American Economic Association

Entry and Exit, Multiproduct Firms, and Allocative Distortions

Author(s): Roberto N. Fattal Jaef

Source: *American Economic Journal: Macroeconomics*, April 2018, Vol. 10, No. 2 (April 2018), pp. 86-112

Published by: American Economic Association

Stable URL: https://www.jstor.org/stable/10.2307/26528435

REFERENCES
Linked references are available on JSTOR for this article:
https://www.jstor.org/stable/10.2307/26528435?seq=1&cid=pdf-reference#references_tab_contents
You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

American Economic Association is collaborating with JSTOR to digitize, preserve and extend access to *American Economic Journal: Macroeconomics*

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

American Economic Journal: Macroeconomics 2018, 10(2): 86–112
https://doi.org/10.1257/mac.20140075

Entry and Exit, Multiproduct Firms, and Allocative Distortions†

By Roberto N. Fattal Jaef*

Most studies quantifying the gains from reversing allocative distortions are static in nature. We propose a model of firm dynamics featuring entry, exit, and multiproduct firms to understand the contribution of these dynamic factors in shaping the welfare and long-run productivity gains from removing distortions. We find that while the entry and exit of firms and their product-portfolio choices exert countervailing forces over long-run total factor productivity (TFP), they reinforce each other in shaping the welfare gains from reversing misallocation. Welfare gains, which account for transition dynamics, become more than twice as high as the long-run changes in TFP. (*JEL* D21, D24, D61, L11, O41)

Resource misallocation across firms is pervasive in developing countries.[1] The evidence shows that an excessive amount of production factors are allocated to less productive firms at the expense of more productive ones, a pattern that deviates from the optimal prescription of equalizing marginal returns across producers.[2] Interpreting the evidence as an outcome of underlying allocative distortions, Hsieh and Klenow (2009), and the many replications of its methodology that followed, found substantial gains in total factor productivity (*TFP*) emerging from efficiently reallocating resources among a given set of producers. However, these quantitative exercises were explicitly static in nature, raising the question of whether dynamic factors such as entry, exit, and product-line choices would magnify or mitigate the static gains. The goal of this paper is to assess the contribution of these dynamic effects for both welfare and *TFP*.

*The World Bank, Development Research Group, Macroeconomics and Growth Division, 1818 H Street NW MC 3-353, Washington, DC 20433 (email: rfattaljaef@gmail.com). We thank two anonymous referees for insightful comments and recommendations. We are also grateful to Andy Atkeson, Francisco Buera, Ariel Burstein, Hugo Hopenhayn, Lee Ohanian, Luis Serven, Venky Venkateswaran, and seminar participants at Stanford University, the Society of Economic Dynamics, and the World Bank's Research Group. A previous version of this paper circulated as Entry, Exit, and Misallocation Frictions. The views presented here do not represent the views of the World Bank Group or any of its member countries. All errors and omissions are the sole responsibility of the author.

† Go to https://doi.org/10.1257/mac.20140075 to visit the article page for additional materials and author disclosure statement or to comment in the online discussion forum.

[1] Salient examples in the literature establishing this fact are Hsieh and Klenow (2009) and Bartelsman et al. (2013).

[2] Our qualification of the prescription to reallocate resources across firms as optimal is accurate only under a restrictive but typically imposed set of assumptions, namely competitive factor markets and costless mobility of resources across firms. Asker, Collard-Wexler, and De Loecker (2014) warn, however, about the potential efficiency in the observation of dispersion in the marginal return to capital in a context where there are adjustment costs to the mobility of capital across firms. Throughout the rest of the paper we shall continue with the initial view of interpreting dispersion in marginal returns as an outcome of distortions.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

To do this, we build a general equilibrium model of firm dynamics featuring entry, exit, and multi-variety production, in which resource misallocation arises as a result of idiosyncratic distortions that are calibrated to match the observed properties of misallocation in the data. We then study the response of the economy to the withdrawal of these distortions, and quantify the contribution of entry, exit, and product choices, both for welfare and long-run changes in *TFP*.

We find that while the entry and exit of firms and their product-portfolio choices exert countervailing forces over long-run *TFP*, they reinforce each other in shaping the welfare gains from reversing misallocation. In the long run, distortions induce a reallocation of products toward low productivity business, which tends to reduce *TFP*. However, these distortions encourage an increase in the number of firms, which tends to increase aggregate productivity. Hence, the overall contribution to the long-run gains from removing distortions is in general ambiguous. However, once the dynamic responses of entry and exit along the transition path are taken into account, these decisions reinforce the static and product-choice channels, and hence magnify the overall gains in welfare, which can become more than twice as high as the gains in long-run productivity.

Our model brings Bernard, Redding, and Schott's (2011) model of multi-variety production into Luttmer's (2007) model of firm dynamics. The economy consists of a unit mass of differentiated products, each of which is a constant elasticity of substitution (CES) composite of a continuum of differentiated varieties. Upon entry, firms are endowed with a firm-wide productivity term and a time-invariant realization of product attributes for the production of a variety of each of the products in the unit measure. Firm-wide productivity evolves stochastically over time, leading to firm growth over the life cycle, conditional on survival. Product attributes are drawn independently and identically, across goods and firms, from a time-invariant probability distribution function. The firms' portfolio of products is limited by a fixed cost of provision. Together with fixed costs of operation, these costs determine whether firms stay or exit operations and the number of products they supply to the market. Misallocation across intermediate producers is created, as in Restuccia and Rogerson (2008), through a distribution of idiosyncratic revenue taxes that are positively correlated with the distribution of firm-wide productivity of the firms. The correlated nature of distortions is intended to make the model consistent with the empirical observation of unproductive firms being too large and productive firms being too small in developing countries relative to the United States.[3] Since idiosyncratic distortions in the data are measured at the firm or establishment level rather than at the product level, we assume the same degree of taxation across all products within a firm.

The ambiguity with which entry, exit, and product choices contribute to long-run *TFP* emphasizes the importance of disentangling the number of varieties from the number of firms for the determination of the long-run implications of misallocation. Distortions that tax the most productive firms more heavily than the least productive

[3] This property of the firm size distribution in developing countries has been extensively documented in the literature. See, for instance, the works of Hsieh and Klenow (2009); Bartelsman, Haltiwanger, and Scarpetta (2013); Neumeyer and Sandleris (2009); Casacuberta and Gandleman (2009); and Camacho and Conover (2010).

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

ones induce a reallocation of products from the former to the latter that magnifies the static (across a given set of producers and a given distribution of products) losses from misallocation. Combined with life-cycle growth of firm-wide productivity, this same pattern of taxation and subsidization creates an increase in entry and in the number of firms, which tends to preserve the number of varieties and thus mitigates the decline in *TFP*. This is because taxing productive firms and subsidizing unproductive ones is equivalent, from the perspective of an entrant, to a redistribution of profits from the future to the present. With positive discounting of future profit streams, this redistribution is less detrimental to the expected profits of entrants than it is to the average profits among incumbents, which requires an increase in entry to restore the equilibrium. Should the equivalence between the number of firms and the number of products be maintained, we would miss a novel layer of misallocation, of products across firms, and would wrongfully undermine the long-run gains from alleviating allocative distortions.

The dynamic nature of entry and exit decisions of firms, which gives a nontrivial role to transitional dynamics, is the key for rationalizing the reinforcing contributions of entry, exit, and product-portfolio choices in shaping the welfare gains. When distortions are removed, the economy converges to an undistorted equilibrium with fewer firms. This constitutes the force through which entry and exit offset the long-run productivity gains. However, along the transition path to such equilibrium, the economy engages in a protracted process of firm decumulation that reallocates resources away from entry and fixed operation costs, and into the production of goods. This productive reallocation gives a temporary boost to consumption that is not captured in measures of long-run welfare gain. Furthermore, the decline in entry and the increase in the exit rate that characterizes the transition imply that more products are produced among the most productive firms, which reinforces the temporary overshoot in consumption that is allowed for by the withdrawal of distortions.

Overall, our analysis dictates that the traditional motivation of thinking about allocative distortions as drivers of cross-country differences in aggregate productivity induces an underestimation of the true magnitude of the welfare gains to be reaped from reforms that alleviate these distortions. Such underestimation did not emerge in the majority of the existing studies because of the static nature of their underlying models, which implied that the convergence from low to high income levels was immediate. With entry, exit, and multi-variety production, the number of firms and their distribution across firm-wide productivity becomes the dynamic variable that gives rise to nontrivial transitional dynamics. Our quantitative analysis suggests that the magnitude of the divergence between the long-run based and the overall welfare gain that is created by the dynamic features of the model is substantial. For instance, calibrating the degree of linear relationship between the logarithm of distortions (*TFPR*) and the logarithm of physical productivity (*TFPP*) to the average regression coefficient between these variables in the data,[4] we find a 10 percent increase in *TFP* in the long run while the welfare gains amount to 32 percent.

[4] Hsieh and Klenow (2007) reports slopes of 0.5, 0.4, and 0.1 for China, India, and the United States, respectively. Chen and Irarrazabal (2014) estimates the elasticity to be between 0.6 and 0.5 in Chile during the period of

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

The rest of the paper is organized as follows. Section I presents the model and characterizes a stationary equilibrium. In Section II, we calibrate parameter values and perform the quantitative experiments. We begin with the welfare analysis and the shift focus to the long-run implications. Concluding remarks are in Section III.

