

Equilibrium Technology Diffusion, Trade, and Growth

Jesse Perla

University of British Columbia

Christopher Tonetti

Stanford Graduate School of Business and NBER

Michael E. Waugh

New York University and NBER

First Draft: November 2012

Revised: March 2020

ABSTRACT

We study how opening to trade affects economic growth in a model where heterogeneous firms can adopt new technologies already in use by other firms in their home country. We characterize the growth rate using a summary statistic of the profit distribution—the mean-min ratio. Opening to trade increases the profit spread through increased export opportunities and foreign competition, induces more rapid technology adoption, and generates faster growth. Quantitatively, these forces produce large welfare gains from trade by increasing an inefficiently low rate of technology adoption and economic growth.

JEL Classification: A1, E1, O4, F00, F43

Keywords: growth, trade, technology diffusion, productivity dynamics

We have benefited from discussions with Treb Allen, Jess Benhabib, Paco Buera, Sebastian Di Tella, Dave Donaldson, Chad Jones, Boyan Jovanovic, Pete Klenow, Bob Lucas, Andrés Rodríguez-Clare, Tom Sargent, Laura Veldkamp, and Gianluca Violante. Excellent research assistance was provided by Chiyoung Ahn, Sev Hou, Arnav Sood, and Dongxiao Zhang. We thank seminar participants at many universities, conferences, and research centers for useful comments. Code associated with this paper can be publicly accessed on our GitHub site at <https://github.com/jlperla/PerlaTonettiWaugh.jl>.
Contact: tonetti@stanford.edu, jesse.perla@ubc.ca, mwaugh@stern.nyu.edu.

1. Introduction

A large body of evidence documents trade-induced productivity effects at the firm level (see, e.g., Pavcnik (2002), Holmes and Schmitz (2010) and Shu and Steinwender (2019)). Why does opening to trade lead to productivity gains at the firm level? What are the consequences of these within-firm productivity gains for aggregate economic growth and welfare?

This paper contributes new answers to these questions. We develop a model where heterogeneous firms choose either to produce with their existing technology or adopt a better technology already in use by other firms in their home country. These choices determine the productivity distribution from which firms can acquire new technologies and, hence, the equilibrium rate of technological diffusion and economic growth. Because firms do not internalize the social benefits of technology adoption, the economy grows at an inefficiently low rate.

We provide a closed form characterization of the economy showing how the reallocation effects of a trade liberalization (i.e., low productivity firms contract, high-productivity exporting firms expand) change firms' incentives to adopt a better technology and lead to faster within-firm productivity gains. Because these choices lead to more adoption and technology diffusion, the aggregate consequence is faster economic growth. Quantitatively, trade-induced increases in technology adoption and aggregate growth lead to large welfare gains from trade because the economy is inefficient.

The starting point of our analysis is a standard heterogeneous firm model in differentiated product markets as in Melitz (2003). Firms are monopolistic competitors who differ in their productivity/technology and have the opportunity to export after paying a fixed cost. There is free entry from a large measure of potentially active firms and firms exit at an exogenous rate. Our model of technology adoption and diffusion builds on Perla and Tonetti (2014), where firms choose to either upgrade their technology or continue to produce with their existing technology in order to maximize expected discounted profits for the infinite horizon. If a firm decides to upgrade its technology, it pays a fixed cost in return for a random productivity draw from the equilibrium distribution of firms that produce in the domestic economy. We interpret this process as technology diffusion, since firms upgrade by adopting technologies already in use by other firms. Economic growth is a result, as firms are continually able to upgrade their technology by imitating other, better firms in the economy. Thus, this is a model of how endogenous technology diffusion contributes to growth—abstracting away from how innovation expands the technology frontier (see Benhabib, Perla, and Tonetti (2019)).

We study how opening to trade affects firms' technology choices and the aggregate consequences for growth and welfare. To do so we first study a simplified economy and characterize the profit and value functions of a firm, the evolution of the productivity distribution, and the growth rate of the economy on the balanced growth path equilibrium. We then study the quantitative model, which features productivity and exit shocks, and explore how changes in iceberg trade costs affect growth rates and the welfare gains from trade along the transition path.

We provide a closed form characterization of the growth rate as a simple, increasing function of a summary statistic of the profit distribution—the ratio of profits between the average and marginal adopting firm. A firm's incentive to adopt depends on two competing forces: the expected benefit of a new pro-

ductivity draw and the opportunity cost of taking that draw. The expected benefit relates to the profits that the average new technology would yield. The opportunity cost of adopting a new technology is the forgone profits from producing with the current technology. Thus, the aggregate growth rate of the economy encodes the trade-off that firms face in a simple and intuitive manner.

Reductions in iceberg trade costs increase the rate of technology adoption and economic growth because they widen the ratio of profits between the average and marginal adopting firm. As trade costs decline, low productivity firms contract as competition from foreign firms reduce their profits; high productivity firms expand and export, increasing their profits. For low productivity firms, this process reduces both the opportunity cost and increases the benefit of a new technology. This leads to more frequent technology adoption at the firm level. Because the frequency at which firms upgrade their technology is intimately tied to aggregate growth, the growth rate is higher in more open economies.

The underlying mechanism in our model is distinct from the standard “market size” effect, i.e., opening to trade increases the size of the market and, hence, raises the value of adoption. We show this by studying a special case of our model with no fixed exporting cost in which all firms export. In this model, growth is the same function of the spread in profits between the average and marginal adopting firm. The difference is that trade has no effect on growth. In this model, opening to trade benefits all firms by increasing firms’ profits and values by the same proportional amount. Consistent with the well understood benefits of a larger market, opening to trade increases the expected value of adopting a new technology. However, a larger market also raises the forgone profits of adoption by the exact same amount. Thus, opening to trade does not affect the *relative* benefit of adoption and, hence, there is no change in growth.

We provide a closed form characterization of the change in welfare from these growth effects. The change in welfare is a weighted sum of the increase in economic growth and the change in the initial level of consumption. We prove that opening to trade reduces the initial level of consumption which is a countervailing force that dampens the gains from faster economic growth. How faster growth competes with the loss in the level of consumption is a quantitative question.

To quantitatively study the gains from lower trade costs, we enrich the theoretical analysis in several directions. First, we extend our model to include firm-specific exit and productivity shocks. Thus, the dynamics of the firm are now driven by two forces: an exogenous stochastic component and the endogenous component that works through the adoption process. This modeling enrichment allows us to calibrate the model by matching moments of micro-level firm dynamics, providing discipline on the welfare gains beyond that imposed by cross-sectional firm heterogeneity moments alone.

We perform two main exercises with the calibrated model. The first quantitative exercise is a (local) decomposition of the welfare gains across steady states. The decomposition identifies the sources of the gains from trade in our economy, how they differ from the gains from trade in an efficient economy, and how they differ from benchmarks in the literature such as Atkeson and Burstein (2010) or Melitz (2003). The decomposition shows that almost all of the gains in our model are due to a highly-valued increase in the growth rate. The increase in growth is valued so highly because of the externality in the economy which makes the decentralized equilibrium have inefficiently low rates of technology adoption and gro-

wth. Even absent lower trade costs, there would be large welfare gains from just reallocating labor away from the production of goods and into adoption to improve growth. In contrast, if the economy were efficient, the gain from reallocating resources into adoption and away from other productive activity would be zero on the margin. Quantitatively, the effect due to this inefficiency is an order of magnitude larger than the direct consumption effect of lower trade costs.

The second quantitative exercise focuses on the welfare gains from a large decrease in trade costs inclusive of transition paths—not across steady states as in our theoretical analysis and local decomposition. The focus on the transition path addresses the concern that the across steady-state analysis might overstate the welfare gains because the benefits of higher growth are in the future and they require costly investment to implement. The quantitative version of the model and the transition path cannot be solved analytically, so we implement numerical methods to solve the PDEs.

We study a ten percent permanent and unanticipated reduction in trade costs. In response to the trade liberalization, the transition path of the economy is nontrivial as productivity growth, the number of varieties, and consumption take time to adjust to the change in openness to trade. On the transition path the level of consumption initially overshoots its new long-run steady state value. Consistent with our theoretical results, the measure of domestically produced varieties is smaller in the long run and, thus, net-exit must occur along the path. Along the transition to the new steady state, the result is that labor is allocated away from the creation of new firms. This labor is reallocated toward adoption activities to facilitate faster growth and toward the increased production of consumption goods. Thus, the level of consumption is elevated along a portion of the path as it converges to a new lower value relative to productivity.

The welfare gains from trade are large and the transition path slightly dampens them. The consumption-equivalent gains, inclusive of the transition path, are about eleven percent. As one point of comparison, the direct consumption effect of lower trade costs in our model are about what the formulas in Atkeson and Burstein (2010) or Arkolakis, Costinot, and Rodriguez-Clare (2012) would predict. The total welfare gains are an order of magnitude larger, however, due to the inefficiency generated by the adoption externality. Finally, we show how the size of the welfare gains are related to the moments used to calibrate the model, with a particular focus on how important the firm dynamics data, and not just data on firm heterogeneity in the cross-section, are in determining the welfare gains from trade.

1.1. Related Literature

We contribute to the theoretical literature on trade and growth. The standard mechanisms creating a relationship between trade and growth typically take two forms. First, openness leads to the cross-country diffusion of new and better ideas. Second, opening to trade increases the size of the market and, hence, raises the value of new idea creation/innovation. Depending on the details of the model, these mechanisms have been shown to increase economic growth as a country opens up to trade (see, for example, the pioneering works of Rivera-Batiz and Romer (1991) and Grossman and Helpman (1993) and their extensions to heterogeneous firm environments in Baldwin and Robert-Nicoud (2008)).

Our model differs from these traditional mechanisms. First, to focus on our distinct mechanism, we

deliberately shut down the cross-country diffusion of new and better ideas. In our model, firms only acquire ideas already present inside their country. Thus, our model delivers growth without any increase in the amount or quality of ideas coming from abroad as a country opens to trade. This distinction is also salient relative to recent work such as Alvarez, Buera, and Lucas (2017), Buera and Oberfield (2019), and Hsieh, Klenow, and Nath (2019).¹ Second, as mentioned above, when only market size effects are present, opening to trade has no effect on economic growth. The relationship between growth and trade in our model is not because a larger market increases the value of adoption; it's due to a change in the value of adoption that arises because a trade liberalization has differential effects on firms with different productivity levels.

Atkeson and Burstein (2010) is an important point of comparison for our welfare gains from trade calculations. They study a model of trade and firm dynamics with both process and product innovation featuring nontrivial transition dynamics. With all of the different margins that could adjust, one might think that there may be new sources of gains from trade. This is not the case, however, as the effects from changes in process and product innovation are second order. This logic essentially follows from the observation that the Atkeson and Burstein (2010) economy is efficient. Process and product innovation require resources. In an efficient economy the welfare gains from reallocating resources into innovation and away from other productive activity should be zero on the margin. Thus, the welfare gain from trade in Atkeson and Burstein (2010) is equal to a direct consumption effect, which is exactly that given in Arkolakis, Costinot, and Rodriguez-Clare (2012) for efficient stationary economies with CES demand and a constant trade elasticity.

In contrast to Atkeson and Burstein (2010), the decentralized equilibrium of our model is not efficient. Firms in our model are not adopting technologies at the socially optimal rate. This inefficiency opens a distinct potential source of welfare gains from trade. Our quantitative results show that the welfare gains are large—an order of magnitude larger than in Atkeson and Burstein (2010). And—most important—virtually all the welfare gains arise because reductions in trade costs increase an inefficiently low rate of adoption.

Sampson (2016) is another important point of comparison. He studies the effects of trade on growth when there is a dynamic complementarity between the ideas of entrants and those of the incumbents: trade induces exit of the worst performing firms and this implies that entrants are able to receive better ideas; because entrants are now better, this induces more selection and so on, leading to faster economic growth.² Most important, dynamic selection creates an externality which firms do not internalize as they make their exit decision. As in our model, the externality in Sampson's model provides extra scope for gains from trade because his economy is inefficient.

¹We study the case of symmetric countries, since absence international technology diffusion, balanced growth paths with asymmetric countries likely exist only for knife-edge conditions. We hope in this paper to study a particular mechanism whose illustration is aided by the analytical tractability obtained via the symmetry assumption. We conjecture that the mechanism we study would still be relevant in a model with international technology diffusion in which a balanced growth path equilibrium would robustly exist even with asymmetric countries.

²The actual source of aggregate productivity growth is very different. Sampson's (2016) model is one in which aggregate productivity growth (and its response to trade) is from the entry margin. In contrast, our model implies that most of aggregate productivity growth comes from within-firm improvements by incumbents, consistent with evidence in Garcia-Macia, Hsieh, and Klenow (2019).

Quantitatively, our welfare gains are larger than those in Sampson (2016) mostly due to our calibration scheme. Our approach is to generalize the Perla and Tonetti (2014) endogenous technology adoption framework by modeling the dynamics of firm productivity as also being shaped by exogenous firm-specific geometric Brownian motion (GBM) shocks. This allows our quantitative model to match the observed dynamics of firms and, hence, provide a consequential (but admittedly indirect) element of discipline on the welfare gains from trade. As we discuss in detail, calibrating the model so that it generates realistic firm dynamics leads to substantially larger gains from trade than those found in Sampson's model or those in our model when calibrated without matching firm dynamic moments.

We also connect with the literature on the relationship between competition and productivity. Arrow's (1962) "replacement effect" is a theoretical explanation for the positive effects of competition on growth. Broadly speaking, Arrow's (1962) idea is that a monopolist will prefer the status quo whereas an entrant or firm with lower market power has less to lose, and hence there is a greater incentive to innovate.

Our mechanism is very similar with the insight that competition reduces the opportunity cost of adoption, i.e., the benefits of the status quo fall with exposure to import competition. As our theoretical results make transparent, the adoption decision and aggregate growth rate depend on the comparison between the potential benefits of adoption versus the forgone profits of operating with the old technology. On its own, the erosion of profits from import competition incentivizes firms to adopt more frequently. In this sense, our model shares similarities with Holmes, Levine, and Schmitz Jr (2012) who show how competition reduces the cost of a switch-over disruption from a new technology and leads to more technology adoption.

Closely related to our work is Bloom, Romer, Terry, and Van Reenen (2014) which focuses explicitly on import competition, within-firm productivity improvements, and aggregate growth. Motivated by the evidence in Bloom, Draca, and Reenen (2016), they show import competition forces firms to innovate more than otherwise. While similar in spirit, the underlying mechanisms are different. Central to their results is the costly adjustment of factors of production within the firm. As firms face import competition, the resources within the firm that are costly to shed are redirected toward innovative activities. Furthermore, their mechanism amplifies, but is not distinct from the traditional market size effects on innovation.

We also contribute to the literature on idea flow models of economic growth in several ways. The most important is the introduction and study of competition effects which are new, additional, forces not present in the economies of Lucas and Moll (2014) and Perla and Tonetti (2014). We extend Perla and Tonetti (2014) in many directions: general equilibrium with labor and goods markets in continuous time with firm-specific geometric Brownian motion shocks to productivity and firm entry and exit, leading to an endogenous measure of varieties. These features allow our model to confront data on the dynamics of firms and shape our quantitative results. Also, our characterization of growth as a function of summary statistics of the profit distribution is a core contribution and is not present in Perla and Tonetti (2014). As our closed form characterization of the growth rate shows, without any relative change in firms' profits, opening to trade has no effect on economic growth. Thus the competition effects that we introduce, which act through a reallocation of profits due to the fixed cost of exporting, are key to delivering interesting relationships between trade, firms' technology choices, and economic growth. We also develop

numerical methods needed to study the transition dynamics of the economy. Finally, we provide an approach to studying the sources of welfare gains using a local decomposition.

1.2. Motivating Evidence: Trade-Induced Productivity Gains

Motivating our work is the empirical evidence that import competition gives rise to within-firm productivity improvements.

Pavcnik (2002) was an important empirical study of the establishment level productivity effects from a trade liberalization using frontier measurement techniques. Pavcnik (2002) studied Chile's trade liberalization in the late 1970s and she found large within-plant productivity improvements for import-competing firms that are attributable to trade. There was no evidence that exporters had any productivity improvements attributable to trade and no evidence of trade induced productivity gains from exit. To be clear, Pavcnik (2002) observed productivity improvements from exit, but there were no *differential* gains from exit across sectors of different trade orientation (i.e., import competing vs. non-traded, etc.). In contrast, import-competing firms had differentially larger within-plant productivity improvements.

Many subsequent studies for different countries and/or data sets have found similar results. In Brazil, Muendler (2004) found import competition led to within-firm productivity gains. Several studies of India's trade liberalization find related results. Topalova and Khandelwal (2011) found large within-firm productivity gains associated with declines in output tariffs which proxy for increases in import competition. Also in India, Sivadasan (2009) finds increases in industry TFP from tariff reductions, with 55 percent of these gains associated with within-firm productivity gains.

Bloom et al. (2016) find within-firm productivity gains in Europe from Chinese import competition. Most importantly, they associate these gains with explicit measures of technical change, e.g., information technology, management practices, and other measures of innovation. Their evidence suggests that firms undertook activities to change the technology with which they operate in response to import competition. Autor, Dorn, Hanson, Pisano, and Shu (2019), however, find that U.S. firms reduced their patenting in response to a rise in Chinese import-competition. In response to tariff reductions in China, Fieler and Harrison (2019) find evidence for larger productivity increases among the initially low-productivity firms and Bombardini, Li, and Wang (2018) find evidence for increases in patenting among the initially high-productivity firms.

Despite the large body of empirical work, theory has lagged.³ Two common explanations for these within-firm productivity gains fall under the category of imperfect measurement. The first explanation is that these gains may reflect changes in the mix of intermediate inputs. For the cases of Indonesia (studied in Amiti and Konings, 2007) and India (Goldberg, Khandelwal, Pavcnik, and Topalova, 2010), there is strong evidence for this mechanism. A second explanation is that they reflect changes in product mix (see, e.g., Bernard, Redding, and Schott, 2011). While these are likely contributing forces, there is evidence they are not the whole story. For example, Bloom et al. (2016) find little evidence that they are

³The standard heterogeneous firm framework of Melitz (2003) does not provide an explanation for these effects. Melitz (2003) deals exclusively with the reallocation of activity across firms; there is no mechanism to generate within-firm productivity growth in response to a trade liberalization. One caveat to this statement is how the overhead costs are treated in the model and, in turn, measured in the data.

the source of the gains in their study.

Non-measurement explanations fall under the guise of “X-efficiency” gains (Leibenstein, 1966). X-efficiency gains can be difficult to understand since it is natural to ask the question: If it was possible for a firm to improve its efficiency after a change in competition, why did it not do it in the first place? One mechanism for X-efficiency gains is that competition relaxes the agency problems within the firm (see, e.g., Raith, 2003; Schmidt, 1997). In our model, increased import competition increases the profitability of technological improvement by lowering the opportunity cost of adoption relative to the returns of adoption. This competition driven increase in the pace of technology adoption leads to within-firm productivity gains that generate faster aggregate economic growth.

Admittedly, there are aspects of firm-level adjustments to trade liberalizations about which we have little to say. In particular, the evidence on firm markups or the productivity enhancing role of becoming an exporter (see, e.g., Bustos (2011) or Marin and Voigtländer (2019)). Theoretical mechanisms studying these aspects of exporters and their quantitative evaluation is far more developed with some examples being, Melitz and Costantini (2007), Atkeson and Burstein (2010), Akcigit, Ates, and Impullitti (2018), and Cavenaile, Roldan-Blanco, and Schmitz (2020). We deliberately focuses on the adoption behavior and how it’s the import exposed firms that respond. Jointly modeling adoption and innovation at the firm level, as in Benhabib et al. (2019), and studying how firm behavior responds to changes in trade and competition is an important area for future research. Further empirical research on the differential impact of changes in trade costs across firms throughout the productivity distribution will be invaluable in this endeavor.

2. Model

2.1. Countries, Time, Consumers

There are N symmetric countries. Time is continuous and evolves for the infinite horizon. Utility of the representative consumer in country i is

$$U_i(t) = \int_t^\infty e^{-\rho(\tau-t)} \log(C_i(\tau)) d\tau. \quad (1)$$

The utility function $U_i(t)$ is the present discounted value of the instantaneous utility of consuming the final good. The discount rate is $\rho > 0$ and instantaneous utility is logarithmic.⁴ The final consumption good is an aggregate bundle of varieties, aggregated with a constant elasticity of substitution (CES) function by a competitive final goods producer.

Consumers supply labor to firms for the production of varieties and the fixed costs of exporting, technology adoption, and entry. Labor is supplied inelastically and the total units of labor in a country are \bar{L}_i . Consumers also own a diversified portfolio of all firms operating within their country. Thus, consumer income is the sum of net profits from firms and total payments to labor. The consumer budget constraint

⁴The model easily generalizes to CRRA power utility, as shown in the Appendix, but analytical characterizations are less sharp.

is

$$C_i(t) \leq \frac{W_i(t)}{P_i(t)} \bar{L}_i + \bar{\Pi}_i(t) - \bar{I}_i(t) \quad (2)$$

where $W_i(t)$ is the nominal wage rate in country i , $P_i(t)$ is the standard CES price index of the aggregate consumption good, $\bar{\Pi}_i(t)$ is aggregate profits, and $\bar{I}_i(t)$ is aggregate investment in entry and technology adoption. See the Appendix for more details.

2.2. Firms

In each country there is a final good producer that supplies the aggregate consumption good competitively. The final good is produced by aggregating an endogenous measure of intermediate varieties produced by monopolistically competitive firms, both domestically and abroad. Variety producing firms are heterogeneous over productivity, Z , with cumulative distribution function $\Phi_i(Z, t)$ and probability density function $\phi_i(Z, t)$ describing how productivity varies across firms, within a country. Each firm alone can supply variety v . As is standard, a final good producer aggregates these individual varieties using a constant elasticity of substitution production function.

2.2.1. Final Good Producer.

Dropping the time index for expositional clarity, a representative final good producer chooses the quantity of intermediate goods to maximize profits:

$$\max_{Q_{ij}(v)} \left[\sum_{j=1}^N \int_{\Omega_{ij}} Q_{ij}(v)^{(\sigma-1)/\sigma} \right]^{\sigma/(\sigma-1)} - \sum_{j=1}^N \int_{\Omega_{ij}} p_{ij}(v) Q_{ij}(v). \quad (3)$$

Y_i is defined to be nominal aggregate expenditure on consumption goods. The parameter $\sigma > 1$ is the elasticity of substitution across varieties. The measure Ω_{ij} defines the endogenous set of varieties consumed in country i produced in country j . Furthermore, the total measure of varieties produced in country i , Ω_i , is also determined in equilibrium, as domestic firms can enter after paying a fixed cost and exit if hit with an exogenous death shock that arrives at rate $\delta \geq 0$.

We will drop the notation carrying around the variety identifier, as it is sufficient to identify each firm with its location and productivity level, Z . Additionally, to focus on the interactions between technology adoption, trade, and growth, we assume that all countries are symmetric in that they have identical parameter values, although each intermediate producer in each country produces a unique good. Because all countries are symmetric, we focus on the results for a typical country and abstract from notation indicating the country's location.

This final good producer problem yields the familiar variety demand and price index equations:

$$Q(Z) = \left(\frac{p(Z)}{P} \right)^{-\sigma} \frac{Y}{P}, \quad (4)$$

$$P^{1-\sigma} = \Omega \left(\int_M^{\infty} p_d(Z)^{1-\sigma} d\Phi(Z) + (N-1) \int_{\hat{Z}}^{\infty} p_x(Z)^{1-\sigma} d\Phi(Z) \right). \quad (5)$$

where p_d and p_x are the prices of domestic and imported varieties, M is the minimum of support of the distribution, and \hat{Z} is an export threshold—all endogenous variables determined in equilibrium.

2.2.2. Individual Variety Producers.

Variety producing firms hire labor, ℓ , to produce quantity Q with a linear production technology: $Q = Z\ell$. Firms can sell freely in their domestic market and also have the ability to export at some cost, controlled by parameter κ . To export, a firm must pay a fixed flow labor cost, $\kappa\bar{L}W/P$, per foreign export market. This fixed cost is paid in market real wages and is proportional to the number of consumers accessed.⁵ Exporting firms also face iceberg trade costs, $d \geq 1$, to ship goods abroad.

Firm productivity evolves according to geometric Brownian motion, which captures the high empirical persistence of firm productivity, while also allowing for changes in firms' relative productivity to occur throughout the productivity distribution. Furthermore, at each instant, any firm can pay a real fixed cost $X(t)$ to adopt a new technology. $X(t)$ represents the cost of hiring labor to upgrade to a higher-efficiency production technology (detailed below). If a firm decides to pay this cost, it receives a random draw from the distribution of producers within its own country, as in Perla and Tonetti (2014).

Given this environment, firms must make choices regarding how much to produce, how to price their product, whether to export, and whether to change their technology. These choices can be separated into problems that are static and dynamic. Below we first describe the more standard static problem of a firm and then describe the dynamic problem of the firm to derive the optimal technology adoption policy.

Firms' Static Problem. Given a firm's location, productivity level, and product demand, the firm's static decision is to choose the amount of labor to hire, the prices to set, and whether to export for each destination to maximize profits each instant. The firm's problem when operating within the domestic market is to choose a price and quantity of labor to maximize profits. Using the standard demand function for individual varieties (equation (4)), the optimal domestic real profit function is

$$\Pi_d(Z) = \frac{1}{\sigma} \left(\frac{\bar{\sigma}W}{ZP} \right)^{1-\sigma} \frac{Y}{P}, \quad \text{where } \bar{\sigma} := \frac{\sigma}{\sigma-1}. \quad (6)$$

$\bar{\sigma}$ is the standard markup over marginal cost.

The decision to (possibly) operate in an export market is similar, but differs in that the firm faces variable

⁵Export costs that are proportional to the number of consumers is consistent with the customer access interpretation featured in Arkolakis (2010). As discussed in Section 5, this influences the population scale effect properties of the model, but has no other impact.

iceberg trade costs and a fixed cost to sell in the foreign market. Optimal per-market real export profits are

$$\Pi_x(Z) = \max \left\{ 0, \frac{1}{\sigma} \left(\frac{\bar{\sigma} dW}{ZP} \right)^{1-\sigma} \frac{Y}{P} - \frac{\kappa \bar{L}W}{P} \right\}. \quad (7)$$

where d is an iceberg trade costs and $\kappa \bar{L}W/P$ is the fixed cost to sell in the foreign market. Given profits described in equation (7), only firms earning positive profits from exporting—those above a productivity threshold \hat{Z} —actually enter a foreign market. Total real firm profits equal the sum of domestic profits plus the sum of exporting profits across export markets,

$$\Pi(Z, t) := \Pi_d(Z, t) + (N - 1)\Pi_x(Z, t). \quad (8)$$

Firms' Dynamic Problem. Given the static profit functions and a perceived law of motion for the productivity distribution and adoption cost, each firm has the choice of when to acquire a new technology, Z . Define $g(t)$ as the growth rate of total expenditures.⁶ Firms choose technology adoption policies to maximize the present discounted expected value of real profits. Since firms are owned by consumers, the firm's effective discount rate, $r(t)$, equals the interest rate derived from the consumer's Euler equation plus the firm's probability of exit, δ . That is, $r(t) = (\rho + g(t) + \delta)$.

