and CES. The second multiplies and divides by marginal entrant's profit. The third uses the entry cutoff in (7) for $\tilde{\pi}_e(\bar{\varphi}_e)$ and the profit expression in (5), which implies that relative operating profits depend only on relative unit costs. We then use the definition of unit cost in (2), definition $s_1 = \left(w/\frac{P}{\phi}\right)^{\delta}$, and the technology assumption that $\alpha_i - \alpha_e = \delta\,(i-e)$ to obtain

$$Y_i = M\sigma f_e \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \left(s_1^{(i-e)} \frac{\varphi}{\bar{\varphi}_e} \right)^{\sigma-1} dG(\varphi). \tag{C.2}$$

Replacing this and $Y = \sum_{i=0}^{n} Y_i$ in (11), and cancelling common terms we have

$$\bar{\alpha} = \sum_{i=0}^{n} \alpha_i \frac{s_1^{(i-e)(\sigma-1)} \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \varphi^{\sigma-1} dG(\varphi)}{\sum_{i=0}^{n} s_1^{(i-e)(\sigma-1)} \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \varphi^{\sigma-1} dG(\varphi)},$$

We then divide both the numerator and denominator by $\sum_{i=0}^{n} \int_{\bar{\varphi}i}^{\bar{\varphi}_{i+1}} \varphi^{\sigma-1} dG(\varphi)$, and defining $\frac{\hat{\varphi}_i}{\hat{\varphi}} \equiv \left[\int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \varphi^{\sigma-1} dG(\varphi) / \sum_{i=0}^{n} \int_{\bar{\varphi}i}^{\bar{\varphi}_{i+1}} \varphi^{\sigma-1} dG(\varphi)\right]^{\frac{1}{\sigma-1}}$ we obtain (12).

C.2.2 $\frac{d\bar{\alpha}}{ds_1} > 0$ iff $s_1 \in s$

1. Necessity. If $s_1 \notin$ s there is homogeneous specialization (Proposition 1) and thus $e = i$ for all i and thus $\bar{\alpha} = \alpha_e$. If $s_1^{\sigma-1} \geq 1 + \hat{f}$ then $\alpha_e = \alpha_n$ (Proposition 1) and is already at the maximum and can't increase with s_1. If $s_1 < 1$ there is no adoption. Alternatively, we see from (13), derived below, that $\frac{d \ln \tilde{\alpha}}{d \ln s_1} = 0$ if $\alpha_i = \alpha_e$ for all i.

2. Sufficiency. We show that both the intensive and extensive margin impacts in (13), derived below, are positive.

C.2.3 Deriving Multiplier Effect in (13)

We define the elasticity components in terms of each of the margins as follows:

$$\frac{d\ln\bar{\alpha}}{d\ln s_1} = \underbrace{\left.\frac{\partial\ln\bar{\alpha}}{\partial\ln s_1}\right|_{\bar{\varphi}_{i\geq e}}}_{\text{intensive}} + \underbrace{\sum_{i=e}^{n}\frac{\partial\ln\bar{\alpha}}{\partial\ln\bar{\varphi}_i}\frac{d\ln\bar{\varphi}_i}{d\ln s_1}}_{\text{extensive}}. \tag{C.3}$$

First, we express the intermediate cost share as

$$\bar{\alpha} \equiv \sum_{i=e}^{n}\lambda_i\alpha_i = \sum_{i=e}^{n}\lambda_i\left[\alpha_e + (i-e)\delta\right] = \alpha_e + \delta\sum_{i=e}^{n}(i-e)\lambda_i, \tag{C.4}$$

where for the last equality we use $\sum_{i=e}^{n}\lambda_i = 1$. Using (C.2), we re-express λ_i as

$$\lambda_i \equiv \frac{Y_i}{\sum_{r=e}^{n}Y_r} = \frac{Y_i\left(\bar{\varphi}_e\right)^{\sigma-1}/(M\sigma F_e)}{\sum_{r=e}^{n}Y_r\left(\bar{\varphi}_e\right)^{\sigma-1}/(M\sigma F_e)} \equiv \frac{\tilde{Y}_i}{\sum_{r=e}^{n}\tilde{Y}_r},$$

where $\tilde{Y}_i \equiv \frac{Y_i(\bar{\varphi}_e)^{\sigma-1}}{M\sigma F_e} = (s_1)^{(i-e)(\sigma-1)}\int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}}\varphi^{\sigma-1}dG(\varphi)$. Using expression (C.4), we can simplify the multiplier effect to

$$\frac{d\ln\bar{\alpha}}{d\ln s_1} = \frac{1}{\bar{\alpha}}\sum_{i=e}^{n}\lambda_i\left(\alpha_i - \bar{\alpha}\right)\frac{d\ln\tilde{Y}_i}{d\ln s_1}, \tag{C.5}$$

and we provide the details in the online appendix. We further decompose the effects on sales into the intensive and extensive margins:

$$\begin{aligned}
\frac{d\ln\tilde{Y}_i}{d\ln s_1} &= \left.\frac{\partial\ln\tilde{Y}_i}{\partial\ln s_1}\right|_{\bar{\varphi}_{i\geq e}} + \sum_{j=e}^{n}\frac{\partial\ln\tilde{Y}_i}{\partial\ln\bar{\varphi}_j}\frac{d\ln\bar{\varphi}_j}{d\ln s_1} \\
&= (\sigma-1)(i-e) + \sum_{j=e}^{n}\frac{\partial\ln\tilde{Y}_i}{\partial\ln\bar{\varphi}_j}\frac{d\ln\bar{\varphi}_j}{d\ln s_1}.
\end{aligned}$$

Substituting the above expression into (C.5) and further simplifying the multiplier effect we obtain (13) in the last equality:

$$\frac{d\ln\bar{\alpha}}{d\ln s_1} = \frac{\sigma-1}{\delta\bar{\alpha}}\sum_{i=e}^{n}\lambda_i\left(\alpha_i-\bar{\alpha}\right)^2 + \frac{1}{\bar{\alpha}}\sum_{i=e}^{n}\lambda_i\left(\alpha_i-\bar{\alpha}\right)\left(\sum_{j=e}^{n}\frac{\partial\ln\tilde{Y}_i}{\partial\ln\bar{\varphi}_j}\frac{d\ln\bar{\varphi}_j}{d\ln s_1}\right),$$

and we provide the details of the derivation in the online appendix.

C.2.4 Positive Intensive Margin in (13):

$\frac{\sigma-1}{\delta\bar{\alpha}}\sum_{i=e}^{n}\lambda_i\left(\alpha_i-\bar{\alpha}\right)^2 > 0$ from inspection since $\sigma > 1$.

C.2.5 Positive Extensive Margin in (13):

We use sales $\tilde{Y}_i$ from (C.2) to simplify the extensive margin as

$$
\begin{aligned}
&\sum_{i=e}^{n}\left(\alpha_i-\bar{\alpha}\right)\lambda_i\sum_{j=e}^{n}\frac{\partial\ln\tilde{Y}_i}{\partial\ln\bar{\varphi}_j}\frac{d\ln\bar{\varphi}_j}{d\ln s_1}\\
=&-\frac{\left(\alpha_e-\bar{\alpha}\right)\left(\bar{\varphi}_e\right)^{\sigma}g\left(\bar{\varphi}_e\right)}{\sum_{i=e}^{n}(s_1)^{(i-e)(\sigma-1)}\int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}}\varphi^{\sigma-1}dG(\varphi)}\frac{d\ln\bar{\varphi}_e}{d\ln s_1}\\
&+\sum_{i=e+1}^{n}\frac{\left(\alpha_{i-1}-\bar{\alpha}\right)(s_1)^{(i-e-1)(\sigma-1)}-\left(\alpha_i-\bar{\alpha}\right)(s_1)^{(i-e)(\sigma-1)}}{\sum_{i=e}^{n}(s_1)^{(i-e)(\sigma-1)}\int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}}\varphi^{\sigma-1}dG(\varphi)}\left(\bar{\varphi}_i\right)^{\sigma}g\left(\bar{\varphi}_i\right)\frac{d\ln\bar{\varphi}_i}{d\ln s_1},
\end{aligned}
$$

and we provide the details of derivation in the online appendix. Observe that the first term is positive as $\alpha_e<\bar{\alpha}$ and $\frac{\partial\ln\bar{\varphi}_e}{\partial\ln s_1}>0$. The second term can be expressed as

$$
\begin{aligned}
&\sum_{i=e+1}^{n}\frac{(s_1)^{(i-e-1)(\sigma-1)}\left(\bar{\varphi}_i\right)^{\sigma}g\left(\bar{\varphi}_i\right)}{\sum_{i=e}^{n}(s_1)^{(\sigma-1)}\int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}}\varphi^{\sigma-1}dG(\varphi)}\left[\left(\alpha_{i-1}-\bar{\alpha}\right)-\left(\alpha_i-\bar{\alpha}\right)(s_1)^{(\sigma-1)}\right]\frac{d\ln\bar{\varphi}_i}{d\ln s_1}\\
=&\sum_{i=e+1}^{n}\frac{\delta(s_1)^{(i-e-1)(\sigma-1)}\left(\bar{\varphi}_i\right)^{\sigma}g\left(\bar{\varphi}_i\right)}{\sum_{i=e}^{n}(s_1)^{(\sigma-1)}\int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}}\varphi^{\sigma-1}dG(\varphi)}\left[\left(i-1-\frac{\bar{\alpha}}{\delta}\right)-\left(i-\frac{\bar{\alpha}}{\delta}\right)(s_1)^{(\sigma-1)}\right]\frac{d\ln\bar{\varphi}_i}{d\ln s_1}\\
=&\sum_{i=e+1}^{n}\frac{\delta(s_1)^{(i-e-1)(\sigma-1)}\left(\bar{\varphi}_i\right)^{\sigma}g\left(\bar{\varphi}_i\right)}{\sum_{i=e}^{n}(s_1)^{(\sigma-1)}\int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}}\varphi^{\sigma-1}dG(\varphi)}\left[\left(\frac{\bar{\alpha}}{\delta}-i+1\right)-\left(\frac{\bar{\alpha}}{\delta}-i\right)(s_1)^{(\sigma-1)}\right]\left(-\frac{d\ln\bar{\varphi}_i}{d\ln s_1}\right)
\end{aligned}
$$

Next, we show that for each $i\geq e+1$,

$$
\left[\left(\frac{\bar{\alpha}}{\delta}-i+1\right)-\left(\frac{\bar{\alpha}}{\delta}-i\right)(s_1)^{\sigma-1}\right]\left(-\frac{d\ln\bar{\varphi}_i}{d\ln s_1}\right)>0. \tag{C.6}
$$

Using the expression for productivity thresholds, the effect of premium on cutoff is

$$
\frac{d\ln\bar{\varphi}_i}{d\ln s_1}=-\frac{(s_1)^{\sigma-1}}{(s_1)^{\sigma-1}-1}-(i-e-1)+\frac{\bar{\alpha}-\alpha_e}{\delta},\quad i\geq e+1,
$$

showing that $\frac{d\ln\bar{\varphi}_i}{d\ln s_1}$ is decreasing in i and $\frac{d\ln\bar{\varphi}_n}{d\ln s_1}<0$. Thus, there exists some $b\in[e,n-1]$ such that $\frac{d\ln\bar{\varphi}_b}{d\ln s_1}>0$ and $\frac{d\ln\bar{\varphi}_{b+1}}{d\ln s_1}<0$. This implies that

$$
\frac{d\ln\bar{\varphi}_i}{d\ln s_1}\begin{cases}>0 & \text{if } i\leq b\\ <0 & \text{if } i\geq b+1\end{cases}. \tag{C.7}
$$