I. The Economic Environment

The model economy embeds Bernard, Redding, and Schott's (2011) theory of multiproduct firms into Luttmer's (2007) closed economy model of firm dynamics, and introduces misallocation frictions in the form of idiosyncratic taxes operating at the level of the aggregate revenue of the firm. Motivated by the empirical pattern of idiosyncratic distortions in the data, these frictions are assumed to be positively and log-linearly related to firm-wide productivity.

A. *Household's Problem*

There is a representative household with preferences of the form $\sum_{t=0}^{\infty}\beta^t[\log(C_t)]$, where C_t is the single final good produced in the economy. Lifetime utility maximization is done subject to a standard inter-temporal budget constraint of the form

$$\sum_{t=0}^{\infty} Q_t[C_t - W_t L - T_t] \leq M_0,$$

where Q_t denotes inter-temporal prices, M_0 is the initial endowment of wealth (claims to the profits and losses from the initial distribution of firms), and T_t represents the lump-sum tax/transfer that balances the deficit or surplus from the collection of idiosyncratic taxes and subsidies. Notice that by rebating these revenue back to (or taking it away from) the household, we are ensuring that all the welfare implications of misallocation frictions manifest solely through their effect on aggregate productivity, rather than from wasteful consumption of goods from the government. Lastly, notice that the inter-temporal accumulation of wealth pins down the economy's interest rate, given by $R_t = Q_{t+1}/Q_t$.

B. *Technologies*

There is a single final good in the economy, which is produced according to the following CES composite of a continuum of measure one of intermediate inputs:

$$Y = \left[\int_0^1 q_k^{\frac{\sigma-1}{\sigma}} dk\right]^{\frac{\sigma}{\sigma-1}}. \tag{1}$$

1980–1996. Cirera et al. (2017) reports values ranging between 0.4 and 0.6 in Ghana, Ethiopia, Kenya, and Côte D'Ivoire. When subtracting the 0.1 estimate for the United States, we get that all of these estimates fall between 0.3 and 0.5.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

We assume that there is a representative producer of the final good that operates under perfect competition.

Each of the intermediate goods is, in turn, another CES composite of a continuum of horizontally differentiated varieties of each product supplied in monopolistically competitive markets by an endogenously determined measure of heterogeneous firms:

$$q_k = \left[\int q_k^d(\omega,\lambda)^{\frac{\rho-1}{\rho}} d\Omega(\omega,\lambda)\right]^{\frac{\rho}{\rho-1}}. \tag{2}$$

Varieties are differentiated by the firm-wide productivity level of the producer supplying it (ω) and the product-specific productivity attribute (λ). We explain later the origin of these idiosyncratic characteristics of the firms, as well as the decision process involved in determining the mass of producers and the product portfolio. For now, it suffices to note that $\Omega(\omega,\lambda)$ denotes the measure of producers with this productivity mix.

Profit maximization yields the following demand function for a variety $e^{\omega+\lambda}$ of a product k:

$$q_k^d(\omega,\lambda) = (P_k)^{\rho-\sigma} p_k(\omega,\lambda)^{-\rho} Y,$$

where the price index of a good k, P_k, and the price of a given variety of the good k, $p_k(\omega,\lambda)$, are in units of final good.

Price indices, in turn, are given by the following expressions:

$$P = \left[\int_0^1 P_k^{1-\sigma} dk\right]^{\frac{1}{1-\sigma}}, \tag{3}$$

$$P_k = \left[\int p_k(\omega,\lambda)^{1-\rho} d\Omega(\omega,\lambda)\right]^{\frac{1}{1-\rho}}, \tag{4}$$

where we are adopting the final good as the numeraire, so $P = 1$.

Product varieties are supplied in monopolistically competitive markets and are produced according to a linear production function with labor as the only input:

$$q_k(\omega,\lambda) = \left[e^{\omega+\lambda}\right]^{\frac{1}{\rho-1}} l_k(\omega,\lambda).$$

Productivity is determined by firm-wide and product-specific terms, e^{ω} and e^{λ}, respectively.[5] Combined with fixed costs of supplying goods, the latter determines the portfolio of products supplied by the firm, while the former generates heterogeneity across firms in factor demands, product provision, and profitability.

[5] Notice that we are treating the product attribute as part of the productivity of the firm. As in Bernard, Redding, and Schott (2011), identical results are obtained if the product attributes enter revenues and profitability through the household's demand functions.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

The distortions that generate misallocation in the model take the form of idiosyncratic taxes and subsidies to the firm's revenue. We adopt the assumption that there is heterogeneity in the degree of taxation across firms, but we assume this rate to be the same across products within a firm. This is the assumption that we consider suits best the nature of the data from which we calibrate the functional form of the distribution of distortions. Since firm-level data are typically reported at the firm or establishment level, reported statistics about distributions of *TFPR* and *TFPQ* can only be informative of firm-wide or establishment-wide—rather than product-specific—distortions. Thus, the combination of misallocation frictions with the multiproduct feature of the model creates another layer of misallocation across firms, namely the distribution of products across producers. However, it forces us to abstract from a channel of misallocation within firms.

While we specify a particular functional form governing the relationship between idiosyncratic distortions and physical productivity once we get to the quantitative analysis, at this stage we require that whichever functional form we impose, it does not create reversals in the positive relationship that exists between productivity and profitability in the undistorted equilibrium. That is, while it is certainly the case that distortions that correlate positively with productivity flatten the positive relationship between productivity and profitability, we rule out cases where a relatively unproductive firm is turned more profitable than a relatively more productive one. We find empirical support for this assumption in that, in the literature, all of the existing estimates of the regression coefficient between the log of *TFPR* and the log of *TFPQ* are far below 1.

Letting τ_ω denote the idiosyncratic distortion corresponding to a firm with firm-wide productivity e^{ω}, the price and the labor demand associated with the provision of variety e^{λ} of a product k solves the following static profit-maximization problem:

$$\max_{l_k(\omega,\lambda)} \left(1-\tau_\omega\right) P_k^{\frac{\rho-\sigma}{\rho}} Y^{\frac{1}{\rho}} \left(e^{\omega+\lambda}\right)^{\frac{1}{\rho}} l_k(\omega,\lambda)^{\frac{\rho-1}{\rho}} - w\, l_k(\omega,\lambda),$$

which yields the following expressions for labor demand, revenues, and profits:

$$l(\omega,\lambda) = e^{(\omega+\lambda)} (1-\tau_\omega)^{\rho} P^{\rho-\sigma} \left(\frac{Y}{w^{\rho}}\right) \left(\frac{\rho-1}{\rho}\right)^{\rho}, \tag{5}$$

$$R(\omega,\lambda) = e^{(\omega+\lambda)} (1-\tau_\omega)^{\rho} P^{\rho-\sigma} \frac{Y}{\left(w^{\rho-1}\right)} \left(\frac{\rho-1}{\rho}\right)^{\rho-1}, \tag{6}$$

$$\pi^{v}(\omega,\lambda) = e^{(\omega+\lambda)} (1-\tau_\omega)^{\rho} P^{\rho-\sigma} \frac{Y}{\left(w^{\rho-1}\right)} \left(\frac{\rho-1}{\rho}\right)^{\rho-1} \frac{1}{\rho}. \tag{7}$$

Notice we have suppressed the k subscript denoting the type of good in the price index of the firms' profit functions, in anticipation of the fact that, as a result of the assumptions about the independence of the distribution of product attributes and the commonality of the fixed cost of provision across products and firms, goods are identical to each other. This will become more evident below.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

C. *Product-Portfolio Determination*

Firms earn positive profits from the provision of each good. To establish a limit to the menu of products actually supplied to the market, we assume that this activity entails a fixed labor-denominated cost of marketing equal to f_p. Furthermore, we assume these fixed costs to be identical across goods and across firms, and constant over time. The total profits earned from the provision of a given variety of a product, then, are equal to

$$\pi(\omega, \lambda) = e^{(\omega+\lambda)}(1-\tau_\omega)^\rho P^{\rho-\sigma}\frac{Y}{(w^{\rho-1})}\left(\frac{\rho-1}{\rho}\right)^{\rho-1}\frac{1}{\rho} - wf_p.$$

Given the assumption of no reversals in the ranking of profits and productivity induced by the introduction of distortions, the total profits of the firms are strictly increasing in firm productivity and product attributes. Therefore, the fixed cost of provision implies that there exists a cutoff level such that varieties with attributes above it are supplied, and varieties with attributes below it are dropped. The cutoff attribute is the one at which the profits of a given product attribute are equal to zero. Solving for it in the expression above yields

$$e^{\bar{\lambda}(\omega)} = (\Pi)^{-1}\left[e^{\omega}(1-\tau_\omega)^\rho\right]^{-1} f_p, \tag{8}$$

where Π, defined below, collects all the aggregate variables and parameters that determine the factor of proportionality of firms' outcomes with respect to idiosyncratic variables:

$$\Pi = \left[P^{\rho-\sigma}\frac{Y}{(w^{\rho})}\left(\frac{\rho-1}{\rho}\right)^{\rho-1}\frac{1}{\rho}\right]. \tag{9}$$

Firms with higher firm-wide productivity confront a lower threshold for the attribute that guarantees positive profitability from the provision of a given product and hence are able to supply a wider range of products to the market. Idiosyncratic distortions interfere with this decision. Firms that are taxed at a positive rate would find it harder to supply products to the market, while firms that are subsidized will find it easier. This mechanism of transmission of misallocation frictions to aggregate productivity is new to the literature and is one of the forces that we are interested in characterizing quantitatively.