The productivity of a non-adopting firm evolves exogenously according to geometric Brownian motion (GBM):

$$dZ_t/Z_t = (\mu + v^2/2)dt + v dW_t, \quad (9)$$

where $\mu \geq 0$ is related to the drift of the productivity process, $v \geq 0$ is the volatility, and W_t is standard Brownian motion. For tractability and clarity of exposition, our theoretical analysis in Sections 5 and 6 focuses on the case in which productivity only changes via adoption (i.e., $\mu = 0, v = 0$). In Section 7 we calibrate the model with GBM to quantify the welfare gains from trade.

The recursive formulation of the firm's problem is as follows. Each instant, a firm with productivity Z chooses between continuing with its existing technology and earning flow profits of $\Pi(Z, t)$ or stopping and adopting a new technology at cost $X(t)$. In a growing economy, adoption opportunities will improve and the firm's profits will erode, decreasing the benefits of continuing to operate its existing technology until the firm enters the stopping region and it chooses to adopt a new technology.

Define the value of the firm in the continuation region as $V(Z, t)$, $M(t)$ as the time dependent productivity boundary between continuation and adoption, and $V_s(t)$ as the expected value of adoption net of costs. $M(t)$ is a reservation productivity function, such that all firms with productivity less than or equal to $M(t)$ choose to adopt and all other firms produce with their existing technology. If a firm chooses to adopt a new technology it pays a cost and immediately receives a new productivity. This new producti-

⁶Since in a balanced growth path equilibrium many growth rates will be equal (e.g., the growth rate of total expenditures, consumption, and the minimum of the productivity distribution), we will abuse notation for the sake of exposition and overload the definition of a single growth rate: $g(t)$.

ivity is a random draw from the cross-sectional productivity distribution of firms.⁷ This distribution will be a function of the optimal policy of all firms, i.e., the firm choice of when to draw a new productivity. Recursively, the optimal policy of firms will depend on the expected evolution of this distribution. With rational expectations, the expected net value of adoption in equilibrium is

$$V_s(t) = \int_{M(t)}^{\infty} V(Z, t) d\Phi(Z, t) - X(t). \quad (10)$$

There are several interpretations of this technology adoption choice. The literal interpretation is that firms are randomly matching and learning from each other. Empirically, this technology choice can be thought of as tangible or intangible investments that manifest themselves as improvements in productivity like improved production practices, work practices, supply-chain and inventory management, etc. that are already in use by other firms (see, for example, the discussions in Holmes and Schmitz (2010) and Syverson (2011)).

Using the connection between optimal stopping and free boundary problems, a set of partial differential equations (PDEs) and boundary conditions characterize the firm's value.⁸ The PDEs and boundary values determining a firm's value are

$$r(t)V(Z, t) = \Pi(Z, t) + \left(\mu + \frac{v^2}{2} \right) Z \frac{\partial V(Z, t)}{\partial Z} + \frac{v^2}{2} Z^2 \frac{\partial^2 V(Z, t)}{\partial Z^2} + \frac{\partial V(Z, t)}{\partial t}, \quad (11)$$

$$V(M(t), t) = \int_{M(t)}^{\infty} V(Z, t) d\Phi(Z, t) - X(t), \quad (12)$$

$$\frac{\partial V(M(t), t)}{\partial Z} = 0. \quad (13)$$

Equation (11) describes how the firm's value function evolves in the continuation region. It says that the flow value of the firm is the flow value of profits (the first term) plus the expected capital gain (the last three terms). The capital gain term is the change in the value of the firm over time. It is comprised of three terms: the first two reflect the expected change in productivity which arise from both the drift and variance in the productivity process; the third term reflects a change in value of producing with a given productivity.

Equation (12) is the value matching condition. It says that at the reservation productivity level, $M(t)$, the firm should be indifferent between continuing to operate with its existing technology and adopting

⁷In discrete time, Perla and Tonetti (2014) presents both draws conditional on only meeting adopters and unconditionally from the whole distribution. Qualitatively, these two environments are very similar and in the limit to continuous time they become identical. See Benhabib et al. (2019) for a discussion and a proof that the unconditional and conditional draw models generate identical equilibrium laws of motion for the productivity distribution.

⁸Standard references and conditions for the equivalence between optimal stopping of stochastic processes and free boundary problems are Dixit and Pindyck (1994) and Peskir and Shiryaev (2006). The deterministic stopping problem is discussed on pages 110-115 of Stokey (2009).

a new technology. That is the definition of the reservation productivity. Equation (13) is the smooth-pasting condition. The smooth pasting condition can be interpreted as an intertemporal no-arbitrage condition that ensures the recursive system of equations is equivalent to the fundamental optimal stopping problem. It ensures that, around the optimal behavior, there is not a discontinuous increase in value associated with an infinitesimal difference in the timing of adoption (or equivalently the reservation productivity).

A couple of comments are in order regarding the economics of this problem.

There are two forces that drive the adoption decision. First, over time the productivity distribution will improve. This eventually gives firms an incentive to adopt a new technology as the benefit of adoption grows over time since a firm's *relative* position in the productivity distribution will drive its adoption behavior. This economic force is the same as in Lucas and Moll (2014) and Perla and Tonetti (2014). Similarly, negative productivity shocks may also drive a firm to abandon its current technology and adopt a new one as its relative position in the distribution will have weakened and the benefits of adoption have risen.

Second, competition and general equilibrium effects are new, additional, forces—not present in Lucas and Moll (2014) and Perla and Tonetti (2014)—which drive the adoption decision. The dependence of the firm's value function (equation 11) on profits (which are time dependent) illustrates this feature. As the economy grows, holding fixed an individual firm's productivity, its profits will erode. The reason is because as other firms become better through adoption, they demand relatively more labor, and this raises wages which reduces profits. This erosion of profits reduces the opportunity cost of continuing to operate and incentivizes adoption. Our paper is about this second force—how equilibrium changes in competition and profits via trade influence adoption and growth.

Finally, there is an externality in this environment. Firms are infinitesimal and do not internalize the effect their technology adoption decisions have on the evolution of the productivity distribution and, in turn, the distribution from which other firms are able to adopt. This externality could be interpreted as a free rider problem, as firms have an incentive to wait before upgrading, and let other firms adopt first, in order to have a better chance of adopting a more productive technology.

Together with the static optimization problem, equations (11), (12), and (13) characterize optimal firm policies given equilibrium prices and a law of motion for the productivity distribution.

2.3. Adoption Costs

Technology adoption is costly. In our baseline specification, this cost takes the form of labor the firm must hire; an adopting firm must hire $\zeta \bar{L}$ workers and pay them each $\frac{W(t)}{P(t)}$. The quantity of labor required to adopt a technology scales with population size to prevent scale effects and depends on the parameter

$\zeta > 0$.⁹ Thus, the real cost of adoption denoted by X , is

$$X(t) := \zeta \bar{L} \frac{W(t)}{P(t)}, \quad (14)$$

The product of this quantity and the real wage determines the real cost of adoption. Note that the specification ensures that adoption cost grows in proportion to the real wage, and, thus, ensures the cost does not become increasingly small as the economy grows.

We define $S(t)$ to be the adoption rate, such that the flow of intermediate firms adopting a new technology is $\Omega(t)S(t)$, i.e., the measure of varieties times the rate of adoption.

2.4. Entry, Exit, and the Measure of Varieties

There is a large pool of non-active firms that may enter the economy by paying an entry cost to gain a draw of an initial productivity from the same distribution from which adopters draw—the cross-section of incumbent productivities. We model the cost of entry as a multiple of the adoption cost for incumbents, $X(t)/\chi$, where $0 < \chi < 1$. Hence, χ is the ratio of adoption to entry costs. The restriction that $\chi \in (0, 1)$ means that the cost to incumbents of upgrading to a better technology is lower than the cost to entrants of starting to producing a new variety from scratch. Thus, the free entry condition is

$$X(t)/\chi \geq \int_{M(t)}^{\infty} V(Z, t) d\Phi(Z, t), \quad (15)$$

which will hold with equality on the BGP, as firms enter until the value of entry equals the cost.

Exit occurs because firms die at an exogenous rate $\delta \geq 0$ that is independent of firm characteristics. For tractability, our theoretical results will focus on the limiting case in which there is no firm death ($\delta = 0$) and, hence, no entry.¹⁰ Even in the limiting case when ($\delta = 0$), the equilibrium number of varieties (firms), Ω , is endogenous and determined by the free entry condition. This allows us to study the growth and welfare implications of lowering trade costs on the endogenous number of varieties.

We define $E(t)$ to be the entry rate, such that the flow of firms entering the market and creating a new variety equals the measure of varieties times the entry rate, $\Omega(t)E(t)$.

2.5. The Productivity Distribution

The final element of the economic environment to describe is the law of motion for the productivity distribution. We highlight the key elements below and provide a detailed derivation in Appendix B.

First, the minimum of the support of the productivity distribution is the adoption reservation productivity $M(t)$. Recall, that when adopting, a firm receives a random productivity draw from the distribution of producers. Except, perhaps, at time 0, the probability of drawing the productivity of a fellow adopting firm is infinitesimal. Therefore, firms adopting at time t will adopt a technology above $M(t)$ almost

⁹Since all analysis in this paper is done holding population constant, the assumption that the cost scales with population size does not affect any result in the paper.

¹⁰In the Appendix, we solve the model with an arbitrary death rate $\delta \geq 0$ and our quantitative analysis of the welfare gains from trade in Section 7 uses the model calibrated to a positive death rate.

certainly. Thus, $M(t)$ is like an absorbing barrier sweeping through the distribution from below and, thus, is the minimum of the support.

The Kolmogorov Forward Equation (KFE) describes the evolution of the productivity distribution for productivities above the minimum of the support. That is, the KFE describes the inflows and outflows of firms for each point in the support of the distribution. The KFE (written in terms of the CDF) is:

$$\begin{aligned} \frac{\partial \Phi(Z, t)}{\partial t} = & \underbrace{\Phi(Z, t) \left(\underbrace{S(t) + E(t)}_{\text{adopt or enter}} \right)}_{\text{distributed below } Z} - \underbrace{S(t)}_{\text{adopt at } M(t)} - \underbrace{\delta \Phi(Z, t)}_{\text{Death}} \dots \\ & - \underbrace{\left(\mu - \frac{v^2}{2} \right) Z \frac{\partial \Phi(Z, t)}{\partial Z}}_{\text{deterministic drift}} + \underbrace{\frac{v^2}{2} Z^2 \frac{\partial^2 \Phi(Z, t)}{\partial Z^2}}_{\text{Brownian motion}}. \end{aligned} \quad (16)$$

Let's walk through each term in (16) carefully. The left-hand side describes how the CDF evaluated at Z evolves over time. The first term reflects the inflows that arise from technology adoption and entry. Since adopting and entering firms' initial productivity is drawn from the existing distribution, $\Phi(Z, t)$, the total measure of firms flowing in below Z is $\Phi(Z, t) (S(t) + E(t))$. The second term reflects the loss of mass in the distribution which arise from the adopting firms. Since all adoption is done by firms at the minimum of support of the distribution, $S(t)$ is subtracted from the CDF for all $Z \geq M(t)$. The third term reflects exit, and since death occurs uniformly across all firms $\delta \Phi(Z, t)$ enters as an outflow. Finally, the last two terms reflect the expected drift due to GBM, which has a deterministic and random component.

Example: No GBM, No Death Case. To provide intuition, consider the small noise and small death rate limiting economy ($\delta = \mu = v = 0$). In this case the KFE simplifies to

$$\frac{\partial \Phi(Z, t)}{\partial t} = \underbrace{\underbrace{\Phi(Z, t)}_{\text{redistribution distribution}} \times \underbrace{S(t)}_{\text{adopters}}}_{\text{distributed below } Z} - \underbrace{S(t)}_{\text{adopters}} \quad (17)$$

The inflow in the fraction of firms with productivity less than Z comes from incumbents who draw a productivity below Z when adopting. This inflow equals the likelihood an adopter draws a productivity below Z time the rate of adoption. The outflow occurs only due to adoption. Since all adoption is done by firms at the minimum of support of the distribution, $S(t)$ is subtracted from the CDF for all $Z \geq M(t)$. The rate of adoption $S(t)$ is determined by

$$S(t) = M'(t) \phi(M(t), t). \quad (18)$$

In words, the rate at which the adoption threshold sweeps across the density, $M'(t)$, and the amount of firms the adoption boundary collects as it sweeps across the density from below, $\phi(M(t), t)$.

In this case, the solution to equation (16), which characterizes the productivity distribution at any date, is the probability density function

$$\phi(Z, t) = \frac{\phi(Z, 0)}{1 - \Phi(M(t), 0)}, \quad Z \geq M(t). \quad (19)$$

That is, the distribution at date t is a truncation of the initial distribution at the reservation adoption productivity at time t , $M(t)$. The solution in equation (19) is rather general. It holds independent of the type of the initial distribution and independent of whether the economy is on a balanced growth path.

3. Computing a Balanced Growth Path Equilibrium

In this section, we define and compute a Balanced Growth Path (BGP) equilibrium. Main results are then discussed in Sections 5-7. The Appendix documents the detailed steps involved in the computation of equilibrium and derivation of our main results.

3.1. Definition of a Balanced Growth Path Equilibrium.

Definition 1. A Balanced Growth Path Equilibrium consists of an initial distribution $\Phi(0)$ with support $[M(0), \infty)$, a sequence of distributions $\{\Phi(Z, t)\}_{t=1}^{\infty}$, firm adoption and export policies $\{M(t), \hat{Z}(t)\}_{t=0}^{\infty}$, firm price and labor policies $\{p_d(Z, t), p_x(Z, t), \ell_d(Z, t), \ell_x(Z, t)\}_{t=0}^{\infty}$, wages $\{W(t)\}_{t=0}^{\infty}$, adoption costs $\{X(t)\}_{t=0}^{\infty}$, measure of varieties Ω , and a growth rate $g > 0$ such that:

- Given aggregate prices, costs, and distributions
 - $M(t)$ is the optimal adoption threshold and $V(Z, t)$ is the continuation value function
 - $\hat{Z}(t)$ is the optimal export threshold
 - Ω is consistent with free entry
 - $p_d(Z, t)$, $p_x(Z, t)$, $\ell_d(Z, t)$, and $\ell_x(Z, t)$ are the optimal firm static policies
- The gross value of adoption and entry equal $\int_{M(t)}^{\infty} V(Z', t)\phi(Z', t)dZ'$
- Markets clear at each date t
- Aggregate expenditure is stationary when re-scaled: $Y(0) = Y(t)e^{-gt}$
- The distribution of productivities is stationary when re-scaled:

$$\phi(Z, t) = e^{-gt}\phi(Ze^{-gt}, 0) \quad \forall t, \quad Z \geq M(0)e^{gt}$$

In order to compute an equilibrium, we proceed in three steps. First, we use the law of motion for the productivity distribution (equation 16) and combine it with an assumption about the initial distribution, given a technology adoption policy $M(t)$. Second, given the law of motion, we solve for the firms' value function and adoption policy. Third, we solve for the growth rate g that ensures consistency between the first two steps.

4. Stationary Productivity Distribution.

To maintain tractability, we assume that the initial distribution $\Phi(0)$ is Pareto

$$\Phi(Z, 0) = 1 - \left(\frac{M(0)}{Z}\right)^\theta, \text{ with density, } \phi(Z, 0) = \theta M(0)^\theta Z^{-\theta-1}, \quad (20)$$

where θ is the shape parameter and $M(0)$ is the initial minimum of support. This assumption has been used in similar models such as Kortum (1997), Jones (2005), and Perla and Tonetti (2014). Furthermore, it is shown to arise from geometric random shocks as a result in the steady-state, rather than as an assumption, in Luttmer (2012) and Benhabib et al. (2019).

With this initial distribution, we obtain a simple stationary solution to the Kolmogorov Forward Equation in equation (16), where the productivity distribution always remains Pareto with shape θ . Specifically, the density at any date t is

$$\phi(Z, t) = \theta M(t)^\theta Z^{-\theta-1}, \quad Z \geq M(t). \quad (21)$$

This stationary result facilitates a solution in two ways. On the static dimension, it allows us to compute the equilibrium relationships, for all time, as one would in a variant of Melitz (2003), e.g., Chaney (2008). On the dynamic dimension, if the technology adoption policy is stationary when re-scaled, then this distribution satisfies the final requirement in Definition 1 that the distribution of productivities is stationary when re-scaled. Thus, it provides us an opportunity to find a balanced growth path.

Perla and Tonetti (2014) and Benhabib et al. (2019) provide detailed discussions on why—in the case of no GBM—an initial distribution with a Pareto tail is necessary to support long run growth via adoption and why the Pareto distribution is the only initial condition consistent with the balanced growth path law of motion for the productivity distribution. The key reason is that power laws have a scale invariance property, which means that as the economy grows geometrically, the distribution's shape remains constant. Economically, because the tail does not get thinner over time, there always remain enough better technologies available for adoption to incentivize sustained investment in adoption in the long run.

Once stochastic shocks are added to idea flow models, the restriction on the initial productivity distribution is not as limiting as it might seem. In the case with GBM, starting from *any* initial distribution with mass arbitrarily close to the lower barrier—including initial distributions with finite support—the stationary distribution is Pareto. In the GBM case, the real content of assuming the initial distribution is Pareto with shape θ is that the shape of the stationarity distribution is the θ estimated from the data. The particular θ associated with a thin-tailed initial distribution would be determined by a combination of initial conditions and the GBM parameters μ and ν . For related results and further discussion see Luttmer (2012) and Benhabib et al. (2019).

4.1. Static and Dynamic Equilibrium Relationships

The second step in computing the equilibrium is to characterize the firm's value function and adoption policy, given the law of motion described above. To do so, we first to normalize the economy and make it stationary. We then describe the important static and dynamic equilibrium relationships on which the firm value function and adoption policy depend.

Normalization. We normalize the economy to be stationary. Roughly speaking, we do this by normalizing all variables by the endogenous reservation productivity threshold $M(t)$; Appendices C and D provide the complete details. Regarding notation, all normalized variables are lower case versions of the relationships described above. For example, lower case z represents $Z/M(t)$, i.e., a firm's productivity relative to the reservation productivity threshold. The normalized productivity distribution relative to the adoption threshold is constant over time, a feature of the Pareto distribution:

$$\begin{aligned} f(z, t) &= M(t)\phi(zM(t), t) \\ f(z) &= \theta z^{-\theta-1}, \quad z \geq 1. \end{aligned} \quad (22)$$

Static Equilibrium Relationships. There are four important static equilibrium relationships that we use repeatedly throughout the rest of the paper. Specifically, the common component to firms' profits, the export productivity threshold, average profits, and the home trade share.

Normalized profits of a firm are

$$\begin{aligned} \pi(z) &= \bar{\pi}_{\min} z^{\sigma-1} + (N-1) (\bar{\pi}_{\min} d^{1-\sigma} z^{\sigma-1} - \kappa) && \text{if } z \geq \hat{z} \\ \pi(z) &= \bar{\pi}_{\min} z^{\sigma-1} && \text{otherwise.} \end{aligned} \quad (23)$$

The common component of firms' profits, defined as $\bar{\pi}_{\min}$, is important for two reasons. First, the value $\bar{\pi}_{\min}$ represents the profits of the *marginal adopting firm*. Given our normalization where z is defined relative to the reservation productivity threshold, a firm with $z = 1$ is the firm that is on the margin between adopting and not. Since, by definition, the choices of the marginal firm determine the adoption decision, how $\bar{\pi}_{\min}$ changes with trade barriers is important to understanding how the incentives to adopt change.

Second, because $\bar{\pi}_{\min}$ is common to all firms, it summarizes how changes to trade barriers affect profits of all firms on the intensive margin. That is holding fixed firms' exporter status, it determines the benefit (or loss) to all firms from opening to trade.

The export productivity threshold, \hat{z} , in equation (23) is the productivity level at which a firm is just indifferent between entering an export market or selling only domestically. This export threshold can be expressed as

$$\hat{z} = d \left(\frac{\kappa}{\bar{\pi}_{\min}} \right)^{\frac{1}{\sigma-1}}, \quad (24)$$

which depends on variable trade costs, fixed trade costs, and the common component of profits.

The profits of the marginal firm and the exporter threshold allow us to compute two summary statistics that are useful in characterizing the relationship between growth, trade, and welfare. The first is the ratio of average profits relative to minimum profits $\bar{\pi}_{\min}$

$$\bar{\pi}_{\text{rat}} := \frac{1}{\bar{\pi}_{\min}} \int_1^{\infty} \pi(z) f(z) dz, \quad (25)$$

which integrates over the normalized profit levels (equation 23) according to the density (equation 22). As we show below, this summary statistic of the profit distribution summarizes the key trade-off for the marginal firm deciding to adopt a new technology and, thus, dictates the rate of economic growth.

The second statistic is a country's home trade share. This is the amount of expenditure a country spends on domestically produced varieties.

$$\lambda_{ii} := \frac{\bar{\pi}_{\min}}{\bar{\pi}_{\min} + (N-1)\hat{z}^{-\theta\kappa}}. \quad (26)$$

This relationship is important because this value summarizes the volume of trade in the economy and, thus, λ_{ii} is a measure of openness. The home trade share connects with the profit ratio in equation (25) to provide a connection between growth and the observed volume of trade.

Dynamic Equilibrium Relationships. On the balanced growth path, the normalized continuation value function, value matching condition, and smooth pasting condition in equations (11–13) simplify to

$$(r-g)v(z) = \pi(z) + \left(\mu + \frac{v^2}{2} - g\right) z \frac{\partial v(z)}{\partial z} + \frac{v^2}{2} z^2 \frac{\partial^2 v(z)}{\partial z^2}, \quad (27)$$

$$v(1) = \int_1^{\infty} v(z) f(z) dz - \zeta, \quad (28)$$

$$\frac{\partial v(1)}{\partial z} = 0. \quad (29)$$

The major advantage of the normalized system is that it reduces the value function to one of z alone, removing the dependence on time. This mirrors the normalization of the productivity distribution. Thus, computing a balanced growth path equilibrium using the normalized system of equations involves solving an ordinary, as opposed to partial, differential equation.

The final, normalized dynamic equilibrium relationship is the free entry condition

$$\frac{\zeta}{\chi} = \int_1^{\infty} v(z) f(z) dz, \quad (30)$$

which relates the normalized entry cost to the gross value of entry (and adoption).

4.2. Algorithm for Computing the BGP Equilibrium

Given the law of motion for the productivity distribution and the normalized static and dynamic equilibrium relationships, we now outline how to solve for the equilibrium growth rate, with details in Appendices E and G.

We first solve the ordinary differential equation describing the firm's value function in equation (27) through the method of undetermined coefficients, using profits and the exporter threshold from equations (23) and (24) and ensuring that the smooth pasting condition in equation (29) is satisfied. The value function depends on a firm's productivity z , the common profit component $\bar{\pi}_{\min}$, the export threshold \hat{z} , and the rate of economic growth g .

We then insert this value function into the value matching condition (equation 28) which, when combined with the free entry condition (equation 30), yields the growth rate g as a function of \hat{z} and $\bar{\pi}_{\min}$. Because the continuation value function of the marginal firm and the expected value of adoption both depend on the rate of economic growth, this boils down to finding a growth rate that makes the marginal firm indifferent between continuing to operate and adopting a new technology. In the small-noise no-death limiting case ($\mu = \nu = \delta = 0$), using the free entry condition and the value function evaluated at the adoption threshold, $\bar{\pi}_{\min}$ can be solved for analytically, yielding g and all other equilibrium objects in closed form. In the general case, the equilibrium reduces to a system of 3 nonlinear algebraic equations in g , Ω , and $\bar{\pi}_{\min}$ that can be solved numerically.

5. Growth and Trade

For the theoretical analysis in Sections 5 and 6, we will focus on the case without GBM and with no death ($\mu = \nu = \delta = 0$). We will return to the general model with GBM and death in Section 7 when we study the quantitative effect of lower trade costs on welfare. The economic forces in the model without GBM and death are qualitatively the same as in the model with GBM and death.

5.1. Growth and Trade

Proposition 1 provides the equilibrium growth rate as a function of parameters, completing the characterization of economic growth in a model with equilibrium technology diffusion, entry and exit, and selection into exporting. For the rest of the paper, we make two substantive parameter restrictions, which are detailed in Proposition 1. First the elasticity of substitution cannot be too large relative to the Pareto tail index to ensure finite aggregate output. Second the elasticity cannot be too small to ensure positive growth; the gain in profits associated with a larger productivity after adoption must not diminish too rapidly relative to the cost of adoption.

Proposition 1 (Growth on the BGP). *If and only if $\theta > \sigma - 1 > \theta\chi > 0$, then there exists a unique Balanced Growth Path Equilibrium with finite and positive growth rate*

$$g = \frac{\rho(1-\chi)}{\chi\theta} \bar{\pi}_{rat} - \frac{\rho}{\chi\theta}, \quad (31)$$

where the ratio of average profits to minimum profits is

$$\bar{\pi}_{rat} = \frac{\left(\theta + (N - 1)(\sigma - 1)d^{-\theta} \left(\frac{\kappa}{\zeta} \frac{\chi}{\rho(1-\chi)} \right)^{1-\frac{\theta}{\sigma-1}} \right)}{(1 + \theta - \sigma)}. \quad (32)$$

Proof. See Appendix G. □

The most interesting feature of Proposition 1 is that the growth rate is an affine function of the ratio of profits between the average and marginal firm. This profit ratio is the key summary statistic in this model and its sensitivity to trade costs will drive many of our main results. The intuition for why the growth rate is a function of the profit ratio is that the incentive to adopt depends on two competing forces: the expected benefit of a new productivity draw and the cost of taking that draw. The opportunity cost of adopting a better technology is the forgone profits from producing with the current technology. The expected benefit relates to the profits that the average new technology would yield. Proposition 1 tells us that a larger spread in the expected benefit relative to the opportunity cost increases the incentives to adopt and, thus, leads to faster economic growth.¹¹

We can go one step further and connect the profit ratio to a country's home trade share. This establishes a connection between economic growth and the volume of trade. After some substitution, a country's home trade share in terms of primitives is

$$\lambda_{ii} = \frac{1}{1 + (N - 1)d^{-\theta} \left(\frac{\kappa\chi}{\zeta\rho(1-\chi)} \right)^{1-\frac{\theta}{\sigma-1}}}. \quad (33)$$

with exporter productivity threshold

$$\hat{z} = d \left(\frac{1}{\zeta} \frac{\kappa\chi}{\rho(1-\chi)} \right)^{\frac{1}{\sigma-1}}. \quad (34)$$

Comparing equation (33) and (32) reveals that the home trade share tightly relates to the profit ratio. This connection allows us summarize growth as a function of the inverse of a country's home trade share in Corollary 1.

Corollary 1 (Growth and the Volume of Trade). *On the Balanced Growth Path, the relationship between growth and the volume of trade is*

$$g = \frac{\rho(1-\chi)}{\chi\theta} \frac{\sigma-1}{\theta-\sigma+1} \lambda_{ii}^{-1} - \frac{\rho}{\theta}, \quad (35)$$

with a country's home trade share given in equation (33).