We derive in the online appendix that the definition of b implies the following two expressions:

$$
\begin{aligned}
\frac{d\ln\bar{\varphi}_b}{d\ln s_1}>0 &\Rightarrow (s_1)^{\sigma-1}>\frac{\frac{\bar{\alpha}}{\delta}-b+1}{\frac{\bar{\alpha}}{\delta}-b};\\
\frac{d\ln\bar{\varphi}_{b+1}}{d\ln s_1}<0 &\Rightarrow \left(\frac{\bar{\alpha}}{\delta}-b-1\right)(s_1)^{\sigma-1}<\frac{\bar{\alpha}}{\delta}-b.
\end{aligned}
$$

last, we divide into two cases to show that (C.6) holds. *Case 1:* If $i \leq b$, then $\frac{\partial \ln \bar{\varphi}_i}{\partial \ln s_1} > 0$. Moreover,

$$\left(\frac{\bar{\alpha}}{\delta} - i + 1\right) - \left(\frac{\bar{\alpha}}{\delta} - i\right)(s_1)^{\sigma-1} < \left(\frac{\bar{\alpha}}{\delta} - i + 1\right) - \left(\frac{\bar{\alpha}}{\delta} - i\right)\left(\frac{\frac{\bar{\alpha}}{\delta} - i + 1}{\frac{\bar{\alpha}}{\delta} - i}\right) = 0,$$

where the inequality uses $\frac{\bar{\alpha}}{\delta} - i \geq \frac{\bar{\alpha}}{\delta} - b > 0$ and $(s_1)^{\sigma-1} > \frac{\frac{\bar{\alpha}}{\delta} - b + 1}{\frac{\bar{\alpha}}{\delta} - b} \geq \frac{\frac{\bar{\alpha}}{\delta} - i + 1}{\frac{\bar{\alpha}}{\delta} - i}$ for $i \leq b$. Thus,

$$\underbrace{\left[\left(\frac{\bar{\alpha}}{\delta} - i + 1\right) - \left(\frac{\bar{\alpha}}{\delta} - i\right)(s_1)^{\sigma-1}\right]}_{<0} \underbrace{\left(-\frac{d \ln \bar{\varphi}_i}{d \ln s_1}\right)}_{<0} > 0.$$

Case 2: If $i \geq b + 1$, then $\frac{\partial \ln \bar{\varphi}_i}{\partial \ln s_1} < 0$. We divide this scenario into two subcases: (a) If $\frac{\bar{\alpha}}{\delta} - i < 0$, then

$$\left(\frac{\bar{\alpha}}{\delta} - i + 1\right) - \left(\frac{\bar{\alpha}}{\delta} - i\right)(s_1)^{\sigma-1} = -\left(\frac{\bar{\alpha}}{\delta} - i\right)\left[(s_1)^{\sigma-1} - 1\right] + 1 > 0$$

as $s_1 > 1$. (b) Else if $\frac{\bar{\alpha}}{\delta} - i \geq 0$, then

$$\left(\frac{\bar{\alpha}}{\delta} - i + 1\right) - \left(\frac{\bar{\alpha}}{\delta} - i\right)(s_1)^{\sigma-1} > \left(\frac{\bar{\alpha}}{\delta} - i + 1\right) - \left(\frac{\bar{\alpha}}{\delta} - i\right)\left(\frac{\frac{\bar{\alpha}}{\delta} - i + 1}{\frac{\bar{\alpha}}{\delta} - i}\right) = 0,$$

where the inequality uses $\frac{\bar{\alpha}}{\delta} - i \geq 0$ and $(s_1)^{\sigma-1} < \frac{\frac{\bar{\alpha}}{\delta} - b}{\frac{\bar{\alpha}}{\delta} - b - 1} \leq \frac{\frac{\bar{\alpha}}{\delta} - i + 1}{\frac{\bar{\alpha}}{\delta} - i}$ for $i \geq b + 1$. Thus,

$$\underbrace{\left[\left(\frac{\bar{\alpha}}{\delta} - i + 1\right) - \left(\frac{\bar{\alpha}}{\delta} - i\right)(s_1)^{\sigma-1}\right]}_{>0} \underbrace{\left(-\frac{d \ln \bar{\varphi}_i}{d \ln s_1}\right)}_{>0} > 0$$

for both subcases. □

C.3 Proof of Proposition 3

C.3.1 Derivation of $\bar{\varphi}_e$ in (16)

Recall that $\tilde{\pi}(\varphi) = \tilde{\pi}_i(\varphi) - f_i$ for $\varphi \geq \varphi_e$ so we rewrite each of these terms in (14) starting with the average profits in each i technology.

$$\begin{aligned}
\int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \tilde{\pi}_i(\varphi) dG(\varphi) &= \tilde{\pi}_e(\bar{\varphi}_e) \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \frac{\tilde{\pi}_i(\varphi)}{\tilde{\pi}_e(\bar{\varphi}_e)} dG(\varphi) \\
&= f_e \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \left(\frac{c_i(\varphi)}{c_e(\bar{\varphi}_e)}\right)^{1-\sigma} dG(\varphi) \\
&= f_e \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \left(\frac{\varphi}{\bar{\varphi}_e} s_1^{(i-e)}\right)^{\sigma-1} dG(\varphi) \\
&= f_e (s_1)^{(i-e)(\sigma-1)} \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \left(\frac{\varphi}{\bar{\varphi}_e}\right)^{\sigma-1} dG(\varphi)
\end{aligned} \tag{C.8}$$

The first line multiplies and divides by marginal entrant's profit. The second uses the entry cutoff in (7) for $\tilde{\pi}_e(\bar{\varphi}_e)$, the profit expression in (5), which implies relative operating profits depend only on relative unit costs. The third uses the definition of unit cost in (2), of $s_1 = \left(w/\frac{P}{\phi}\right)^{\delta}$ and the technology assumption that $\alpha_i - \alpha_e = \delta\,(i-e)$. The last line factors out s_1. Second, using part 2 of technology assumption 1, $\frac{\Delta f_{i+1}}{f_i} = \hat{f}$ for all i, we have $f_i = f_e(1+\hat{f})^{i-e}$. Thus, replacing this expression for f_i and substitute expression (C.8) in (14) we obtain

$$\sum_{i=e}^{e+I}\left[(s_1)^{(i-e)(\sigma-1)}\int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}}\left(\frac{\varphi}{\bar{\varphi}_e}\right)^{\sigma-1}dG(\varphi) - \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}}(1+\hat{f})^{i-e}dG(\varphi)\right] = f_E/f_e \tag{C.9}$$

Further replacing $\bar{\varphi}_{i>e} \equiv \bar{\varphi}_{i>e}(\bar{\varphi}_e, \hat{f}, s_1)$ in the above expression we obtain the implicit solution for the entry cutoff in (16).

C.3.2 Proof of $\frac{d\bar{\varphi}_e}{ds_1} > 0$ iff $s_1 \in s$

1. Necessity. If $s_1 \notin$ s there is homogeneous specialization (proposition 1) and thus $e = i$ for all i. Expression (C.9) reduces to $\int_{\bar{\varphi}_e}^{\infty}\left[\left(\frac{\varphi}{\bar{\varphi}_e}\right)^{\sigma-1} - 1\right]dG(\varphi) = \frac{f_E}{f_e}$, which is independent of s_1.

2. Sufficiency. We can apply the implicit function theorem to (C.9) along with the equilibrium adoption cutoffs to establish $\frac{d\bar{\varphi}_e}{ds_1} > 0$ for $s_1 \in$ s. We take an alternative approach below: to derive and differentiate the equivalent expression, (15).

C.3.3 Derivation of (15) and elasticity $\frac{d\ln\bar{\varphi}_e}{d\ln s_1}$ in (17)

We proceed in three steps:

1. Rewrite free entry in (14) as a function of $\bar{\varphi}_i$ to obtain (15);

2. Take the derivative of (15) with respect to $\bar{\varphi}_i$.

3. Use the solutions for $\bar{\varphi}_i$ and their change in terms of s_1 and $\bar{\varphi}_e$, then simplify to obtain $\frac{d\ln\bar{\varphi}_e}{d\ln s_1}$.

To derive (15), we first rewrite the aggregate fixed cost component in (14). Defining $F_{i+1} \equiv f_{i+1} - f_i$,and the operating cost for non-adopters as $F_e = f_e$ we have

$$\sum_{i=e}^{n}\int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} f_i dG(\varphi) = \sum_{i=e}^{n}\int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}}\sum_{r=e}^{i} F_r dG(\varphi) = \sum_{i=e}^{n}\int_{\bar{\varphi}_i}^{\infty} F_i dG(\varphi), \tag{C.10}$$

where the first equality uses $f_i = \sum_{r=e}^{i} F_r$ and $n = e + I$. Using the expression for relative cutoffs, in the online appendix we show that

$$\left(\frac{\bar{\varphi}_i}{\bar{\varphi}_e}\right)^{\sigma-1} = \left(s_1^{\sigma-1} - 1\right)^{-1}\left(s_1^{-1}\right)^{(i-e-1)(\sigma-1)}\left(\frac{F_i}{F_e}\right). \tag{C.11}$$

Replacing this in the average profit component derived in (C.8) for $i > e$ and simplifying we obtain the first equality below

$$
\begin{aligned}
& f_e \sum_{i=e}^{n} (s_1)^{(i-e)(\sigma-1)} \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \left(\frac{\varphi}{\bar{\varphi}_e}\right)^{\sigma-1} dG(\varphi) \\
=& f_e \int_{\bar{\varphi}_e}^{\bar{\varphi}_{e+1}} \left(\frac{\varphi}{\bar{\varphi}_e}\right)^{\sigma-1} dG(\varphi) + \frac{{s_1}^{\sigma-1}}{s_1^{\sigma-1}-1} \sum_{i=e+1}^{n} F_i \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \left(\frac{\varphi}{\bar{\varphi}_i}\right)^{\sigma-1} dG(\varphi) \\
=& \sum_{i=e}^{n} F_i \int_{\bar{\varphi}_i}^{\infty} \left(\frac{\varphi}{\bar{\varphi}_i}\right)^{\sigma-1} dG(\varphi)
\end{aligned}
\tag{C.12}
$$

The second equality requires additional algebra (provided in the online appendix). It relies on rewriting the expressions so that we compute the average incremental profits if all firms above a certain $\bar{\varphi}_i$ upgrade to the next technology, $F_i \int_{\bar{\varphi}_i}^{\infty} \left(\frac{\varphi}{\bar{\varphi}_i}\right)^{\sigma-1} dG(\varphi)$. Subtracting (C.10) from (C.12) and equating to f_E we have the version of free entry in (**15**). Then we differentiate (15) with respect to each $\bar{\varphi}_i$ and obtain

$$
\sum_{i=e}^{n} \left[\frac{F_i}{\bar{\varphi}_i} \int_{\bar{\varphi}_i}^{\infty} \left(\frac{\varphi}{\bar{\varphi}_i}\right)^{\sigma-1} dG(\varphi) \right] d\bar{\varphi}_i = 0,
\tag{C.13}
$$

where we use the Leibniz integral rule for each technology i to obtain

$$
\begin{aligned}
& d\left[F_i \int_{\bar{\varphi}_i}^{\infty} \left(\left(\frac{\varphi}{\bar{\varphi}_i}\right)^{\sigma-1} - 1 \right) dG(\varphi) \right] \\
=& F_i \left[- \left(\left(\frac{\bar{\varphi}_i}{\bar{\varphi}_i}\right)^{\sigma-1} - 1 \right) g(\bar{\varphi}_i) - \frac{\sigma-1}{\bar{\varphi}_i} \int_{\bar{\varphi}_i}^{\infty} \left(\frac{\varphi}{\bar{\varphi}_i}\right)^{\sigma-1} dG(\varphi) \right] d\bar{\varphi}_i \\
=& - \left[F_i \frac{\sigma-1}{\bar{\varphi}_i} \int_{\bar{\varphi}_i}^{\infty} \left(\frac{\varphi}{\bar{\varphi}_i}\right)^{\sigma-1} dG(\varphi) \right] d\bar{\varphi}_i .
\end{aligned}
$$