D. *Distribution and Evolution of Idiosyncratic Productivities*

Firms are endowed, upon entry, with a realization of firm-wide productivity and a continuum of product attributes for each of the product types in the unit measure. We assume that entrants start off at a firm-wide productivity level consistent with the ratio of the average size of entrants to incumbents in the US data, and then fan out over time according to a stochastic process. Attributes for each product, however, are drawn independently and identically from a known distribution

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

$F(e^{\lambda})$. Unlike firm-wide productivity, we assume that the product attribute is constant throughout the life cycle of the firm.

Firm dynamics are driven by the stochastic process for firm-wide productivity. I parameterize this process taking a discrete-time random walk approximation to a Brownian motion with drift μ and variance σ^2. Following Stokey (2008), we assume that given current log productivity ω, next period's log productivity could give a jump of size h, either upward with probability α, or downward with probability $(1-\alpha)$. The discrete time approximation of the drift and variance of the process is

$$\mu \Delta t = (2\alpha - 1) h,$$

$$\sigma^2 \Delta t = 4\alpha(1-\alpha) h^2.$$

The appeal of the discrete-time random walk approximation of the Brownian motion is that it easily maps into the model of Luttmer (2007), who shows that the resulting stationary distribution from this stochastic process displays a right tail of the cumulative distribution function that is of the Pareto type, which constitutes an accurate characterization of the right tail of the firm-size distribution in the United States. In our model with multiproduct firms, the mapping from firm-wide productivity to size is mediated by the product-attribute distribution; however, given the latter, it is still true that there is a tight link between the parameters of the binomial process and the tail of the size distribution.

E. *Aggregation within Firms*

Let us first characterize the aggregation of product-specific outcomes within firms. This is greatly simplified by the i.i.d. assumption of the product-attribute distribution. Together with the law of large numbers, they imply that averages across the unit continuum of products are equal to the corresponding average of each product. Aggregate productive employment, revenue, and profits within the firm are thus given by

$$L(\omega) = \Pi(\rho - 1) e^{\omega} (1 - \tau_{\omega})^{\rho} \left[\int_{e^{\bar{\lambda}(\omega)}} e^{\lambda} dF(e^{\lambda}) \right], \tag{10}$$

$$R(\omega) = w\Pi(\rho - 1) e^{\omega} (1 - \tau_{\omega})^{\rho} \left[\int_{e^{\bar{\lambda}(\omega)}} e^{\lambda} dF(e^{\lambda}) \right], \tag{11}$$

$$\pi(\omega) = w\Pi(\rho - 1) e^{\omega} (1 - \tau_{\omega})^{\rho} \left[\int_{e^{\bar{\lambda}(\omega)}} e^{\lambda} dF(e^{\lambda}) \right] - w \times f_c. \tag{12}$$

The expressions show the transparent way in which firm-wide productivity interacts with the product-scope decision of the firm in turning into observable variables at the firm level, such as employment and sales. Conditional on a product portfolio, determined by the cutoff attribute of the marginal product supplied to the market $e^{\bar{\lambda}(\omega)}$, more productive firms employ more workers, generate more revenue, and are more profitable. This is the standard relationship between productivity and size implied by single product models. With multiproduct firms, the elasticity of

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

firm-level variables with respect to firm-wide productivity gets magnified by the endogeneity of the product scope of the firms. More productive firms also get to produce a wider range of products, which, in turn, expands the demand for labor, sales, and profits. By the same token, this same force operates to magnify the distortive effect of allocative distortions. Not only will firms with higher taxes produce and earn less from each product, but they shall also provide fewer varieties to the market. This is a novel layer of misallocation that we are able to take into account by means of our multiproduct structure.

F. *Entry and Exit*

Until now, we have dispensed from carrying a time subscript in the description of the model, since we were characterizing static decisions for which a time dimension was irrelevant. To decide about entry and exit, however, firms are looking ahead into future profitability, so we must bring a time subscript into the model.

The value of an operating producer with productivity e^{ω} is characterized by the following equation:

$$v_t^o(\omega) \;=\; \pi_t(\omega) - w_t f_c + R_t(1-\delta)\, E_t\big[v_{t+1}(\omega') \,|\, \omega\big],$$

where $\pi_t(\omega)$ is defined in equation (12), with the addition of a time subscript to the aggregate variables absorbed in Π_t. Firms discount the future at the market interest factor, R_t, augmented by an exogenous death probability δ, orthogonal to firm characteristics. The main goal of this shock is to account for the exit of firms from across the entire size distribution, a feature that we see in the data but would be absent in the model if exit were to be purely endogenous.

The value of the firm, then, is given by

$$v_t(\omega) \;=\; \max_{x_t(\omega)} \big\{0, v_t^o(\omega)\big\}$$

with $x_t(\omega)$ encoding the exit decision of the firm, being equal to 1 if it stays in operation, and 0 otherwise.

In terms of entry, prospective entrants compare the value of a labor-denominated entry cost f_e with the value under the entrant's productivity e^{ω_e}. Free entry from an infinite pool of producers ensures that entry costs and expected valuation are equalized in equilibrium:

$$w_t f_e \;=\; R_t(1-\delta)\, v_{t+1}(\omega_e).$$

Notice that we are assuming a one period time to build before the entrant starts operations.

G. *Equilibrium and Macroeconomic Variables*

The remaining step in the characterization of the equilibrium is to aggregate the outcomes of the firm up to the level of the macroeconomy. Key for the aggregation

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

is the distribution of firms across firm-wide productivities, which we denote with $M_t(\omega)$. This distribution evolves according to the following law of motion:

$$M_{t+1}(\omega') = (1-\delta)\alpha M_t(\omega' - h) + (1-\delta)(1-\alpha) M_t(\omega' + h) + (1-\delta) M_{e,t} I(\omega_e). \tag{13}$$

The expression establishes that a fraction $(1-\delta)\alpha$ of firms with productivity less than or equal to $(\omega' - h)$ survives the exogenous exit shock and transitions to a productivity level that is less than or equal to ω'. A fraction $(1-\delta)(1-\alpha)$ of the mass of firms with productivity between ω' and $(\omega' + h)$ survives the exit shock and jumps downward to have productivity less than or equal to ω'. There is also an inflow of new firms into this group, which is given by the mass of entrants, conditional on productivity being equal to the productivity level assumed for entry, e^{ω_e}. Endogenous exit will be driven by the mass of firms that transition downward from the productivity cutoff, $(1-\delta)\alpha M_t(\underline{\omega} + h)$.

A *competitive equilibrium* in this economy is: (i) a sequence of aggregate consumption decisions from the household $\{C_t\}_{t=0}^{\infty}$; (ii) sequences of prices, labor demands, value functions, product cutoffs, and exit cutoffs for the producers of varieties, $\{p_{k,t}(\omega,\lambda), l_t(\omega,\lambda), V_t(\omega), e^{\overline{\lambda}_t(\omega)}, \underline{\omega}_t\}_{t=0}^{\infty}$; (iii) a sequence of final good quantities and demand functions for intermediate variety $\{Y_t, q_{k,t}^d(\omega,\lambda)\}_{t=0}^{\infty}$; (iv) a sequence of measures of firms $\{M_t(\omega)\}_{t=0}^{\infty}$ and its law of motion (equation (13)); (v) a sequence of entrants $\{M_{e,t}\}_{t=0}^{\infty}$; (vi) a sequence of prices and transfers $\{w_t, R_t, P_t, Q_t, T_t\}$; (vii) a distortion profile $G(\tau|\omega)$, a distribution of productivity at entry $G(\omega)$, a distribution of product attributes $F(\lambda)$, and a stochastic process for firm-wide productivity; and (viii) an initial wealth of the household M_0 such that: (a) given (vi), (iv), and (viii), (i) solves household's optimization problem; (b) given (vi), (iii), and (vii), (ii) solves the incumbents' dynamic optimization problem; (c) given $p_{k,t}(\omega)$, (iii) solves the final good sector's profit maximization problem; (d) $M_{e,t}$ is such that the free entry condition is satisfied in every period; and (e) markets clear in every period:

$$L = L_{p,t} + L_{fc,t} + L_{fp,t} + f_e M_{e,t},$$

$$C_t = Y_t,$$

where $L_{p,t}$, $L_{fc,t}$, and $L_{fp,t}$ are the aggregate demands for labor in production, fixed costs of operation, and fixed cost of provision of products, defined by

$$L_{p,t} = \Pi_t(\rho - 1)\Lambda_t, \tag{14}$$

$$L_{fc,t} = f_c \int dM_t(\omega),$$

$$L_{fp,t} = f_p \int \left[1 - F\left(e^{\overline{\lambda}_t(\omega)}\right)\right] dM_t(\omega).$$

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

The term Λ_t stands for the statistic of the joint distribution of productivity and product attributes that characterizes aggregates across firm-level outcomes in the economy, and is given by

$$\Lambda_t = \int e^{\omega}(1-\tau_{\omega})^{\rho}\left[\int_{e^{\bar{\lambda}_t(\omega)}} e^{\lambda}\, dF(\lambda)\right] dM_t(\omega).$$

A *stationary competitive equilibrium* is one in which the distribution of productivity has become stationary and where aggregate variables and prices have become constant.

Aggregate Variables.—Prior to the quantitative analysis, it is instructive to derive an analytical characterization of aggregate output and aggregate productivity.[6] To derive it, we start substituting the aggregate component of firms' variable profits defined in equation (9) into equation (14) for aggregate labor demand in production. Then, solving for aggregate output we get

$$Y = \frac{w^{\rho}}{P^{\rho-\sigma}}\left(\frac{\rho}{\rho-1}\right)^{\rho}\frac{L_p}{\Lambda}. \tag{15}$$

As it is common in models of monopolistic competition with CES demand systems, the price of a variety of product κ produced by a firm with productivity e^{ω}, product attribute e^{λ}, and idiosyncratic distortion τ_{ω} is given by

$$p(\omega,\lambda) = \frac{\rho}{\rho-1}\frac{w}{\left[e^{(\omega+\lambda)}\right]^{\frac{1}{\rho-1}}(1-\tau_w)}.$$

Firms with higher productivity and higher product attributes offer lower prices. Revenue taxes and subsidies are passed on to the consumer, to the extent allowed by the elasticity of substitution.