¹¹This result is closely related to Hornstein, Krusell, and Violante (2011). In a McCall (1970) labor-search model they establish a relationship between the frequency of search and a summary statistic of wage dispersion—the ratio of the average wage to the minimum wage. In our model, growth is related to the frequency of firms searching to adopt new technologies and equation (31) shows how this depends on the ratio of profits between the average and marginal firm; similar to Hornstein et al. (2011).

Proof. See Appendix G. □

Corollary 1 relates to the results in Arkolakis et al. (2012) which connects the *level* of real wages to a country's home trade share.¹² An interpretation of their results in the Melitz (2003) model is that the trade induced welfare gains from reallocation are completely summarized by the change in a country's home trade share. In our model, Proposition 1 tells us that the incentives to adopt new technologies are driven by the spread in profits between the average and the marginal firm. Similar to Arkolakis et al.'s (2012) findings, Corollary 1 says that these distributional effects are summarized by the aggregate volume of trade.

How do changes in trade costs affect growth? A quick examination of equations (32) or (33) shows that a decrease in variable trade costs will increase the spread in profits between the average and marginal firm, reduce a country's home trade share, and increase the rate of economic growth. Proposition 2 summarizes these effects in the form of elasticities for a world-wide reduction in variable trade costs. Moreover, we provide a "sufficient statistic" representation of these elasticities that depends only on several parameters and observable trade statistics.

Proposition 2 (Comparative Statics: Trade, Profits, and Growth). *A decrease in variable trade costs*

1. *Decreases the exporter productivity threshold*

$$\varepsilon_{\hat{z},d} := \frac{d \log \hat{z}(d)}{d \log(d)} = 1. \quad (36)$$

2. *Decreases a country's home trade share.*

$$\varepsilon_{\lambda_{ii},d} := \frac{d \log \lambda_{ii}(d)}{d \log(d)} = \theta(1 - \lambda_{ii}) > 0. \quad (37)$$

3. *Increases the spread in profits between the average and marginal firm.*

$$\varepsilon_{\bar{\pi}_{rat},d} := \frac{d \log \bar{\pi}_{rat}(d)}{d \log(d)} = \frac{-(\sigma - 1)\varepsilon_{\lambda_{ii},d}}{1 + \lambda_{ii}(\theta - 1)} < 0. \quad (38)$$

4. *Increases economic growth.*

$$\varepsilon_{g,d} := \frac{d \log g(d)}{d \log(d)} = \left(\frac{\chi(1 + \theta - \sigma)}{(\sigma - 1)(1 - \chi)} \lambda_{ii} - 1 \right)^{-1} \varepsilon_{\lambda_{ii},d} < 0. \quad (39)$$

Proof. See Appendix G. □

The first result defines what we will call the domestic consumption elasticity $\varepsilon_{\lambda_{ii},d}$, which is the amount the home trade share declines (and volume of trade increases) as variable trade costs decline. This elasticity takes a simple form which depends on the volume of trade and the parameter θ . As is typical in the Melitz (2003) class of models, θ is the trade elasticity.

¹²Sampson (2016) finds a similar affine relationship between growth and parameters which determine the volume of trade.

The second result connects the domestic consumption elasticity with reallocation effects. A reduction in trade costs increases the difference in profits between the average firm and the marginal firm. And the spreading of profits is tied to the domestic consumption elasticity. The basic idea is that lower trade costs induce high productivity exporting firms to expand and low productivity firms to contract as competition from foreign firms reduce their profits. Given that these distributional effects are summarized by the home trade share (Corollary 1), the change in the profit spread is proportional to the domestic consumption elasticity.

The third result in Proposition 2 shows that reductions in trade costs increase economic growth. The sensitivity of growth to changes in trade costs crucially depends on the domestic consumption elasticity. The intuition for this connection lies in the previous result. The reallocation effects from reductions in trade costs incentivizes more frequent technology adoption by both lowering the opportunity cost of adoption and increasing the expected benefit. For low productivity (non-exiting firms), the loss of market share and profits reduces the opportunity cost of adoption. Because exporting now has a larger return, this increases the expected benefit of a new technology. Together this leads firms to more frequently improve their technology. Because the frequency at which firms upgrade their technology is intimately tied to aggregate growth, growth increases as variable trade costs decrease.

The sufficient statistic representation of these elasticities isolates which parameters are useful in quantifying how growth responds to reductions in trade costs. In particular, it shows that the important factors are the: level of trade flows λ_{ii} , extent of firm heterogeneity θ , curvature on the demand for varieties σ , and size of adoption costs relative to entry costs χ . The trade share is observed. There is an array of estimates in the literature for θ and σ . The only difficult parameter to discipline is χ , but this can be determined as a residual from equation (35) (along with information on the discount rate) to target a particular growth rate. While these results are for the analytically tractable calibration of the model ($\mu = v = \delta = 0$), we show in Section 7.5 that in the full quantitative model, the calibrated values of μ , v , and δ largely affect the welfare gains from trade by affecting the calibrated value of χ . In this sense, the sufficient statistic representations of these elasticities in the simple model provide useful intuition for the quantitative model.

This representation also illuminates the role of scale in our economy. First, notice that in both the growth rate and elasticity formulas, population size does not appear at all. This is largely by construction as the fixed costs of exporting and adoption are scaled by population. Second, examination of equations (31) and (32) in Proposition 1 shows that the number of countries in the economy affects economic growth. What Corollary 1 and Proposition 2 show, however, is that what ultimately matters is the volume of trade, and not the number of countries or the trade costs per se. As our discussion of market size effects in Section 5.3 makes clear, growth depends on the ratio of profits between the average and marginal firm—not the level of profits. That is, controlling for the amount of trade as summarized by λ_{ii} , the number of countries does not affect the relative benefit of adoption and, hence, the incentives of firms to adopt a new technology.

Our model has the prediction that changes in trade costs leads to permanent growth effects. The empirical evidence is inconclusive as to whether there are growth or level effects associated with changes in trade costs. We would also caution against using evidence from cross-country regressions (e.g., Dollar

and Kraay, 2004; Frankel and Romer, 1999) to make inferences about our results for several reasons: our results imply a non-linear relationship between growth and the volume of trade, as Corollary 1 makes clear; our results are in the context of a symmetric country equilibrium; our results also abstract from important mechanisms such as cross-country technology diffusion; these results also are across BGP results not inclusive of transition paths which we find to be important in Section 7. We view our main contribution as providing a better understanding of how trade affects the relative benefits of technology adoption across firms.

5.2. Firms, Trade, and Technology Adoption

Below we focus on a firm's adoption policy and how it relates to the aggregate environment. Proposition 3 summarizes our results.

Proposition 3 (Firms and Technology Adoption). *Given an aggregate growth rate g ,*

1. *The time $\tau(z)$ until an individual firm with productivity z adopts a new technology is*

$$\tau(z) = \frac{\log(z)}{g}. \quad (40)$$

2. *The average time until adoption is*

$$\bar{\tau} = \frac{1}{\theta g}. \quad (41)$$

3. *Over a Δ length of time, the measure of firms that adopt is*

$$S\Delta = \frac{\Delta}{\bar{\tau}}. \quad (42)$$

Proof. See Appendix G. □

The first part of the proposition focuses on an individual firm and computes the time until it changes its technology. The second and third part of the proposition aggregate. Across all firms, the average time until adoption depends inversely on the growth rate and the Pareto shape parameter. Over an increment of time, the number of firms adopting is the flow of adopters times the length of time, which turns out to take a very simple form: the time increment multiplied by the inverse of the average time until adoption.

From a firm's perspective, more rapid economic growth means that it optimally waits a shorter amount of time before upgrading its technology. This effect of faster growth holds for all firms and, thus, the average time across all firms is shorter. Furthermore, this result implies that over a given time increment, more firms adopt.

Connecting these observations with the aggregate growth effects of trade in Proposition 2 we have predictions about firms' responses to a trade liberalization. Proposition 3 predicts that we should see more firms adopting new technologies in an open economy relative to a closed economy. More specifically, average, within-firm productivity growth is larger in the open economy relative to a closed economy.

These results connect with the motivating evidence discussed in Section 1.2. The predictions in Proposition 3 imply that an empirical specification which projects changes in firm level measures of technology on measures of openness should display a positive relationship. This is exactly what empirical papers using specifications of this type find in the data (see, e.g., Bloom et al., 2016; Pavcnik, 2002; Topalova and Khandelwal, 2011). These papers use data from an economy in transition, not necessarily at a new steady state. Even in the full quantitative model, the time to adoption is decreasing in g . Thus, if in the calibrated quantitative model below we find growth increasing along the transition path, then this prediction would also hold along the transition.

Our model makes distinct predictions about who adopts. In our model, average, within-firm productivity growth hides heterogeneity in who adopts and who does not. In particular, firms at the top of the productivity distribution continue to operate with their existing technology while firms at the bottom adopt and experience productivity growth. In contrast, in Atkeson and Burstein (2010) or Bustos (2011), it is the large, exporting firms that innovate more; small import-competing firms decrease their innovation.

5.3. Reallocation vs. Market Size Effects

Heterogeneity in firms' incentives and actions are the essential ingredients driving the relationship between trade and growth in our model. The heterogeneous incentives induce a reallocation effect across firms that is distinct from the "market size" mechanisms emphasized in the previous literature, i.e., the ability to spread the same cost of adoption across a larger market resulting in growth effects from openness. In this section we highlight the key mechanism active in our model by removing the type of heterogeneity that links trade and growth. See Burstein and Melitz (2013) who highlight the importance of the reallocation effect that is introduced by selection into exporting in a model with firm dynamics.

To focus on traditional market size forces and abstract from the role of distributional effects, we set the fixed cost of exporting equal to zero to study an environment in which all firms export. Equilibrium objects in this environment are labeled with a superscript k , since this model resembles a heterogeneous firm version of Krugman (1980). Proposition 4 summarizes growth in this model.

Proposition 4 (Growth with No Selection into Exporting). *In the model with $\kappa = 0$ and all firms selling internationally, the growth rate is*

$$g^k = \frac{\rho(1-\chi)}{\chi\theta} \bar{\pi}_{rat}^k - \frac{\rho}{\chi\theta}, \quad (43)$$

where the ratio of average profits to minimum profits is

$$\bar{\pi}_{rat}^k = \frac{(1 + (N-1)d^{1-\sigma})\bar{\pi}_{min}E[z^{\sigma-1}]}{(1 + (N-1)d^{1-\sigma})\bar{\pi}_{min}} = \frac{\theta}{1 + \theta - \sigma}. \quad (44)$$

Proof. See Appendix F. □

The expression for the growth rate in equation (43) takes the same form as that in Proposition 1: Growth is an affine function of the ratio of profits between the average firm and the marginal firm. Even

though the details of international trade differ, the technology adoption choice is the same and, hence, the aggregate growth rate takes the same form.

The only difference between Proposition 4 and Proposition 1 is that the profit ratio is now independent of trade costs. This implies that growth is independent of trade costs. When there is no selection and all firms export, reductions in trade costs do not induce reallocation—all firms' profits increase by the same proportion in response to lower trade costs.

The intuition for this result is the following: Technology adoption in our model depends on a comparison between the expected value of adoption versus the value of continuing to operate with the existing technology. Lower trade costs have two effects on a firm's incentive to adopt a new technology. Just as in the model with selection into exporting, lower trade costs increase the expected value of adopting a new technology. However, when all firms export, lower trade costs also increase the value of continuing to operate the old technology for the marginal firm. The profit ratio summarizes this comparison. Because all firms' profits scale up by the same proportion, lower trade costs do not affect the *relative* benefit of adoption.¹³ Thus, changes in trade costs do not change the rate of economic growth.

In contrast to this neutrality result, with selection into exporting, the differential exposure of firms to trade opportunities affects the distribution of profits and, hence, the incentives to adopt and economic growth. To illustrate these mechanics, we decompose the profit ratio in equation (32) from Proposition 1 in our baseline model in the following way:

$$\bar{\pi}_{\text{rat}} = \underbrace{\frac{\theta}{1 + \theta - \sigma}}_{= \bar{\pi}_{\text{rat}}^k} + \frac{(N - 1)(\sigma - 1)d^{-\theta} \left(\frac{\kappa}{\zeta} \frac{\chi}{\rho(1-\chi)} \right)^{1 - \frac{\theta}{\sigma-1}}}{1 + \theta - \sigma}. \quad (45)$$

The profit ratio in our baseline economy is now written as the sum of two terms. The first term is the same as the profit ratio above in the No Selection into Exporting economy. The second term isolates the affects from trade-induced reallocation. This second term shows how reductions in either variable or fixed trade costs "spread" the distribution of profits relative to the No Selection into Exporting economy, incentivize faster adoption, and lead to increases in economic growth.

6. The Welfare Gains From Trade: Theoretical Analysis

This section studies the welfare effects of lower barriers where with the focus on the case without GBM and with no death ($\mu = \nu = \delta = 0$) like in the previous section. Analytically we show how changes in trade costs affect the initial level of consumption and how consumption and growth rate effects combine to determine welfare. Section 7 studies study the quantitative effect of lower trade costs on welfare.

¹³This result is similar to one in Eaton and Kortum (2001) who highlight that there are two competing forces from a larger market: a larger market makes an innovation more valuable, but also makes it more costly to achieve a new innovation. In their model, these two mechanisms cancel out leaving a result similar to Proposition 4.

6.1. Consumption and Welfare

How does welfare change with trade costs? While increased trade leads to faster economic growth, we show that there are more subtle effects of opening to trade on welfare. In particular, the gains from trade are a race between the positive dynamic growth effects, the positive static reallocation effects, and the negative static effects of less varieties produced in each country, and reallocation of workers away from goods production to adoption activity.

Time zero utility and the associated initial level of consumption are

$$\bar{U} = \frac{\rho \log(c) + g}{\rho^2}, \quad (46)$$

$$c = (1 - \tilde{L}) \Omega^{\frac{1}{\sigma-1}} \lambda_{ii}^{\frac{1}{1-\sigma}} (\mathbb{E}[z^{\sigma-1}])^{\frac{1}{\sigma-1}}. \quad (47)$$

The level of consumption depends on several factors. The $(1 - \tilde{L})$ term is the amount of labor devoted to the production of goods—as opposed to adoption and entry activity. The second term is the measure of varieties. The third term is the home trade share and the fourth term is just the $\sigma - 1$ moment of the productivity distribution. Again, recall from the discussion of Corollary 1, that the trade share is a summary statistic for the distribution of activity across producers.

With changes in trade costs, the initial level of consumption changes for several reasons: changes in the home trade share, the measure of varieties, and the share of labor engaged in goods production. Proposition 2 showed how the home trade share changes. The proposition below formalizes how the measure of varieties and the share of labor engaged in goods production changes with trade costs.

Proposition 5 (Comparative Statics: Variety and Labor Allocations). *As a function of the home trade share,*

1. *The amount of varieties is*

$$\Omega = \frac{\chi}{\zeta \rho} \left(\frac{(1 - \chi)\theta\sigma}{1 + \theta - \sigma} \lambda_{ii}^{-1} - 1 \right)^{-1}. \quad (48)$$

2. *A decrease in variable trade costs reduces the measure of varieties produced in each country.*

$$\varepsilon_{\Omega,d} := \frac{d \log \Omega(d)}{d \log(d)} = \left(1 - \frac{1 + \theta - \sigma}{\theta\sigma(1 - \chi)} \lambda_{ii} \right)^{-1} \varepsilon_{\lambda_{ii},d} > 0. \quad (49)$$

3. *The share of workers in goods production is*

$$1 - \tilde{L} = (\sigma - 1) \left(\sigma - \frac{1 + \theta - \sigma}{\theta(1 - \chi)} \lambda_{ii} \right)^{-1}. \quad (50)$$

4. A decrease in variable trade costs reduces the share of workers in goods production.

$$\varepsilon_{\tilde{L},d} := \frac{d \log(1 - \tilde{L}(d))}{d \log(d)} = \left(\frac{\theta\sigma(1 - \chi)}{1 + \theta - \sigma} \lambda_{ii}^{-1} - 1 \right)^{-1} \varepsilon_{\lambda_{ii},d} > 0. \quad (51)$$

Proof. See Appendix G. □

Lower trade costs decrease the measure varieties produced in each country. This is consistent with the Melitz (2003) logic that reductions in trade costs induce exit. Lower trade costs also result in less labor allocated to the production of goods and more labor allocated to the fixed costs of adoption and exporting. In a sense, there is more “investment” in the adoption of technology and exporting—a force towards a lower level of consumption.

Proposition 6 (Consumption Effects). *A decrease in variable trade costs decreases the initial level of consumption.*

$$\varepsilon_{c,d} = \varepsilon_{\tilde{L},d} + \frac{\varepsilon_{\Omega,d} - \varepsilon_{\lambda_{ii},d}}{\sigma - 1} > 0. \quad (52)$$

Proof. See Appendix G. □

Proposition 6 shows that there are two drags on the initial level of consumption and one offsetting gain. Reductions in trade costs reallocate labor away from the production of goods and, hence, the consumption of goods ($\varepsilon_{\tilde{L},d} > 0$). Reductions in trade costs lead to less production of domestic varieties ($\varepsilon_{\Omega,d} > 0$), which reduces consumption due to the love for variety CES final goods aggregator. Counteracting these forces are gains from an increase in the measure of foreign varieties consumed and that these foreign varieties are produced by firms that are relatively highly productive. These forces are summarized by a decrease in the home trade share ($\varepsilon_{\lambda_{ii},d} > 0$).

Which of these forces wins? The initial consumption level decreases with lower trade costs. To see this, note that from equation (49), the loss in domestic varieties is larger than the gain from importing foreign goods ($\varepsilon_{\Omega,d} > \varepsilon_{\lambda_{ii},d}$). Given that labor is always reallocated away from production as trade costs decrease, equation (52) shows that the level of initial consumption will decrease with lower trade costs.

This result deserves some discussion. First, while the initial level of consumption decreases, there is an inter-temporal gain as consumption in future periods is higher (Proposition 2).

Second, our model’s static gains from trade need not correspond with those in purely static models of trade. In Melitz (2003), the typical parameterization finds that domestic variety falls, but the gains from foreign variety more than compensate for this loss.¹⁴ In our model, dynamics lead to additional effects on the free-entry condition. Wages increase due to increased demand for labor, which increases the cost of entry. However, the expected value of entry does not increase in equal proportion, as profits are reallocated across firms and these profits are discounted at a higher rate because of faster growth. Thus, to satisfy the free-entry condition the reduction in domestic variety is larger (relative to static

¹⁴As our phrasing suggests, this need not be the case. See the discussion on page 1713 of Melitz (2003).

models). Atkeson and Burstein (2010) have a similar quantitative finding that the gains due to increased innovation by firms are offset by losses in variety and reallocation of resources away from production for consumption.

Finally, we should note that the race between these welfare reducing and welfare enhancing effects from trade cost changes is an important normative feature of our model relative to other idea-flow models studying trade and growth. For example, in Alvarez et al. (2017), there are no resource costs associated with acquiring ideas. As a result, when trade facilitates faster idea acquisition and economic growth, there is no corresponding increase in costs. In contrast, in our model, faster idea acquisition comes with the cost of more labor being allocated away from production.

Using equations (46) and (47) and the results in Proposition 2, 5 and 6, the elasticity of utility with respect to a change in trade costs—the welfare gain from trade—is summarized below in Proposition 7.

Proposition 7 (Welfare Effects). *The welfare change from a decrease in variable trade costs is*

$$\varepsilon_{\bar{U},d} = \frac{\rho^2}{\bar{U}} (\rho \varepsilon_{c,d} + g \varepsilon_{g,d}). \quad (53)$$

Proof. See Appendix G. □

Comparing across BGPs, the welfare gain from trade is proportional to the change in the initial level of consumption, $\varepsilon_{c,d}$, and the change in the growth rate, $\varepsilon_{g,d}$.¹⁵ Proposition 2 tells us that the benefit from growth is positive and Proposition 6 shows that the effect on the level of consumption is negative. Conventional wisdom suggests that a decrease in trade costs should increase utility, but conventional wisdom may be misleading in inefficient economies. Furthermore, this comparative static analysis ignores transition dynamics. There exist parameter values (when $\sigma - 1$ is close to its lower bound of $\theta\chi$) such that lower trade costs are associated with lower utility according to this steady-state measure (see equation G.51). Thus, understanding the welfare effects of lower trade costs requires solving for the economy along the transition path at calibrated parameter values. We perform such an analysis in Section 7.

6.2. Consumption and Welfare with No Selection into Exporting

To further isolate the race between the competing welfare reducing and welfare enhancing effects, consider the welfare effects in our model when all firms export. Proposition 8 details how trade, growth, varieties, labor allocations, and welfare are affected by trade costs in the economy in which there is no selection into exporting.

Proposition 8 (Comparative Statics with No Selection into Exporting). *In the model with $\kappa = 0$ and all firms exporting, a decrease in variable trade costs*

¹⁵Note that since $d > 0$, the sign of the elasticity and the derivative of utility with respect to trade are equal if and only if $\bar{U} > 0$. With log utility the sign of \bar{U} depends on initial conditions on the mean of the productivity distribution and the population size. Most importantly, the sign of the derivative of utility with respect to trade costs is independent of these initial conditions and the sign of utility.

1. Reduces the home trade share.

$$\varepsilon_{\lambda_{ii},d}^k = \frac{d \log \lambda_{ii}^k(d)}{d \log(d)} = (\sigma - 1)(1 - \lambda_{ii}^k) > 0, \quad (54)$$

where,

$$\lambda_{ii}^k = (1 + (N - 1)d^{1-\sigma})^{-1}. \quad (55)$$

2. Does not change the growth rate, the measure of domestic varieties, or the share of workers in goods production.

$$\varepsilon_{g,d}^k = 0, \quad \varepsilon_{\Omega,d}^k = 0, \quad \varepsilon_{\tilde{L},d}^k = 0. \quad (56)$$

3. Increases the initial level of consumption.

$$\varepsilon_{c,d}^k = -\frac{\varepsilon_{\lambda_{ii},d}^k}{\sigma - 1} < 0, \quad (57)$$

4. Increases welfare.

$$\varepsilon_{\tilde{U},d}^k = \frac{\rho^2}{\bar{U}} \rho \varepsilon_{c,d}^k < 0. \quad (58)$$

Proof. See Appendix F. □

Lower trade costs lower the home trade share, although now moderated by $(\sigma - 1)$ as opposed to θ . In addition to growth being independent of trade costs, the share of workers allocated to goods production does not change with trade costs either. Thus, while this model does not have dynamic gains from trade, it does not experience some of the losses associated with dynamics, such as labor reallocation.

The measure of varieties does not change either. The reason is that when all firms export, all firms' profits scale up by the same amount as the real wage and hence the entry cost. As trade costs decline, all profits increase by the same amount and the expected value of entry increases by this amount, but the cost of entry increases in the same proportion as well.¹⁶ Thus, the measure of varieties need not change for the free-entry condition to be satisfied.

These observations help explain why growth effects are associated with a loss of variety in Proposition 5. When there is selection into exporting, in response to a decrease in trade costs, the normalized expected value of entry increases less than the cost of entry because of the reallocation of profits across firms. This implies that a reduction in varieties must occur for the free-entry condition to be satisfied. Thus, the same exact reallocation effects which lead to faster economic growth, generate the loss of varieties.

¹⁶While the independence of the growth rate and trade costs is true independent of whether the cost of adoption is in labor or goods, that the elasticity of varieties is independent of trade costs relies on the cost of adoption being in labor only. This result is similar to Atkeson and Burstein's (2010) Proposition 2 that states a similar neutrality result on the measure of varieties when the research good in their model uses only labor.

The final line of Proposition 8 computes the total gains from trade. These gains are purely static and these static level effects only operate through the component associated with the home trade share. This result is similar to the welfare gain calculations in Arkolakis et al. (2012). Unlike the finding in Arkolakis et al. (2012) that the gains from trade are equivalent in models with selection or without, a comparison of Proposition 7 to Proposition 8 shows that the essential element to delivering dynamic gains from trade is the reallocation effect introduced by selection into exporting.

7. The Welfare Gains From Trade: Quantitative Analysis

This section extends the theoretical analysis in several quantitative directions. First, we calibrate the full model with GBM productivity shocks and firm exit. While we lose the analytical tractability of the simple model analyzed above, these features of the model allow us to discipline parameter values using micro-level moments of firm dynamics.

We then perform two quantitative exercises with the calibrated model. The first exercise is a (local) decomposition of the welfare gains across steady states. The decomposition identifies the sources of the gains from trade in our economy, how they differ from the gains from trade in an efficient economy, and how they differ from benchmarks in the literature. The second exercise focuses on the welfare gains from a large decrease in trade costs inclusive of the transition path—not across steady states as in our theoretical analysis and local decomposition. The focus on the transition path addresses the concern that the across steady-state analysis might overstate the welfare gains because the benefits of higher growth are in the future and they require costly investment to implement.

Since the quantitative model cannot be solved analytically, we use numerical methods to compute the BGP for calibration and to solve for transition dynamics. The key steps are to discretize the spatial dimension (i.e., productivity) of the Hamilton-Jacobi-Bellman PDE to turn it into ODEs and then to add in all necessary equilibrium conditions, which includes the value matching integral equation, to create a differential-algebraic system of equations (DAEs). This approach of solving for the transition dynamics of heterogeneous agent models by representing the economy as a differential-algebraic system of equations and using high-performance DAE/ODE solvers is applicable to a large class of models. Online Appendix Numerical Methods describes the algorithm used to compute the equilibrium of the quantitative model with GBM and exit shocks.

7.1. Calibration

We choose parameters by using a mix of normalization and calibration to match outcomes of our model on the BGP with moments in the data.

The normalizations are described in the top panel of Table 1. We set the number of countries equal to ten. Per the implications of Corollary 1 and Proposition 2 a different value of N has no impact on how growth responds to trade as long as the home-expenditure share λ_{ii} is matched to its empirical value. The technology adoption cost ζ is normalized to the value one and discussed below.

Our calibration procedure uses the simulated method of moments to determine the remaining parameter values. The parameter vector we are choosing is $\{\rho, \theta, \sigma, \mu, v^2, \delta, d, \kappa, \chi\}$, in words: the discount rate,

Table 1: Calibration Results: Parameters and Values

Parameter	Value	Parameter	Value
Technology Adoption Cost, ζ	1.0	Number of Countries, N	10
Discount Rate, ρ	0.0203	Pareto Shape Parameter, θ	4.99
Variety Elasticity of Substitution, σ	3.17	Drift of GBM Process μ	-0.031
Variance of GBM Process v^2	0.048	Death Rate of Firms δ	0.020
Iceberg Trade Cost, d	3.02	Export Fixed Cost, κ	0.104
Entry Cost Relative to Adoption Cost $1/\chi$	7.88		

Pareto shape parameter, variety elasticity of substitution, drift and variance of the GBM process, exit shock, iceberg trade cost, exporter fixed cost, and (relative) entry cost. The bottom panel of Table 1 presents a description of these nine parameters and their calibrated values.

These nine parameter values are jointly chosen so that the model best fits 14 moments. Because the economy with GBM shocks does not have closed-form solutions, we compute the BGP equilibrium using numerical methods and then construct sample moments in the model to compare to sample moments in the data.

The data sources are a mix of aggregate statistics, measures from other papers, and micro data. Most moments that we target are constructed for the U.S. over the time period from 1977-2000. We choose this time period because it is the span for which we have access to the micro-level firm dynamics data.