After that, we obtain $d\bar{\varphi}_i$ by using $\left(\frac{\bar{\varphi}_i}{\bar{\varphi}_e}\right)^{\sigma-1}$ in (C.11):

$$
\begin{aligned}
d\bar{\varphi}_i &= \frac{\partial \bar{\varphi}_i}{\partial \bar{\varphi}_e} d\bar{\varphi}_e + \frac{\partial \bar{\varphi}_i}{\partial s_1} ds_1 \\
&= \left(s_1^{\sigma-1} - 1\right)^{\frac{1}{1-\sigma}} \left(s_1^{-1}\right)^{(i-e-1)} \left(\frac{F_i}{F_e}\right)^{\frac{1}{\sigma-1}} \left[d\bar{\varphi}_e - \bar{\varphi}_e s_1^{-1} \left(\frac{s_1^{\sigma-1}}{s_1^{\sigma-1}-1} + (i-e-1) \right) ds_1 \right].
\end{aligned}
\tag{C.14}
$$

Substituting $d\bar{\varphi}_i$ into (C.13), and after some algebraic manipulation (provided in the online appendix), we obtain

$$
\frac{d \ln \bar{\varphi}_e}{d \ln s_1} = \frac{\sum_{i=e+1}^{n} F_e (i-e)(s_1)^{(i-e)(\sigma-1)} \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \left(\frac{\varphi}{\bar{\varphi}_e}\right)^{\sigma-1} dG(\varphi)}{\sum_{i=e}^{n} \left[F_i \int_{\bar{\varphi}_i}^{\infty} \left(\frac{\varphi}{\bar{\varphi}_i}\right)^{\sigma-1} dG(\varphi) \right]}.
\tag{C.15}
$$

Finally, to obtain the elasticity in (17) we use the expression for aggregate sales of technology i firms, Y_i, derived in (C.2). Substituting it into (C.15) we can further simplify the selection effect to

$$\begin{aligned}\frac{d\ln\bar{\varphi}_e}{d\ln s_1} &= \frac{\sum_{i=e+1}^{n}(i-e)Y_i}{\sum_{i=e}^{n}Y_i} = \frac{1}{\delta}\frac{\sum_{i=e}^{n}\delta(i-e)Y_i}{\sum_{i=e}^{n}Y_i} \\ &= \frac{\bar{\alpha}-\alpha_e}{\delta},\end{aligned}$$

where the last equality uses technology assumption 1: $\alpha_i - \alpha_e = (i-e)\delta$, and $\bar{\alpha} \equiv \frac{\sum_{i=e}^{n}\alpha_i Y_i}{Y}$. $\square$

C.4 Proof of Proposition 4

1. Existence requires $P_s(s_1) = \tilde{P}(s_1)$ for some s_1 and follows from (i) the continuity of P_s (guaranteed by its definition) and of $\tilde{P}(s_1)$ over all s_1 (proved in lemma C.4) (ii) $P_s \in (0,\infty)$ and steeper than $\tilde{P}(s_1)$ at both extremes of s_1 i.e. for $s_1 \notin \mathbf{s}$, which implies some $\tilde{s}$ s.t. $\tilde{P}(\tilde{s},x) = P_s(\tilde{s},x)$. If there are multiple $\tilde{s}$ then define $\tilde{s}^{\min} = \min\tilde{s}$ and $\tilde{s}^{\max} = \max\tilde{s}$ and note that the slope condition below implies that

$$\tilde{P}(s,x) < P_s(s,x) \quad \text{if } s < \tilde{s}^{\min}, \tag{C.16}$$

$$\tilde{P}(s,x) > P_s(s,x) \quad \text{if } s > \tilde{s}^{\max}. \tag{C.17}$$

   We prove it by deriving the relationship between the slopes:

$$\frac{d\ln P_s}{d\ln s_1} = -\frac{1}{\delta} < -\frac{\alpha_n}{\delta} \leq \frac{d\ln\tilde{P}(s_1 \notin \mathbf{s})}{d\ln s_1} \tag{C.18}$$

   where the first equality uses (22), the second uses $\alpha_n < 1$. The third relies on the results in lemma C.4 (below) showing that in the range of s_1 with homogenous specialization there is no selection or endogenous multiplier effect of s_1. Thus, using (20) we obtain $\frac{d\ln\tilde{P}(s_1\leq 1)}{d\ln s_1} = -e$, the marginal producer's technology. Lemma C.4 (below) shows this is equal to the minimum available technology share unless there is full specialization in which case it is $\frac{d\ln\tilde{P}(s_1\notin\mathbf{s}\ \&\ s_1>1)}{d\ln s_1} = -n = -\frac{\alpha_n}{\delta}$. If the minimum share is zero then $\frac{d\ln\tilde{P}(s_1\leq 1)}{d\ln s_1} = 0$, as depicted in Figure 4.

2. (a) The argument is similar for either parameter x. An equilibrium with $\bar{\alpha} = \alpha_0$ exists if there is some x such that

$$\tilde{P}(s_1,x) = P_s(s_1,x),\ s_1 \leq 1. \tag{C.19}$$

   If it exists then it must be unique since in part 1 expression (C.18) we have shown that P_s is steeper in that range than $\tilde{P}$, which implies stability, defined as starting from any other s_1 in that range of x and returning to the same equilibrium. If $s_1 \leq 1$ then $\bar{\alpha} = \alpha_0$, and $\bar{\varphi}_e(s_1) = \bar{\varphi}_e$ is independent of s_1, L, and ϕ (it depends only on fixed costs, productivity distribution, and σ from the free entry condition $f_e\int_{\bar{\varphi}_e}^{\infty}\left[\left(\frac{\varphi}{\bar{\varphi}_e}\right)^{\sigma-1} - 1\right]dG(\varphi) = f_E$). Thus, we obtain a critical $\tilde{x}_0$ (L or ϕ) below which an equilibrium exists.

$$\left(\frac{f_e}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}}\left(1-\frac{\sigma-1}{\sigma}\bar{\alpha}(s_1)\right)^{\frac{1}{\sigma-1}}\left(s_1^e\bar{\varphi}_e(s_1)\right)^{-1}=\phi s_1^{-\frac{1}{\delta}}$$
$$\left(\frac{f_e}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}}\left(1-\frac{\sigma-1}{\sigma}\alpha_0\right)^{\frac{1}{\sigma-1}}\left(s_1^e\bar{\varphi}_e\right)^{-1}=\phi s_1^{-\frac{1}{\delta}}$$
$$\left(\frac{f_e}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}}\left(1-\frac{\sigma-1}{\sigma}\alpha_0\right)^{\frac{1}{\sigma-1}}\left(\phi\bar{\varphi}_e\right)^{-1}=s_1^{e-\frac{1}{\delta}}$$
$$\left(\frac{f_e}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}}\left(1-\frac{\sigma-1}{\sigma}\alpha_0\right)^{\frac{1}{\sigma-1}}\left(\phi\bar{\varphi}_e\right)^{-1}\geq 1$$
$$\tilde{L}_0\equiv\frac{f_e}{\tilde{\sigma}}\left(1-\frac{\sigma-1}{\sigma}\alpha_0\right)\left(\phi\bar{\varphi}_e\right)^{1-\sigma}\geq L$$

where we use $s_1\leq 1$ and $e-\frac{1}{\delta}<e-\frac{\alpha_n}{\delta}=e-n<0$. A similar definition can be obtained for $\tilde{\phi}_0\equiv\left(\frac{f_e}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}}\left(1-\frac{\sigma-1}{\sigma}\alpha_0\right)^{\frac{1}{\sigma-1}}\left(\bar{\varphi}_e\right)^{-1}$ and so we denote either by $\tilde{x}_0$. If $\frac{d\ln\tilde{P}(s_1\in\mathbf{s})}{d\ln s_1}>-\frac{1}{\delta}$ then the critical value in 2(a) is $x_0=\tilde{x}_0$ since at $s_1=1$ we have $\tilde{P}\left(1,\tilde{x}_0\right)\geq P_s\left(1,\tilde{x}_0\right)$ and the continuity along with this slope condition implies that

$$\tilde{P}\left(s_1,\tilde{x}_0\right)>P_s\left(s_1,\tilde{x}_0\right),\ s_1>1. \tag{C.20}$$

However, if $\frac{d\ln\tilde{P}(s_1\in\mathbf{s})}{d\ln s_1}\leq-\frac{1}{\delta}$ then there may exist other equilibria with $\bar{\alpha}>\alpha_0$ and $s_1>1$ such that

$$\tilde{P}\left(s_1,\tilde{x}_0\right)=P_s\left(s_1,\tilde{x}_0\right),\ s_1>1.$$

We can then lower either x untill we satisfy expression (C.20) with equality and we denote the value of either x as $\hat{x}_0$. Such $\hat{x}_0$ exists because (i) Decreases in L leave P_s schedule unchanged but shift $\tilde{P}$ schedule up proportionally ($\frac{\partial\ln\tilde{P}(s_1,L)}{\partial\ln L}=\frac{1}{1-\sigma}$ since neither $\bar{\alpha}\left(s_1\right)$ nor $\bar{\varphi}_e\left(s_1\right)$ depends on L except via s) (ii) Decreases in ϕ leave $\tilde{P}$ schedule unchanged but shift P_s schedule down proportionally ($\frac{\partial\ln P_s}{\partial\ln\phi}=1$) The above implies that $\hat{x}_0$ still satisfies $\tilde{P}\left(s_1,\hat{x}_0\right)=P_s\left(s_1,\hat{x}_0\right)$ with $s_1\leq 1$. Thus, we have $x_0=\min\left\{\hat{x}_0,\tilde{x}_0\right\}$.

(b) The argument is similar to 2(a) but the critical values are now defined by

$$\left(\frac{f_e}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}}\left(1-\frac{\sigma-1}{\sigma}\bar{\alpha}(s_1)\right)^{\frac{1}{\sigma-1}}\left(s_1^e\bar{\varphi}_e(s_1)\right)^{-1}=\phi s_1^{-\frac{1}{\delta}}$$
$$\left(\frac{f_n}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}}\left(1-\frac{\sigma-1}{\sigma}\alpha_n\right)^{\frac{1}{\sigma-1}}\left(s_1^n\bar{\varphi}_n\right)^{-1}=\phi s_1^{-\frac{1}{\delta}}$$
$$\left(\frac{f_n}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}}\left(1-\frac{\sigma-1}{\sigma}\alpha_n\right)^{\frac{1}{\sigma-1}}\left(\phi\bar{\varphi}_n\right)^{-1}=s_1^{n-\frac{1}{\delta}}$$
$$\left(\frac{f_n}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}}\left(1-\frac{\sigma-1}{\sigma}\alpha_n\right)^{\frac{1}{\sigma-1}}\left(\phi\bar{\varphi}_n\right)^{-1}\leq\left(1+\hat{f}\right)^{\frac{n-1/\delta}{\sigma-1}}$$
$$\tilde{L}_n\equiv\frac{f_n}{\tilde{\sigma}}\left(1-\frac{\sigma-1}{\sigma}\alpha_n\right)\left(\phi\bar{\varphi}_n\right)^{1-\sigma}\left(1+\hat{f}\right)^{\frac{1-\alpha_n}{\delta}}\leq L,$$

where $\bar{\varphi}_n$ is from the free entry condition $f_n \int_{\bar{\varphi}_n}^{\infty} \left[\left(\frac{\varphi}{\bar{\varphi}_n} \right)^{\sigma-1} - 1 \right] dG(\varphi) = f_E$, while the inequality uses $s_1 \geq \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$ under full specialization and $n - \frac{1}{\delta} < n - \frac{\alpha_n}{\delta} = n - n = 0$. Similarly, we obtain $\tilde{\phi}_n \equiv \left(\frac{f_n}{\tilde{\sigma}L}\right)^{\frac{1}{\sigma-1}} \left(1 - \frac{\sigma-1}{\sigma}\alpha_n\right)^{\frac{1}{\sigma-1}} (\bar{\varphi}_n)^{-1} \left(1+\hat{f}\right)^{\frac{1-\alpha_n}{\delta(\sigma-1)}}$ and so we denote either by $\tilde{x}_n$. If $\frac{d \ln \tilde{P}(s_1 \in \mathbf{s})}{d \ln s_1} > -\frac{1}{\delta}$ then the critical value in 2(b) is $x_n = \tilde{x}_n$ since $\tilde{P}(s_1, \tilde{x}_n) \leq P_s(s_1, \tilde{x}_n)$ when $s_1 = \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$ and the continuity along with this slope condition implies that $\tilde{P}$ schedule remains below P_s schedule for all $s_1 < \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$.