Plugging the individual prices into the price index for good k, we get

$$P = \frac{\rho}{\rho-1} w\left[\int\int_{e^{\bar{\lambda}(\omega)}} e^{(\omega+\lambda)}(1-\tau_w)^{\rho-1}\, dF(\lambda)\, dM(\omega)\right]^{\frac{1}{1-\rho}}. \tag{16}$$

Notice that the expression justifies the earlier claim that price indices are symmetric across products. This is guaranteed by the assumption of an independent and identical distribution of product attributes and the assumption of a common fixed cost of provision across products and firms, which together imply that $F(\lambda)$ and $e^{\overline{\lambda}(\omega)}$ are identical for all k.

Exploiting this symmetry and recalling that the price of the final good is chosen to be the numeraire, it follows that the price index of each product k is also equal

[6] Once again, to simplify notation, we avoid carrying time subscripts in the notation with the understanding that the aggregate variables just defined are static in the stationary equilibrium and time-varying in the analysis of transitions.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

to unity, $P = 1$. Also, substituting into the price index of the final good in equation (3), we can solve for the equilibrium wage rate as

$$w = \frac{\rho - 1}{\rho}\left(\Lambda^{w}\right)^{\frac{1}{\rho-1}}, \tag{17}$$

$$\Lambda^{w} = \left[\int\int_{e^{\bar{\lambda}(\omega)}} e^{\omega+\lambda}\left(1-\tau_{\omega}\right)^{\rho-1} dF(\lambda)\, dM(\omega)\right], \tag{18}$$

where we have subsumed the aggregation of productivities and product attributes that determine marginal costs into the term Λ^{ω}.

We can go back to equation (15) and get the following expression for *GDP* and *TFP* in the model:

$$Y = \frac{\left(\Lambda^{\omega}\right)^{\frac{\rho}{\rho-1}}}{\Lambda} \times L_p, \tag{19}$$

$$TFP = \frac{\left(\Lambda^{\omega}\right)^{\frac{\rho}{\rho-1}}}{\Lambda} \times \frac{L_p}{L}. \tag{20}$$

Changes in the entry and exit of firms affect *TFP* through the consequent change in the number of firms and through the demand for labor that goes into entry and fixed operation costs, which determines L_p. Changes in the firms' portfolio of products also affect output through changes in L_p, due to the labor-intensive nature of the fixed costs of supplying goods. Furthermore, it affects *TFP* by shaping the distribution of products across firms with different productivities, controlled by $e^{\overline{\lambda(\omega)}}$.

II. Quantitative Analysis

We turn now to the quantitative evaluation of the model. As emphasized throughout the paper, we are interested in measuring the welfare gains from liberalizations that eliminate misallocation frictions. We seek to provide a quantitative answer accompanied with an understanding of how the various forces at work in the model contribute to shaping these gains.

A. *Calibration*

We must choose parameter values for the elasticity of substitutions σ and ρ, the subjective discount factor of the household β, the size of the labor force L, and the set of parameters governing the process of firm dynamics, entry, and exit: entry and fixed operation costs f_e and f_c, the size and probability of the jump in the binomial process h and α, and the exogenous exit rate δ. Furthermore, we must specify values for the fixed cost of supplying varieties to the market, f_p, as well as the shape parameter of the distribution of product attributes η. We calibrate these parameters working with the undistorted stationary allocation of the model, taking the United States as an empirical target. Table 1 summarizes the parameter values.

The strategy of the calibration is as follows. For the elasticity of substitution, we set $\rho = \sigma = 3$, which lies in the middle of estimates of substitutability found in the trade and industrial organization literature, and is the value chosen by Hsieh and

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

Table 1—Parameter Values and Calibration Targets

Parameter	Value	Target
$\rho = \sigma$	3	Hsieh and Klenow (2009), Broda and Weinstein (2006)
β	$\frac{1}{1.05}$	Interest rate of 5%
δ	0.02	Employment-based exit rate of large firms of 2%
α	0.467	Slope log of right tail of employment-based size distribution $= -0.2$
h	$\frac{0.25}{\eta}$	Standard deviation of employment growth of large firms
ω_e	$e^{\omega_e} = 1$	Size of entrants = 6% of median incumbent (Luttmer 2010)
$\frac{f_c}{f_e}$	0.1	Exit rate of 5%
f_p	210.95	Average fraction of products per firm = 0.0024
η	2	Distribution of output-share by product in multiproduct firms

Klenow (2009) in their measurement of misallocation in the United States, China, and India.[7] Since our assumptions imply no heterogeneity of prices across products, distortions will not create changes in relative prices across products; hence, only the elasticity of substitution across varieties within a product category matters for aggregate variables. Accordingly, we set the elasticities of substitution to be identical to each other. We choose the discount factor so that $\frac{1}{\beta} - 1$ equals a real interest rate of 5 percent, and normalize the size of the labor force to be equal to 1.

In terms of product attributes, we assume these are distributed Pareto, with shape parameter η and lower bound $e^{\lambda_{\min}} = 1$. Besides its analytical tractability, this parameterization is consistent with the behavior of multiproduct firms in the United States—in particular, the distribution of output-shares across products within multiproduct firms. Bernard, Redding, and Schott (2010) reports the output-share by rank of the product for firms of various degrees of multiple-good production (4, 6, 8, and 10 products). They run a regression of log of product rank and the log of the output share of the product. They find the coefficient of this regression to be equal to 0.5. In the Appendix, we show that this target maps into a parameter value of $\eta = 2$ for the Pareto distribution of product-attributes.

Our strategy to calibrate the fixed cost of provision of goods $\left(f_p\right)$ also relies on data from Bernard, Redding, and Schott (2010).[8] They show that out of a potential of 1,440 products, the average number of products supplied by a multiproduct firm is 3.5. Given that our product space lives in the continuum of measure 1, this gives us a target of 3.5/1,440 to be matched by the product range of the average multiproduct firm in the model. Formally, f_p is chosen so that:

$$\frac{\int \left[1 - F\left(e^{\bar{\lambda}(\omega)}\right)\right] dM(\omega)}{\int dM(\omega)} = 0.0024,$$

which amounts to devoting 17 percent of the labor force to the task of confronting fixed costs involved in the supply of products to the market.

[7] See Broda and Weinstein (2006) for a range of estimates of the elasticity of substitution for US imports at a four-digit disaggregation level.

[8] Data taken from table 1 in Bernard, Redding, and Schott (2010).

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

Despite the empirical merits and the analytical tractability of the Pareto distribution, the lower limit in its support carries the implication that some sufficiently productive multiproduct firms in the model will find it profitable to supply all of the products in the continuum. This is problematic for two reasons. First, according to Bernard, Redding, and Schott (2010), there is no evidence of such firms in the data for the manufacturing sector in the United States. Second, it introduces a discontinuity in the elasticity of firm size with respect to firm-wide productivity. To see this, consider the expression for firm-wide employment in the undistorted economy that results from aggregating across products within a firm with firm-wide productivity e^{ω} under a Pareto distribution of product attributes and an arbitrary lower bound $e^{\lambda_{\min}}$:

$$L(\omega) = (\Pi)^{\eta} \frac{(\rho - 1)}{\left[f_p^{(\eta-1)}\right]} \frac{\eta}{\eta - 1} \left(e^{\lambda_{\min}}\right)^{\eta} \begin{cases} \left[e^{\omega}(1 - \tau_w)^{\rho}\right]^{\eta} & \text{if } e^{\bar{\lambda}(\omega)} \geq e^{\lambda_{\min}} \\ e^{\omega}(1 - \tau_w)^{\rho}\left(e^{\lambda_{\min}}\right)^{1-\eta} & \text{if } e^{\bar{\lambda}(\omega)} < e^{\lambda_{\min}} \end{cases}.$$

Since $e^{\bar{\lambda}(\omega)}$ is decreasing in firm-wide productivity, it can be shown that there exists a sufficiently large but finite value of e^{ω} such that the cutoff attribute is lower than the lower bound of the Pareto distribution. In this case, once the full product-space has been covered, the firm behaves as if it was a single product firm, with an elasticity with respect to firm-wide productivity that switches from η to 1.

In spite of these complications, we find that a negligible fraction of firms (0.02 percent) in the undistorted stationary equilibrium hits the lower bound and supplies all the products to the market under our final calibration. Thus, the gains in tractability from the adoption of a Pareto distribution—in particular, the transparent way in which the product-portfolio decision magnifies the elasticity of size with respect to productivity—more than compensates for its cost.

Having characterized more sharply the aggregate employment level within a firm, we can proceed with the description of the strategy to calibrate the parameters governing the stochastic process of firm-wide productivity. Taking logs of firm-wide employment in the region where the product range is less than one, we get

$$\log\left[L(\omega)\right] = \eta \log(\Pi) + \Psi - \eta \log\left(e^{\omega}\right),$$

where $\Psi = \frac{(\rho - 1)}{\left[f_p^{(\eta-1)}\right]} \frac{\eta}{\eta - 1}$ collects parameters related to the fixed cost of provision of products, the elasticity of substitution across varieties within a product, and the tail of the Pareto distribution of product attributes.