The micro-level data that we use is from the Synthetic Longitudinal Business Database (SLBD) (U.S. Census Bureau (2011)). The SLBD is a public access database that synthesizes information on establishments' employment, payroll, industry, and birth and death year. It is synthesized from the confidential, restricted-access Longitudinal Business Database that is maintained by the U.S. Census Bureau. We construct SLBD moments using only establishments with at least 20 employees which is the cutoff value used in Hurst and Pugsley (2011). Per the observations of Hurst and Pugsley (2011), the motivation for this sample restriction is that many of these small establishments appear to have no intention to grow or innovate and, thus, their motives do not correspond with the motives of firms in our model. In Section 7.5, we illustrate how alternative parameterizations of the GBM process (and, hence, firm dynamics) shape our results.¹⁷

¹⁷When calibrating to all firms in the SLBD most parameters are unchanged. The key difference is a larger χ , which generates slightly lower welfare gains from trade than our baseline calibration—consistent with Figures 6 and 7 in Section 7.5.

Table 2: Aggregate Moments: Model and Data

Moment Description	Data	Model
U.S. Real Interest Rate	2.83	2.83
U.S. Productivity Growth	0.79	0.79
U.S. Import/GDP	10.63	10.63
Share of Exporting Establishments	3.3	3.3
Relative Size of Exporting Establishments	4.8	4.8

Details of moment construction are provided in the body of the paper. The real interest rate, productivity growth rate, and import/GDP ratio are in percent and averages over the 1977–2000 time period.

Below we describe the specific moments that we target, how we measure them, and a brief description of the parameters about which these moments are most informative.

Real interest rate. We construct the real interest rate as the difference between the rate of return on a one year U.S. treasury constant maturity nominal bond and the realized inflation rate. We use the U.S. consumer price index, all items as our measure of realized inflation. Over the 1977–2000 time period, the average real interest rate was 2.83 percent. This value, along with the growth rate of the economy, pins down the consumers discount rate ρ through the fact that the consumer’s Euler equation implies that the real interest rate equals $\rho + g$.

Aggregate total factor productivity growth. To measure U.S. productivity growth, we use the Bureau of Labor and Statistics Multifactor Productivity Database. We focus on the private non-farm business sector and measure labor-augmenting productivity growth. Over the 1977–2000 time period, average productivity growth was 0.79 percent per year. Conditional on all other parameter values, the rate of economic growth is connected with the frequency of adoption and, hence, its cost χ . In Section 7.5, we discuss the relationship between g and χ in more depth.

Aggregate import share. We construct the aggregate import share as the value of imports of goods and services divided by GDP. Over the 1977–2000 time period, the average import share was 10.63 percent and quite stable. As in most models of trade, this moment is informative about the costs of trade and specifically the iceberg trade cost, d .

Share of exporting establishments. This moment is constructed from the finding of Bernard, Eaton, Jensen, and Kortum (2003), who use the 1992 U.S. Census of Manufactures to report that 20 percent of manufacturing establishments export. Because we associate each variety in our model with the output of an establishment within the entire U.S. aggregate economy, we adjust the Bernard et al. (2003) number

Table 3: Establishment Size Dynamics, Data and Model

Transition Matrix of Relative Size: largest, quartile 1; smallest, quartile 4									
Data					Model				
	1	2	3	4		1	2	3	4
1	50.0	27.0	13.6	9.4	1	45.7	31.5	14.2	8.6
2	16.2	28.6	29.1	26.1	2	17.1	28.8	28.2	25.9

Corr(Data, Model) = 0.98

Employment Share of New Establishments			
Data:	0.02	Model:	0.02

Note: In the top panel, rows represent the establishment-size quartile in period t ; columns represent the establishment-size quartile in period $t + 5$. Data Source: U.S. Census Bureau (2011).

by the relative size of the manufacturing sector. In the SLBD data, only 16.5 percent of establishments are in manufacturing and we assume, as a lower bound, that only manufactures are exported. This implies that the fraction of all establishments that export is $0.2 \times 0.165 = 3.3$ percent. This moment is particularly informative about the fixed cost of exporting, κ .

Relative size of exporting establishments. We target the finding of Bernard et al. (2003) that exporters' domestic shipments are 4.8 times larger than non-exporters' shipments. This moment is particularly informative about parameters controlling the size of the firm, and thus, θ and σ .

Size dynamics of establishments. We target the observed transition probabilities of establishments as they move across different quartiles of the size distribution. We focus on the dynamics of firms at the top of the size distribution because in the model these transition probabilities are governed by the GBM process, since they are away from the adoption threshold.

We construct these moments in the following way. For each year, we sort establishments into quartiles by employment size. We then compute the probabilities that establishments transition across size-quartiles at a five-year horizon, conditional on continuing to operate. We then average these transition probabilities across all years in the SLBD, i.e., from 1977–2000. The focus on the five-year horizons is to sweep out any transient fluctuations that may occur at higher frequencies due to shocks that are outside of our model. To give some perspective on the overall size of establishments, establishments larger than about 40 employees fall into the top two quartiles; establishments larger than about 100 employees fall into the top quartile.

The left panel of Table 3 reports the observed transition probabilities. For example, 50 percent of establishments in the top quartile remain in the top quartile after 5 years, 27 percent fall to the second

quartile, and the remainder are sprinkled between the bottom two quartiles.¹⁸ These moments show that there is a large degree of persistence in firm size and that the probability of moving to a nearby quartile tends to be larger than the probability of moving to a distant quartile. We target all 8 transition probabilities.

As mentioned, these transition probabilities are very informative about parameters which control the drift and variance of the GBM process. In addition, they also place restrictions on the size of firms, and thus, inform the values of θ and σ .

Employment share of new establishments. This moment is measured as the share of employment (out of total employment) that is employed in new establishments each year, in the SLBD. Averaging across all time periods, we find this value to be two percent. On the balanced growth path, the entry rate exactly corresponds with the exit rate and, hence, this moment directly pins down the exit shock faced by firms, δ . We choose to target a moment related to entry since entry dynamics are key to the behavior of the economy on the transition path.

To summarize, 9 parameters values are chosen to target these 14 moments: three aggregate moments (real interest rate, import share, productivity growth), two exporter facts (share of establishments, relative size of exporters), and nine firm dynamics moments (eight transition probabilities and the share of employment of entrants).

7.2. Calibration Results

The bottom panel of Table 1 reports the calibrated parameter values. Many of these parameter values are easily compared and are close to alternative estimates of these values in the literature.

As stated earlier, we set $\zeta = 1$ as a normalization. Several comments are in order regarding this normalization. First, calibrating the model with a different normalization results in identical parameter values for all parameters except κ . In the theoretical analysis, equations (33) and (35) show that ζ only appears in the form $\frac{\zeta}{\kappa}$. Thus, a different choice of ζ results in a different κ to match the g and λ_{ii} moments. Different normalizations do lead to a different amount of variety as equation (48) shows. However, as equation (53) shows, any combination of ζ and κ results in the same welfare gains from trade as long as λ_{ii} and all other parameter values are equal. It is in this sense that $\zeta = 1$ is a normalization. Numerical analysis of the quantitative model confirms these observations both for steady-state comparative statics and for the analysis of transition dynamics.¹⁹

The Pareto shape parameter value of five is exactly in line with the trade evidence discussed in Head and Mayer (2014). More broadly, the value is a bit larger than Simonovska and Waugh's (2014) point estimate for the Melitz model who use both trade and price data. The value is a bit larger than usually found when models are matched to firm-level size moments, e.g., Bernard et al. (2003) and Eaton, Kortum, and

¹⁸These transition probabilities are quite similar when restricting the sample to only manufacturing establishments, with the only difference being a bit more persistence. Moreover, they look similar to the productivity transitions reported in Baily, Hulten, and Campbell (1992).

¹⁹An alternative approach to normalizing $\zeta = 1$ would be to include an empirical moment for the number of varieties. It is difficult to define and measure varieties (one firm one variety?; one plant one variety?; one UPC code one variety?). Since it does not affect the main results of this paper on the welfare gains from trade we choose to normalize $\zeta = 1$.

Kramarz (2011) find point estimates of around 3.6 for the Pareto shape parameter.

We find the elasticity of substitution across varieties to be 3.17, which is right in the range of the median estimates of Broda and Weinstein (2006). For example, their estimates at the most disaggregate level fall between 3.1 and 3.7 depending upon the time period considered. The values of both θ and σ determine the size distribution of firms. In our model, the size distribution is Pareto with shape parameter $\frac{\theta}{\sigma-1}$. The empirical literature suggests that this shape parameter is just slightly above one, see, e.g. Axtell (2001) or Luttmer (2007). Together our estimate of θ and σ generates a tail index of 2.3, which is larger than the near one value found in the data.

The GBM variance parameter value is 0.048 and the value for the drift is -0.031. These values are close to those used in Arkolakis (2015) and Luttmer (2007). In particular, Arkolakis (2015) uses moments related to our calibration procedure from Dunne, Roberts, and Samuelson (1988) and infers a variance of 0.067 and a drift of -0.0194. Similarly, Luttmer (2007) infers a GBM variance of 0.048.

We infer a death shock of two percent. This is in line with a wide range of papers who have a similar shock structure. Two examples (among many) that cover a range inside which our value falls are: Arkolakis, Papageorgiou, and Timoshenko (2018), who calibrates this value to be 2.5 percent to match exit rates of large firms in the Colombian economy and Atkeson and Burstein (2010), who calibrate the value to be 0.55 percent to match the share of employment displaced by large firms (larger than 500 employees) in the U.S. economy.

The iceberg trade cost is large, but by no means abnormal relative to those found in the “gravity” literature that infers these costs from trade flows. For example, Waugh and Ravikumar (2016) employs a similar calibration strategy (one trade friction to match a country’s import/GDP ratio) and infers the iceberg trade cost for developed countries to be around five. The large value of five (relative to our trade costs of about three) is partially a result of the lower trade elasticity θ in that paper and partially a result of not modeling non-tradable goods. At the lower end of the range is Waugh (2010) who finds iceberg trade costs of around two for OECD countries—again, in part because Waugh (2010) employs a larger trade elasticity and has a model of both tradable and non-tradable goods.

In terms of fit, the aggregate moments, the exporter facts, and the employment share of entrants match the data exactly. Table 2 reports these aggregate moments in the model and the data. The right panel of Table 3 reports the model’s fit for the eight firm dynamic moments. While not perfect, the model replicates the dynamics of firms well. The correlation between the model and the data is extremely high at 0.98.

The final parameter to discuss is the entry cost. Here the interpretation is that the cost of entry is about eight times larger than the cost of adoption. Because this parameter value is hard to externally validate, we discuss its implications in Section 7.5 below.

7.3. The Sources of the Welfare Gains from Trade—A Quantitative Decomposition

As a first step to understanding the welfare gains from trade in the model, we perform a quantitative decomposition of the welfare change induced by a small change in trade costs. This local analysis helps

to set the stage for understanding the gains from a large change in trade costs that we present in Section 7.4 and how they differ from benchmarks in the literature such as Atkeson and Burstein (2010) or Arkolakis, Costinot, and Rodriguez-Clare (2012) (hereafter ACR). We compare welfare across steady states for two reasons. First, this across steady-state decomposition links to the theoretical analysis of the simple model performed above. Second, the steady-state analysis aids in comparing our results to the welfare gains in stationary models like Melitz (2003). We compute the full welfare gains including transition dynamics in the next section.

Decomposing the Change in Welfare due to a Change in Trade Costs. Equation (46) gives steady state utility as $\bar{U}(c, g)$. We totally differentiate utility with respect to the iceberg trade cost parameter, which provides a decomposition of the total change in utility into economically meaningful components.

To totally differentiate utility, notice that a feasible allocation for our economy in steady state can be described as a zero of a system of equations $\Gamma(\Omega, \hat{z}, g; d) = 0$.²⁰ That is, given the economic environment (resource constraint and production functions) and parameter values there is a set of feasible allocations and each allocation can be uniquely identified by a triple of values: the amount of varieties, the exporter productivity threshold, and the growth rate. Therefore, there exists a function $f_c = (\Omega, \hat{z}, g; d)$ whose output is the level of consumption c that can be produced given particular values for Ω, \hat{z}, g, d . Furthermore, the decentralized equilibrium is a particular feasible allocation and, given d , there are unique equilibrium values of $\Omega = f_\Omega(d)$, $\hat{z} = f_{\hat{z}}(d)$, and $g = f_g(d)$. Using these functions, the total derivative for the change in utility when changing the trade cost is given by the chain rule

$$\frac{d\bar{U}(f_c(\Omega, \hat{z}, g, d), f_g(d))}{dd} = \bar{U}_1 \frac{\partial f_c}{\partial d} + \bar{U}_1 \left[\frac{\partial f_c}{\partial \Omega} \frac{df_\Omega}{dd} + \frac{\partial f_c}{\partial \hat{z}} \frac{df_{\hat{z}}}{dd} + \frac{\partial f_c}{\partial g} \frac{df_g}{dd} \right] + \bar{U}_2 \frac{df_g}{dd}, \quad (59)$$

and can be grouped into three terms of interest. The first term is the direct effect of a change in the trade cost on consumption. This direct effect measures the increase in the level of consumption that would arise from a lower trade cost holding fixed the amount of labor allocated to adoption, entry, and the fixed cost of exporting. The second term in the brackets contains the indirect effects of trade costs on consumption through changes in the measure of varieties, the exporter threshold, and the growth rate. The final term is the direct effect of a change in the trade cost on growth.²¹ With log utility, equation (59) can be transformed from utils to consumption equivalent units by multiplying by the discount rate ρ .

Quantitative Decomposition. Starting from the equilibrium steady state allocation of the calibrated model, we first compute the change in welfare from a very small reduction in trade costs (the LHS of equation (59)). We then compute the derivatives on the RHS of equation (59) evaluated at the equilibrium allocations of our calibrated model. The contribution of each of the direct and indirect terms to the total

²⁰For exposition we explicitly include the dependence of allocations on just the trade cost parameter d because we focus on a decomposition of the utility change that is induced by a change in d holding fixed all other parameter values, even though of course allocations depend on all parameter values.

²¹Equation (59) is closely related to Equation (53) from Proposition (7), which provides the elasticity of utility with respect to trade costs. Equation (59), however, provides a decomposition of the total effect of a change in consumption on welfare into direct and indirect effects. Propositions 2 and 5 provide $\frac{df_{\hat{z}}}{dd}$, $\frac{df_g}{dd}$, and $\frac{d\Omega}{dd}$ analytically in elasticity form for the simple model case of no GBM and no death ($\mu = \nu = \delta = 0$).

change in welfare is

$$\underbrace{\frac{d\bar{U}(c, g)}{dd}}_{100\%} = \underbrace{\bar{U}_1 \frac{\partial f_c}{\partial d}}_{8.32\%} + \underbrace{\bar{U}_1 \frac{\partial f_c}{\partial \Omega} \frac{d\Omega}{dd}}_{-7.28\%} + \underbrace{\bar{U}_1 \frac{\partial f_c}{\partial \hat{z}} \frac{d\hat{z}}{dd}}_{0\%} + \underbrace{\bar{U}_1 \frac{\partial f_c}{\partial g} \frac{dg}{dd}}_{-9.90\%} + \underbrace{\bar{U}_2 \frac{df_g}{dd}}_{108.86\%}, \quad (60)$$

presented as a percentage of the total change. Working from left to right, first note that the direct effect of a change in trade costs on the consumption level accounts for 8.32 percent of the gains from trade. This number is important because, as we discuss below, it closely corresponds to the welfare gains from trade in Melitz (2003), Atkeson and Burstein (2010), or Arkolakis, Costinot, and Rodriguez-Clare (2012).

The next three components are the indirect effects of trade costs on consumption through changes in variety, the exporter threshold, and growth. The indirect effects on consumption of changes in variety (-7.28%) and growth (-9.90%) both reduce the gains from trade. The exporter margin effect is zero. The final term represents the main source of the gains from trade—the direct effect on growth (108.86%).

Comparison to the Efficient Allocation. The decomposition above shows that changes in economic growth are the key contributing factor to the welfare gains from trade. By using conditions that necessarily hold in an efficient economy, we next show that the gains from trade in our model predominantly arise because the decentralized equilibrium is inefficient—in particular, with growth that is too slow.

To make this point, we consider the allocation a social planner would choose in order to maximize welfare if they were subject to the same technological and resource constraints faced by firms in the decentralized equilibrium. Specifically, we consider a constrained planner who chooses the amount of labor allocated to adoption and the amount of labor allocated to fixed export costs, but faces a constraint that the amount of labor allocated to entry is fixed (making Ω constant). With Ω constant, the planner would choose a g and \hat{z} to maximize steady-state utility, since Ω is the only source of slow-moving transition dynamics. This constrained planner problem is useful to consider for two reasons. First, it allows us to better understand the steady-state comparative statics derived in the theory section above. Second, this problem helps to relate our paper to the literature, since it is closely related to a planner problem in a stationary environment without transition dynamics, which are common in the trade literature.

If the constrained planner were choosing g efficiently, then

$$\bar{U}_1 \frac{\partial f_c}{\partial g} + \bar{U}_2 = 0. \quad (61)$$

That is, if g were chosen optimally, there should be no welfare gain from increasing g a little bit, after taking into account that an increase in g will require resources, and thus lower c , all else constant.

The final two terms in equation (60) capture the growth effects on welfare and can be rearranged as

$$\left[\bar{U}_1 \frac{\partial f_c}{\partial g} + \bar{U}_2 \right] \frac{df_g}{dd}. \quad (62)$$

The optimality condition in equation (61) shows that this term is nonzero only if the growth rate varies with trade costs and if the economy is inefficient. Thus, the fact that this term contributes the large

majority of the total welfare change, means that the key source of the gains from trade in our model is due to inefficiency.

The growth rate in the decentralized equilibrium is inefficiently low due to an externality. Firms do not internalize that they improve the productivity distribution for future adopting firms when they pay the cost to adopt a better technology themselves. Lower trade costs and their effects on firms' incentives to adopt lead to more adoption and a higher growth rate. Thus, lower trade costs are partially correcting the suboptimal level of adoption and growth in the economy, which generates welfare gains above and beyond the direct consumption effect.

There could be other sources of inefficiency in our economy as well. If a planner optimally chose the exporter threshold, then there should be no welfare gain from changing that on the margin either:

$$\bar{U}_1 \frac{\partial f_c}{\partial \hat{z}} = 0. \quad (63)$$

This condition implies that the third component in equation (60) should equal zero. While we do not provide a proof that the exporter threshold is efficient in our economy, indeed the exporter threshold effect in our quantitative decomposition is zero.

Because the constrained planner cannot choose Ω and because Ω would be a state variable in the unconstrained problem, there is no way to use this analysis to compare $\bar{U}_1 \frac{\partial f_c}{\partial \Omega}$ to the value it should take in an efficient economy. In the quantitative decomposition of gains across steady states in our decentralized equilibrium, the change in the measure of varieties has a modest negative contribution to the overall welfare gain.

How Inefficient is the Decentralized Equilibrium? The term in equation (62) contributes about 100 percent to the welfare gains from trade. This term could be large for two reasons: (i) the economy is very inefficient relative to the direct consumption effect or (ii) because the economy has a large elasticity of growth to trade costs. In our calibrated economy, the inefficiency term in consumption equivalent units is $\rho \left[\bar{U}_1 \frac{\partial f_c}{\partial g} + \bar{U}_2 \right] = 44.7$. This means that if growth increases by one percentage point, consumption equivalent welfare increases by 44.7 percent. While of course faster growth—all else equal—raises welfare, in an efficient economy this one percent change in growth would generate zero welfare gain to a first order. This is because growth would have to come from a reallocation of labor into adoption, and in an efficient economy labor is equally valuable on the margin in adoption as in other activities.

A useful benchmark to discuss the size of this inefficiency is a comparison with an exogenous increase in growth that does not require resources to be drawn away from other activities. In steady state, an exogenous increase in growth of dg —leaving the level of consumption unchanged—would increase consumption equivalent utility by $\frac{dg}{\rho}$. Using the calibrated value of $\rho = 0.0203$, a one percentage point increase in “free” growth generates a 49.2 percent increase in consumption equivalent welfare. Remarkably, the 44.7 percent welfare gain that can be achieved in the decentralized equilibrium just by reallocating labor is almost the same as that which can be achieved by adding more labor to adoption at no cost. In this sense, the size of the inefficiency seems large.

If growth did not change with trade costs, then the inefficiency term in equation (62) would not con-

tribute to welfare changes.²² In our model, growth is sensitive to trade costs, with a semi-elasticity of $d \times \frac{df_g}{dd} = -0.028$. Extrapolating using this local elasticity, the ten percent reduction in trade costs that we explore in Section 7.4 increases growth by approximately 0.19 percentage points. We do not have strong evidence regarding the plausibility of this semi-elasticity value. As a point of comparison, we compute the average productivity growth for non-overlapping decades in the U.S. from 1950-2017. The standard deviation of those decadal averages is 0.78. Thus, movements in the growth rate that we find and study are well within the range of historical variation in the long-run aggregate productivity growth rate.

To summarize, when the economy is efficient, the direct effect of lower trade costs on consumption, i.e., the first term in equation (60) is the only source of gains. The decomposition in equation (60) shows that the total gains are an order of magnitude larger than those from the direct consumption effect due to a large inefficiency component $\left[\bar{U}_1 \frac{\partial f_c}{\partial g} + \bar{U}_2 \right]$ and not because growth rates are extremely sensitive to trade costs $\left(\frac{df_g}{dd} \right)$.

This observation is consistent with Atkeson and Burstein's (2019) conclusion that the elasticity of growth with respect to innovation policy is likely small, but the welfare gain from small increases in growth can be large. This is also consistent with the conclusions of Baqaee and Farhi (2020), who find that in inefficient economies the effects of reallocation on welfare can be orders of magnitude larger than direct effects.

The size of growth externalities and the degree of inefficiency at the country or world level is still an open question in economics. This question is especially salient in the field of economic growth because most of the prominent models of growth feature externalities that make the decentralized equilibrium inefficient. Jones and Williams (1998, 2000) and Bloom, Schankerman, and Van Reenen (2013) estimate the social rate of return to investment in growth and find very large numbers relative to estimates of the private rate of return, suggesting large externalities and very inefficient economies.

Relation to the Literature. This analysis delivers a bridge between the results in this paper and prominent benchmarks in the trade literature. For example, Atkeson and Burstein (2010) argue that the only first order effect lower trade costs have on welfare is the direct consumption effect and that indirect effects such as changes in variety or innovation are second order. The logic above conforms with Atkeson and Burstein (2010) because their argument is applied to an efficient economy (see their footnotes 11 and 24). The decentralized equilibrium in our paper is not efficient and, hence, the gains from trade will be different than those in Atkeson and Burstein (2010).

Similarly, our model with $g = 0$ is very similar to the Melitz (2003) model, except the Melitz model has selection on exit that induces instantaneous transitions. Because the Melitz model is efficient (see Dhingra and Morrow (2019)), the total welfare gain in the Melitz model is equal to the gain from the direct consumption effect. That is, the Melitz model is always in steady state and efficiency means $\bar{U}_1 \frac{\partial f_c}{\partial \Omega} = 0$. The constrained planner problem we considered above, which held varieties fixed, is comparable to the Melitz economy case because it enforced by construction zero gains from changes in varieties. The wel-

²²In fact, this is exactly what is occurring in the No Selection into Exporting model. While there is an inefficiency because of the externality in that model, growth does not move with changes in trade costs and, hence, the inefficiency does not matter for the welfare gains from trade.

fare gain from lower trade costs in our model is larger than that in the Melitz model not only because we have nontrivial transition dynamics or because $g > 0$ in our model, but crucially because the way we model growth introduces an externality such that the decentralized equilibrium is not efficient.

Quantitatively, we find that the direct effect on consumption (the first term in equation (60)) corresponds almost exactly to the gains implied by the ACR formula. That is, the direct consumption effect of a change in trade costs on consumption equivalent welfare $\left(\rho \bar{U}_1 \frac{\partial f_c}{\partial d}\right)$ is 99 percent of the value given by the ACR formula. The ACR formula does not, however, characterize the total welfare change mostly because our economy is not efficient and the total welfare change is dominated by the inefficient growth rate effect. Arkolakis, Costinot, Donaldson, and Rodriguez-Clare (2019) explore non-CES demand functions and non-Pareto productivity distributions and conclude that empirically plausible deviations from CES and Pareto are unlikely to deliver much larger gains from trade. Our paper shows that inefficient growth—arising from externalities similar to those that are central to most prominent models of endogenous growth—can provide an alternative mechanism through which there may be large gains from trade.

7.4. The Welfare Effects of a Reduction in Trade Costs

We now study how the economy is affected by a larger reduction in trade costs inclusive of transition dynamics. We start our quantitative experiment from the baseline economy on a BGP and then shock the economy with an unanticipated ten percent permanent reduction in ad valorem trade costs ($d_T = 1 + 0.9 * (d_0 - 1)$). We study how the economy transits from the baseline BGP equilibrium to the new low-trade-cost BGP equilibrium.

7.4.1. Trade, Growth, and Entry Dynamics

The reduction in the trade cost at date zero causes an increase in imports relative to GDP. Figure 1 shows that imports instantly jump almost entirely to the new steady state level and then slowly converges to its final value. Part of this rise in trade simply comes from the reduction in trade costs, i.e., the cost of importing declines, and thus existing exporters expand and sell more to foreign markets. The second force that increases trade comes from the extensive margin with the entry of new exporters, i.e., the lower iceberg trade cost decreases the exporter productivity threshold, \hat{z} , leading to an expansion of exports and imports. Table 4 reports that across steady states, the ten percent reduction in trade costs causes about a 3.8 percentage point increase in imports to GDP. This is closely related, but not exactly equal, to the trade elasticity θ due to general equilibrium effects.

Unlike the rapid change in the volume of trade, the measure of domestically produced varieties takes time to adjust. Figure 2 illustrates this process, showing that the measure of domestic varieties falls gradually after the reduction of trade costs. The measure of varieties is lower in the new steady state because import competition decreases revenue for domestic non-exporting firms and because domestic exporting firms bid up the local wage as they hire more labor in order to increase their exports. Altogether, this leads to a decrease in the expected value of entry relative to the cost of entry and the net-exit of firms. This mirrors the result in the simple analytical model that the measure of domestically produced varieties is smaller in steady states with lower trade costs (see Proposition 5).