$$\tilde{P}(s_1, \tilde{x}_n) < P_s(s_1, \tilde{x}_n), \ s_1 < \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}. \tag{C.21}$$

However, if $\frac{d \ln \tilde{P}(s_1 \in \mathbf{s})}{d \ln s_1} \leq -\frac{1}{\delta}$ then there may exist other equilibria with $\bar{\alpha} < \alpha_n$. Suppose that is the case, so we have some s_1 such that

$$\tilde{P}(s_1, \tilde{x}_n) = P_s(s_1, \tilde{x}_n), \ s_1 < \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}.$$

Following similar reasoning as in part 2(a), we can then increase either x until we satisfy (C.21) with equality and we let $\hat{x}_n$ denote the value of either x that does so. Thus, we have $x_n = \max\{\hat{x}_n, \tilde{x}_n\}$.

3. Start with an economy where $x = \tilde{x}_n$ as defined in 2(b) so we have an equilibrium with $\bar{\alpha} = \alpha_n$, $\tilde{s}_1 = \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$. If as $s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$ we have

   (a) $\frac{d \ln \tilde{P}(x, s_1)}{d \ln s_1} > (<) - \frac{1}{\delta}$ then a decrease (increase) in x generates an equilibrium with $s_1 \in \mathbf{s}$ by shifting $\tilde{P}$ schedule up (down) or P_s schedule down (up).

   (b) $\frac{d \ln \tilde{P}(x, s_1)}{d \ln s_1} = -\frac{1}{\delta}$ then at $x = \tilde{x}_n$ there are multiple equilibria with $s_1 \in \mathbf{s}$ since $\tilde{P}$ and P_s overlap over some range

   A similar argument applies if we start at $x = \tilde{x}_0$, where now the slope is evaluated at $s_1^+ \to 1$: if $\frac{d \ln \tilde{P}(x, s_1)}{d \ln s_1} > (<) - \frac{1}{\delta}$, then an increase (decrease) in x yields a heterogeneous specialization equilibrium. To show the range where it exists is $x \in [x_0, x_n]$ we first define the two endpoints. Let x_0 be the point such that there is a unique no specialization equilibrium for $x \leq x_0$, and x_n be the point such that there is a unique full specialization equilibrium for $x \geq x_n$. The knife edge cases with $\frac{d \ln \tilde{P}(x, s_1)}{d \ln s_1} = -\frac{1}{\delta}$ at $x = \tilde{x}_n$ are in this range. To see that there exists a heterogeneous specialization equilibrium under $\frac{d \ln \tilde{P}(x, s_1)}{d \ln s_1} < -\frac{1}{\delta}$: If $\frac{d \ln \tilde{P}(x, s_1)}{d \ln s_1} < -\frac{1}{\delta}$ at $s_1^+ \to 1$ then from 2(a) there is some other equilibrium with $s > 1$ for $x \in [x_0, \tilde{x}_0)$. On the other hand if $\frac{d \ln \tilde{P}(x, s_1)}{d \ln s_1} < -\frac{1}{\delta}$ at $s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$ then from 2(b) there is some other equilibrium with $s < \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$ for $x \in (\tilde{x}_n, x_n]$. Finally if $\frac{d \ln \tilde{P}(x, s_1)}{d \ln s_1} > -\frac{1}{\delta}$ at both limits then a heterogeneous specialization equilibrium exists for some $x \in (x_0, x_n)$

4. In part 1 we have shown that $\frac{d \ln \tilde{P}}{d \ln s_1} > -\frac{1}{\delta}$ for $s_1 \notin \mathbf{s}$ and in 2(a) and 2(b) that $\bar{\alpha}$ is independent

of x and thus non-decreasing in it. If $\frac{d\ln\tilde{P}}{d\ln s_1} > -\frac{1}{\delta}$ for $s_1 \in \mathbf{s}$ then from part 2 we have $x_0 = \tilde{x}_0$ and $x_n = \tilde{x}_n$ and the equilibrium starting at either value is unique. We now show that is also the case for values in between. Suppose we start at $\tilde{x}_n$, then reductions in x reduce $\tilde{s}_1$ (part 3) and thus $\bar{\alpha}$ locally, we can apply same argument at new equilibrium as long as $\left.\frac{d\ln\tilde{P}}{d\ln s_1}\right|_{\tilde{s}_1} > -\frac{1}{\delta}$ and thus $\bar{\alpha}$ continues to fall until $\tilde{x}_0$. □

The price schedule is continuous for all $s_1 \in (0,\infty)$.

Proof. We show first that the price schedule is continuous under homogenous specialization: $s_1 \notin \mathbf{s}$. We can express the price schedule as

$$\tilde{P}(s_1) = \begin{cases} \left[L\frac{\tilde{\sigma}}{f_e}\left(1-\frac{\sigma-1}{\sigma}\alpha_e\right)^{-1}\right]^{1/(1-\sigma)} \cdot \left(s_1^e\bar{\varphi}_e\left(s_1\right)\right)^{-1} & \text{if } s_1 \in (0,1) \\ \left[L\frac{\tilde{\sigma}}{f_n}\left(1-\frac{\sigma-1}{\sigma}\alpha\right)^{-1}\right]^{1/(1-\sigma)} \cdot \left(s_1^n\bar{\varphi}_e\left(s_1\right)\right)^{-1} & \text{if } s_1 \in \left(\left(1+\hat{f}\right)^{\frac{1}{\sigma-1}},\infty\right) \end{cases}$$

The derivative with respect to specialization premium is

$$\frac{d\ln\tilde{P}(s_1)}{d\ln s_1} = \begin{cases} -e-\frac{d\ln\bar{\varphi}_e(s_1)}{d\ln s_1} & \text{if } s_1 \in (0,1) \\ -n-\frac{d\ln\bar{\varphi}_e(s_1)}{d\ln s_1} & \text{if } s_1 \in \left(\left(1+\hat{f}\right)^{\frac{1}{\sigma-1}},\infty\right) \end{cases}$$

$$\frac{d\ln\tilde{P}(s_1)}{d\ln s_1} = \begin{cases} -e & \text{if } s_1 \in (0,1) \\ -n & \text{if } s_1 \in \left(\left(1+\hat{f}\right)^{\frac{1}{\sigma-1}},\infty\right) \end{cases},$$

where the last step observes that under both no specialization and full specialization there are no selection effect so that $\frac{d\ln\bar{\varphi}_e(s_1)}{d\ln s_1} = 0$. Thus, $\tilde{P}(s_1)$ is continuous under homogeneous specialization. Under heterogeneous specialization $s_1 \in \left(1,\left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}\right)$, differentiate the price index expression with respect to s_1 we have

$$\begin{aligned} \frac{d\ln\tilde{P}(s_1)}{d\ln s_1} &= \frac{d\ln}{d\ln s_1}\left[\left(1-\frac{\sigma-1}{\sigma}\bar{\alpha}\right)^{\frac{1}{\sigma-1}}\right] - e - \frac{d\ln\bar{\varphi}_e\left(s_1\right)}{d\ln s_1} \\ \frac{d\ln\tilde{P}(s_1)}{d\ln s_1} &= -\frac{1}{\sigma-1}\frac{\frac{\sigma-1}{\sigma}\bar{\alpha}}{1-\frac{\sigma-1}{\sigma}\bar{\alpha}}\frac{d\ln\bar{\alpha}}{d\ln s_1} - e - \frac{\bar{\alpha}-\alpha_e}{\delta} \\ \frac{d\ln\tilde{P}(s_1)}{d\ln s_1} &= -\frac{\bar{\alpha}}{\sigma-(\sigma-1)\bar{\alpha}}\frac{d\ln\bar{\alpha}}{d\ln s_1} - \frac{\bar{\alpha}}{\delta}, \end{aligned} \tag{C.22}$$

where the second step uses $\frac{d\ln\bar{\varphi}_e(s_1)}{d\ln s_1} = \frac{\bar{\alpha}-\alpha_e}{\delta}$. Because the intermediate cost share is increasing in specialization premium, $\frac{d\ln\bar{\alpha}}{d\ln s_1} > 0$, the price schedule is decreasing in s_1 and continuous over $s_1 \in \mathbf{s}$. Last, the price schedule is continuous at $s_1 = 1$ (from no specialization to heterogeneous specialization) and $s_1 = \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$ (from heterogeneous specialization to full specialization). The intuition is that when $s_1^+ \to 1$, all firms use the least specialized technology as no firms have incentive to adopt, so it collapses to the no specialization equilibrium. Conversely, when $s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$, all firms adopt the most specialized technology so that it becomes the full specialization equilibrium. To prove this it is sufficient to show that both $\bar{\alpha}(s_1)$ and $\bar{\varphi}_e(s_1)$ are continuous at those points.

Continuity of $\bar{\alpha}(s_1)$ at $s_1 = 1$ and $s_1 = \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$

First note that $\lim_{s_1^- \to 1} \bar{\alpha}(s_1) = \alpha_e$ and $\lim_{s_1^+ \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}} \bar{\alpha}(s_1) = \alpha$. Under heterogeneous specialization,

$$\left(\frac{\bar{\varphi}_i}{\bar{\varphi}_e}\right)^{\sigma-1} = \frac{\hat{f}}{(s_1)^{\sigma-1}-1}\left[\frac{1+\hat{f}}{(s_1)^{\sigma-1}}\right]^{i-e-1}, \quad i > e.$$

Therefore, $\lim_{s_1^+ \to 1} \frac{\bar{\varphi}_i}{\bar{\varphi}_e} = \infty$ so that no firms adopt more specialized technology. This implies that $\lim_{s_1^+ \to 1} \lambda_i(s_1) = 0$ for $i > e$ and $\lim_{s_1^+ \to 1} \lambda_e(s_1) = 1$ so that $\lim_{s_1^+ \to 1} \bar{\alpha}(s_1) = \lim_{s_1^+ \to 1} \sum_{i=e}^{n} \lambda_i(s_1)\alpha_i = \alpha_e$. Thus, $\lim_{s_1^+ \to 1} \bar{\alpha}(s_1) = \lim_{s_1^- \to 1} \bar{\alpha}(s_1) = \alpha_e$ so that $\bar{\alpha}(s_1)$ is continuous at $s_1 = 1$.

Similarly, $\lim_{s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}} \frac{\bar{\varphi}_i}{\bar{\varphi}_e} = 1$ so that all firms adopt the most specialized technology. This implies that $\lim_{s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}} \lambda_i(s_1) = 0$ for $i < n$ and $\lim_{s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}} \lambda_n(s_1) = 1$ so that $\lim_{s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}} \bar{\alpha}(s_1) = \lim_{s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}} \sum_{i=e}^{n} \lambda_i(s_1)\alpha_i = \alpha_n$. Thus, $\lim_{s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}} \bar{\alpha}(s_1) = \lim_{s_1^+ \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}} \bar{\alpha}(s_1) = \alpha_n$ so that $\bar{\alpha}(s_1)$ is continuous at $s_1 = \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$.