The moment in the data that we target to pin down the step size h is the standard deviation of the distribution of employment growth rates for large firms, which in the US economy is equal to 0.25. It is easy to show that in the stationary equilibrium of our model, this variance is approximately equal to[9]

$$\text{var}\left(\hat{L}\right) \cong (\eta h)^2 = 0.25^2. \tag{21}$$

[9] The exact value for the variance is $\text{var}\left(\hat{L}\right) = (\eta h)^2 - [\eta h]^2(2p - 1)^2$. The approximation is exact in the case of a stochastic process with zero drift, namely $p = 0.5$. We show below that our calibrated value of p is 0.467,

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

Thus, for a given value of η, the standard deviation of employment growth rates in the cross section of large firms in the United States implies that $h = 0.25/\eta$. We take into account that there exists a small fraction of large firms supplying all the products, for which η no longer plays a role. For these, the variance of employment growth rates is driven entirely by the step size in the process for productivity, so we set $h = 0.25$ for such group.

In terms of the probability of technological upgrading, α, we set it to match properties of the right tail of the US employment-based size distribution. Luttmer (2010) highlights the linearity of the right tail of the US establishment and firm size distribution across employment, as well as the stationarity of the distribution over time, using various sources of US micro data.[10] For a given value of the productivity step size h and a given value of the exogenous exit probability δ, the slope of the right tail of the firm-size distribution in the model is determined by α. To ensure that we are capturing the linear portion of the size distribution, we focus on the slope implied by the ratio of the change in the logarithm of the fraction of employment accounted for by firms with 1,000 or more and 5,000 or more employees, relative to the log-difference of the employment levels, which is equal to -0.2 in the data. Formally, we set p so as to match

$$\frac{\log\left[1 - G(5{,}000)\right] - \log\left[1 - G(1{,}000)\right]}{\log(5{,}000) - \log(1{,}000)} = -0.2,$$

where $G(\cdot)$ denotes the cumulative distribution function of the employment-weighted firm-size distribution in the United States. Targeting this slope translates into a 72 percent share of employment accounted for by the top 10 percent largest firms, which is consistent with the data for the United States' manufacturing sector.

Regarding the probability of the exogenous exit shock, this parameter controls the exit rate among the largest firms. Thus, we calibrate it to replicate the exit rate among large firms in the employment-weighted size distribution, which, according to the Small Business Administration (SBA),[11] is equal to 0.02.

With respect to the realization of firm-wide productivity for entering firms, we assume that all entrants get a productivity draw of $e^{\omega_e} = 1$, and then fan out over the space of idiosyncratic productivity according to its stochastic process. The target driving this choice is the average size of an entrant relative to the average incumbent in the United States' size distribution, which Luttmer (2007) reports to be equal to 6 percent. The entry and fixed production costs are set at $f_e = 1$ and $f_c = 0.1$, respectively. Together with the parameters of the stochastic process for firm-wide productivity—particularly, the standard deviation of the shocks to idiosyncratic productivity—the choice of f_c and f_e imply an exit rate of 5 percent.

which allows us to confirm the approximated value of the variance as a close approximation. The sole advantage of it is that we can independently identify the values of the parameters of the stochastic process that match their counterparts in the data.

[10] County Business Patterns Database, statistics from the Small Business Administration, and the Business Dynamics Statistics from the census.

[11] This is the exit rate that applies to firms larger than or equal to 500 workers in the year 2002.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

The last portion of the calibration refers to the distribution of distortions. A pervasive feature in developing countries is that idiosyncratic distortions exhibit a strong degree of correlation with the distribution of physical productivities across firms. This is a key property of the data, since it is the one that creates a nontrivial response in the dynamic decisions of the firms, such as the decision to enter and exit the economy.[12] To capture this relationship between idiosyncratic distortions and idiosyncratic firm-wide productivities, we propose the following log-linear functional form for the relationship between idiosyncratic distortions and the idiosyncratic firm-wide productivities:

$$\left(1-\tau(\omega)\right) = \left[\left(e^{\omega}\right)^{\frac{1}{\rho-1}}\right]^{-\gamma}.$$

The appeal of the expression is that it provides a transparent mapping between γ, the slope, and the coefficient of a regression between the log of *TFPR* and the log of *TFPQ*, which is a statistic that is normally reported in empirical papers about misallocation. In order to understand the strength of the mechanism at various degrees of distortions, we experiment with a range of values of γ. In choosing this range, we restrict ourselves to values that do not reverse the positive relationship between productivity and profitability. Doing so would carry the counterfactual prediction of having the most productive firms be the ones exiting the market endogenously.

We consider the following values: $\gamma \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6\}$. As a useful reference point, according to Hsieh and Klenow (2007), γ would be roughly equal to 0.1 for the United States in 1997, around 0.5 for China in 1998, and about 0.4 for India in 1987.

B. *Welfare Analysis*

This subsection begins the quantitative exploration measuring the welfare gains to be reaped from liberalizations that eliminate idiosyncratic distortions. We follow the tradition of thinking about permanent consumption compensations that should be given to households in order to make them indifferent among different allocations. The benchmark comparison is between keeping the household at a stationary equilibrium with distortions and transitioning towards a stationary equilibrium without frictions. When we investigate long-run gains only, the comparison is between the level of consumption in the distorted steady state and the undistorted steady state one. The results are reported in Figure 1. The black line reports the welfare gains accounting for both transitional and long-run gains, while the gray line reports the latter only.

Given the Pareto optimality of the undistorted allocation, welfare gains are expectedly increasing throughout the space of distortion slopes. The gains are also sizable, requiring an increase of up to 55 percent in permanent consumption in order to

[12] It can be shown that uncorrelated distortions, although damaging for allocative efficiency in terms of products and labor, have neutral effects on entry and exit. We resume this point once we explore the quantitative performance of the model.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

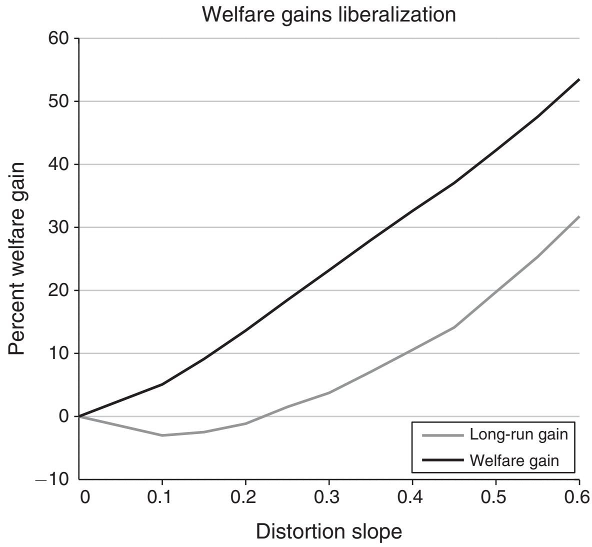

Figure 1. Welfare Gains from Liberalizing Distortions

compensate the household to remain at the stationary equilibrium with distortions. For instance, at the level of distortions consistent with China in 1998, $\gamma = 0.5$; the welfare gains from liberalization amount to almost 40 percent. For India in 1987, with $\gamma = 0.4$, the gains reach 32 percent.[13]

The figure also reveals a significant divergence between the welfare gains with and without consideration of transition dynamics. The gray line, which reports the welfare gains from steady state to steady state improvements in consumption, shows that ignoring transition dynamics significantly underestimates the overall potential gains in welfare. Continuing with the examples of China and India, long-run gains are only one-third of the overall gains.

The divergence between the two measures of welfare gain arises as a result of the temporary gain in consumption that the liberalization of distortions allows in the early years of the transition path. We show this in Figure 2, which reproduces the time series of consumption along the transition path from liberalizations at the bottom, middle, and top of the range of distortions that we consider: $\gamma \in \{0.1, 0.3, 0.6\}$. We can see that consumption overshoots upon reform and persists for several periods above the level to which it converges in the undistorted steady state.

In order to uncover the forces that allow for this temporary boost, Figure 3 shows the time series of the total number of firms in the economy, the shares of labor allocated to the production of goods, entry costs, fixed costs of operation and the fixed costs of supplying goods, and the total number of products per firm. The number

[13] All elasticities are read off table 4 in Hsieh and Klenow (2007). Albeit at a weaker degree, there is also evidence of correlated distortions in the United States, which we use as the efficient benchmark in our calibration. Had we subtracted the estimated elasticity for the United States from the estimates for China and India, the corresponding slopes of distortions would have been 0.3 and 0.2, respectively, leading to welfare gains of 32 percent and 23 percent.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