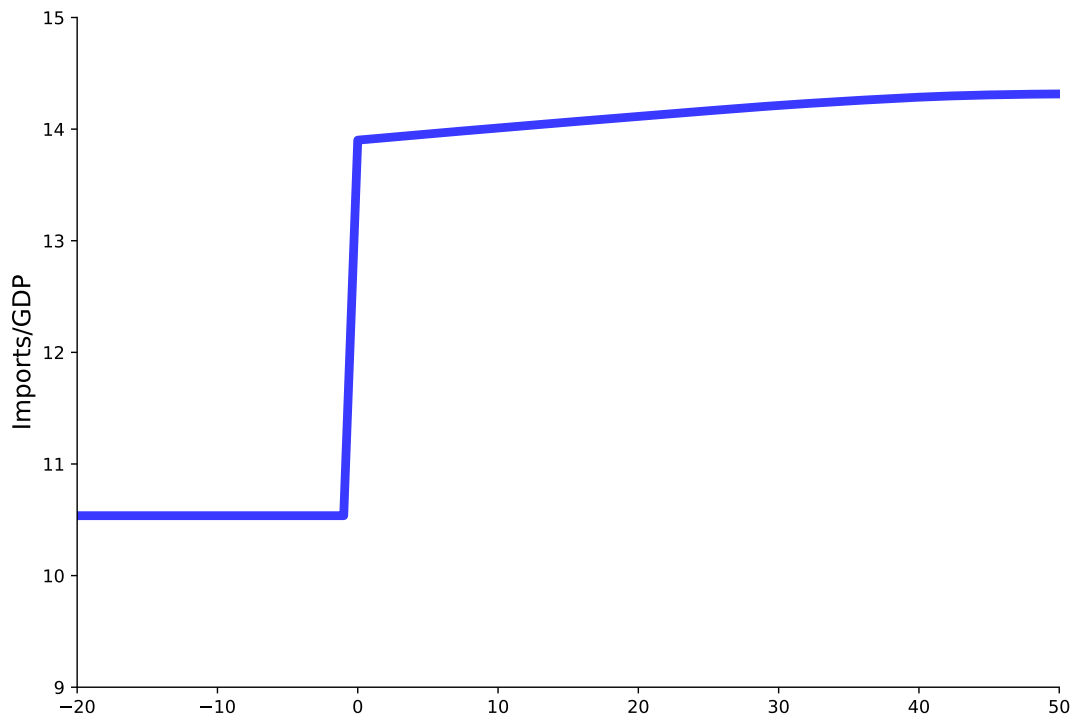


Figure 1: Imports/GDP: $1 - \lambda_{ii}(t)$

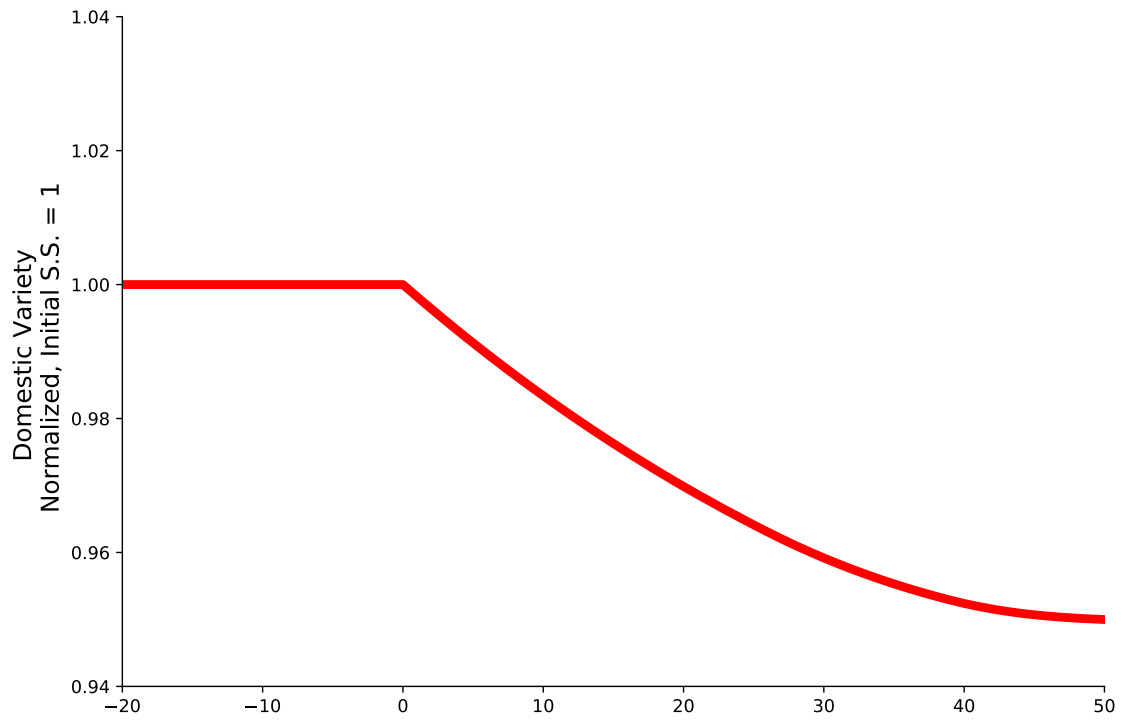


Figure 2: Domestic Variety: $\Omega(t)$

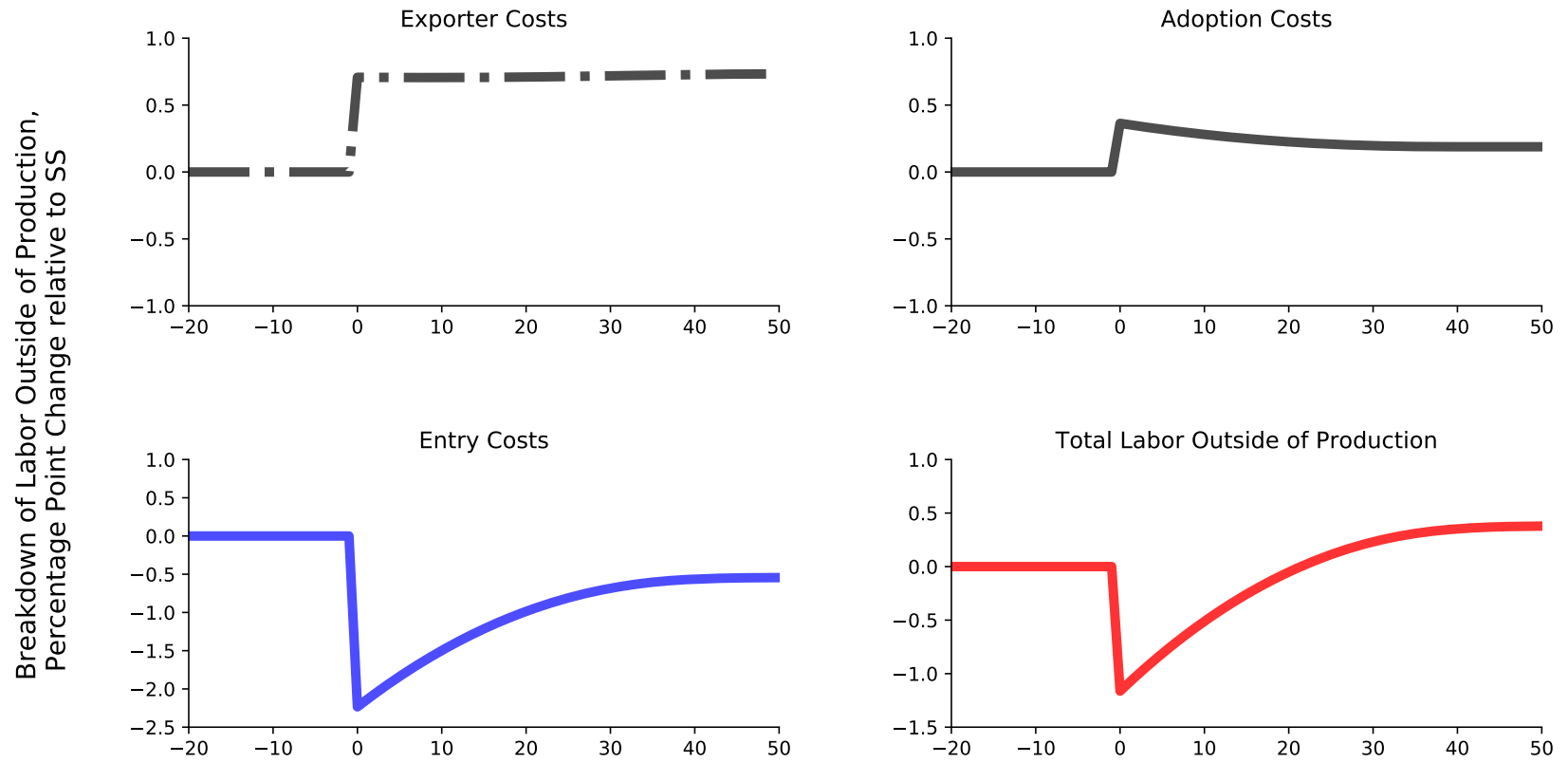


Figure 3: Reallocation Effects in Fixed Costs

Two forces are behind the gradual exit of firms. The first force is mechanical: since exit is exogenous it takes time for the measure of varieties to adjust. The second, endogenous force, relates to consumption smoothing and leads to a positive entry rate along the transition path. Thus, net-exit is less than the mechanical exit of firms would deliver while the economy is transitioning to its new steady state $\left(\left|\frac{\partial\Omega(t)}{\partial t}\right| < \delta\right)$.

The consumption smoothing motive is just like that in the neoclassical growth model when the capital stock is too large relative to its steady state value. In the neoclassical growth model, along the optimal transition path, investment is still positive because consumers want to smooth consumption, i.e., avoid large differences in consumption across time periods. Analogously, in our model it is feasible to enjoy initially higher consumption and have zero entry on the transition path, but the intertemporal smoothing motives dictate that the economy forgo this path to satisfy the desired consumption plan as given by the Euler equation.

The desired path of consumption across time must be supported by the allocation of labor across activities, over time. The increase in trade means more labor is allocated to the fixed costs of exporting. The decrease in entry, however, means less labor is needed for entry costs. Quantitatively, the decrease in labor allocated to entry is much larger than the increase in labor allocated to export costs, so in an accounting sense there is “excess labor” to be allocated. The two other tasks that use labor are adoption and the production of goods. Allocating all of the excess labor to the production of goods would maximize short-run consumption. In contrast, allocating all of the excess labor to adoption would increase growth rates and maximize long-run consumption. Due to the intertemporal smoothing motive, the economy does not wind up at a corner solution on the transition path, but rather the allocation of labor balances out more consumption today (production) with more consumption in the future (adoption).

Figure 3 illustrates the change in the allocation of labor across activities. Note the interpretation of the y-axis is literally the change in labor allocated to that activity because the total labor endowment is normalized to one. The top left panel illustrates the 0.75 percentage point rise in labor allocated to the fixed costs of exporting, consistent with the jump in trade. The top right panel depicts the increase in adoption activity by a bit less than 0.50 percentage point. In contrast, the bottom left panel in Figure 3 shows that the amount of labor dedicated to entry falls dramatically—nearly 2.5 percentage points. The bottom right panel sums everything up: On impact, on net, the amount of labor allocated to these investment-like activities falls by a little more than one percentage point. After the decline, the total amount of labor in these activities steadily rises to a new higher level than in the initial steady state.

These changes in the allocation of labor have two implications. First, consumption overshoots along the transition path. Figure 4 illustrates that this effect is large. As the inset in Figure 4 shows, the trade liberalization induces an initial increase in consumption C_t of more than three percent. Along the transition path, the level of consumption $C(t)$ always lies above its previous path. Productivity grows faster than consumption, however, so that normalized consumption $c_t := C_t/M(t)$ steadily declines to its new lower level of about one percent below the initial level. This effect can be seen in Figure 4 by noting that the level of normalized consumption ends up lying below the level of productivity in period 50. Proposition 6 stated that normalized consumption is lower in BGPs with lower trade costs, but the transition dynamics analysis illustrates that consumption first overshoots before eventually converging

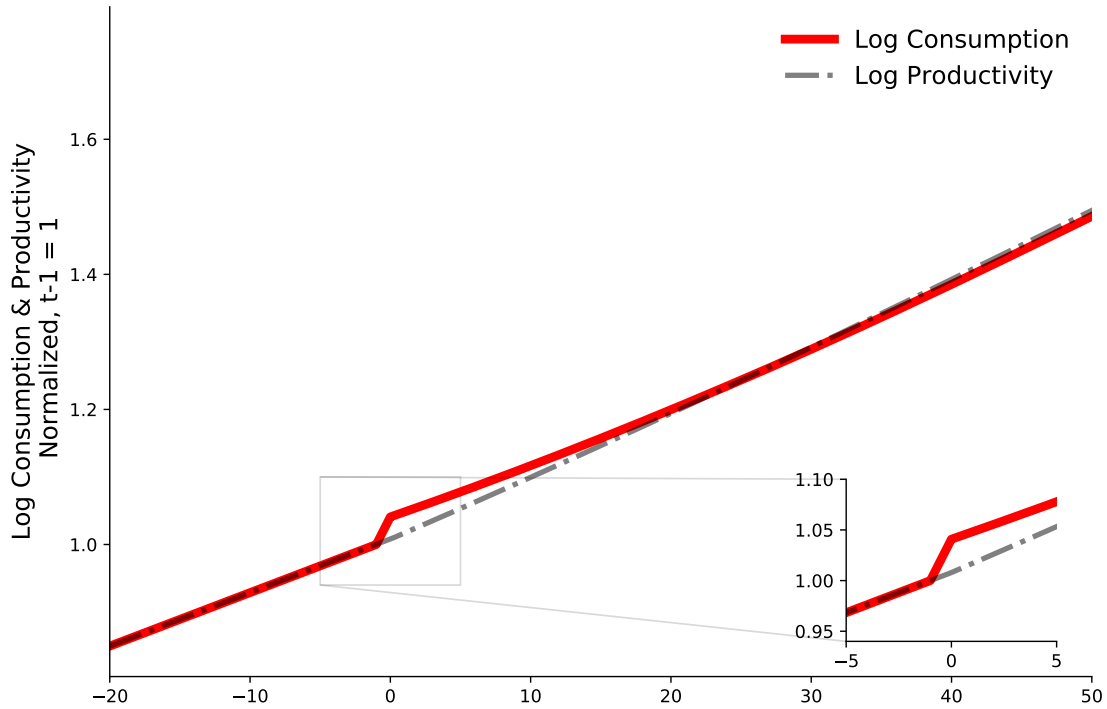


Figure 4: Log Consumption: $C(t)$

to the new lower relative level. The dynamics of consumption and varieties during the transition are closely related to those discussed in Burstein and Melitz (2013).²³

The second implication is that productivity growth slowly rises to its new steady state rate. Figure 5 shows that after the change in trade costs, there is an initial jump in productivity growth from the baseline of 0.79 percent to 0.90 percent. From this point on, productivity growth gradually rises at a decelerating rate as it moves towards the new BGP growth rate of 1.03 percent—0.24 percentage points higher than in the baseline. In other words, a little less than one half of the increase is instantaneous and the remaining half of the rise in growth plays out over the next 30 years. To emphasize this point, this pattern of growth and the corresponding adoption behavior is intimately connected with the pattern of consumption. Since adoption entails a tradeoff of paying a cost today for gains that will payout in the future, the firm’s discount rate which reflects the consumer’s desired path of consumption must support investment in adoption and rising productivity growth.

7.4.2. The Welfare Gains from Trade

From a welfare perspective, the reduction in trade costs leads to several competing effects (as Proposition 7 shows for our simple economy). Productivity growth increases due to higher adoption rates. Comparing across steady states, more adoption is associated with a lower level of normalized consumption, as firms are using more labor to invest in technology upgrading. The reduction in variety along the transition path, however, more than offsets the effects from more adoption, leading to consumption

²³This logic is also related to the intuition for consumption overshooting in Alessandria and Choi (2014). Alessandria, Choi, and Ruhl (2014) study a decrease in tariffs and find amplified welfare gains from trade, in part because tariffs generate inefficiency.

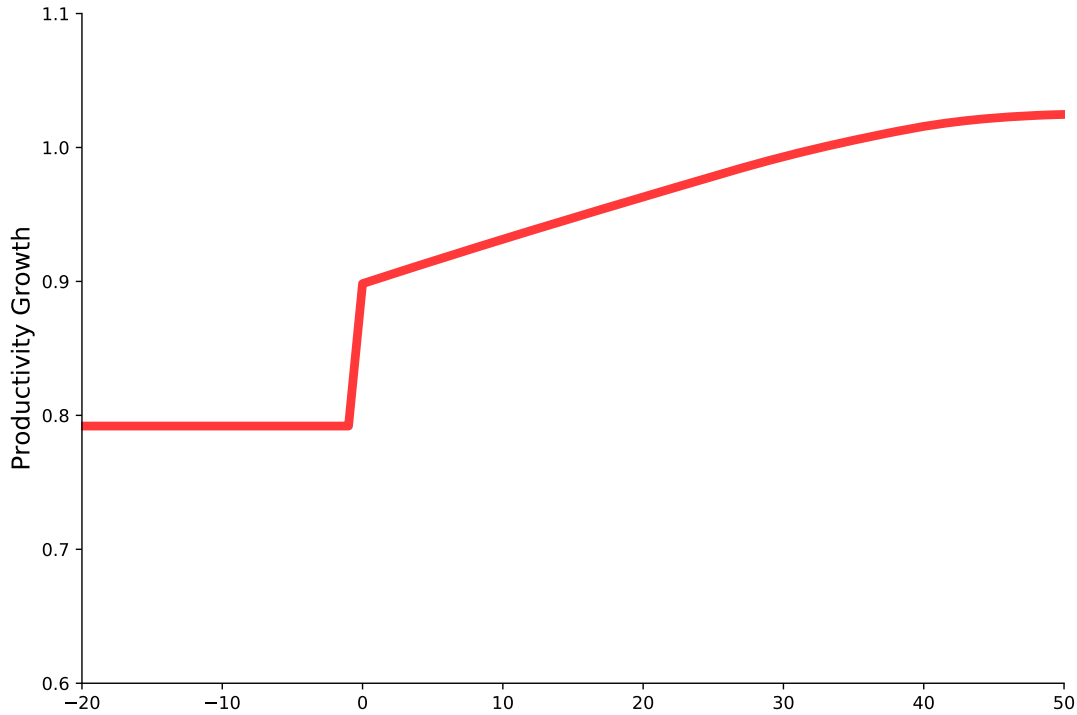


Figure 5: Productivity Growth: $g(t)$

overshooting.

To measure these effects on welfare, the final two rows of Table 4 report consumption-equivalent gains associated with the ten percent reduction in trade costs. Consumption-equivalent gains are defined as the permanent percent increase in consumption a household requires in the old regime to be indifferent between the new and old regimes.

The third row reports that the consumption equivalent welfare gain inclusive of the transition path is 10.8 percent. Because this is inclusive of the transition path, it includes the net-effect of two different forces: (i) the increase in growth is not immediate which tends to lower welfare due to discounting and (ii) the path of consumption overshoots which tends to increase welfare gains. The fourth row of Table 4 provides a sense of how much the transition path matters by computing the consumption-equivalent gain across steady states. While either force can dominate depending on parameter values, in our calibrated model the gains from consumption overshooting are not enough to offset the delay of faster growth. Thus, the net effect of the transition path is to modestly diminish the welfare gains from trade.

Returning to the decomposition of welfare gains in Section 7.3, we can use the local results to learn about the sources of this large change in trade costs. The local approximation for the consumption equivalent welfare gain across steady states from a 10 percent reduction in trade costs is 8.5 percent and the predicted change in the growth rate is 0.19 percentage points. Solving for the gains globally, we find an 11.2 percent gain and a 0.24 percentage point increase in growth, showing that the local approximation is decent, if not perfect. From this we conjecture that the insights on the sources of the welfare gains

Table 4: 10 Percent Reduction in Trade Costs: Growth, Trade, and Welfare

	Baseline BGP	New BGP
Growth	0.79	1.03
Imports/GDP	10.6	14.4
Welfare		
Consumption Equivalent Gain (Transition Path):		10.8
Consumption Equivalent Gain (SS to SS):		11.2

Note: All values are in percent. Consumption-equivalent is the permanent percent increase in consumption a household requires in the old regime to be indifferent between the new and old regimes.

derived from the local analysis are likely to apply globally. That is, the consumption gains are large relative to benchmark measures in the literature predominantly because the equilibrium growth rate is inefficiently low. It is not enough to observe that the growth rate increased by 0.24 percentage points to conclude that welfare gains should be large. It is this increase in growth multiplied by the size of the inefficiency that mostly determines the welfare gains.

Relation to the Literature. To illustrate that these gains are large compared to prominent benchmarks in the literature, we use the ACR formula to compute the gains from trade implied by our trade elasticity (Pareto shape parameter θ) and the model-predicted change in trade. Using their formula the welfare gains are simply $100 \times -\frac{1}{\theta} \log \left(\frac{1-0.144}{1-0.106} \right) = 0.87\%$. This is an order of magnitude smaller than our welfare gains.²⁴ As we discuss above in Section 7.3, the ACR formula corresponds closely to the direct consumption effect of a change in trade costs on welfare, which would be equal to the total gains if our economy were efficient. In our model the welfare gains are an order of magnitude larger because lower trade costs increase an inefficiently low rate of adoption and growth.

Another point of comparison is Atkeson and Burstein (2010) who study a dynamic model with both process and product innovation featuring nontrivial transition dynamics. However, the welfare gains from trade in their model are an order of magnitude smaller than ours. To illustrate this point, we recalibrated their economy according to our calibration scheme.²⁵ In the Atkeson-Burstein model with our calibrated parameter values, a 10 percent reduction in trade costs generates a 0.85 percent increase in consumption equivalent units—almost exactly equal to the ACR gain reported above for our economy.

²⁴Large reductions in trade flows lead to large welfare losses in our model, but welfare does not fall as much in relative terms as the ACR formula would predict. Comparing steady states of our model, a move to autarky would generate a welfare loss of about 22 percent. In other words, starting from baseline, the loss from a move to autarky is about two times the gains from a ten percent reduction in trade costs. In contrast, the ACR formula would predict a 2.2 percent loss from autarky—more than two and a half times larger than the gains from a ten percent reduction in trade costs.

²⁵Specifically, we set the discount rate, elasticity of substitution, the exit rate of firms, and the aggregate import share to be the same as in our calibration. We focused on the intermediate innovation elasticity case of $b = 30$.

This result conforms with the discussion above—that the Atkeson and Burstein (2010) economy is an efficient economy with only the direct consumption effect driving welfare gains.

Hsieh et al. (2019) is a related paper that also features an inefficient economy due to growth externalities. They focus on the dynamic gains from trade and international technology diffusion in a model of creative destruction along the lines of Klette and Kortum (2004) and Garcia-Macia et al. (2019). Compared to our paper, Hsieh et al. (2019) have even larger welfare gains from trade. In their baseline economy, consumption equivalent welfare is 30 percent higher than in autarky.

Sampson (2016) is perhaps the closest paper to ours. Sampson studies the role of dynamic selection in generating dynamic gains from trade. The idea is that trade induces endogenous exit which makes the pool of ideas from which entering firms draw technologies better and, thus, facilitates faster economic growth. Most important, dynamic selection creates an externality which firms do not internalize as they make their exit decision. As in our model, Sampson's provides scope for large gains from trade because his economy is not efficient.

A direct quantitative comparison of our results to those of Sampson (2016) is complicated by the use of different calibration strategies and welfare metrics (we study transition paths). Therefore, we recalibrate our model and compute new counterfactuals to facilitate a comparison of the quantitative predictions of the two papers. Specifically, we shut down the GBM process, and then re-calibrate the remaining parameters of our model to match Sampson's aggregate moments and his discount rate.²⁶

After making these calibration adjustments, we study the experiment of Sampson (2016) by computing the equilibrium growth rates associated with the observed level of trade and autarky. By construction, in our model and in Sampson's the steady state growth rate at observed trade levels is 1.56 percent. In our model, the autarky growth rate is 1.34 percent, while in Sampson's it is 1.41 percent (see Table 2 of Sampson (2016)). There are also similarities in terms of welfare, as the welfare loss from autarky in our re-calibrated model is -3.16 percent, while it is -3.6 percent in Sampson. Thus, when calibrating our model to mimic Sampson's model, we find very similar steady state changes in the aggregate growth rate and in the welfare gains from trade.

These results show that the calibration scheme is the key difference in generating the different gains from trade between Sampson (2016) and our model. More specifically, it is the role that firm dynamics play in shaping outcomes. For example, just shutting down GBM and recalibrating our model to our aggregate moments yields smaller changes in growth and welfare compared to our baseline model, similar to those in Sampson (2016). The growth rate drops from 0.79 percent to 0.63 percent and the welfare loss from autarky is -5.67 percent. Thus, calibrating the model so that it generates realistic firm dynamics has a quantitatively significant effect on the response of growth and welfare to changes in trade costs. We now turn to a more detailed exploration of the relationship between firm dynamics and the welfare gains from trade.

²⁶We set μ and v^2 to zero and fix σ equal to 3.17 (our calibrated value). We then choose the remaining parameter values to target our exporter moments, his aggregate growth rate of 1.56 percent, import/GDP ratio of 0.081, and set our discount rate equal to 0.04 (See Table 1 and 2 of Sampson (2016)).

7.5. The Role of Firm Dynamics and Adoption Costs

This section discusses how some of the parameters that are absent from our simple analytical model shape the growth effects and welfare gains from trade. Relative to our theoretical results that are derived only in the case of ($\mu = v = \delta = 0$), the open question is how these firm productivity and exit shock parameters (and hence firm dynamics data) affect trade, consumption, and growth.

In this section we show that the elasticity of growth to trade costs is sensitive to χ (the relative cost of adoption). The parameter χ is hard to discipline directly from firm-level data; thus, it is calibrated indirectly, primarily influenced by the aggregate growth rate moment. Here we show that these firm dynamics moments affect the computed welfare gains from lower trade costs mostly by affecting the calibrated value of χ . It would be ideal to directly target moments related to adoption and entry costs to pin down the size of the inefficiency in the economy, but absent such a measured moment to target, we believe it is better to include firm dynamics data than to only target cross-sectional and aggregate moments.

Focusing on the importance of the GBM variance parameter v^2 , the left panel in Figure 6 plots the change in growth rates in response to a ten percent reduction in trade costs on the vertical axis and the variance of the GBM process on the horizontal axis. Each line is for a different value of the adoption cost parameter χ — the large χ value is ten percent larger than the baseline calibrated χ value and the small χ value is ten percent smaller than the baseline calibrated value. All other parameter values are fixed, i.e., we do not re-calibrate the model when changing these parameter values. The middle blue line is drawn for the baseline value of χ and the intersecting dashed line indicates the baseline variance. Consistent with our numerical results above, the left panel of Figure 6 reports that the aggregate growth rate changes by about 0.24 percentage points across steady states.

The first thing to observe from Figure 6 is that the percentage point change in productivity is nearly constant across different values of the variance parameter. In other words, the variance v^2 does not much affect the response of growth to a change in trade costs.

The parameter which does influence the change in growth is the adoption cost parameter, χ . The three different lines on the left panel in Figure 6 illustrate this point. A small value of χ (top black line) corresponds to small costs of adoption. When adoption costs are small, growth is more responsive to changes in trade costs. In contrast, a large value of χ (bottom red line) corresponds to large adoption costs and a smaller response of growth to trade costs. The closed form equations available in the non-GBM version of the model deliver some insight. Equation 35 shows that the change in the growth rate for a given change in trade costs is larger when adoption costs are smaller.²⁷

²⁷In all the cases in this figure, the change in trade costs is held constant, but the change in the amount of trade is slightly different as the χ parameter shows up in the trade share. In total, the effect that χ has on the change in growth is $\left(\frac{\rho}{\chi} - \rho\right)^{\frac{\theta}{\sigma-1}}$.