Continuity of $\bar{\varphi}_e(s_1)$ at $s_1 = 1$ and $s_1 = \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$.

First note that $\lim_{s_1^- \to 1} \frac{d \ln \bar{\varphi}_e(s_1)}{d \ln s_1} = 0$ and under heterogeneous specialization, we have shown that $\frac{d \ln \bar{\varphi}_e(s_1)}{d \ln s_1} = \frac{\bar{\alpha}(s_1)-\alpha_e}{\delta}$. Applying the continuity of $\bar{\alpha}(s_1)$ we get differentiability of $\bar{\varphi}_e(s_1)$ at $s_1 = 1$, therefore, $\bar{\varphi}_e(s_1)$ is continuous at $s_1 = 1$. For the continuity of $\bar{\varphi}_e(s_1)$ at $s_1 = \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$ we use the free entry condition. First note that for $s_1^+ \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$ the free entry condition is the full specialization case

$$f_n \int_{\bar{\varphi}_n}^{\infty} \left[\left(\frac{\varphi}{\bar{\varphi}_n}\right)^{\sigma-1} - 1\right] dG(\varphi) = f_E,$$

and for $s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$ the free entry condition is the heterogeneous specialization case

$$\sum_{i=e}^{n} F_i \int_{\bar{\varphi}_i}^{\infty} \left[\left(\frac{\varphi}{\bar{\varphi}_i}\right)^{\sigma-1} - 1\right] dG(\varphi) = f_E.$$

We have shown that $\lim_{s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}} \bar{\varphi}_i = \bar{\varphi}_n$ for $i < n$. Therefore, taking the limit of the free entry condition we have

$$
\begin{aligned}
\lim_{s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}} \sum_{i=e}^{n} F_i \int_{\bar{\varphi}_i}^{\infty} \left[\left(\frac{\varphi}{\bar{\varphi}_i} \right)^{\sigma-1} - 1 \right] dG(\varphi) &= f_E \\
\int_{\bar{\varphi}_n}^{\infty} \left[\left(\frac{\varphi}{\bar{\varphi}_n} \right)^{\sigma-1} - 1 \right] dG(\varphi) \sum_{i=e}^{n} F_i &= f_E \\
f_n \int_{\bar{\varphi}_n}^{\infty} \left[\left(\frac{\varphi}{\bar{\varphi}_n} \right)^{\sigma-1} - 1 \right] dG(\varphi) &= f_E,
\end{aligned}
$$

so the two free entry conditions with $s_1^- \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$ and $s_1^+ \to \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$ are the same and therefore yield the same $\bar{\varphi}_n$, indicating that $\bar{\varphi}_n$ is continuous at $s_1 = \left(1+\hat{f}\right)^{\frac{1}{\sigma-1}}$. $\square$

In an equilibrium with heterogeneous specialization the marginal producer uses the least intermediate intensive technology available, i_0.

Proof. From proposition 1 we have heterogeneous specialization iff $s^{\sigma-1} \in (1, 1+\hat{f})$. The marginal producer has productivity threshold given by

$$
(\bar{\varphi}_{i,e})^{\sigma-1} = \frac{w f_i}{\tilde{\sigma} X} \cdot P^{1-\sigma} \cdot \left(w s_1^{-i} \right)^{\sigma-1}
$$

suppose that the producer is just indifferent between technology i and $i-1$. This implies a full specialization equilibrium because

$$
(\bar{\varphi}_{i,e})^{\sigma-1} = (\bar{\varphi}_{i-1,e})^{\sigma-1} \Leftrightarrow f_i \left(s_1^{-i} \right)^{\sigma-1} = f_{i-1} \left(s_1^{-i-1} \right)^{\sigma-1} \Leftrightarrow (s_1)^{\sigma-1} = \left(1+\hat{f} \right).
$$

We can continue applying this argument until we reach $i = i_0 + 1$ to derive that under heterogeneous specialization $e = i_0$. $\square$

C.5 Inefficiency

We provide details on how Figure 6 is constructed. We first show that at $s_1 = s_{\max}$ the price schedule is steeper at the left limit. Second, we derive the price schedule when only the most specialized technology is available. In the proof of Lemma C.4 we show that the price schedule and aggregate intermediate share ($\bar{\alpha}$) are continuous at $s_1 = s_{\max}$, and the slope of the price schedule is

$$
\frac{d \ln \tilde{P}(s_1)}{d \ln s_1} = \begin{cases} -\frac{\bar{\alpha}}{\sigma-(\sigma-1)\bar{\alpha}} \frac{d \ln \bar{\alpha}}{d \ln s_1} - \frac{\bar{\alpha}}{\delta} & \text{if } s_1 \in (1, s_{\max}) \\ -n & \text{if } s_1 > s_{\max}. \end{cases}
$$

Therefore,

$$
\lim_{s_1 \to s_{\max}^-} \frac{d \ln \tilde{P}(s_1)}{d \ln s_1} = - \frac{\alpha}{\sigma - (\sigma-1)\alpha} \left. \frac{d \ln \bar{\alpha}}{d \ln s_1} \right|_{s_1 = s_{\max}^-} - \frac{\alpha}{\delta} < -n = \lim_{s_1 \to s_{\max}^+} \frac{d \ln \tilde{P}(s_1)}{d \ln s_1},
$$

where the inequality uses the fact that the multiplier effect is positive under heterogeneous specialization $\frac{d \ln \bar{\alpha}}{d \ln s_1} > 0$, and by definition $\alpha \equiv n\delta$. Thus, the slope of price schedule is steeper left of $s_{\max}$ for some s_1.

Next we derive the price schedule when there is only the most specialized technology available and show that it coincides with the one under endogeneous specialization with $s_1 > s_{\max}$. Specifically, we express the equilibrium of this single-technology economy in terms of the specialization premium defined in our model. Note that all firms in this single-technology economy have the following unit cost function:

$$c(\varphi) = \frac{w^{1-\alpha}P^{\alpha}}{\varphi\phi^{\alpha}} = \frac{P^{\alpha}}{\varphi\phi^{\alpha}}, \tag{C.23}$$

with fixed cost f_n and wage normalized to $w = 1$. The price schedule can be derived from the zero profit condition for entrants ($\bar{\varphi}^{spec}$)

$$\bar{\varphi}^{spec} = \frac{P^{\alpha}}{P\phi^{\alpha}} \left(\frac{f_n}{\tilde{\sigma}Y^{spec}}\right)^{\frac{1}{\sigma-1}}, \tag{C.24}$$

where $Y^{spec} = L\left(1 - \frac{\sigma-1}{\sigma}\alpha\right)^{-1} = L\bar{a}$ is total sales. Thus, the price schedule is

$$(\tilde{P}^{spec})^{1-\alpha} = \frac{1}{\bar{\varphi}^{spec}\phi^{\alpha}} \left(\frac{f_n}{\tilde{\sigma}L\bar{a}}\right)^{\frac{1}{\sigma-1}}. \tag{C.25}$$

Using the definition of specialization premium, $s_1 \equiv \left(\phi\frac{1}{P}\right)^{\alpha/n}$, (C.25) becomes

$$P^{spec}(s_1) = \frac{1}{(s_1)^n\bar{\varphi}^{spec}} \left(\frac{f_n}{\tilde{\sigma}L\bar{a}}\right)^{\frac{1}{\sigma-1}}, \tag{C.26}$$

which is the same as the price schedule with full specialization, as shown in the proof of Lemma C.4. Thus, $\tilde{P}^{spec}(s_1)$ coincides with $\tilde{P}(s_1)$ as illustrated in Figure 6.

C.6 Proof of Proposition 5

To obtain (25) we first totally differentiate the equilibrium price expression (21) and noting that within an heterogeneous equilibrium range the marginal entrant always adopts the minimum technology, $\alpha_e = \alpha_0$, we have

$$\begin{aligned}
(1-\alpha_0)\frac{d\ln\bar{P}}{d\ln L} &= -\frac{1}{\sigma-1} + \frac{d}{d\ln L}\left(\ln\bar{a}\left(s_1\left(\bar{P}\right)\right)^{\frac{-1}{\sigma-1}} + \ln\left[\bar{\varphi}_e\left(s_1\left(\bar{P}\right)\right)\right)^{-1}\right] \\
\underbrace{-\frac{d\ln\bar{P}}{d\ln L}}_{\epsilon_W^L} &= \underbrace{\left[(\sigma-1)(1-\alpha_0)\right]^{-1}}_{\bar{\epsilon}_W^L} \\
&+ \underbrace{(1-\alpha_0)^{-1}\frac{d}{d\ln s_1}\left[\ln\bar{a}\left(s_1\left(\bar{P}\right)\right)^{\frac{1}{\sigma-1}} + \ln\bar{\varphi}_e\left(s_1\left(\bar{P}\right)\right)\right]}_{\epsilon_W^s}\cdot\frac{d\ln s_1}{d\ln L},
\end{aligned} \tag{C.27}$$

where we use the chain rule and the results that $\bar{\varphi}_e$ in (16) and $\tilde{\alpha}$ in (12) do not depend directly on L. Using the definition of s_1 we obtain $\frac{d\ln s_1}{d\ln L} = -\frac{d\ln\bar{P}}{d\ln L}\delta$; replacing above and solving for $-\frac{d\ln\bar{P}}{d\ln L} = \frac{d\ln W}{d\ln L}$ yields the expression for ϵ_W^L in (25). Applying a similar approach for ϕ

$$
\begin{aligned}
(1-\alpha_0)\frac{d\ln\bar{P}}{d\ln\phi} &= -\alpha_0 - \frac{d}{d\ln\phi}\left[\ln\bar{a}\left(s_1\left(\bar{P}\right)\right)^{\frac{1}{\sigma-1}} + \ln\left(\bar{\varphi}_e\left(s_1\left(\bar{P}\right)\right)\right)\right] \\
\underbrace{-\frac{d\ln\bar{P}}{d\ln\phi}}_{\epsilon_W^{\phi}} &= \underbrace{\frac{\alpha_0}{1-\alpha_0}}_{\bar{\epsilon}_W^{\phi}} + \epsilon_W^s \cdot \frac{d\ln s_1}{d\ln\phi},
\end{aligned}
$$

where we use the chain rule and the results that $\tilde{\varphi}_e$ in (16) and $\tilde{\alpha}$ in (12) do not depend directly on ϕ. Using the definition of s_1 we obtain $\frac{d\ln s_1}{d\ln\phi} = \delta\left(1 - \frac{d\ln\bar{P}}{d\ln\phi}\right)$; replacing above and solving for ϵ_W^{ϕ} yields the expression in (25).

Selection and multiplier decomposition in (26)

Using the definition of ϵ_W^s in (C.27) and the assumption that the minimum is $\alpha_0 = 0$

$$
\begin{aligned}
\epsilon_W^s &= \frac{1}{\sigma-1}\frac{d\ln\bar{a}\left(s_1\left(\bar{P}\right)\right)}{d\ln s_1} + \frac{d\ln\bar{\varphi}_e\left(s_1\left(\bar{P}\right)\right)}{d\ln s_1} \\
&= \frac{\bar{\alpha}}{\sigma-\bar{\alpha}\left(\sigma-1\right)}\frac{d\ln\bar{\alpha}}{d\ln s_1} + \frac{\bar{\alpha}}{\delta}
\end{aligned}
$$

and multiplying by δ we obtain (26). The second line uses the multiplier definition $\bar{a} = \left[1 - \frac{\sigma-1}{\sigma}\bar{\alpha}\left(s_1\right)\right]^{-1}$ and the expression for selection effect in (17) with $\alpha_e = \alpha_0 = 0$.