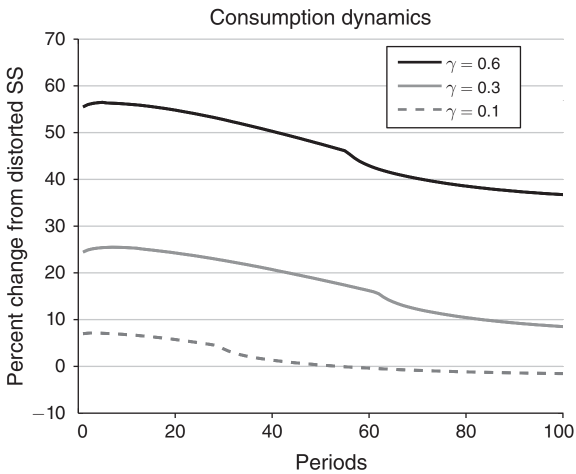

FIGURE 2. CONSUMPTION DYNAMICS ALONG TRANSITION PATHS

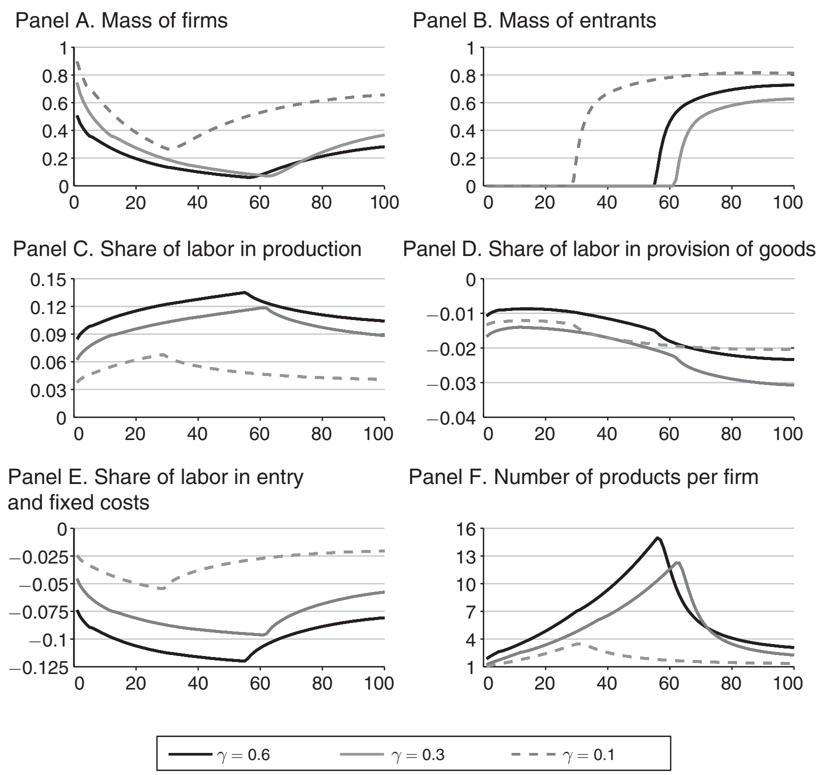

FIGURE 3. UNDERSTANDING SOURCES OF WELFARE GAIN: TRANSITION DYNAMICS

Notes: The number of firms, the mass of entrants, and the number of products per firm are measured as ratios with respect to their distorted steady-state values. Labor shares are expressed as absolute deviations from the initial shares.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

of firms, the mass of entrants, and the number of products per firm are measured as ratios relative to their distorted steady state values, while the labor shares are expressed as absolute deviations from the initial shares.

The salient feature of the figure is the process of decumulation of firms that takes place in the transition toward the undistorted stationary equilibrium. As the economy converges to an allocation with fewer firms, a result that we shall explain in greater detail below, resources are reallocated away from entry and fixed costs of operation and into the production of goods. The number of products per firm increases sharply, in reflection of the efficient reallocation of products toward highly productive firms that takes place in response to the reform. However, since the number of firms is falling, the total number of varieties and the labor required to cover fixed costs of provision also fall during the transition, which further reinforces the reallocation of labor towards productive activities and further contributes to sustaining consumption above the undistorted stationary level.

As mentioned above, the increase in the number of firms in the economy's distorted stationary equilibrium is a key property of the model's response to allocative distortions that explains the divergence in results between measures of welfare gain. As can be observed in Figure 3, the number of firms is between 30 percent and 70 percent lower in the undistorted than in the distorted stationary allocations.

The rise in the number of firms in the distorted allocation is a result of the following two forces: an increase in the measure of entrants and a decrease in the rate of exit. The rise in firm entry emerges as an outcome of the interaction between two features of the calibration that are grounded in the data: the stochastic process for productivity upon entry, which, conditional on survival, gives firms an increasing pattern of productivity growth over the life cycle; and the positive relationship between distortions and idiosyncratic productivity. From the perspective of an entrant, a profile of distortions that taxes high productivity firms and subsidizes low productivity ones creates a redistribution of profits from the future to the present. In the cross section of incumbents, this pattern of taxation redistributes profits from high to low productivity firms. Since firms discount future profit streams at a positive interest rate, the overall decline in expected profitability of entrants will be lower than the average decline in profitability among incumbents. In order to restore the equilibrium in the labor market and the balance in the free entry condition, an increase in entry will follow. It is easy to show that, had any of the above mentioned properties of the calibration been absent, distortions would not have had any effect on the number of firms.[14]

The decline in the exit rate, in turn, is also attributable to the positive relationship between distortions and productivity. By taxing the high productivity firms, labor demand is relaxed and factor prices fall. This allows for marginal firms that would have been left out of the market to sustain operations profitably. In addition, these firms are aided by subsidies to producers in the left tail of the productivity distribution. Even though the decline in the exit rate carries a decline in aggregate

[14] That was the case, for instance, in Restuccia and Rogerson (2008), where the missing ingredient justifying the neutrality of entry in their model is the lack of firm dynamics upon entry.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

productivity per firm, it also contributes to increasing the number of firms since, at a given number of entrants, it sustains a larger number of producers in the stationary equilibrium.

A last feature of the transition dynamics that is worth highlighting is the non-monotonicity of the convergence. Two properties of the model play a key role in generating this pattern: the non-negativity of entry and the productivity distribution of entrants relative to incumbents. The non-negativity of entry prevents a sudden adjustment in the number of firms. The fastest the economy can reduce the number of firms is by shutting down entry, increasing the exit productivity cutoff, and waiting for the negative drift in incumbents' productivity and the exogenous exit shocks to drive firms out of the market. This explains the smooth decline of the number of firms and the protracted increase in the shares of labor allocated to production. The point of entry of new firms in the productivity distribution, in turn, is the key ingredient to understand the eventual recovery in the number of firms. Newcomers to the market enter at a lower productivity level than the average incumbent. This means that during the periods where there is zero entry, production gets concentrated at the top of the productivity distribution, leaving a lower mass of firms around the marginal productivity cutoff. When entry resumes, it creates an inflow of firms that takes time to populate the left tail. Until this happens, entry exceeds exit; thus, the number of firms reverses its trajectory and converges to the new stationary level from below.

Summarizing, our analysis shows that accounting for the dynamic response in the economy's number of varieties through changes in the number and allocation of products, and the entry and exit of firms, magnifies the welfare gains from alleviating misallocation far beyond the gains that can be captured by focusing on long-run measures of consumption growth. Despite the merit that resource misallocation has in the context of explaining cross-country differences in *TFP* (a merit that we reassess below with our richer framework), our results motivate the adoption of a broader measure of welfare when discussing benefits and costs of reforms that are aimed at alleviating allocative inefficiencies.

C. *Long-Run Implications of Misallocation Frictions*

In this section, we focus on understanding the long-run implications of misallocation frictions in the context of our model. While the previous section emphasized the importance of transition dynamics for welfare assessments, the steady-state consequences of distortions were at the core of the initial motivation for thinking about misallocation as a potential driver of the large differences in income and productivity that exist across countries. For this reason, we here connect with the bulk of the literature by providing a quantitative evaluation of the contribution of entry, exit, and multi-variety production to the ability of idiosyncratic distortions to play this role.

In order to transparently identify the contributions of the new margins, we compare the long-run changes in *TFP* in three alternative long-run allocations of our

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

model, where entry, exit, and multi-variety firms are at play sequentially. More specifically, recall the expression that defines measured *TFP* in the model:

$$TFP = \frac{\left(\Lambda^{\omega}\right)^{\frac{\rho}{\rho-1}}}{\Lambda} \times \frac{L_p}{L},$$

$$\Lambda^{w} = \left[\int\int_{e^{\bar{\lambda}(\omega)}} e^{\omega+\lambda}\left(1-\tau_{\omega}\right)^{\rho-1} dF(\lambda)\, dM(\omega)\right],$$

$$\Lambda = \left[\int\int_{e^{\bar{\lambda}(\omega)}} e^{\omega+\lambda}\left(1-\tau_{\omega}\right)^{\rho} dF(\lambda)\, dM(\omega)\right].$$

We construct a benchmark allocation where all the gains from liberalizations accrue due to the efficient reallocation of resources among a given set of producers, and among a given distribution of products across firms. This allocation, which we label *static*, is the one that resembles the type of gains that are captured in Hsieh and Klenow (2009) and the many replications thereafter, and the one that we consider an adequate benchmark to gauge the contributions of entry, exit, and multi-variety production. To construct it, we set all taxes and subsidies to zero while keeping the distribution of cutoff attributes $e^{\bar{\lambda}(\omega)}$ for the provision of a product, the number and distribution of firms $M(\omega)$, and the share of labor allocated to the production of goods L_p/L, fixed at their distorted allocation's level. Then, in order to assess the contribution of the multiproduct structure of production, we construct an allocation where we allow the distribution of products across firms $e^{\bar{\lambda}(\omega)}$ and the labor share in production L_p/L to adjust to the efficient level, while keeping constant the number of firms and the distribution of firms across firm-wide productivity at the distorted economy's allocation. We label this allocation as *product-channel.* Lastly, we show the long-run gains in the *full* model which, when compared to the previous two, identifies the contribution of entry and exit. Results are shown in Figure 4.