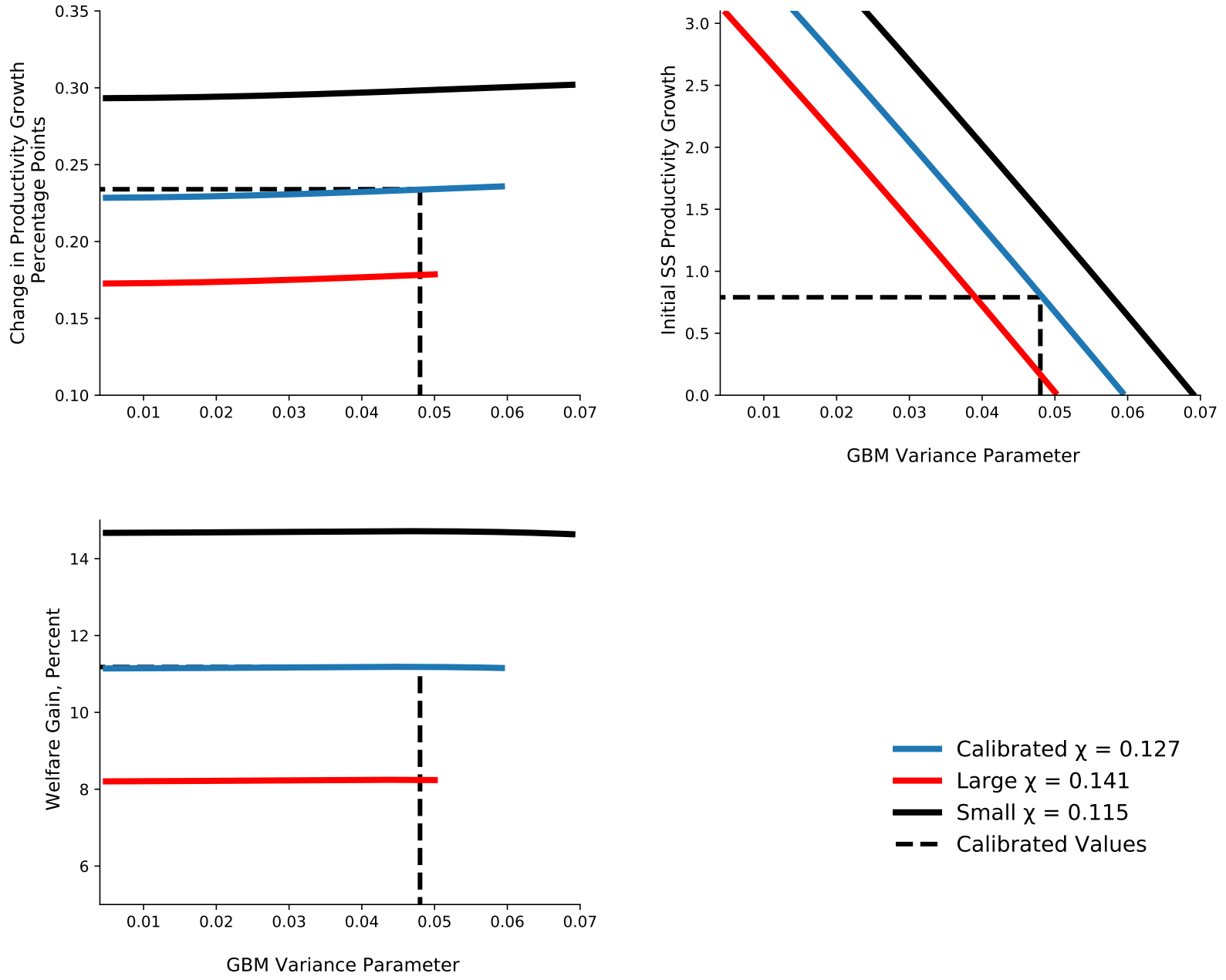


Figure 6: GBM variance, Adoption Costs, and Economic Growth

Even though the elasticity of growth to trade costs is not sensitive to the value of v^2 holding adoption costs constant, the value of v^2 —and, thus, the firm dynamics data—strongly influences the calibrated value of the adoption cost. The right panel in Figure 6 illustrates this point by tracing out how the growth rate in the initial steady state varies with v^2 . For a given χ value, there is a near linear decrease in the steady state growth rate as the variance increases. Across χ values, the slope is essentially the same, but the intercept shifts, with smaller χ values leading to higher growth rates. This is intuitive—lower adoption costs lead to more adoption and faster economic growth.

The implication of these observations is that data on firm dynamics influences the inferred adoption cost and, thus, the elasticity of growth to trade costs. For example, holding fixed our target of an aggregate growth rate of 0.79 percent, if the transition matrix of relative size (Table 3) had pushed for us to find a smaller value for v^2 , then the right panel of Figure 6 shows this would have led us to calibrate a larger value for χ . Combining this observation with the left panel of Figure 6, our calibration strategy would have then led to a smaller increase in the growth rate for the same decrease in trade costs.

The lower panel of Figure 6 shows that the welfare gains from trade (comparing BGPs) are nearly constant across values of v^2 , but sensitive to the value of χ , just like the elasticity of growth to trade costs. Thus, the value of χ is crucial for determining both the change in growth and the welfare gains from trade. Even though v^2 does not much affect the welfare gains from trade when holding all other parameters constant, different values of v^2 (which are associated with different firm dynamics moments) affect the calibration of χ . It is in this sense that not just firm heterogeneity, but firm dynamics, matter for the welfare gains from trade in our model.

Our discussion above, which compares our gains from trade to those in Sampson (2016), strongly suggests this point as well. When the GBM process is shut down and the model is re-calibrated, the gains from trade are still larger than what the ACR formula would imply, but they are far more modest and in line with what Sampson finds. Recall from Section 7.3 that much of the welfare gains arise because the equilibrium has an inefficiently low growth rate and that changes in trade costs change the growth rate. Using the decomposition from Section 7.3, we find that the different values of χ associated with different values of v^2 affect the welfare gains from trade almost completely because of a change in the sensitivity of growth to the trade cost $\left(\frac{df_g}{dd}\right)$ and not because of different levels of inefficiency $\left(U_1 \frac{\partial f_c}{\partial g} + U_2\right)$.

The other parameter introduced in the quantitative model is δ , which controls the rate of entry and death. Figure 7 is analogous to Figure 6, except that the horizontal axis varies δ . The top left panel shows that, holding all else equal, lower values of δ lead to less responsive changes in the growth. This is completely different from the analysis of the GBM variance presented in Figure 6.

Similarly to the GBM variance case, the δ parameter interacts with the adoption cost parameter to affect the calibrated value of χ . The right panel in Figure 6 illustrates this point by tracing out how the growth rate in the initial steady state varies with δ . For a given χ value, the steady state growth rate increases with δ ; across χ values, smaller χ values (lower adoption costs) lead to higher growth rates. Figure 6 shows that larger δ values (i.e., more entry observed in the data) would induce the calibration to infer larger χ values. But because these two parameters have opposite effects on economic growth, the change in parameter values generates offsetting effects and leaves the model's elasticity of growth to trade costs

relatively unchanged.

The welfare gains from trade display a similar pattern. The bottom panel of Figure 7 shows that the welfare gains from trade increase with the value of δ , holding all else fixed. Again, however, larger values of δ generate larger calibrated values of χ , which offset to keep the welfare gains from trade largely unchanged. Re-calibrating the model holding fixed different values for δ verifies this observation—welfare only increases slightly as δ increases.

In summary, Figure 6 shows that if the empirical transition matrix of relative size (Table 3) had pushed for us to calibrate a smaller value for v^2 , it would have resulted in a larger value for χ and, thus, a smaller response of growth to trade costs and smaller welfare gains from trade. In contrast, Figure 7 shows that the responsiveness of growth to trade costs and the welfare gains from trade would largely have been the same if the data featured more or less entry. In both contexts, an important lesson is that the adoption cost parameter, χ , has a large quantitative impact on the welfare gains from trade.

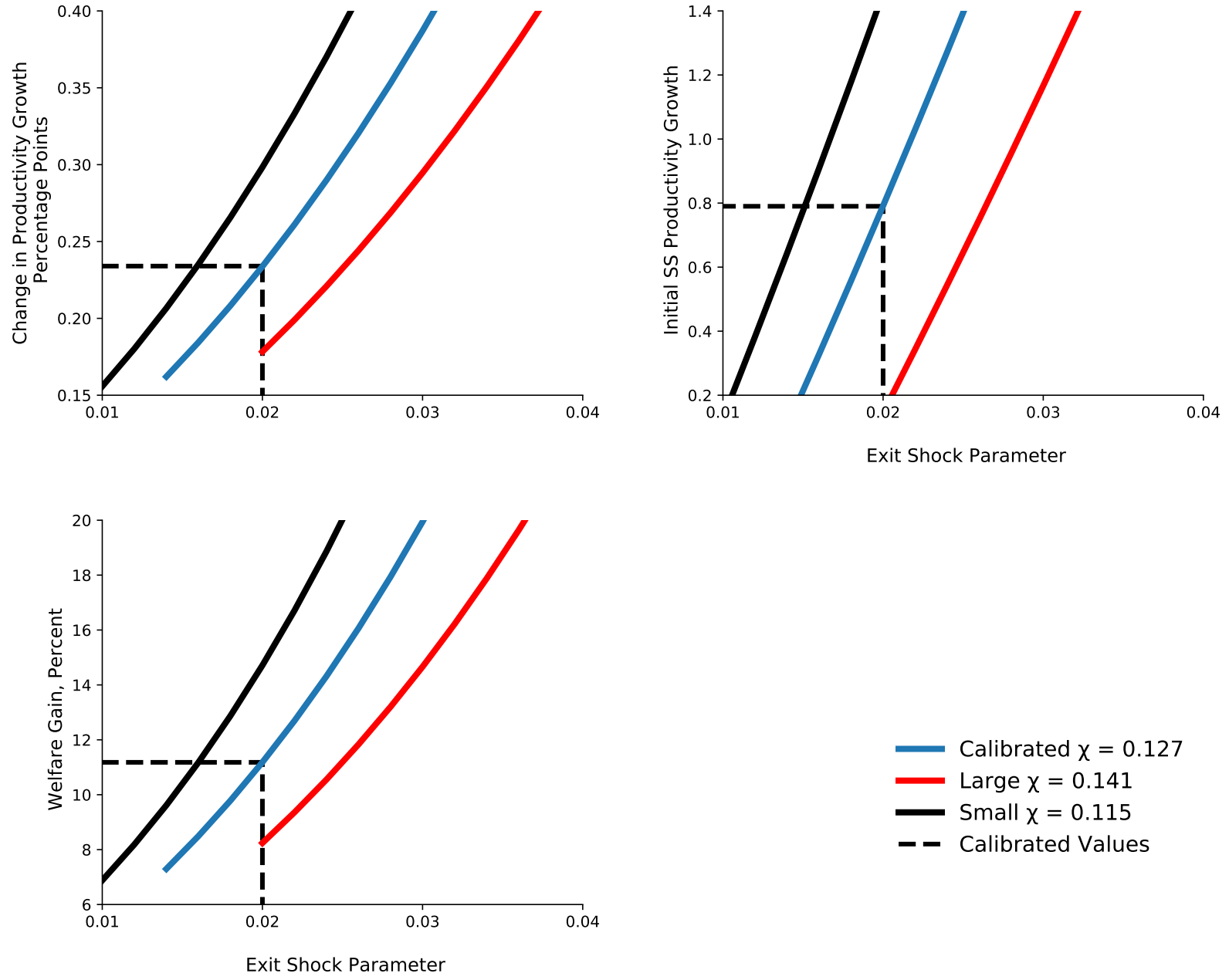


Figure 7: Exit Shock, Adoption Costs, and Economic Growth

8. Conclusion

This paper contributes a dynamic model of international trade and growth, driven by domestic technology diffusion. Firms choose to upgrade their productivity by adopting better technologies in order to remain competitive and profitable. There is an externality, in that firms do not internalize that by adopting better technologies they improve the adoption opportunities for future adopters. In response to a decline in trade costs, highly productive firms benefit because they are the exporters who can take advantage of increased sales abroad. Low productivity firms only sell domestically and are hurt by the increased competition from foreign firms. This paper provides a mechanism that links this increase in import competition to an increase in firm-level technology adoption and aggregate growth and studies the welfare gains from trade. Quantitatively, calibrating the model to match U.S. aggregate and firm-level moments, we find large welfare gains from trade—an order of magnitude larger than standard calculations like Atkeson and Burstein (2010) and ACR deliver—largely because the economy is inefficient.

We want to share two broad takeaways regarding these results and directions for future research. Our first broad takeaway is that increases in international trade have the potential to deliver large welfare gains by addressing inefficiencies. It is not a new idea that large welfare effects can occur in inefficient economies, but our results emphasize this in a transparent nontrivial economic environment with international trade and growth. In doing so, hopefully we shed light on the sources of the gains from trade in efficient economies such as Melitz (2003) or Atkeson and Burstein (2010). One lesson is that studying the gains from trade may be inherently tied to studying the degree of inefficiency in aggregate economies.

Our second takeaway is that many papers studying the welfare gains of policy interventions or technological change reach the conclusions they do because of the size of the inefficiency in their baseline calibrated economies. This shows the great importance of quantifying the size of such inefficiencies (this is a point raised by Atkeson and Burstein (2019)). Our paper, and others like Garcia-Macia, Hsieh, and Klenow (2019) and much of the research of Ufuk Akcigit and coauthors (see e.g., Akcigit, Aghion, and Howitt (2014); Acemoglu, Akcigit, Alp, Bloom, and Kerr (2018)) are examples of research attempting to learn about the size and importance of knowledge externalities using data on firm dynamics combined with growth theory. While admittedly indirect, these approaches suggest that externalities are large, and in our case they lead to large welfare gains to trade. One interesting path forward is to combine model based and direct approaches to measuring the size of knowledge externalities, such as using estimates of the social rate of return from Bloom et al. (2013); Jones and Williams (1998, 2000).

References

- ACEMOGLU, D., U. AKCIGIT, H. ALP, N. BLOOM, AND W. KERR (2018): "Innovation, Reallocation and Growth," *American Economic Review*, 108, 3450–3491.
- AKCIGIT, U., P. AGHION, AND P. HOWITT (2014): "What Do We Learn from Schumpeterian Growth Theory?" in *Handbook of Economic Growth*, ed. by P. Aghion and S. N. Durlauf, Elsevier, vol. 2, 515–563.
- AKCIGIT, U., S. T. ATES, AND G. IMPULLITTI (2018): "Innovation and Trade Policy in a Globalized World," Tech. rep., National Bureau of Economic Research.
- ALESSANDRIA, G. AND H. CHOI (2014): "Establishment heterogeneity, exporter dynamics, and the effects of trade liberalization," *Journal of International Economics*, 94, 207–223.
- ALESSANDRIA, G., H. CHOI, AND K. RUHL (2014): "Trade adjustment dynamics and the welfare gains from trade," Tech. rep., National Bureau of Economic Research.
- ALVAREZ, F. E., F. J. BUERA, AND R. E. LUCAS (2017): "Idea Flows, Economic Growth, and Trade," *working paper*.
- AMITI, M. AND J. KONINGS (2007): "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia," *The American Economic Review*, 1611–1638.
- ARKOLAKIS, C. (2010): "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy*, 118, 1151–1199.
- (2015): "A Unified Theory of Firm Selection and Growth," *Quarterly Journal of Economics*, forthcoming.
- ARKOLAKIS, C., A. COSTINOT, D. DONALDSON, AND A. RODRIGUEZ-CLARE (2019): "The Elusive Pro-Competitive Effects of Trade," *Review of Economic Studies*, 86, 46–80.
- ARKOLAKIS, C., A. COSTINOT, AND A. RODRIGUEZ-CLARE (2012): "New Trade Models, Same Old Gains?" *American Economic Review*, 102, 94–130.
- ARKOLAKIS, C., T. PAPAGEORGIOU, AND O. A. TIMOSHENKO (2018): "Firm learning and growth," *Review of Economic Dynamics*, 27, 146–168.
- ARROW, K. (1962): "Economic Welfare and the Allocation of Resources for Invention," in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton University Press, 609–626.
- ATKESON, A. AND A. BURSTEIN (2019): "Aggregate Implications of Innovation Policy," *Journal of Political Economy*, 127, 2625–2683.
- ATKESON, A. AND A. T. BURSTEIN (2010): "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 118, 433–484.
- AUTOR, D., D. DORN, G. H. HANSON, G. PISANO, AND P. SHU (2019): "Foreign Competition and Domestic Innovation: Evidence from US Patents," *American Economic Review: Insights*, forthcoming.
- AXTELL, R. L. (2001): "Zipf distribution of US firm sizes," *science*, 293, 1818–1820.
- BAILY, M. N., C. HULTEN, AND D. CAMPBELL (1992): "Productivity dynamics in manufacturing plants," *Brookings papers on economic activity*, 187–267.
- BALDWIN, R. E. AND F. ROBERT-NICOUD (2008): "Trade and Growth with Heterogeneous Firms," *Journal of International Economics*, 74, 21–34.

- BAQAEE, D. R. AND E. FARHI (2020): "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 135, 105–163.
- BENHABIB, J., J. PERLA, AND C. TONETTI (2019): "Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution and Technology Frontier," *working paper*.
- BERNARD, A., J. EATON, J. B. JENSEN, AND S. KORTUM (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93, 1268–1290.
- BERNARD, A. B., S. J. REDDING, AND P. K. SCHOTT (2011): "Multiproduct Firms and Trade Liberalization," *The Quarterly Journal of Economics*, 126, 1271–1318.
- BLOOM, N., M. DRACA, AND J. V. REENEN (2016): "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity," *Review of Economic Studies*, 83, 87–117.
- BLOOM, N., P. M. ROMER, S. J. TERRY, AND J. VAN REENEN (2014): "Trapped Factors and China's Impact on Global Growth," *NBER Working Paper 19951*.
- BLOOM, N., M. SCHANKERMAN, AND J. VAN REENEN (2013): "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 81, 1347–1393.
- BOLLARD, A., P. KLENOW, AND H. LI (2014): "Entry Costs Rise with Development," *working paper*.
- BOMBARDINI, M., B. LI, AND R. WANG (2018): "Import Competition and Innovation: Evidence from China," *working paper*.
- BRODA, C. AND D. WEINSTEIN (2006): "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121, 541–585.
- BUERA, F. J. AND E. OBERFIELD (2019): "The Global Diffusion of Ideas," *working paper*.
- BURSTEIN, A. AND M. MELITZ (2013): "Trade Liberalization and Firm Dynamics," in *Advances in Economics and Econometrics, Tenth World Congress*, ed. by D. Acemoglu, M. Arellano, and E. Dekel, Cambridge University Press, vol. II of *Econometric Society Monographs*.
- BUSTOS, P. (2011): "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of Mercosur on Argentinian Firms," *The American Economic Review*, 101, 304–340.
- CAVENAILE, L., P. ROLDAN-BLANCO, AND T. SCHMITZ (2020): "International Trade and Innovation Dynamics with Endogenous Markups," *working paper*.
- CHANEY, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98, 1707–1721.
- DAVID, J. M. (2017): "The Aggregate Implications of Mergers and Acquisitions," *working paper*.
- DHINGRA, S. AND J. MORROW (2019): "Monopolistic competition and optimum product diversity under firm heterogeneity," *Journal of Political Economy*, 127, 196–232.
- DIXIT, A. K. AND R. S. PINDYCK (1994): *Investment Under Uncertainty*, Princeton.
- DOLLAR, D. AND A. KRAAY (2004): "Trade, Growth, and Poverty," *The Economic Journal*, 114, F22–F49.
- DUNNE, T., M. J. ROBERTS, AND L. SAMUELSON (1988): "Patterns of firm entry and exit in US manufacturing industries," *The RAND journal of Economics*, 495–515.

- EATON, J. AND S. KORTUM (2001): "Technology, Trade, and Growth: A Unified Framework," *European Economic Review: Papers and Proceedings*, 45, 742–755.
- EATON, J., S. KORTUM, AND F. KRAMARZ (2011): "An Anatomy of International Trade: Evidence from French Firms," *Econometrica*, 79, 1453–1498.
- FIELER, A. C. AND A. HARRISON (2019): "Escaping Import Competition in China," *working paper*.
- FOSTER, L., J. HALTIWANGER, AND C. SYVERSON (2008): "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98, 394–425.
- FRANKEL, J. AND D. ROMER (1999): "Does Trade Cause Growth?" *American Economic Review*, 89, 379–399.
- GABAIX, X. (2009): "Power Laws in Economics and Finance," *Annual Review of Economics*, 1, 255–294.
- GARCIA-MACIA, D., C.-T. HSIEH, AND P. J. KLENOW (2019): "How Destructive is Innovation?" *Econometrica*, forthcoming.
- GOLDBERG, P. K., A. K. KHANDELWAL, N. PAVCNIK, AND P. TOPALOVA (2010): "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India," *The Quarterly Journal of Economics*, 125, 1727–1767.
- GROSSMAN, G. AND E. HELPMAN (1993): *Innovation and Growth in the Global Environment*, MIT press.
- GROSSMAN, G. M. AND E. HELPMAN (1991): "Quality Ladders in the Theory of Growth," *The Review of Economic Studies*, 58, 43–61.
- HEAD, K. AND T. MAYER (2014): "Gravity equations: Workhorse, toolkit, and cookbook," in *Handbook of international economics*, Elsevier, vol. 4, 131–195.
- HOLMES, T. AND J. SCHMITZ (2010): "Competition and Productivity: A Review of Evidence," *Annual Reviews of Economics*, 2, 619–642.
- HOLMES, T. J., D. K. LEVINE, AND J. A. SCHMITZ JR (2012): "Monopoly and the Incentive to Innovate When Adoption Involves Switchover Disruptions," *American Economic Journal: Microeconomics*, 4, 1–33.
- HOPENHAYN, H. A. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60, 1127–50.
- HORNSTEIN, A., P. KRUSELL, AND G. VIOLANTE (2011): "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *American Economic Review*, 101, 2873–2898.
- HSIEH, C.-T., P. J. KLENOW, AND I. NATH (2019): "A Global View of Creative Destruction," *working paper*.
- HURST, E. AND B. W. PUGSLEY (2011): "What Do Small Businesses Do?" *Brookings Papers on Economic Activity*, 2011, 73–118.
- JONES, C. I. (2005): "The Shape of Production Functions and the Direction of Technical Change," *The Quarterly Journal of Economics*, 120, 517–549.
- JONES, C. I. AND J. WILLIAMS (1998): "Measuring the Social Return to R&D," *Quarterly Journal of Economics*, 113, 1119–1135.
- (2000): "Too Much of a Good Thing? The Economics of Investment in R&D," *Journal of Economic Growth*, 5, 65–85.

- KLETTE, T. J. AND S. KORTUM (2004): "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 112, 986–1018.
- KORTUM, S. S. (1997): "Research, Patenting, and Technological Change," *Econometrica*, 65, 1389–1420.
- KRUGMAN, P. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 70, 950–959.
- LEIBENSTEIN, H. (1966): "Allocative Efficiency vs. "X-efficiency"," *The American Economic Review*, 56, 392–415.
- LUCAS, R. E. (1978): "On the Size Distribution of Business Firms," *The Bell Journal of Economics*, 9, 508–523.
- LUCAS, R. E. AND B. MOLL (2014): "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 122, 1–51.
- LUTTMER, E. G. J. (2007): "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics*, 122, 1103–1144.
- (2012): "Technology Diffusion and Growth," *Journal of Economic Theory*, 147, 602–622.
- MARIN, A. G. AND N. VOIGTLÄNDER (2019): "Exporting and Plant-Level Efficiency Gains: It's in the Measure," *Journal of Political Economy*, forthcoming.
- MCCALL, J. J. (1970): "Economics of Information and Job Search," *The Quarterly Journal of Economics*, 84, 113–126.
- MELITZ, M. AND J. COSTANTINI (2007): "The Dynamics of Firm-Level Adjustment to Trade Liberalization," in *The Organization of Firms in a Global Economy*, ed. by E. Helpman, D. Marin, and T. Verdier, Harvard University Press.
- MELITZ, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695 – 1725.
- MUENDLER, M.-A. (2004): "Trade, Technology, and Productivity: A Study of Brazilian Manufacturers 1986-1998," *CESifo Working Paper 1148*.
- PAVCNIK, N. (2002): "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *The Review of Economic Studies*, 69, 245–276.
- PERLA, J. AND C. TONETTI (2014): "Equilibrium Imitation and Growth," *Journal of Political Economy*, 122, 52–76.
- PESKIR, G. AND A. N. SHIRYAEV (2006): *Optimal Stopping and Free-Boundary Problems*, Lectures in Mathematics. ETH Zürich, Birkhäuser Basel.
- RAITH, M. (2003): "Competition, Risk, and Managerial Incentives," *American Economic Review*, 93, 1425–1436.
- RIVERA-BATIZ, L. A. AND P. M. ROMER (1991): "Economic Integration and Endogenous Growth," *The Quarterly Journal of Economics*, 106, 531–555.
- ROMER, P. M. (1994): "The Origins of Endogenous Growth," *The Journal of Economic Perspectives*, 8, 3–22.
- SAMPSON, T. (2016): "Dynamic Selection: An Idea Flows Theory of Entry, Trade and Growth," *Quarterly Journal of Economics*, 131, 315–380.
- SCHMIDT, K. M. (1997): "Managerial incentives and product market competition," *The Review of Economic Studies*, 64, 191–213.

- SHU, P. AND C. STEINWENDER (2019): "The Impact of Trade Liberalization on Firm Productivity and Innovation," in *Innovation Policy and the Economy*, ed. by J. Lerner and S. Stern, University of Chicago Press, vol. 19 of *NBER Book Series*, 39–68.
- SIMONOVSKA, I. AND M. E. WAUGH (2014): "Trade Models, Trade Elasticities, and the Gains from Trade," *NBER Working Paper 20495*.
- SIVADASAN, J. (2009): "Barriers to Competition and Productivity: Evidence from India," *The BE Journal of Economic Analysis & Policy*, 9.
- STOKEY, N. (2009): *The Economics of Inaction*, Princeton Press.
- SYVERSON, C. (2011): "What Determines Productivity?" *Journal of Economic Literature*, 49, 326–365.
- TOPALOVA, P. AND A. KHANDELWAL (2011): "Trade Liberalization and Firm Productivity: The Case of India," *Review of Economics and Statistics*, 93, 995–1009.
- U.S. CENSUS BUREAU (2011): "Synthetic LBD Beta Version 2.0," [computer file], U.S. Census Bureau and Cornell University, Synthetic Data Server [distributor], Washington,DC and Ithaca, NY, USA.
- WAUGH, M. E. (2010): "International Trade and Income Differences," *American Economic Review*, 100, 2093–2124.
- WAUGH, M. E. AND B. RAVIKUMAR (2016): "Measuring Openness to Trade," *Journal of Economic Dynamics and Control*, 72, 29–41.

Appendix (For Online Publication)

A. Environment and Optimization Problems

To demonstrate the extensibility of the model, in this appendix we derive conditions for a version of the model in which there is CRRA power utility, a weakly positive probability of death δ , costs of adoption and entry can be a convex combination of labor and goods, and exogenous shocks to TFP that follow a geometric Brownian motion with drift. For expositional clarity, the body of the paper studies the special case of log utility, no exogenous productivity shocks, costs of adoption and entry in labor only, and studies the limiting economy as the death rate, and thus the BGP equilibrium entry rate, is zero.