Positive elasticities

We have $\epsilon_W^L > 0$ since $\bar{\epsilon}_W^L > 0$ and $1 - \delta\epsilon_W^s > 0$ where the latter holds as

$$
1 - \delta\epsilon_W^s = 1 - \bar{\alpha} - \left(\frac{\delta\bar{\alpha}}{\sigma-\bar{\alpha}\left(\sigma-1\right)}\frac{d\ln\bar{\alpha}}{d\ln s_1}\right) > 0
$$

from the stability condition, which is equivalent to $\frac{\delta\bar{\alpha}}{\sigma-\bar{\alpha}(\sigma-1)}\frac{d\ln\bar{\alpha}}{d\ln s_1} < (1-\bar{\alpha})$ by simplifying (23). We have $\epsilon_W^{\phi} > 0$ since $\bar{\epsilon}_W^{\phi} + \delta\epsilon_W^s > 0$ and $1 - \delta\epsilon_W^s > 0$ under stability.

Comparing elasticities in (27) and (28)

The elasticity in an alternative model where firms must use a common technology $\bar{\alpha}_e$ is obtained by diferentiating (21), holding s_1 fixed. So for L it is $\epsilon_W^L|_{\bar{\alpha}_e \leq \bar{\alpha}, s_1} = \left[\left(\sigma-1\right)\left(1-\bar{\alpha}_e\right)\right]^{-1}$. Comparing to $\epsilon_W^L|_{s_1 \in \mathbf{s}}$ and using the result that $\alpha_e = \alpha_0$ and setting the latter to 0 we have the inequality in (27), which is satisfied if $\delta\epsilon_W^s\left(\bar{\alpha}\right) > \bar{\alpha}_e$. Since $\delta\epsilon_W^s\left(\bar{\alpha}\right) \geq \bar{\alpha}$ the inequality holds for any fixed intermediate share model where $\bar{\alpha}_e \leq \bar{\alpha}$. The result also applies to ϕ. Thus the welfare elasticity is higher and if we fix $\bar{\alpha}_e = \bar{\alpha}$ the difference is due to the multiplier effect. □

C.7 Proof of Proposition 6

To derive (30) we use (24) to write the size elasticities as $\epsilon_o^L = \bar{\epsilon}_o^L + \epsilon_o^s \cdot \frac{d\ln s}{d\ln L}$ and note that $\bar{\epsilon}_o^L = 0$ since conditional on $\bar{\alpha}$ all shares are independent of L and the same is true for $\bar{\alpha}$ conditional on s_1 (from $\tilde{\alpha}$ in (12) and $\bar{\varphi}_e$ in (16)). Also, $\epsilon_{\bar{\alpha}}^s \equiv \frac{d\ln\bar{\alpha}}{d\ln s_1}$ and $\epsilon_{l_{sc}}^s = \frac{\partial\ln(1-\bar{\alpha})}{\partial\ln\bar{\alpha}}\frac{d\ln\bar{\alpha}}{d\ln s_1} = -\frac{\bar{\alpha}}{1-\bar{\alpha}}\frac{d\ln\bar{\alpha}}{d\ln s_1}$ and $\epsilon_{l_{sv}}^s = \frac{\partial\ln\left(\frac{(\sigma-1)(1-\bar{\alpha})}{1+(\sigma-1)(1-\bar{\alpha})}\right)}{\partial\ln\bar{\alpha}}\frac{d\ln\bar{\alpha}}{d\ln s_1} = -\frac{\bar{\alpha}}{1-\bar{\alpha}} \cdot \frac{1}{1+(\sigma-1)(1-\bar{\alpha})}\frac{d\ln\bar{\alpha}}{d\ln s_1}$. Rewriting the latter two as functions of $\epsilon_{\bar{\alpha}}^L$ we obtain (30). Under homogeneous specialization $\frac{d\ln\bar{\alpha}}{d\ln s_1} = 0$ (proposition 2) thus so are all elasticities given in (30). Otherwise $\frac{d\ln\bar{\alpha}}{d\ln s_1} > 0$. In a stable equilibrium $\frac{d\ln s_1}{d\ln L} \geq 0$, with strict inequality under heterogeneous specialization (proposition 4),so $\epsilon^L_{\bar{\alpha}} > 0$ and $\epsilon^L_{l_{sc}}, \epsilon^L_{l_{sv}} < 0$. □

C.8 Proof of Proposition 7

By definition, the variable labor productivity in terms of output is

$$\begin{aligned}
\bar{\varphi}_Q &\equiv \frac{Y}{PL_v} = \frac{Y}{P\frac{\sigma-1}{\sigma}(1-\bar{\alpha})Y} = \frac{1}{P\frac{\sigma-1}{\sigma}(1-\bar{\alpha})} \\
\frac{d\ln\bar{\varphi}_Q}{d\ln L} &= \frac{d\ln W}{d\ln L} - \frac{d\ln l_{sc}}{d\ln L} \\
\epsilon^L_{\bar{\varphi}_Q} &= \epsilon^L_W - \epsilon^L_{l_{sc}},
\end{aligned}$$

where L_v is variable labor in production and the first line uses the fact that wage is normalized so that L_v equals variable labor cost. Similarly, the variable labor productivity in terms of value-added is

$$\begin{aligned}
\bar{\varphi}_{VA} &\equiv \frac{Y\left(1-\frac{\sigma-1}{\sigma}\bar{\alpha}\right)}{PL_v} = \frac{Y\left(1-\frac{\sigma-1}{\sigma}\bar{\alpha}\right)}{P\frac{\sigma-1}{\sigma}(1-\bar{\alpha})Y} = \frac{1}{Pl_{sv}} \\
\frac{d\ln\bar{\varphi}_{VA}}{d\ln L} &= \frac{d\ln W}{d\ln L} - \frac{d\ln l_{sv}}{d\ln L} \\
\epsilon^L_{\bar{\varphi}_{VA}} &= \epsilon^L_W - \epsilon^L_{l_{sv}}.
\end{aligned}$$

□

C.9 Proof of Proposition 8

In parts 1–3, selection into production, $\frac{M_a}{M}$, depends on L only via s (via $\bar{\varphi}_e$ as shown in proposition 3), hence $\bar{\epsilon}^L_{M_a/M} = 0$. We obtain

$$\epsilon^s_{M_a/M} = \frac{d\ln\left[1-G(\bar{\varphi}_e)\right]}{d\ln s_1} = -\frac{\bar{\varphi}_e g(\bar{\varphi}_e)}{1-G(\bar{\varphi}_e)}\frac{d\ln\bar{\varphi}_e}{d\ln s_1} = -\frac{\bar{\varphi}_e g(\bar{\varphi}_e)}{1-G(\bar{\varphi}_e)}\frac{\bar{\alpha}-\alpha_e}{\delta},$$

where the first equality differentiates and second equality uses proposition 3. Under homogeneous specialization then $\bar{\alpha} = \alpha_e$, otherwise $\epsilon^s_{M_a/M} < 0$ since $\bar{\alpha} > \alpha_e$. Moreover, in a stable equilibrium $\frac{d\ln s_1}{d\ln L} \geq 0$ iff under heterogeneous specialization (0 otherwise), so we need only show $\epsilon^s_{M_a/M} < 0$ for part 3. For parts 1 and 2, using (18) we have $\bar{\epsilon}^L_M = 1$, same as under the fixed premium. To obtain ϵ^s_M note that other effects occur through $\bar{a}$ and $\bar{F}$

$$\begin{aligned}
\epsilon^s_M &= \frac{\partial\ln M}{\partial\ln\bar{a}}\frac{d\ln\bar{a}}{d\ln s_1} + \frac{\partial\ln M}{\partial\ln\bar{F}}\frac{d\ln\bar{F}}{d\ln s_1} \\
&= \frac{\partial\ln\bar{a}}{\partial\ln\bar{\alpha}}\frac{d\ln\bar{\alpha}}{d\ln s_1} - \frac{\bar{F}}{f_E+\bar{F}}\frac{d\ln\bar{F}}{d\ln s_1},
\end{aligned}$$

using expression for $\bar{a}$ we have $\frac{\partial\ln\bar{a}}{\partial\ln\bar{\alpha}}$=$(\bar{a}-1)$, so we obtain ϵ^s_M in (33). Using $M_a \equiv \left[1-G(\bar{\varphi}_e)\right]M$ we obtain $\epsilon^L_{M_a} = \epsilon^L_{M_a/M} + \epsilon^L_M$.

Part 4: Existence of $g(\cdot)$ s.t. $\epsilon^s_M > 0$

From proposition 2, $\frac{d\ln\bar{\alpha}}{d\ln s_1} > 0$ iff under heterogeneous specialization and in that case $\bar{a} > 1$. Thus, $\epsilon^s_M > 0$ for any $g(\cdot)$ s.t. $\frac{d\ln\bar{F}}{d\ln s_1} \leq 0$. We show existence using a special but useful case. Under any $g_k(\cdot)$

s.t. $\frac{(\bar{\varphi}_i)^\sigma g_k(\bar{\varphi}_i)}{\int_{\bar{\varphi}_i}^\infty \varphi^{\sigma-1} dG_k(\varphi)} = \tilde{K}(\sigma)$ for any $\bar{\varphi}_i > 0$, $\sigma \geq 1$ and $\tilde{K}(\sigma) > 0$ we have $\frac{d\ln \bar{F}}{d\ln s_1} = 0$ and thus $\epsilon^s_M > 0$ under heterogeneous specialization.

Proof. See the online appendix. $\square$

An example of such $g_k(\cdot)$ is the unbounded Pareto with $G(\varphi) = 1 - \left(\frac{\varphi_{\min}}{\varphi}\right)^k$. In this case $\frac{(\bar{\varphi}_i)^\sigma g_k(\bar{\varphi}_i)}{\int_{\bar{\varphi}_i}^\infty \varphi^{\sigma-1} dG_k(\varphi)} = \tilde{K}(\sigma) = k - (\sigma - 1)$, which represents the sales distribution dispersion parameter. More generally, the result requires $\frac{d}{d\varphi}\left(\frac{\bar{\varphi}^\sigma g(\bar{\varphi})}{\int_{\bar{\varphi}}^\infty \varphi^{\sigma-1} dG(\varphi)}\right) = \frac{d}{d\varphi}\tilde{K}(\sigma) = 0$ for all φ, which we can solve for $g(\cdot)$ and show it requires $\ln g_k(\varphi) = -\left(\tilde{K}(\sigma) + \sigma\right)\ln\varphi$. Note also that since the condition above can be applied to all $\sigma \geq 1$, the result can be applied to other moments, e.g. if $\frac{\bar{\varphi}^z g_k(\bar{\varphi}_i)}{\int_{\bar{\varphi}_i}^\infty \varphi^{z-1} dG_k(\varphi)} = \tilde{K}(z)$, so it can be applied to $z = 1$ and in the case of Pareto $\tilde{K}(1) = k$. Next we show that there exist $g(\cdot)$ satisfying lemma C.9 that implies that $\epsilon^s_M > 0$ and $\epsilon^s_M + \epsilon^s_{M_a/M} < 0$ for an economy with sufficiently large L and α_n From (32) and (33) we see that the active firm elasticity due to specialization is simply $\epsilon^s_M + \epsilon^s_{M_a/M}$. In lemma C.9 we show that $\epsilon^s_M|_{g_k} > 0$, so for existence we show when $\left(\epsilon^s_M + \epsilon^s_{M_a/M}\right)|_{g_k} < 0$

$$
\begin{aligned}
\left(\epsilon^s_M + \epsilon^s_{M_a/M}\right)|_{g_k} &= (\bar{a} - 1)\frac{d\ln\bar{\alpha}}{d\ln s_1} - \tilde{K}(1)\frac{\bar{\alpha} - \alpha_e}{\delta} \\
&< (\bar{a} - 1)\frac{\sigma\left[1 - \frac{\sigma-1}{\sigma}\bar{\alpha}\right](1 - \bar{\alpha})}{\delta\bar{\alpha}} - \tilde{K}(1)\frac{\bar{\alpha} - \alpha_e}{\delta} \\
&= \left[(1 - \bar{\alpha})(\sigma - 1) - \tilde{K}(1)(\bar{\alpha} - \alpha_e)\right]\frac{1}{\delta}
\end{aligned}
$$

where second line uses stability condition and simplifies. Thus, a sufficient condition for an equilibrium with heterogeneous specialization (where $\alpha_e = \alpha_0 = 0$) is

$$
\frac{1 - \bar{\alpha}}{\bar{\alpha}} \leq \frac{\tilde{K}(1)}{\sigma - 1} \Leftrightarrow \bar{\alpha} \geq \frac{1}{1 + \tilde{K}(1)/(\sigma - 1)}
$$