A comparison between the solid black line, corresponding to the static benchmark, and the dashed line, corresponding to the gains with consideration of the product-reallocation channel, shows that the latter margin significantly magnifies the improvements in *TFP* that would emerge in the long-run from a liberalization of misallocation frictions.[15] Subsidies to unproductive firms and taxes to productive ones reallocate products from the latter to the former. Because the favored firms are relatively less productive, a lower amount of output per product is produced under a given aggregate allocation of labor to production, reinforcing the decline in *TFP* in the distorted allocation and enhancing the gains when these are reversed.[16]

[15] The non-monotonicity in the static gains is related to the existence of a small fraction of firms that supply the whole range of products. We develop this intuition in the next subsection, when we compare the multiproduct with the single-product model.

[16] Given that there is a small fraction of firms that cannot expand beyond the full range of products, the total number of products falls slightly in the *product-channel* allocation when removing allocative distortions. While this force tends to weaken the *TFP* gains, it is partially compensated by the increase in share of labor allocated to the production of goods that is allowed for by the lower demand for labor that is required to confront the fixed costs of provision. It can be shown, though, that the predominant force driving the productivity gains in this allocation is given by the efficient reallocation of products to more productive firms.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

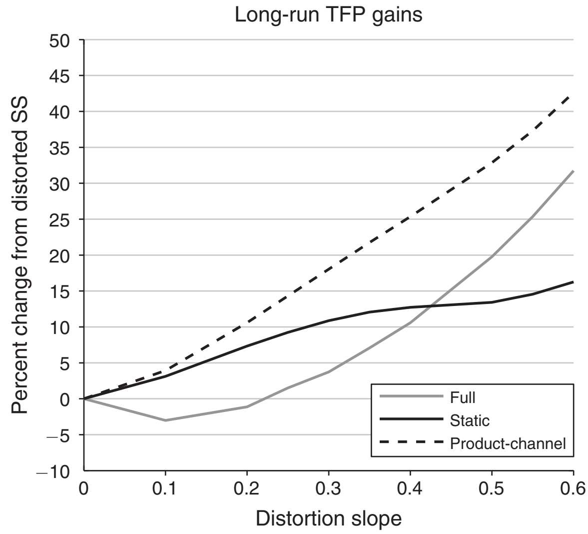

Figure 4. Long-Run TFP Gains from Liberalizing Distortions: The Role of the Number of Products and the Number of Firms

The consideration of firm entry and exit, illustrated by the gray line in Figure 4, brings a countervailing force to the analysis, which makes the overall contribution of the response in the number of varieties, in general, ambiguous. As explained earlier in the context of understanding transition dynamics, distortions that correlate positively with firm-wide productivity induce an increase in entry and a decline in the exit rate, which increases the total number of firms. This increase is reversed in the undistorted long-run equilibrium, which under a *CES* production structure constitutes a drag on aggregate productivity.

The non-monotonicity of the response in *TFP* in the full model as a function of the underlying degree of distortions is worth exploring. Figure 4 shows that the countervailing effect of entry and exit prevails at liberalizations from low to middle values of the slope between distortions and idiosyncratic productivity, while it is dominated by the labor and product reallocation channels at high values.[17] To understand this feature of the results, recall that the rise in entry in the distorted allocation was critically determined by the interaction between an increasing trajectory of profitability upon entry (conditional on survival), a profile of distortions that taxed productive firms more heavily, and discounting of future profit streams. As we move along the range of distortion slopes, the redistribution of profits from higher to lower productivity values flattens the upward-sloping trajectory of profits. Hence, an additional increase in distortions carries a weaker effect than when distortions were first put in place. However, the damaging effect of distortions on allocative efficiency increases together with the severity of distortions. Thus, there exists a distortion level beyond

[17] Notice that there are levels of distortion slopes where the undistorted *TFP* is even lower than the distorted one. This feature stresses even further the need to account for transitional dynamics, given the knowledge that the undistorted allocation is Pareto optimal and, hence, transitioning to it should deliver higher welfare to the household.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

which the latter force dominates the former, so that liberalizing distortions becomes more productivity-enhancing than implied by the static benchmark.[18]

Going back to the question of how much cross-country differences in allocative distortions contribute to understanding cross-country differences in *TFP*, our findings indicate that entry, exit, and multi-variety production can enhance the role attributed to resource misallocation provided we are considering countries with sufficiently high degrees of correlation between idiosyncratic distortions and idiosyncratic productivity. Based on existing studies, the majority of the estimates of the elasticity between distortions $(TFPR)$ and productivity $(TFPQ)$ of developing countries fall in the range where the magnifying forces and the countervailing ones roughly balance out.[19] Had it not been for the consideration of the dynamic implications of reforms, the long-run analysis would have mistakenly led us to conclude that there is not much to be lost in terms of capturing the full extent of the costs of misallocation, from ignoring the effect of distortions over the number of varieties in the economy. The findings from this section, then, serve to reinforce our message of motivating the adoption of a broader measure of welfare that internalizes these dynamic gains.

The Role of Multi-variety Production.—While the analysis above was designed to isolate the marginal effect of entry, exit, and multiproduct firms, it was not well suited to establish a comparison between the results of our model with multiproduct firms against those stemming from a model with single-product units. The reason is that the models are not comparable unless the calibration of the shock process for firm-wide productivity is adjusted to make sure that both models match the same targets in the data. To appreciate this, recall that the relationship between productivity and size in the multiproduct model is intermediated by the product scope of the firm. This source of magnification in the mapping from productivity to employment dispersion is absent in the single-product model; hence, it must be compensated through an increase in the calibrated value of the variance of the shock process. This leads us to set $h = 0.25$.

We proceed by constructing the same *full* and *static* allocations that we did in the context of the analysis of the benchmark model, comparing the overall magnitude of the *TFP* gains across models, and identifying the contribution of entry and exit relative to the static gains, depending on whether the multiproduct channel is at play or not. The results are illustrated in Figure 5.

The salient property of the figure is the remarkable difference in the contribution of the static gains in shaping the overall improvements in productivity across models. While the gains under the full model are slightly lower in the multiproduct than

[18] Notice that the argument relies in a sense on continuity in the evolution of the strengths of the two forces affecting *TFP* as a function of distortions, a feature that we have only explored numerically.

[19] Hsieh and Klenow (2007) reports slopes of 0.5 and 0.4 for China and India, respectively. Chen and Irarrazabal (2015) estimates the elasticity to be between 0.6 and 0.5 in Chile during the period of 1980–1996. Cirera et al. (2017) reports values ranging between 0.4 and 0.6 in Ghana, Ethiopia, Kenya, and Côte D'Ivoire.

However, not every study of misallocation reports statistics about the degree of relationship between distortions $(TFPR)$ and productivity $(TFPQ)$, so there are much fewer estimates of this elasticity than there are estimates of the dispersion.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

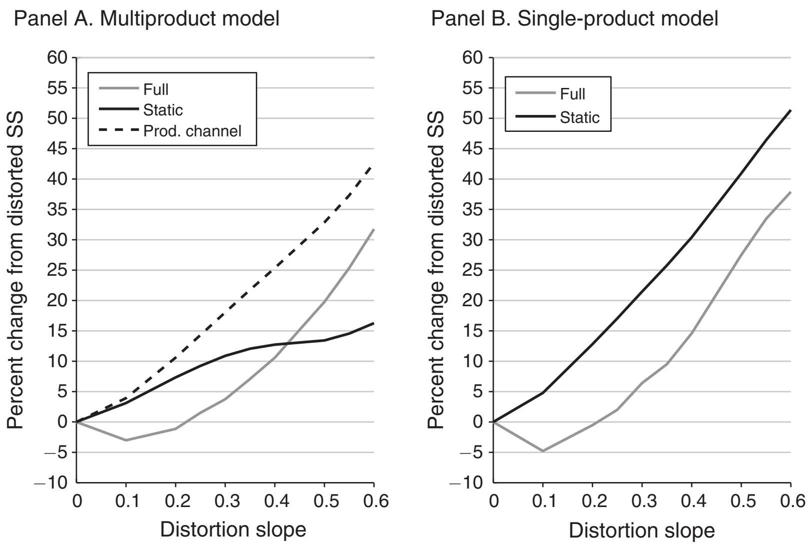

Figure 5. Long-Run Gains from Liberalizing Distortions: Multiproduct versus Single-Product Models

in the single-product version, the static gains only reach 15 percent in the former, significantly below the 50 percent gain that can be reached in the latter.

The key driver of the difference in the magnitude of the static gains lies in the change in the distribution of *TFPQ* that is induced, in the multi-variety model, by the firms' incentives to adjust their portfolio of products. When firms are multiproduct, *TFPQ* at the firm level is a combination of the exogenous firm-wide productivity level and the endogenously determined number of products supplied by each firm. Allocative distortions that correlate positively with firm-wide productivity induce a reallocation of products from high to low productivity firms that compresses the distribution of *TFPQ* in the distorted stationary equilibrium. Therefore, when computing the gains from reallocation keeping the distribution of products across firms fixed, as in the *static* allocation, we are holding back the scope for efficiency gains by reallocating labor among firms that are not too different from each other in terms of their measured physical productivity.[20] This is in stark contrast with the single-product version of the model, where the dispersion of firms across the space of *TFPQ* is not affected by the existence of allocative distortions (except only through changes in the exit cutoff), and hence, the scope for resource misallocation is not minimized. Notice, then, the importance of having recalibrated the variance of the idiosyncratic productivity shocks.

The point of inflection in the shape of the static gains in the multiproduct model is also intricately related to the endogeneity of the *TFPQ* distribution in this model. In particular, it is an outcome of the interaction between the reallocation of products towards unproductive firms, and the existence of a small fraction of firms that supplies the whole range of products in the undistorted equilibrium. At low levels

[20] The intuition that the properties of the underlying distribution of *TFPQ* determine the detrimental effects of a given distribution of distortions on aggregate productivity is formalized in Hopenhayn (2014).