A.1. Consumer Preferences and Labor Supply

All countries are symmetric. In each country there exists a representative consumer of measure \bar{L} . The utility of the consumer is given by a constant relative risk aversion (CRRA) function in final goods consumption (C), given an inelastic supply of labor (\bar{L}). The coefficient of relative risk aversion is $\gamma \geq 0$, and the time discount rate is ρ .

Final goods are produced through CES aggregation of an endogenous number of intermediate varieties, including those produced domestically and those imported from abroad. There is an endogenous measure $\Omega(t)$ of firms operating in each country. The flow of intermediate firms adopting a new technology is $\Omega(t)S(t)$ and the flow of firms entering the market and creating a new variety is $\Omega(t)E(t)$. The consumer purchases the final consumption good, invests in technology adoption with a real cost of $X(t)$ per upgrading intermediate firm, and invests in firm entry with a real cost of $X(t)/\chi$.²⁸ Consumers income consists of labor earnings paid at wage $W(t)$ and profits from their ownership of the domestic firms. Aggregate profits from selling domestically are $\bar{\Pi}_d(t)$ and aggregate profits from exporting to the $N - 1$ foreign countries is $(N - 1)\bar{\Pi}_x(t)$. Thus, welfare at time \tilde{t} is

$$\begin{aligned} \bar{U}(\tilde{t}) &= \int_{\tilde{t}}^{\infty} U(C(t))e^{-\rho(t-\tilde{t})} dt \\ \text{s.t. } C(t) + \Omega(t)X(t)(S(t) + E(t)/\chi) &= \frac{W(t)}{P(t)}\bar{L} + \bar{\Pi}_d(t) + (N - 1)\bar{\Pi}_x(t). \end{aligned} \quad (\text{A.1})$$

With this, the period real profits as $\bar{\Pi}_i(t) = \bar{\Pi}_d(t) + (N - 1)\bar{\Pi}_x(t)$ and real investment as $\bar{I}_i(t) = \Omega(t)X(t)(S(t) + E(t)/\chi)$.

A.2. The Static Firm Problem

Intermediate Goods Demand. There is a measure $\Omega(t)$ of intermediate firms in each country that are monopolistically competitive, and the final goods sector is perfectly competitive. The final goods sector takes prices as given and aggregates intermediate goods with a CES production function, with $\sigma > 1$ the elasticity of substitution between all available products.

²⁸Firms are maximizing real profits, discounting using the interest rate determined by the consumer's marginal rate of substitution plus the death rate. Hence, the investment choice of the consumer and firm is aligned, and consumers will finance upgrades to their existing firms through equity financing. As consumers own a perfectly diversified portfolio of domestic firms, they are only diluting their own equity with this financing method.

The CDF of the productivity distribution at time t is $\Phi(Z, t)$, normalized such that $\Phi(\infty, t) = 1$ for all t . Therefore, the total measure of firms with productivity below Z at time t is $\Omega(t)\Phi(Z, t)$.

Drop the t subscript for clarity. The standard solutions follow from maximizing the following final goods production problem,

$$\max_{Q_d, Q_x} \left[\Omega \int_M^\infty Q_d(Z)^{(\sigma-1)/\sigma} d\Phi(Z) + (N-1)\Omega \int_{\dot{Z}}^\infty Q_x(Z)^{(\sigma-1)/\sigma} d\Phi(Z) \right]^{\sigma/(\sigma-1)} - \quad (\text{A.2})$$

$$\Omega \int_M^\infty p_d(Z)Q_d(Z)d\Phi(Z) + (N-1)\Omega \int_{\dot{Z}}^\infty p_x(Z)Q_x(Z)d\Phi(Z)$$

Since the final good market is competitive, nominal aggregate income is

$$Y = \Omega \int_M^\infty p_d(Z)Q_d(Z)d\Phi(Z) + (N-1)\Omega \int_{\dot{Z}}^\infty p_x(Z)Q_x(Z)d\Phi(Z). \quad (\text{A.3})$$

Defining a price index P , the demand for each intermediate product is,

$$Q_d(Z) = \left(\frac{p_d(Z)}{P} \right)^{-\sigma} \frac{Y}{P}, \quad Q_x(Z) = \left(\frac{p_x(Z)}{P} \right)^{-\sigma} \frac{Y}{P} \quad (\text{A.4})$$

$$P^{1-\sigma} = \Omega \left(\int_M^\infty p_d(Z)^{1-\sigma} d\Phi(Z) + (N-1) \int_{\dot{Z}}^\infty p_x(Z)^{1-\sigma} d\Phi(Z) \right). \quad (\text{A.5})$$

Static Profits. A monopolist operating domestically chooses each instant prices and labor demand to maximize profits, subject to the demand function given in equation (A.4),

$$P\Pi_d(Z) := \max_{p_d, \ell_d} \{ (p_d Z \ell_d - W \ell_d) \} \text{ s.t. equation (A.4)}. \quad (\text{A.6})$$

Where $\Pi_d(Z)$ is the real profits from domestic production.

Firms face a fixed cost of exporting, $\kappa \geq 0$. To export, a firm must hire labor in the foreign country to gain access to foreign consumers. This fixed cost is paid in market wages, and is proportional to the number of consumers accessed. Additionally, exports are subject to a variable iceberg trade cost, $d \geq 1$, so that firm profits from exporting to a single country (i.e., export profits per market) are

$$P\Pi_x(Z) := \max_{p_x, \ell_x} \{ (p_x \frac{Z}{d} \ell_x - W \ell_x - \bar{L}\kappa W) \} \text{ s.t. equation (A.4)}. \quad (\text{A.7})$$

Optimal firm policies consist of $p_d(Z)$, $p_x(Z)$, $\ell_d(Z)$, and $\ell_x(Z)$ and determine $\Pi_d(Z)$ and $\Pi_x(Z)$. As is standard, it is optimal for firms to charge a constant markup over marginal cost, $\bar{\sigma} := (\sigma/\sigma - 1)$.

$$p_d(Z) = \bar{\sigma} \frac{W}{Z}, \quad (\text{A.8})$$

$$p_x(Z) = \bar{\sigma} d \frac{W}{Z}, \quad (\text{A.9})$$

$$\ell_d(Z) = \frac{Q_d(Z)}{Z}, \quad \ell_x(Z) = d \frac{Q_x(Z)}{Z}. \quad (\text{A.10})$$

To derive firm profits, take equation (A.6) and divide by P to get $\Pi_d(Z) = \frac{p_d(Z)}{P} Q_d(Z) - \frac{Q_d(Z)}{Z} \frac{W}{P}$. Substitute from equation (A.8), as $W = \frac{Z p_d(Z)}{\bar{\sigma}}$, to yield $\Pi_d(Z) = \frac{1}{\sigma} \frac{p_d(Z)}{P} Q_d(Z)$. Finally, use $Q_d(Z)$ from equation (A.4) to show

$$\Pi_d(Z) = \frac{1}{\sigma} \left(\frac{p_d(Z)}{P} \right)^{1-\sigma} \frac{Y}{P}. \quad (\text{A.11})$$

Using similar techniques, export profits per market are

$$\Pi_x(Z) = \max \left\{ 0, \frac{1}{\sigma} \left(\frac{p_x(Z)}{P} \right)^{1-\sigma} \frac{Y}{P} - \bar{L} \kappa \frac{W}{P} \right\}. \quad (\text{A.12})$$

Since there is a fixed cost to export, only firms with sufficiently high productivity will find it profitable to export. Solving equation (A.12) for the productivity that earns zero profits gives the export productivity threshold. That is, a firm will export iff $Z \geq \hat{Z}$, where \hat{Z} satisfies

$$\left(\frac{p_x(\hat{Z})}{P} \right)^{1-\sigma} = \sigma \bar{L} \kappa \frac{W}{Y}, \quad (\text{A.13})$$

$$\hat{Z} = \bar{\sigma} d (\sigma \bar{L} \kappa)^{\frac{1}{\sigma-1}} \left(\frac{W}{P} \right) \left(\frac{W}{Y} \right)^{\frac{1}{\sigma-1}}. \quad (\text{A.14})$$

Let aggregate profits from domestic production be $\bar{\Pi}_d$ and aggregate export profits per market be $\bar{\Pi}_x$.

$$\bar{\Pi}_d := \Omega \int_M^\infty \Pi_d(Z) d\Phi(Z). \quad (\text{A.15})$$

$$\bar{\Pi}_x := \Omega \int_{\hat{Z}}^\infty \Pi_x(Z) d\Phi(Z). \quad (\text{A.16})$$

The trade share for a particular market, λ , is

$$\lambda = \Omega \int_{\hat{Z}}^\infty \left(\frac{p_x(Z)}{P} \right)^{1-\sigma} d\Phi(Z). \quad (\text{A.17})$$

A.3. Firms Dynamic Problem

Stochastic Process for Productivity Assume that operating firms have (potentially) stochastic productivity following the stochastic differential equation for geometric Brownian motion (GBM),

$$dZ_t/Z_t = (\mu + v^2/2)dt + v dW_t, \quad (\text{A.18})$$

where $\mu \geq 0$ is related to the drift of the productivity process, $v \geq 0$ is the volatility, and W_t is standard Brownian motion.

At any instant in time, a firm will exit if hit by a death shock, which follows a Poisson process with arrival rate $\delta \geq 0$. Thus, all firms have the same probability of exiting and the probability of exiting is independent of time.

The case of $\mu = v = \delta = 0$, is the baseline model studied in the body of the paper.

Firm's Problem. Define a firm's total real profits as

$$\Pi(Z, t) := \Pi_d(Z, t) + (N - 1)\Pi_x(Z, t), \quad (\text{A.19})$$

where from equation (A.12), $\Pi_x(Z, t) = 0$ for firms who do not export. Let $V(Z, t)$ be the value of a firm with productivity Z at time t . The firm's effective discount rate, $r(t)$, is the sum of the consumer's intertemporal marginal rate of substitution and the firm death rate δ .

Given the standard Bellman equation for the GBM of equation (A.18) and optimal static policies,

$$r(t)V(Z, t) = \Pi(Z, t) + \left(\mu + \frac{v^2}{2} \right) Z \frac{\partial V(Z, t)}{\partial Z} + \frac{v^2}{2} Z^2 \frac{\partial^2 V(Z, t)}{\partial Z^2} + \frac{\partial V(Z, t)}{\partial t}, \quad (\text{A.20})$$

$$V(M(t), t) = \int_{M(t)}^{\infty} V(Z, t) d\Phi(Z, t) - X(t), \quad (\text{A.21})$$

$$\frac{\partial V(M(t), t)}{\partial Z} = 0. \quad (\text{A.22})$$

Equation (A.20) is the bellman equation for a firm continuing to produce with its existing technology. It receives instantaneous profits and the value of a firm where productivity may change over time. Equation (A.21) is the value matching condition, which states that the marginal adopter must be indifferent between adopting and not adopting. $M(t)$ is the endogenous productivity threshold that defines the marginal firm. Equation (A.22) is the smooth-pasting condition.²⁹

A.4. Adoption Costs

In order to upgrade its technology a firm must buy some goods and hire some labor. These costs are in proportion to the population size, reflecting market access costs as in Arkolakis (2010) or Arkolakis (2015). Whether costs are in terms of goods or labor is a common issue in growth models, with many papers specifying goods costs and many specifying labor costs.³⁰ The growth literature often uses labor costs, since new ideas cannot simply be purchased, and instead must be the result of innovators doing R&D. In the technology diffusion context, in which low productivity firms are adopting already existing technologies, an adoption cost that has some nontrivial component denominated in goods is more reasonable than in the innovation case. Since there is a paucity of empirical evidence to guide our decision in

²⁹Using the standard relationship between free boundary and optimal stopping time problems, the firm's problem could equivalently be written as the firm choosing a stopping time at which it would upgrade. If $v > 0$ this stopping time is a random variable, otherwise it is deterministic.

³⁰As discussed in Bollard, Klenow, and Li (2014), for examples see Hopenhayn (1992), Romer (1994), Foster, Haltiwanger, and Syverson (2008) and David (2017) for papers using output costs and Lucas (1978), Grossman and Helpman (1991), Melitz (2003), Klette and Kortum (2004), Luttmer (2007), and Acemoglu et al. (2018) for papers that use labor costs.

the adoption context, we model the adoption cost in a way that nests the costs being in labor exclusively, in goods exclusively, or as a mix of labor and goods. Although we solve the model for this general case, our baseline is that costs are purely labor denominated.

The amount of labor needed is parameterized by ζ , which is constant. The labor component of the adoption cost, however, increases in equilibrium in proportion to the real wage, ensuring the cost does not become increasingly small as the economy grows. The amount of goods that needs to be purchased to adopt a technology increases with the scale of the economy—otherwise the relative costs of goods would become infinitesimal in the long-run. Θ parameterizes the amount of goods required to adopt a technology, with the cost, $M(t)\Theta$, growing as the economy grows. Essentially, ζ controls the overall cost of technology adoption, while $\eta \in [0, 1]$ controls how much of the costs are to hire labor versus buy goods. We model the mix of labor and goods as additive in order to permit a balanced growth path equilibrium.

The real cost of adopting a technology is

$$X(t) := \bar{L}\zeta \left[(1 - \eta) \frac{W(t)}{P(t)} + \eta M(t)\Theta \right]. \quad (\text{A.23})$$

A.5. Entry and Exit

There is a large pool of non-active firms that may enter the economy by paying an entry cost—equity financed by the representative consumer—to gain a draw of an initial productivity from the same distribution from which adopters draw. Since entry and adoption deliver similar gains, we model the cost of entry as a multiple of the adoption cost for incumbents, $X(t)/\chi$, where $0 < \chi < 1$. Hence, χ is the ratio of adoption to entry costs and $\chi \in (0, 1)$ reflects that incumbents have a lower cost of upgrading to a better technology than entrants have to start producing a new variety from scratch.

Thus, the free entry condition that equates the cost of entry to the value of entry is

$$X(t)/\chi = \int_{M(t)}^{\infty} V(Z, t) d\Phi(Z, t). \quad (\text{A.24})$$

If a flow $E(t)$ of firms enter, and a flow δ exit, then the differential equation for the number of firms is $\Omega'(t) = (E(t) - \delta)\Omega(t)$. Since we study a stationary equilibrium, on a BGP Ω will be constant and determined by free entry, and $E(t) = \delta$ for all t .

The costs of entry will determine the number of varieties in equilibrium, and for $\delta > 0$, there is gross entry on a balanced growth path. This model of entry and exit is very different from those in Luttmer (2007) and Sampson (2016). Here, exit is exogenous, whereas a key model mechanism studied in those papers is the endogenous selection of exit induced by fixed costs of operations. We have not modeled a fixed cost to domestic production in order to isolate our distinct mechanism. We have introduced entry and exit in our model to generate an endogenous number of varieties so that we can analyze the effect of our mechanism on welfare, taking into account changes in incumbent technology adoption behavior and changes on the extensive margin in the number of varieties produced. Given the exogenous death

shock, the effect of $\delta > 0$ is only to change the firm's discount rate. For the most part, the economics are qualitatively identical to the $\delta = 0$ case and there is no discontinuity in the limit as $\delta \rightarrow 0$.³¹

A.6. Market Clearing and the Resource Constraint

Total labor demand is the sum of labor used for domestic production, export production, the fixed cost of exporting, technology adoption, and entry. Equating labor supply and demand yields

$$\bar{L} = \Omega \int_M^\infty \ell_d(Z) d\Phi(Z) + (N-1)\Omega \int_{\hat{Z}}^\infty \ell_x(Z) d\Phi(Z) + (N-1)\Omega(1 - \Phi(\hat{Z}))\kappa\bar{L} + \bar{L}(1 - \eta)\zeta\Omega(S + \delta/\chi). \quad (\text{A.25})$$

The quantity of final goods must equal the sum of consumption and investment in technology adoption. Thus, the resource constraint is

$$\frac{Y}{P} = C + \Omega\bar{L}\eta\zeta M\Theta(S + \delta/\chi). \quad (\text{A.26})$$

B. Deriving the Productivity Distribution Law of Motion and Flow of Adopters

This section describes the details of deriving the law of motion for the productivity distribution.

The Productivity Distribution Law of Motion. At points of continuity of $M(t)$, there exists a flow of adopters during each infinitesimal time period.³² The Kolmogorov Forward Equation (KFE) for $Z > M(t)$, describes the evolution of the CDF. The KFE contains standard components accounting for the drift and Brownian motion of the exogenous GBM process detailed in equation (A.18). Furthermore, it includes the flow of adopters (source) times the density they draw from (redistribution CDF). Determining the flow of adopters is the fact that the adoption boundary $M(t)$ sweeps across the density from below at rate $M'(t)$. As adoption boundary acts as an absorbing barrier, and as it sweeps from below it collects $\phi(M(t), t)$ amount firms. The cdf that the flow of adopters is redistributed into is determined by two features of the environment. In the stationary equilibrium, $M(t)$ is the minimum of support of $\Phi(Z, t)$, so the adopters are redistributed across the entire support of $\phi(Z, t)$.³³ Since adopters draw directly from the productivity density, they are redistributed throughout the distribution in proportion to the density. Thus, the flow of adopters $S(t)$ multiplies the cdf, $\Phi(Z, t)$.

³¹For special cases where $\delta = 0$ and the initial Ω is large relative to that which would be achieved on a BGP from a relatively small initial Ω , the free entry condition could hold as an inequality (i.e. $X(t)/\chi > \int_{M(t)}^\infty V(Z, t) d\Phi(Z)$). In that case, the lack of exit would prevent Ω from decreasing so that the free entry condition held with equality. In our baseline with $\delta = 0$, we ignore this special one-sided case, as it is economically uninteresting; it is unreasonable for the number of varieties to require major decreases in a growing economy. Also, as there is no discontinuity in the solution when taking $\delta \rightarrow 0$, we will consider our baseline economy a small δ approximation.

³²The evolution of $M(t)$, and hence the distribution itself, can only be discontinuous at time 0 or in response to unanticipated shocks. Since this paper analyzes balanced growth path equilibria, we omit derivation of the law of motion for the distribution with discontinuous $M(t)$.

³³Unlike in discrete time, the distinction between drawing from the unconditional distribution or the conditional distribution of non-adopting incumbents is irrelevant. The number of adopting firms is a flow, and hence measure 0, which leads to identical conditional vs. unconditional distributions.

Since there is a constant death rate, $\delta \geq 0$, a normalized measure of $\delta\Phi(Z, t)$ exit with productivity below Z , but as new entrants of normalized flow $E(t)$ adopt a productivity through the same process as incumbents, they are added to the flow entering with measure below Z .³⁴

$$\begin{aligned} \frac{\partial\Phi(Z, t)}{\partial t} = & \underbrace{\Phi(Z, t) \left(\underbrace{S(t) + E(t)}_{\text{Adopt or Enter}} \right)}_{\text{Distributed below } Z} - \underbrace{S(t)}_{\text{Adopt at } M(t)} - \underbrace{\delta\Phi(Z, t)}_{\text{Death}} \\ & - \underbrace{\left(\mu - \frac{v^2}{2} \right) Z \frac{\partial\Phi(Z, t)}{\partial Z}}_{\text{Deterministic Drift}} + \underbrace{\frac{v^2}{2} Z^2 \frac{\partial^2\Phi(Z, t)}{\partial Z^2}}_{\text{Brownian Motion}}. \end{aligned} \quad (\text{B.3})$$

In the baseline $\delta = \mu = v = 0$ case, this simplifies to equation (17).³⁵

Normalized Productivity Distribution. Define the change of variables $z := Z/M(t)$, $g(t) := M'(t)/M(t)$, and

$$\Phi(Z, t) =: F(Z/M(t), t). \quad (\text{B.4})$$

Differentiating,

$$\phi(Z, t) = \frac{1}{M(t)} f(Z/M(t), t). \quad (\text{B.5})$$

This normalization generates an adoption threshold that is stationary at $z = M(t)/M(t) = 1$ for all t .

Law of Motion for the Normalized Distribution. To characterize the normalized KFE, first differentiate the cdf with respect to t , yielding

$$\frac{\partial\Phi(Z, t)}{\partial t} = \frac{\partial F(Z/M(t), t)}{\partial t} - \frac{Z}{M(t)} \frac{M'(t)}{M(t)} \frac{\partial F(Z/M(t), t)}{\partial z}. \quad (\text{B.6})$$

³⁴To derive from the more common KFE written in PDFs: use the standard KFE for the pdf $\phi(Z, t)$, integrate this with respect to Z to convert into cdf $\Phi(Z, t)$, use the fundamental theorem of calculus on all terms, then the chain rule on the last term, and rearrange,

$$\frac{\partial\phi(Z, t)}{\partial t} = -\frac{\partial}{\partial Z} \left[\left(\mu + \frac{v^2}{2} \right) Z \phi(Z, t) \right] + \frac{\partial}{\partial Z^2} \left[\frac{v^2}{2} Z^2 \phi(Z, t) \right] + \dots \quad (\text{B.1})$$

$$\frac{\partial\Phi(Z, t)}{\partial t} = \left(\frac{v^2}{2} - \mu \right) Z \frac{\partial\Phi(Z, t)}{\partial Z} + \frac{v^2}{2} Z^2 \frac{\partial^2\Phi(Z, t)}{\partial Z^2} + \dots \quad (\text{B.2})$$

³⁵While *conditional on an optimal policy* the law of motion in equation (17) and the more general equation (16) are identical to that in Luttmer (2007), this is a mechanical result of any distribution evolution with resetting of agents through direct sampling of the distribution. Economically, the forces which determine the endogenous policy are completely different, as we concentrate on the choices of incumbents rather than entrants/exit. Some of the key differences are evident in the connection to search models as discussed in Section 5.

Differentiating the cdf with respect to Z yields

$$\frac{\partial \Phi(Z, t)}{\partial Z} = \frac{1}{M(t)} \frac{\partial F(Z/M(t), t)}{\partial z}, \quad (\text{B.7})$$

$$\frac{\partial^2 \Phi(Z, t)}{\partial Z^2} = \frac{1}{M(t)^2} \frac{\partial^2 F(Z/M(t), t)}{\partial z^2}. \quad (\text{B.8})$$

Given that $z := \frac{Z}{M(t)}$ and $g(t) := M'(t)/M(t)$, combining equations (B.3), (B.6), (B.7), and (B.8) provides the KFE in cdfs of the normalized distribution:

$$\frac{\partial F(z, t)}{\partial t} = (S(t) + E(t) - \delta) F(z, t) + (g(t) - \mu + v^2/2) z \frac{\partial F(z, t)}{\partial z} + \frac{v^2}{2} z^2 \frac{\partial^2 F(z, t)}{\partial z^2} - S(t). \quad (\text{B.9})$$

The interpretation of this KFE is that while non-adopting incumbent firms are on average not moving in absolute terms, they are moving backwards at rate $g(t)$ relative to $M(t)$ (adjusted for the growth rate of the stochastic process). As the minimum of support is $z = M(t)/M(t) = 1$ for all t , a necessary condition is that $F(1, t) = 0$ for all t , and therefore $\frac{\partial F(1, t)}{\partial t} = 0$. Thus, evaluating equation (B.9) at $z = 1$ gives an expression for $S(t)$,³⁶

$$S(t) = \left(g(t) - \mu + \frac{v^2}{2} \right) \frac{\partial F(1, t)}{\partial z} + \frac{v^2}{2} \frac{\partial^2 F(1, t)}{\partial z^2}. \quad (\text{B.10})$$

This expression includes adopters caught by the boundary moving relative to their drift, as well as the flux from the GBM pushing some of them over the endogenously determined threshold. If $\mu = v = \delta = 0$, then a truncation at $M(t)$ solves equation (B.9) for any t and for any initial condition, as in Perla and Tonetti (2014).

$$\phi(Z, t) = \frac{\phi(Z, 0)}{1 - \Phi(M(t), 0)}, \quad Z \geq M(t). \quad (\text{B.11})$$

The only non-degenerate stationary $F(z)$ consistent with equation (B.11) is given by equation (B.13)—the same form as that with $v > 0$.

The Stationary Normalized Productivity Distribution. From equation (B.9), the stationary KFE is

$$0 = SF(z) + \left(g - \mu + \frac{v^2}{2} \right) z F'(z) + \frac{v^2}{2} z^2 F''(z) - S. \quad (\text{B.12})$$

subject to $F(1) = 0$ and $F(\infty) = 1$.

Moreover, for any strictly positive $v > 0$, the KFE will asymptotically generate a stationary Pareto distribution for some tail parameter from any initial condition. While many $\theta > 1$ could solve this differential equation, the particular θ tail parameter is determined by initial conditions and the evolution of $M(t)$. As in Luttmer (2007) and Gabaix (2009), the geometric random shocks leads to an endogenously determined

³⁶Equivalently, the flow of adopters can be derived as the net flow of the probability current through the adoption threshold.

power-law distribution.

$$F(z) = 1 - z^{-\theta}, \quad z \geq 1. \quad (\text{B.13})$$

From equations B.12 and B.13, evaluating at $z = 1$ for a given g and θ ,

$$S = \theta \left(g - \mu - \theta \frac{v^2}{2} \right) \quad (\text{B.14})$$

Therefore, given an equilibrium g and θ , the CDF in equation (B.13) and S from equation (B.14) characterize the stationary distribution. The relationship between g and θ is determined by the firms' decisions given S and $F(z)$. It is independent of δ on a BGP since the exit rate is constant and uniform across firms (in contrast to Luttmer (2007), where selection into exit is generated by fixed costs of operations and is not independent of firm productivity).

The Stationary Distribution with no GBM. For $v = 0$, the lack of random shocks means that the stationary distribution will not necessarily become a Pareto distribution from arbitrary initial conditions. However, if the initial distribution is Pareto, the normalized distribution will be constant. If the initial distribution is a power-law, then the stationary distribution is asymptotically Pareto.

Since a Pareto distribution is attained for any $v > 0$, we consider our baseline ($v = 0, \mu = 0$) case with an initial Pareto distribution as a small noise limit of the full model with GBM from some arbitrary initial condition. Beyond endogenously determining the tail index and changing the expected time to execute the adoption option, exogenous productivity volatility of incumbent firms has qualitatively little impact on the model. For the baseline case, from equation (B.14), the flow of adopters is $S = \theta g$.

C. Normalized Static Equilibrium Conditions

To aid in computing a balanced growth path equilibrium, in this section we transform the problem and derive normalized static equilibrium conditions.