Since $\frac{1}{1+\tilde{K}(1)/(\sigma-1)} < 1$ we know it is possible to have $\alpha_n > \frac{1}{1+\tilde{K}(\sigma)/(\sigma-1)}$ and in that case there is some large enough economy with heterogeneous specialization such that $\bar{\alpha} \in [\frac{1}{1+\tilde{K}(\sigma)/(\sigma-1)}, \alpha_n)$. If $g_k(\cdot)$ is Pareto then $\tilde{K}(1) = k$, and the sufficient condition is $\alpha_n \geq 1/2$ since $k/(\sigma - 1) > 1$ for a finite first moment of sales distribution. $\square$

C.10 Proof of Proposition 9

C.10.1 Necessity of Change in Specialization

If there is no change in specialization such that $e = 0$ (or $e = n$) before and after the shock then (15) implies that φ_e is unchanged. Moreover, homogeneous specialization implies relative profits across any two firms depend only on their fundamental productivity, so relative profits and the distribution $\Phi(\cdot)$ of $\tilde{\pi}(\cdot)$ remains unchanged.

C.10.2 Sufficiency of Change in Specialization for Increase in Concentration

$\tilde{\pi}(\varphi, L)$ is a continuous increasing function of φ and there is some x^* s.t. $\tilde{\pi}(\varphi^*, L) = x^* = \tilde{\pi}(\varphi^*, L')$ with $\tilde{\pi}(\varphi, L) \leq \tilde{\pi}(\varphi, L')$ for $\varphi > \varphi^*$ and $\tilde{\pi}(\varphi, L) \geq \tilde{\pi}(\varphi, L')$ otherwise. Moreover, φ^* is unique iff there is a change in specialization, otherwise $\tilde{\pi}(\varphi, L) = \tilde{\pi}(\varphi, L')$ for all φ. Continuity is straightforward given the continuum of productivities and the optimal choice of technology that maximizes $\tilde{\pi}(\cdot)$. The size increase implies selection: $\bar{\varphi}'_e > \bar{\varphi}_e$ and thus some $\varphi^* > \bar{\varphi}'_e$ such that $\tilde{\pi}(\varphi, L) \geq \tilde{\pi}(\varphi, L')$ for $\varphi \leq \varphi^*$, together with continuity ensures at least one intersection from below. The free entry condition implies that there is some φ^{**} such that $\tilde{\pi}(\varphi, L) \leq \tilde{\pi}(\varphi, L')$ for $\varphi \geq \varphi^{**}$. If there is no change in specialization then $\varphi^* = \varphi^{**}$ and they are not unique since the profit distribution remains unchanged as shown in C.10.1. If there is a change in specialization we show that $\varphi^* = \varphi^{**}$ at a unique point. This is simple to illustrate if the economy switches from no specialization to full specialization as shown in Figure 7. The proof for heterogeneous specialization is provided in the online appendix. In the following we use lemma C.10.2 to show the sufficiency

1. Increase in concentration with respect to $\Pi(\cdot)$ under $L' > L$: $\Pi(\bar{\varphi}, L) \leq \Pi(\bar{\varphi}, L')$ for any $\bar{\varphi}$ (strictly for some φ).

   Proof. Trivially, the cumulative shares are $\Pi(\bar{\varphi}, L') = \Pi(\bar{\varphi}, L') = 1$ for $\bar{\varphi} \in [0, \bar{\varphi}_e^L]$ where $\bar{\varphi}_e^L$ is the lowest entry threshold, which occurs under L. Using lemma C.10.2 there is some φ^* such that $\tilde{\pi}(\varphi, L') \geq \tilde{\pi}(\varphi, L)$ for all $\varphi \geq \varphi^*$ with inequality above φ^* when the latter is unique. Thus,

   $$\Pi(\bar{\varphi}, L') = \frac{\int_{\bar{\varphi}}^{\infty} \tilde{\pi}(\varphi, L')\, dG(\varphi)}{f_E} > \frac{\int_{\bar{\varphi}}^{\infty} \tilde{\pi}(\varphi, L)\, dG(\varphi)}{f_E} = \Pi(\bar{\varphi}, L) \quad \text{for all } \bar{\varphi} \geq \varphi^*$$

   Using lemma 9 $\tilde{\pi}(\varphi, L') \leq \tilde{\pi}(\varphi, L)$ for all $\varphi < \varphi^*$ with strict inequality for $\varphi \in (\bar{\varphi}_e^L, \varphi^*)$. Thus

   $$\Pi(\bar{\varphi}, L') = 1 - \frac{\int_0^{\bar{\varphi}} \tilde{\pi}(\varphi, L')\, dG(\varphi)}{f_E} > 1 - \frac{\int_0^{\bar{\varphi}} \tilde{\pi}(\varphi, L)\, dG(\varphi)}{f_E} = \Pi(\bar{\varphi}, L)$$

   for all $\bar{\varphi} \in (\bar{\varphi}_e^L, \varphi^*)$, where the first equality uses the free entry condition:

   $$\int_0^{\bar{\varphi}} \tilde{\pi}(\varphi, L')\, dG(\varphi) + \int_{\bar{\varphi}}^{\infty} \tilde{\pi}(\varphi, L')\, dG(\varphi) = f_E.$$

   □

2. Increase in concentration with respect to $\Phi(\cdot)$: $\Phi(x, L)$ SSD $\Phi(x, L')$.

   Proof. Free entry in (14) implies equal mean profits, thus we show that $\Phi(x, L')$ is a MPS of $\Phi(x, L)$, which implies SSD. A sufficient condition is that $\Phi(x, L)$ intersects $\Phi(x, L')$ only once from below. Thus, we use lemma C.10.2 to show that there exists a x^* such that $\Phi(x \leq x^*, L) \leq \Phi(x \leq x^*, L')$ and $\Phi(x \geq x^*, L) \geq \Phi(x \geq x^*, L')$. Under lemma C.10.2 we can invert $\tilde{\pi}(\varphi, L)$ to obtain

   $$\begin{aligned} \Phi(x, L) &= \Pr\left(\varphi \leq \tilde{\pi}^{-1}(x, L)\right) = G\left(\tilde{\pi}^{-1}(x, L)\right) \\ &\leq G\left(\tilde{\pi}^{-1}(x, L')\right) = \Pr\left(\varphi \leq \tilde{\pi}^{-1}(x, L')\right) \\ &= \Phi(x, L') \text{ if } x \leq x^*, \end{aligned}$$

where the inequality in the second line is from evaluating the constant productivity distribution, $G(\varphi)$, and noting there is a higher proportion of firms under L' with profits below a given level implied by $x \leq x^*$ (from lemma C.10.2). A similar argument implies that $\Phi(x, L) \geq \Phi(x, L')$ if $x \geq x^*$. A continuous $G(\cdot)$ implies that $\Phi(\cdot)$ is also continuous, and a unique x^* implies a unique intersection: $\Phi(x^*, L) = \Phi(x^*, L')$. Thus, $\Phi(x, L')$ is a MPS of $\Phi(x, L)$. If there is no change in specialization then $\tilde{\pi}^{-1}(x, L) = \tilde{\pi}^{-1}(x, L')$ for all x and so $\Phi(x, L) = \Phi(x, L')$. $\square$

C.11 Proof of Proposition 10

Necessity of change in specialization. If specialization is unchanged then $\bar{F}$ is constant and so are mean sales, $Y/M = \sigma\left(f_E + \bar{F}\right)$. Moreover, the cumulative shares are also unchanged for any $\bar{\varphi}$: $\lambda(\bar{\varphi}, L) = \int_{\bar{\varphi}}^{\infty} \varphi^{\sigma-1} dG(\varphi) / \int_{\bar{\varphi}_e}^{\infty} \varphi^{\sigma-1} dG(\varphi)$ since $\bar{\varphi}_e$ is unchanged as shown in section 4.1.2. Thus we require a change in specialization *Conditions for* $\lambda(\bar{\varphi}, L') \geq \lambda(\bar{\varphi}, L)$

$$\begin{aligned}
\lambda(\bar{\varphi}, L') &= \frac{\int_{\bar{\varphi}}^{\infty} \left[\tilde{\pi}(\varphi, L') + f(\varphi, L')\right] dG(\varphi)}{\int_{\varphi_{\min}}^{\infty} \left[\tilde{\pi}(\varphi, L') + f(\varphi, L')\right] dG(\varphi)} \qquad \text{(C.28)} \\
&= \frac{\int_{\bar{\varphi}}^{\infty} \tilde{\pi}(\varphi, L') dG(\varphi) + \bar{F}(\bar{\varphi}, L')}{f_E + \bar{F}(L')} \\
&> \frac{\int_{\bar{\varphi}}^{\infty} \tilde{\pi}(\varphi, L) dG(\varphi) + \bar{F}(\bar{\varphi}, L')}{f_E + \bar{F}(L')} \\
&= \frac{\int_{\bar{\varphi}}^{\infty} \tilde{\pi}(\varphi, L) dG(\varphi) + \bar{F}(\bar{\varphi}, L) + \bar{F}(\bar{\varphi}, L') - \bar{F}(\bar{\varphi}, L)}{f_E + \bar{F}(L) + \bar{F}(L') - \bar{F}(L)} \\
&> \frac{\int_{\bar{\varphi}}^{\infty} \tilde{\pi}(\varphi, L) dG(\varphi) + \bar{F}(\bar{\varphi}, L) + \bar{F}(L') - \bar{F}(L)}{f_E + \bar{F}(L) + \bar{F}(L') - \bar{F}(L)} \\
&= \frac{\int_{\bar{\varphi}}^{\infty} \tilde{\pi}(\varphi, L) dG(\varphi) + \bar{F}(\bar{\varphi}, L)}{f_E + \bar{F}(L)} \\
&= \lambda(\bar{\varphi}, L)
\end{aligned}$$

The first line uses the definition of λ in (37) and the relation of sales and profits $\tilde{y}(\varphi)/\sigma = \tilde{\pi}(\varphi) + f(\varphi, L')$. The second line uses the free entry condition and definition of integrals of $f(\varphi, L')$. The inequality in the third line reflects the higher cumulative profits shown in proposition 9 for all $\bar{\varphi} > \bar{\varphi}_e^L$. The fourth line adds and subtracts similar terms. The inequality in line 5 is from

$$\Delta\bar{F} \equiv \bar{F}(\bar{\varphi}, L') - \bar{F}(\bar{\varphi}, L) - \left[\bar{F}(L') - \bar{F}(L)\right] > 0, \qquad \text{(C.29)}$$

which holds for all $\bar{\varphi} > \bar{\varphi}_e^L$ if $\frac{d\gamma(\bar{\varphi})}{d\bar{\varphi}} = 0$, as we prove in the online appendix. Line 6 holds with equality if $\frac{d\gamma(\bar{\varphi})}{d\bar{\varphi}} = 0$ since in that case $\bar{F}(L') = \bar{F}(L)$. Moreover, if two distributions have the same mean and $\lambda(\bar{\varphi}, L') \geq \lambda(\bar{\varphi}, L)$ for all $\bar{\varphi}$ (with some inequality) then the new distribution is Lorenz-dominated by the original, which is equivalent to it being a mean-preserving spread (Atkinson, 1970); The last line in (C.28) uses the same results for sales and profits in the first two lines.