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

of distortions, there is still a remaining fraction of firms that supply all products. These firms do not contribute to the compression of the distribution of *TFPQ* that happens among the firms that provide only a subset of the available goods. Since this compression was the key for the relatively lower magnitude of static gains in the multiproduct model, the survival of full-range firms tends to keep the static gains high. However, as we move along the space of distortions slopes, the fraction of these firms shrinks and, hence, so does the magnitude of the static gains, which explains the concavity of its shape at relatively milder levels of distortions. Once the economy reaches the point where no such firms exist in the distorted equilibrium, between $\gamma = 0.3$ and $\gamma = 0.4$, there is no such compensating force, so distortions become increasingly costly in terms of forgone static productivity gains.

Overall, the results of this section shed light on the importance of jointly considering margins of adjustment that contribute to aggregate productivity through changes in the number of varieties. It highlights that focusing only on entry and exit would mistakenly attribute a negative contribution to the variety channel that may become overturned once changes in varieties that operate through multiproduct incumbents are taken into account. Furthermore, it emphasizes the importance of recalibrating the properties of the distribution of idiosyncratic productivity when comparing the effect of allocative distortions in models with endogenous *TFPQ* against models where this object is exogenous.

III. Concluding Remarks

Allocative distortions have acquired a pivotal role among economists in their thinking about growth. The central hypothesis is that by inducing a suboptimal allocation of resources across firms, this type of distortion could become an important driver of the large productivity gaps that we observe across countries. While this hypothesis led to promising results, it focused mostly on accounting for a static allocative channel through which misallocation can be detrimental to productivity. However, less attention was given to the goal of understanding the implications of these distortions for other firm-level decisions that, besides being empirically relevant margins of adjustment, have direct effects over welfare and productivity. In this paper, we provided an integrated investigation of the effects of allocative distortions for three particular channels: the entry, exit, and product-provision decisions of firms.

We found an ambiguous result with respect to the contributions of these channels to the long-run gains in productivity from removing distortions, but showed that these forces have a significant reinforcing contribution when it comes to shaping the improvement in welfare. Quantitatively, our results showed that even if the calibration of the distortions led to a modest gain in long-run *TFP*, reforms that dismantle these distortions can generate substantially higher improvements in welfare, more than three times as high as the improvement in long-run aggregate productivity.

The paper's approach to modeling misallocation was very stylized. Even though the resulting tractability allowed us to accurately account for the nature and extent of allocative distortions in the data, the large welfare gains associated with their removal calls for a deeper investigation of what the sources of misallocation really

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

are. Several candidates have already been explored in the literature, such as financial frictions, firing costs, and heterogeneous markups,[21] but none of these seem to be able to account for most of the misallocation that has been documented in the data. Our results provide a different motivation for why thinking about the actual drivers of misallocation is important. Even if these were to leave a large portion of the development gaps unexplained, their welfare implications far exceed their contributions to economic growth.

Appendix: Calibrating the Distribution of Product Attributes

This Appendix shows how to use the data about the distribution of output shares in multiproduct firms to calibrate the shape parameter of the Pareto distribution of product attributes. We restrict attention to the case where firms are in the interior of the product space.

In the model, the output of a firm with productivity e^{ω}, producing a good with product attribute e^{λ}, is given by

$$q(\omega,\lambda) \propto \left(e^{\omega+\lambda}\right)^{\frac{\rho}{\rho-1}},$$

where the factor of proportionality is a function of prices and other parameters in the model.

Under Pareto distribution of product attributes, the total output of the firm is

$$Q(\omega) \propto \left(e^{\omega}\right)^{\frac{\rho}{\rho-1}}\left(e^{\bar{\lambda}(\omega)}\right)^{-\left(\eta-\frac{\rho}{\rho-1}\right)}\frac{\eta}{\left(\eta-\frac{\rho}{\rho-1}\right)}.$$

Therefore, the output share of a product $e^{\omega+\lambda}$ is

$$x(\omega,\lambda) = \frac{q(\omega,\lambda)}{Q(\omega)} = \frac{\left[\eta-\frac{\rho}{\rho-1}\right]}{\eta}\frac{\left(e^{\lambda}\right)^{\frac{\rho}{\rho-1}}}{\left(e^{\bar{\lambda}(\omega)}\right)^{-\left(\eta-\frac{\rho}{\rho-1}\right)}}.$$

To determine the shape of the distribution of this share, its cumulative distribution function will be given by

$$\int_{x(\omega,\bar{\lambda})}^{x'(\omega,\lambda)} x(\omega,\lambda)\,dF(x) = \frac{\left[\eta-\frac{\rho}{\rho-1}\right]}{\eta}\left(e^{\bar{\lambda}(\omega)}\right)^{\left(\eta-\frac{\rho}{\rho-1}\right)}\int_{e^{\bar{\lambda}(\omega)}}^{e^{\lambda'}}\eta\left(e^{\lambda}\right)^{-\left[\left(\eta-\frac{\rho}{\rho-1}-1\right)\right]}d\left(e^{\lambda}\right)$$

$$\Leftrightarrow$$

$$\int_{x(\omega,\bar{\lambda})}^{x'(\omega,\lambda)} x(\omega,\lambda)\,dF(x) = \int_{e^{\bar{\lambda}(\omega)}}^{e^{\lambda'}}\left(\eta-\frac{\rho}{\rho-1}-1\right)\frac{\left(e^{\bar{\lambda}(\omega)}\right)^{\left(\eta-\frac{\rho}{\rho-1}\right)}}{\left(e^{\lambda}\right)^{-\left[\left(\eta-\frac{\rho}{\rho-1}-1\right)\right]}}d\left(e^{\lambda}\right),$$

[21] See, for instance, Veracierto (2001), Buera and Shin (2013), and Peters (2016).

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms

which implies that the share of a given product's output over total firm output is distributed Pareto, with shape parameter $\left(\eta - \frac{\rho}{\rho - 1}\right)$, and lower bound of the support equal to $e^{\bar{\lambda}(\omega)}$. Therefore, the 0.5 estimate Bernard, Redding, and Schott (2010) finds for the regression between log of rank and log of output share establishes that

$$0.5 = \eta - \frac{\rho}{\rho - 1}.$$

Given our calibrated value of $\rho = 3$, we infer that $\eta \cong 2$.

REFERENCES

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic Inputs and Resource (Mis)Allocation." *Journal of Political Economy* 122 (5): 1013–63.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review* 103 (1): 305–34.

Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott. 2010. "Multiple-Product Firms and Product Switching." *American Economic Review* 100 (1): 70–97.

Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott. 2011. "Multiproduct Firms and Trade Liberalization." *Quarterly Journal of Economics* 126 (3): 1271–1318.

Broda, Christian, and David E. Weinstein. 2006. "Globalization and the Gains From Variety." *Quarterly Journal of Economics* 121 (2): 541–85.

Buera, Francisco J., and Yongseok Shin. 2013. "Financial Frictions and the Persistence of History: A Quantitative Exploration." *Journal of Political Economy* 121 (2): 221–72.

Camacho, Adriana, and Emily Conover. 2010. "Misallocation and Productivity in Colombia's Manufacturing Industries." Inter-American Development Bank (IDB) Working Paper IDB-WP-123.

Casacuberta, Carlos, and Néstor Gandelman. 2009. "Productivity, exit and crisis in Uruguayan manufacturing and services sectors." http://www.bcu.gub.uy/Comunicaciones/Jornadas%20de%20Economa/iees03j3541010.pdf.

Chen, Kaiji, and Alfonso Irarrazabal. 2015. "The role of allocative efficiency in a decade of recovery." *Review of Economic Dynamics* 18 (3): 523–50.

Cirera, Xavier, Roberto N. Fattal Jaef, and Hibret B. Maemir. 2017. "Taxing the Good? Distortions, Misallocation, and Productivity in Sub-Saharan Africa." Policy Research Working Paper 7949.

Hopenhayn, Hugo A. 2014. "On the Measure of Distortions." National Bureau of Economic Research (NBER) Working Paper 20404.

Hsieh, Chang-Tai, and Peter J. Klenow. 2007. "Misallocation and Manufacturing TFP in China and India." Stanford Institute for Economic Policy Research Discussion Paper 07-06.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics* 124 (4): 1403–48.

Luttmer, Erzo G. J. 2007. "Selection, Growth, and the Size Distribution of Firms." *Quarterly Journal of Economics* 122 (3): 1103–44.

Luttmer, Erzo G. J. 2010. *On the Mechanics of Firm Growth*. Federal Reserve Bank of Minneapolis. Minneapolis, March.

Neumeyer, Andres, and Guido Sandleris. 2009. "Productivity and resource misallocation in the Argentine manufacturing sector, 1997–2002." Unpublished.

Peters, Michael. "Heterogeneous Markups, Growth and Endogenous Misallocation." http://mipeters.weebly.com/uploads/1/4/6/5/14651240/markupsmisallocation.pdf.

Restuccia, Diego, and Richard Rogerson. 2008. "Policy distortions and aggregate productivity with heterogeneous establishments." *Review of Economic Dynamics* 11 (4): 707–20.

Stokey, Nancy L. 2008. *The Economics of Inaction: Stochastic Control Models with Fixed Costs*. Princeton: Princeton University Press.

Veracierto, Marcelo. 2001. "Employment Flows, Capital Mobility, and Policy Analysis." *International Economic Review* 42 (3): 571–96.

This content downloaded from
104.234.151.63 on Fri, 14 Feb 2025 07:19:13 UTC
All use subject to https://about.jstor.org/terms