C.12 Alternative Productivity Measures

How are alternative weighted averages of TFPQ affected by size? We can weigh by relative quantities as done in Melitz (2003) to define a harmonic mean of firm TFPQ. Doing so yields[44]

$$\tilde{C}^{-1} = \left[\sum_{i=e}^{n} \Lambda_i \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \varphi^{\sigma-1} \mu(\varphi, L)\, d\varphi\right]^{1/(\sigma-1)}. \tag{C.30}$$

If $\Lambda_i = 1$ we obtain the expression in Melitz (2003). More generally, we have an adoption weight increasing in i defined by

$$\Lambda_i \equiv \left[\frac{(P/\phi)^{\alpha_i}}{\tilde{c}/\tilde{C}}\right]^{-\sigma},$$

which measures the relative unit cost of intermediates and labor in a segment i (recall $w = 1$) relative to the average unit costs in the economy ($\tilde{c}/\tilde{C}$ is an index of input costs gross of productivity terms since it aggregates all unit cost and divides by an aggregate index of the productivity terms). Under homogeneous specialization $\Lambda_i = 1$ and L works only via selection so again a change in specialization is necessary. Moreover, a large enough increases in L from no specialization to full specialization equilibria implies an increase in $\tilde{C}^{-1}$. Together with the fact that (C.30) is continuous in L implies that it increases TFP average over some heterogeneous specialization range. Finally, we can consider a sales share average: $\sum_{i=e}^{n} \lambda_i \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \varphi \mu(\varphi, L)\, d\varphi$. This changes iff specialization changes (otherwise λ_i is unchanged). The measure increases with a shock to size that moves from no specialization to full specialization, because of selection, thus a large enough increase in size eventually increases it. Moreover, under heterogeneous specialization this measure is monotonically increasing in size whenever the sales cumulative shares behave as in proposition 10, e.g. in case G is Pareto.

D Quantification

D.1 Measurement of Data Moments

Moments from NBER CES database, sic classification

- Intermediates cost share: measured as

$$\bar{\alpha} = \frac{\text{matcost} - \text{energy}}{\text{matcost} + \text{payroll} + \text{invest}} \tag{D.1}$$

  at the industry level from NBERCES sic data. We then aggregate to the whole US manufacturing using year 1997 industry total cost as fixed weights. The reason we do not use yearly variable weights is that our framework captures within-industry adoption so we want to tease out effects from between-industry reallocations.

- Log relative factor price: measured as

$$\ln\left(\frac{w}{P}\right) = \ln\left(\frac{\text{payroll/employment}}{\text{material price index}}\right),$$

[44] Specifically, $\tilde{c} \equiv \sum_{i=e}^{n} \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} c_i \frac{q(c)}{q(\tilde{c})} \mu(\varphi, L)\, d\varphi$ so $\tilde{c} = \left[\sum_{i=e}^{n} \left((P/\phi)^{\alpha_i}\right)^{1-\sigma} \int_{\bar{\varphi}_i}^{\bar{\varphi}_{i+1}} \varphi^{\sigma-1} \mu(\varphi, L)\, d\varphi\right]^{1/(1-\sigma)}$

at the industry level, and we measure wage rate as payroll/employment. We then aggregate to the whole US manufacturing using year 1997 industry total cost as fixed weights.

As our calibration procedure captures long-run change, we take the geometric average of the above moments between year 1987 and 1989 as the initial period, and the geometric average of the year 2005 to 2007 as the final period, in order to smooth out data fluctuations. This gives us $\bar{\alpha}_0 = 0.699$, $\bar{\alpha}_T = 0.743$, and log relative factor growth during the period as $\Delta \ln \left(\frac{w}{P}\right) = 0.383$. Moments from Census of Manufacturing: year 1987 and 2007

- Trade targets: this is the same as the literature

$$\text{Export intensity} = \frac{\text{value of exports}}{\text{total sales of exporters}}.$$

From Census data, we have $\text{Intenstity}_0 = 10.0\%$ and $\text{Intenstity}_T = 16.3\%$. Year 1987 Census moments come from Bernard et al. (1995).

- Sales share targets: we use the census data on sales share of top 20 firms in each naics industry to compute the sales share of top $20V$ largest firms at the aggregate US manufacturing in year 1987 using industry sales as weights, where V is the number of naics industries with more than 100 firms. The value is 64.5%. The implicit assumption here is that the top 20 sales firm in each industry still remains top $20V$ firms in the aggregate. We can interpret the measure as some sort of the industry average. Furthermore, with data on number of firms in each industry, we can calculate the fraction of those $20V$ firms relative to whole manufacturing, $\chi_{20V} = 2.27\%$.

D.2 Calibration Procedure

In the following we discuss in detail the steps to calibrate the endogenous specialization framework to the US manufacturing between the initial and final periods. The first step is to solve for the variable trade cost τ_0 and τ_T by using data on export intensity. The model implies that $\text{Intenstity}_t = \frac{(N-1)(\tau_t)^{1-\sigma}}{1+(N-1)(\tau_t)^{1-\sigma}}$ so we can invert the expression to get τ by using the trade targets directly. We calibrate the rest of the model parameters while solving the equilibria for both periods. The calibration procedure consists of two loops. In the outer loop, we guess the maximum intensity α so that the resulting sales share of top $20V$ firms matches the data moments, 64.5%. In the inner loop, we calibrate the technology and trade cost parameters by matching the intermediates cost share in both periods, and the log relative factor price growth during the period. Specifically, by assuming the equilibrium is of the heterogeneous specialization type, we start with a guess for the specialization premium for both periods, $s_{1,0}$ and $s_{1,T}$, as well as the value of $\hat{f}$. First, given the value of specialization premium, the relative adoption cost, and variable trade cost, and the preassigned parameters, we calculate the fractions of two types of adopters, $\chi_1 = \left(\frac{\bar{\varphi}_1}{\bar{\varphi}_0}\right)^{-k}$ and $\chi_2 = \left(\frac{\bar{\varphi}_2}{\bar{\varphi}_0}\right)^{-k}$, which we know are functions of specialization premium and adoption cost only from expressions of relative cutoffs. Second, given fractions of adopters and the initial guess, we calculate the implied value of maximum intermediates intensity α_t in both periods using data on intermediates cost share:

$$\bar{\alpha}_t = \frac{F_1\chi_{1,t} + 2F_2\chi_{2,t} + \frac{1}{(s_{1,t})^{\sigma-1}-1}\left(F_1\chi_{1,t} + F_2\chi_{2,t}\right)}{F_0 + F_1\chi_{1,t} + F_2\chi_{2,t}}\left(\frac{\alpha_t}{2}\right), \tag{D.2}$$

where $F_0 = f_0 = 1$, $F_1 = \hat{f} F_0$, and $F_2 = \hat{f}(1+\hat{f})F_0$. Third, given the fractions of adopters, we use the free entry condition to calculate the entry cutoff $\bar{\varphi}_0$:

$$\frac{(\sigma-1)\varphi_{\min}^k}{k-\sigma+1}\left(\bar{\varphi}_0\right)^{-k}\left(F_0 + F_1\chi_1 + F_2\chi_2\right) = f_E. \tag{D.3}$$

Fourth, we use the goods market clearing condition to calculate total expenditure X, and then use the entry cutoff expression to calculate the price index. Finally, we use the definition of specialization premium to calculate the value of intermediates technology ϕ_t for both periods.

$$\phi_t = \left(s_{1,t}\right)^{\frac{2}{\alpha}} P_t. \tag{D.4}$$

Fifth, we calculate the model-implied change in log relative factor price:

$$\Delta \ln\left(\frac{w}{P}\right)^m = \frac{\Delta \ln s_{1,t}}{\alpha_0} - \Delta \ln \phi_t. \tag{D.5}$$

We also calculate the model-implied sales share of top $20V$ firms in the initial period, $\lambda_{20V,0}$. In the inner loop, if the model-implied maximum intermediates intensity equals the initial guess and the changes in relative factor price equal the observed change

$$\alpha_0 = \alpha_T = \alpha, \quad \Delta \ln\left(\frac{w}{P}\right)^m = \Delta \ln\left(\frac{w}{P}\right)^{data}, \tag{D.6}$$

we retrieve the values of technology parameters, ϕ_0, ϕ_T, and $\hat{f}$, and go to the outer loop. If not, we go back to the first step with a new guess. In the outer loop, if the model-implied sales share of top $20V$ firms equal the observed sales share, we get the maximum intermediates intensity α. If not, we go back with a new guess of α and start the inner loop again.

D.3 Model Versus Data Regressions: Specialization and Concentration

We provide details for the validation exercise in section 7.2.5. To obtain variation in initial sales concentration, we shock the cost of adoption, $\hat{f}$, in the initial equilibrium while holding all other parameters constant. We draw lognormal shocks such that the mean of $\ln \hat{f}_V$ reflects the aggregate calibrated value ($\ln(11.3)$) and its standard deviation is 0.1.[45] We compute the model sales share of the percentile x of operating firms in the model, $\lambda\left(x, \hat{f}_V\right)$. We employ $x = 2.27\%$ since in the initial calibration this ensures the model accounts for the (aggregated) top 20 firm sales share in the data (i.e. $\lambda^{data} = \lambda(2.27\%, 11.3)$). If we observed the full sales distribution then we could compute the share using the same percentile for each industry in the data. What we observe is a proxy with some measurement error: the top 20 sales given by $\lambda^{data}(20/N_V) = \lambda\left(x, \hat{f}_V\right) + e_V$, which suggests the elasticity using the data will be attenuated relative to the one using the model. We then take the new initial equilibrium for a given $\hat{f}_V$ and apply a log normal trade cost shock: $\Delta_t \ln \tau \tilde{} N(-0.14, 0.23)$. The shock distribution reflects the aggregate mean decrease in the calibration and its standard deviation reflects industry variation in observed changes in

[45]The small variation in $\hat{f}_V$ allows us to remain close to the original calibration; it also generates smaller variation in top-20 sales share than in the data but we can account for this by comparing the impact of a standard deviation in that share in the data vs. the model.

trade costs.[46] In the data, the correlation between the log trade cost shock and initial top 20 firms sales share is only 0.04 so we assume the two distributions of shocks are independent. After applying the trade shock we compute the log change in intermediate cost share and relative factor price (our measure of specialization premium in the empirical exercise). We take 500 draws to approximately match the number of industries in the data and use the model generated data to run a regression similar to Table 1 but in changes instead of a panel. As noted section 7.2.5 the differential elasticity using the empirical estimation is similar to but smaller than the model's, possibly due to the attenuation error mentioned above. We find some support for this attenuation by re-computing shares in the model to try to capture the different percentiles in the data. Specifically we use percentiles $x^m = x * u^m$ where $\ln u_V$ is normally distributed with mean zero and consider alternative standard deviations. The elasticity differentials for the model is 0.10 when the standard deviation is 1/4 and 0.04 when it is 2/3.[47]

[46]Specifically, we first compute the initial and final trade share as the geometric avarage of import penetration and export sales share for 1990–1992 and 2005–2007, respectively. Then we use the mapping in the model that trade share is $\frac{\tau_{i,t}^{1-\sigma}}{1+\tau_{i,t}^{1-\sigma}}$ to get the changes in trade cost, $\tau_{i,t}$, and compute the standard deviation of the log change.

[47]Higher standard deviations induce more error and attenuation. We chose values lower than the observed *sd(* $\ln 20/N_V$ *)* in the data, which is about unity for the industries used.