Expectations using the Normalized Distribution. If an integral of the following form exists, for some unary function $\Psi(\cdot)$, substitute for $f(z)$ from B.5, then do a change of variables of $z = \frac{Z}{M}$ to obtain a useful transformation of the integral.

$$\int_M^\infty \Psi\left(\frac{Z}{M}\right) \phi(Z) dZ = \int_M^\infty \Psi\left(\frac{Z}{M}\right) f\left(\frac{Z}{M}\right) \frac{1}{M} dZ = \int_{M/M}^\infty \Psi(z) f(z) dz. \quad (\text{C.1})$$

The key to this transformation is that moving from ϕ to f introduces a $1/M$ term. Thus, abusing notation by using an expectation of the normalized variable,

$$\int_M^\infty \Psi\left(\frac{Z}{M}\right) \phi(Z) dZ = \mathbb{E}[\Psi(z)]. \quad (\text{C.2})$$

C.1. Normalizing the Static Equilibrium

Define the following normalized, real, per-capita values:

$$\hat{z} := \frac{\hat{Z}}{\bar{M}} \quad (\text{C.3})$$

$$y := \frac{Y}{\bar{LMP}} \quad (\text{C.4})$$

$$c := \frac{C}{\bar{LM}} \quad (\text{C.5})$$

$$q_d(Z) := \frac{Q_d(Z)}{\bar{LM}} \quad (\text{C.6})$$

$$x := \frac{X}{\bar{LM}w} \quad (\text{C.7})$$

$$w := \frac{W}{MP} \quad (\text{C.8})$$

$$\pi_d(Z) := \frac{\Pi_d(Z)}{\bar{LM}w}. \quad (\text{C.9})$$

In order to simplify algebra and notation, we normalize profits and adoption costs relative to real, normalized wages. This, for example, means the normalized cost of adoption $x = \zeta$.

Combining the normalized variables with equations (A.8) and (A.9) provides the real prices in terms of real, normalized wages.

$$\frac{p_d(Z)}{P} = \bar{\sigma} \frac{w}{Z/M}, \quad (\text{C.10})$$

$$\frac{p_x(Z)}{P} = \bar{\sigma} d \frac{w}{Z/M}. \quad (\text{C.11})$$

Substituting equations (C.10) and (C.11) into equation (A.4) and dividing by \bar{LM} yields normalized quantities,

$$q_d(Z) = \bar{\sigma}^{-\sigma} w^{-\sigma} y \left(\frac{Z}{M}\right)^{\sigma}, \quad (\text{C.12})$$

$$q_x(Z) = \bar{\sigma}^{-\sigma} w^{-\sigma} d^{-\sigma} y \left(\frac{Z}{M}\right)^{\sigma}. \quad (\text{C.13})$$

Divide equation (A.10) by \bar{L} , then substitute from equations (A.4) and (C.10). Finally, divide the top and bottom by M to obtain normalized demand for production labor

$$\ell_d(Z)/\bar{L} = \bar{\sigma}^{-\sigma} w^{-\sigma} y \left(\frac{Z}{M}\right)^{\sigma-1}, \quad (\text{C.14})$$

$$\ell_x(Z)/\bar{L} = \bar{\sigma}^{-\sigma} w^{-\sigma} y d^{1-\sigma} \left(\frac{Z}{M}\right)^{\sigma-1}. \quad (\text{C.15})$$

Divide equation (A.5) by $P^{1-\sigma}$ and then substitute from equation (C.10) for $p_d(Z)/P$ to obtain

$$1 = \Omega \bar{\sigma}^{1-\sigma} w^{1-\sigma} \left[\int_M^{\infty} \left(\frac{Z}{M}\right)^{\sigma-1} d\Phi(Z) + (N-1) \int_{\hat{Z}}^{\infty} d^{1-\sigma} \left(\frac{Z}{M}\right)^{\sigma-1} d\Phi(Z) \right]. \quad (\text{C.16})$$

Simplify equation (C.16) by defining \bar{z} , a measure of effective aggregate productivity. Then use equation (C.2) to give normalized real wages in terms of parameters, \hat{z} , and the productivity distribution

$$\bar{z} := \left[\Omega \left(\mathbb{E} [z^{\sigma-1}] + (N-1)(1-F(\hat{z}))d^{1-\sigma} \mathbb{E} [z^{\sigma-1} \mid z > \hat{z}] \right) \right]^{\frac{1}{\sigma-1}}, \quad (\text{C.17})$$

$$w^{\sigma-1} = \bar{\sigma}^{1-\sigma} \bar{z}^{\sigma-1}, \quad (\text{C.18})$$

$$w = \frac{1}{\bar{\sigma}} \bar{z}. \quad (\text{C.19})$$

Note that if $d = 1$ and $\hat{z} = 1$, then $w = \frac{1}{\bar{\sigma}} (\Omega N \mathbb{E} [z^{1-\sigma}])^{1/(\sigma-1)}$. Divide equations (A.11) and (A.12) by $\bar{L}Mw$ and substitute with equation (C.18) to obtain normalized profits,

$$\pi_d(Z) = \frac{1}{\sigma} \left(\frac{p(Z)}{P} \right)^{1-\sigma} \frac{y}{w} = \frac{1}{\bar{\sigma} \bar{z}^{\sigma-1}} \frac{y}{w} \left(\frac{Z}{M} \right)^{\sigma-1}, \quad (\text{C.20})$$

$$\pi_x(Z) = \frac{1}{\bar{\sigma} \bar{z}^{\sigma-1}} \frac{y}{w} d^{1-\sigma} \left(\frac{Z}{M} \right)^{\sigma-1} - \kappa. \quad (\text{C.21})$$

Divide equations (A.15) and (A.16) by $\bar{L}Mw$ and use equations (C.20) and (C.21) to find aggregate profits from domestic production and from exporting to one country,

$$\bar{\pi}_d = \Omega \frac{1}{\bar{\sigma} \bar{z}^{\sigma-1}} \frac{y}{w} \mathbb{E} [z^{\sigma-1}], \quad (\text{C.22})$$

$$\bar{\pi}_x = \Omega \left(\frac{1}{\bar{\sigma} \bar{z}^{\sigma-1}} \frac{y}{w} d^{1-\sigma} (1-F(\hat{z})) \mathbb{E} [z^{\sigma-1} \mid z > \hat{z}] - (1-F(\hat{z}))\kappa \right). \quad (\text{C.23})$$

Divide equation (A.25) by \bar{L} , and aggregate the total labor demand from equations (C.14) and (C.15) to obtain normalized aggregate labor demand

$$1 = \Omega \bar{\sigma}^{-\sigma} w^{-\sigma} y \left(\mathbb{E} [z^{\sigma-1}] + (N-1)(1-F(\hat{z}))d^{1-\sigma} \mathbb{E} [z^{\sigma-1} \mid z > \hat{z}] \right) + \Omega(N-1)(1-F(\hat{z}))\kappa + \Omega(1-\eta)\zeta S \quad (\text{C.24})$$

$$= \bar{\sigma}^{-\sigma} w^{-\sigma} y \bar{z}^{\sigma-1} + \Omega((N-1)(1-F(\hat{z}))\kappa + (1-\eta)\zeta(S + \delta/\chi)). \quad (\text{C.25})$$

Define \tilde{L} as a normalized quantity of labor used outside of variable production. Multiply equation (C.25) by w , and use equation (C.18) to show that

$$\tilde{L} := \Omega[(N-1)(1-F(\hat{z}))\kappa + (1-\eta)\zeta(S + \delta/\chi)], \quad (\text{C.26})$$

$$w = \frac{1}{\bar{\sigma}} y + \tilde{L}w, \quad (\text{C.27})$$

$$1 = \frac{1}{\bar{\sigma}} \frac{y}{w} + \tilde{L}. \quad (\text{C.28})$$

Reorganize to find real output as a function of the productivity distribution and labor supply (net of labor used for the fixed costs of exporting and adopting technology)

$$\frac{y}{w} = \bar{\sigma} (1 - \tilde{L}), \quad (\text{C.29})$$

$$y = (1 - \tilde{L}) \bar{z}. \quad (\text{C.30})$$

This equation lends interpretation to \bar{z} as being related to the aggregate productivity. Substituting equation (C.29) into equations (C.20) and (C.21) to obtain a useful formulation of firm profits

$$\pi_d(Z) = \frac{1-\tilde{L}}{(\sigma-1)\bar{z}^{\sigma-1}} \left(\frac{Z}{M}\right)^{\sigma-1}, \quad (\text{C.31})$$

$$\pi_x(Z) = \frac{1-\tilde{L}}{(\sigma-1)\bar{z}^{\sigma-1}} d^{1-\sigma} \left(\frac{Z}{M}\right)^{\sigma-1} - \kappa. \quad (\text{C.32})$$

Define the common profit multiplier $\bar{\pi}_{\min}$ as

$$\bar{\pi}_{\min} := \frac{1-\tilde{L}}{(\sigma-1)\bar{z}^{\sigma-1}} = \frac{1-\tilde{L}}{(\sigma-1)(\bar{\sigma}w)^{\sigma-1}} = \frac{(\bar{\sigma}w)^{1-\sigma}y}{\sigma w}, \quad (\text{C.33})$$

$$\pi_d(Z) = \bar{\pi}_{\min} \left(\frac{Z}{M}\right)^{\sigma-1}, \quad (\text{C.34})$$

$$\pi_x(Z) = \bar{\pi}_{\min} d^{1-\sigma} \left(\frac{Z}{M}\right)^{\sigma-1} - \kappa. \quad (\text{C.35})$$

Use equation (C.35) set to zero to solve for \hat{z} . This is an implicit equation as $\bar{\pi}_{\min}$ is a function of \hat{z} through \bar{z}

$$\hat{z} = d \left(\frac{\kappa}{\bar{\pi}_{\min}} \right)^{\frac{1}{\sigma-1}}. \quad (\text{C.36})$$

Substitute equations (C.34) and (C.35) into equations (C.22) and (C.23) to obtain a useful formulation for aggregate profits

$$\bar{\pi}_d = \Omega \bar{\pi}_{\min} \mathbb{E} [z^{\sigma-1}], \quad (\text{C.37})$$

$$\bar{\pi}_x = \Omega \left((1 - F(\hat{z})) (\bar{\pi}_{\min} d^{1-\sigma} \mathbb{E} [z^{\sigma-1} \mid z > \hat{z}] - \kappa) \right). \quad (\text{C.38})$$

Combine to calculate aggregate total profits

$$\begin{aligned} \bar{\pi}_d + (N-1)\bar{\pi}_x &= \Omega \bar{\pi}_{\min} \left[\mathbb{E} [z^{\sigma-1}] + (N-1)(1 - F(\hat{z})) d^{1-\sigma} \mathbb{E} [z^{\sigma-1} \mid z > \hat{z}] \right] \\ &\quad - \Omega(N-1)(1 - F(\hat{z}))\kappa. \end{aligned} \quad (\text{C.39})$$

Rewriting aggregate total profits using the definition of \bar{z} yields

$$\bar{\pi}_{\text{agg}} := \bar{\pi}_{\min} \bar{z}^{\sigma-1} - \Omega(N-1)(1 - F(\hat{z}))\kappa. \quad (\text{C.40})$$

Note that in a closed economy, $\bar{z} = (\Omega \mathbb{E} [z^{\sigma-1}])^{1/(\sigma-1)}$ and therefore aggregate profits relative to wage are a markup dependent fraction of normalized output relative to wage $\bar{\pi}_d = \frac{1-\tilde{L}}{\sigma-1}$. Take the resource constraint in equation (A.26) and divide by $M\bar{L}w$ and then use equation (C.29) to get an equation for normalized, per-capita consumption

$$\frac{c}{w} = \frac{y}{w} - \Omega \eta \zeta \Theta (S + \delta/\chi) / w = \bar{\sigma} \left(1 - \tilde{L} \right) - \Omega \eta \zeta \frac{\Theta}{w} (S + \delta/\chi), \quad (\text{C.41})$$

$$c = \left(1 - \tilde{L} \right) \bar{z} - \eta \zeta \Omega \Theta (S + \delta/\chi). \quad (\text{C.42})$$

Normalize the cost in equation (A.23) by dividing by $\bar{LPM}w$. This is implicitly a function of \hat{z} through w

$$x = \zeta (1 - \eta + \eta\Theta/w). \quad (\text{C.43})$$

Normalize the trade share in equation (A.17) by substituting from equations (C.11) and (C.18)

$$\lambda = (1 - F(\hat{z}))d^{1-\sigma} \frac{\Omega \mathbb{E} [z^{\sigma-1} \mid z > \hat{z}]}{\bar{z}^{\sigma-1}}. \quad (\text{C.44})$$

Starting from C.40, use C.33, and C.26, to derive that

$$\bar{\pi}_{agg} = \frac{1}{\sigma-1} (1 - \Omega [(1 - \eta)\zeta (S + \delta/\chi) - \sigma(N - 1)(1 - F(\hat{z}))\kappa]). \quad (\text{C.45})$$

Stationary Trade Shares and Profits. Using the stationary distribution in equation (B.13), calculate average profits from C.40,

$$\frac{\bar{\pi}_{agg}}{\Omega} = \frac{\bar{\pi}_{min}\theta}{1 + \theta - \sigma} + \frac{(\sigma - 1)(N - 1)\kappa\hat{z}^{-\theta}}{1 + \theta - \sigma}. \quad (\text{C.46})$$

Note that the minimum profits are at $z = 1$ and equal to $\bar{\pi}_{min}$ as long as $\hat{z} > 1$. Using this to define the profit spread between the average and worst firm in the economy,

$$\frac{\bar{\pi}_{agg}}{\Omega} - \bar{\pi}_{min} = \underbrace{\frac{\bar{\pi}_{min}\theta}{1 + \theta - \sigma} + \frac{(\sigma - 1)(N - 1)\kappa\hat{z}^{-\theta}}{1 + \theta - \sigma}}_{\text{ProfitSpread}} - \bar{\pi}_{min} = \frac{(\sigma - 1)\bar{\pi}_{min}}{1 + \theta - \sigma} + \frac{(\sigma - 1)(N - 1)\kappa\hat{z}^{-\theta}}{1 + \theta - \sigma}. \quad (\text{C.47})$$

Define the ratio of mean to minimum profits as $\bar{\pi}_{rat} := \frac{\bar{\pi}_{agg}/\Omega}{\bar{\pi}_{min}}$. From C.47 and C.36 find that

$$\bar{\pi}_{rat} = \frac{\theta}{1 + \theta - \sigma} + (N - 1)d^{1-\sigma} \frac{(\sigma - 1)\hat{z}^{\sigma-1-\theta}}{1 + \theta - \sigma}. \quad (\text{C.48})$$

Take equations (C.44) and (C.17) to find

$$\bar{z}^{\sigma-1} = \Omega \mathbb{E} [z^{\sigma-1}] + (N - 1)\lambda \bar{z}^{\sigma-1}. \quad (\text{C.49})$$

Solving gives an expression for a function of aggregate productivity in terms of underlying productivity and trade shares. Defining the home trade share as $\lambda_{ii} := 1 - (N - 1)\lambda$,

$$\bar{z}^{\sigma-1} = \Omega \frac{\mathbb{E} [z^{\sigma-1}]}{1 - (N - 1)\lambda} = \Omega \frac{\mathbb{E} [z^{\sigma-1}]}{\lambda_{ii}}. \quad (\text{C.50})$$

From equations (C.19) and (C.50),

$$w = \frac{1}{\bar{\sigma}} \Omega^{\frac{1}{\sigma-1}} \mathbb{E} [z^{\sigma-1}]^{\frac{1}{\sigma-1}} \lambda_{ii}^{\frac{1}{1-\sigma}}. \quad (\text{C.51})$$

This relates the real normalized wage to the aggregate productivity, the home trade share, and the number of varieties. Given that $\sigma > 1$, this expression implies that the larger the share of goods purchased at home, the lower the real wage is.

From C.44 and C.50 ,

$$\lambda = (1 - F(\hat{z})) d^{1-\sigma} \frac{\mathbb{E} [z^{\sigma-1} \mid z > \hat{z}] \lambda_{ii}}{\mathbb{E} [z^{\sigma-1}]}. \quad (\text{C.52})$$

Using the stationary distribution,

$$\lambda = \hat{z}^{-\theta} d^{1-\sigma} \frac{\hat{z}^{\sigma-1} \frac{\theta}{\theta-(\sigma-1)} \lambda_{ii}}{\frac{\theta}{\theta-(\sigma-1)}} = \hat{z}^{-\theta} d^{1-\sigma} \hat{z}^{\sigma-1} \lambda_{ii}. \quad (\text{C.53})$$

Using the definition of the home trade share,

$$\lambda_{ii} = \frac{1}{1 + (N-1) \hat{z}^{\sigma-1-\theta} d^{1-\sigma}}. \quad (\text{C.54})$$

Furthermore, multiplying the numerator and denominator by $\bar{\pi}_{\min}$ and using equation (C.36),

$$\lambda_{ii} = \frac{\bar{\pi}_{\min}}{\bar{\pi}_{\min} + (N-1) \hat{z}^{-\theta} \kappa}. \quad (\text{C.55})$$

Which gives an alternative expression for $\bar{\pi}_{\min}$ when $\kappa > 0$,

$$\bar{\pi}_{\min} = (N-1) \hat{z}^{-\theta} \kappa \frac{\lambda_{ii}}{1 - \lambda_{ii}}. \quad (\text{C.56})$$

D. Normalized and Stationary Dynamic Equilibrium Conditions

This section derives normalized stationary dynamic balanced growth path equilibrium conditions.

D.1. Utility and Welfare on a BGP

Using the substitution $C(t) = c \bar{L}M(t)$ shows time 0 welfare \bar{U} as a function of c and g is

$$\bar{U}(c, g) = \frac{1}{1-\gamma} \frac{(c \bar{L}M(0))^{1-\gamma}}{\rho + (\gamma-1)g}. \quad (\text{D.1})$$

With log utility

$$\bar{U}(c, g) = \frac{\rho \log(c\bar{L}M(0)) + g}{\rho^2}. \quad (\text{D.2})$$

Using the standard IES of the consumer, adjusted for stochastic death of the firm, the firms' effective discount rate on a BGP is,

$$r = \rho + \gamma g + \delta. \quad (\text{D.3})$$

With log utility

$$r = \rho + g + \delta. \quad (\text{D.4})$$

We restrict parameters such that $g(1 - \gamma) < \rho$ in equilibrium to ensure finite utility. In the log utility case, this is simply $\rho > 0$.

D.2. Normalization of the Firm's Dynamic Problem

We proceed to derive the normalized continuation value function, value matching condition, and smooth pasting condition originally specified in equations (A.20)–(A.22). Define the normalized real value of the firm relative to normalized wages as

$$v(z, t) := \frac{V(Z, t)}{LM(t)w(t)}. \quad (\text{D.5})$$

Rearranging

$$V(Z, t) = \bar{L}w(t)M(t)v(Z/M(t), t). \quad (\text{D.6})$$

First differentiate the continuation value $V(Z, t)$ with respect to t in equation (D.6) and divide by $w(t)M(t)\bar{L}$, using the chain and product rule. This gives,

$$\frac{1}{w(t)M(t)\bar{L}} \frac{\partial V(Z, t)}{\partial t} = \frac{M'(t)}{M(t)} v(z, t) - \frac{M'(t)}{M(t)} \frac{Z}{M(t)} \frac{\partial v(z, t)}{\partial z} + \frac{M(t)}{M(t)} \frac{\partial v(z, t)}{\partial t} + \frac{w'(t)}{w(t)} v(z, t). \quad (\text{D.7})$$

Defining the growth rate of $g(t) := M'(t)/M(t)$ and $g_w(t) := w'(t)/w(t)$. Substitute these into equation (D.7), cancel out $M(t)$, and group $z = Z/M(t)$ to give

$$\frac{1}{w(t)M(t)\bar{L}} \frac{\partial V(Z, t)}{\partial t} = (g(t) + g_w(t))v(z, t) - g(t)z \frac{\partial v(z, t)}{\partial z} + \frac{\partial v(z, t)}{\partial t}. \quad (\text{D.8})$$

Differentiating equation (D.6) with respect to Z yields

$$\frac{\partial V(Z, t)}{\partial Z} = \frac{\bar{L}M(t)w(t)}{M(t)} \frac{\partial v(Z/M(t), t)}{\partial z} = \bar{L}w(t) \frac{\partial v(z, t)}{\partial z}. \quad (\text{D.9})$$

Similarly,

$$\frac{\partial^2 V(Z, t)}{\partial Z^2} = \frac{\bar{L}w(t)}{M(t)} \frac{\partial^2 v(z, t)}{\partial z^2}. \quad (\text{D.10})$$

Define the normalized profits from equation (A.19) as

$$\pi(z, t) := \frac{\Pi(zM(t), t)}{w(t)M(t)\bar{L}}. \quad (\text{D.11})$$

Divide equation (A.20) by $M(t)w(t)\bar{L}$, then substitute for $\frac{\partial V(Z, t)}{\partial t}$, $\frac{\partial V(Z, t)}{\partial Z}$, and $\frac{\partial^2 V(Z, t)}{\partial Z^2}$ from eqs. (D.8), (D.9), and (D.10) in to (A.20). Finally, group the normalized profits using equation (D.11):

$$(r(t) - g(t) - g_w(t))v(z, t) = \pi(z, t) + \left(\mu + \frac{v^2}{2} - g(t) \right) z \frac{\partial v(z, t)}{\partial z} + \frac{v^2}{2} z^2 \frac{\partial^2 v(z, t)}{\partial z^2} + \frac{\partial v(z, t)}{\partial t}. \quad (\text{D.12})$$

Equation (D.12) is the normalized version of the value function of the firm in the continuation region. The stationary version of this equation is,

$$(r - g)v(z) = \pi(z) + \left(\mu + \frac{v^2}{2} - g \right) zv'(z) + \frac{v^2}{2} z^2 v''(z). \quad (\text{D.13})$$

To derive the normalized smooth pasting condition, use equation (A.22) to show that equation (D.9) evaluated at $Z = M(t)$ equals 0, delivering

$$\frac{\partial v(1, t)}{\partial z} = 0. \quad (\text{D.14})$$

To arrive at the normalized value matching condition, divide equation (A.21) by $M(t)w(t)\bar{L}$ to obtain

$$\frac{V(M(t), t)}{M(t)w(t)\bar{L}} = \int_{M(t)}^{\infty} \frac{V(Z, t)}{M(t)w(t)\bar{L}} \phi(Z, t) dZ - \frac{X(t)}{M(t)w(t)\bar{L}} \quad (\text{D.15})$$

Substituting using equation (D.5) and the definition of $x(t)$ yields

$$v(M(t)/M(t), t) = \int_M^{\infty} v(Z/M, t) \phi(Z, t) dZ - x(t). \quad (\text{D.16})$$

Finally, normalize the integral, realizing it is of the form discussed in equation (C.2), to obtain the normalized value matching condition:

$$v(1, t) = \int_1^{\infty} v(z, t) f(z, t) dz - x(t). \quad (\text{D.17})$$

D.3. Normalization of the Free Entry Condition

Normalizing the free entry condition given in equation (A.24), following similar steps that delivered the normalized value matching condition in equation (D.17), gives

$$x(t)/\chi = \int_1^{\infty} v(z, t) f(z, t) dz. \quad (\text{D.18})$$

Relating this to the value-matching condition of the adopting firm given in equation (D.17) provides a simple formulation of the stationary free entry condition that is useful in determining Ω and g :

$$v(1) = x \frac{1 - \chi}{\chi}. \quad (\text{D.19})$$

E. Solving for the Continuation Value Function

Although our baseline model does not feature exogenous productivity shocks, in this section we solve for the value function of a more general model that has GBM with $v \geq 0$. Our baseline case of $v = 0$ is nested in this formulation. The differential equation for the value function is solved using the method of undetermined coefficients. The goal of this section is to solve for the value function as a function of parameters, g , Ω , and \hat{z} (sometimes implicitly through $\bar{\pi}_{min}$).

Selection into Exporting. If $\kappa > 0$, generically some firms will choose to be exporters and some firms will only sell domestically. The value function will have a region of productivities representing the value of firms that only sell domestically and a region representing firms that also export. That is,

$$v(z) = \begin{cases} v_d(z) & \text{if } z \leq \hat{z} \\ v_x(z) & \text{if } z \geq \hat{z}. \end{cases}$$

We guess the value function is of the following form, with undetermined constants a , ν , and b ,³⁷

$$v_d(z) = a \bar{\pi}_{min} \left(z^{\sigma-1} + \frac{\sigma-1}{\nu} z^{-\nu} \right), \quad (\text{E.1})$$

$$v_x(z) = a \bar{\pi}_{min} \left(\left(1 + (N-1) d^{1-\sigma} \right) z^{\sigma-1} + \frac{\sigma-1}{\nu} z^{-\nu} + (N-1) \frac{1}{a(r-g)} \left(b z^{-\nu} - \frac{\kappa}{\bar{\pi}_{min}} \right) \right). \quad (\text{E.2})$$

³⁷This guess is also applying a standard transversality condition to eliminate an explosive root.

The value of a firm can be decomposed into the value of operating with its current productivity forever and the option value of adopting a better technology. The constant a is a discounting term on the value of earning the profits from producing with productivity z in perpetuity. The constant ν reflects the rate at which the option value of technology adoption goes to zero as productivity increases. The constant b is an adjustment to the perpetuity profits that reflects a firm with productivity z will eventually switch from exporting to being a domestic producer if z is constant in a growing economy.

By construction, the form of these guesses ensures that value matching and smooth pasting are satisfied, both at the adoption threshold ($z = 1$) and the exporter threshold ($z = \hat{z}$). To solve for a and ν , substitute equation (E.1) into the continuation value function in equation (D.13) using $\pi_d(z)$ from equation (C.34). This generates an ODE and, grouping terms, the method of undetermined coefficients provides a system of 2 equations in the 2 unknowns.³⁸

Solving the system gives,

$$\nu = \frac{\mu - g}{v^2} + \sqrt{\left(\frac{g - \mu}{v^2}\right)^2 + \frac{r - g}{v^2/2}}, \quad (\text{E.3})$$

$$a = \frac{1}{r - g - (\sigma - 1)(\mu - g + (\sigma - 1)v^2/2)}. \quad (\text{E.4})$$

To solve for b , plug equation (E.2) into the continuation value function in equation (D.13) using $\pi_x(z)$ from equation (C.35). This generates an ODE and, grouping terms, the method of undetermined coefficients provides a system of 3 equations in the 3 unknowns. By construction, the a and ν terms match those previously found, giving a consistent solution for b :

$$b = (1 - a(r - g)) d^{1-\sigma} \hat{z}^{\nu+\sigma-1}. \quad (\text{E.5})$$

Note, the main effect of the GBM is to modify the ν constant to reflect changes in the expected execution time of the option value of technology diffusion, and hence the exponent for discounting.

As will be useful in solving for g , evaluating at the adoption threshold yields,

$$v(1) = a\bar{\pi}_{\min} \left(1 + \frac{\sigma - 1}{\nu}\right). \quad (\text{E.6})$$

For the baseline case of $\mu = \nu = 0$,

$$a = \frac{1}{r + (\sigma - 2)g}, \quad (\text{E.7})$$

$$\nu = \frac{r}{g} - 1, \quad (\text{E.8})$$

$$b = \frac{\sigma - 1}{\nu + \sigma - 1} d^{1-\sigma} \hat{z}^{\nu+\sigma-1}. \quad (\text{E.9})$$

³⁸Instead of the method of undetermined coefficients, a direct solution approach would be to solve the continuation value function ODEs in the domestic sales and exporter regions, using the smooth pasting condition as the boundary value.

F. Computing the BGP Equilibrium when All Firms Export

In the case of $\kappa = 0$ all firms export (given d s.t. the economy is not in autarky), and the value function has only one region. We guess that the value function will take the following form,

$$v(z) = a\bar{\pi}_{\min} (1 + (N - 1)d^{1-\sigma}) \left(z^{\sigma-1} + \frac{\sigma-1}{\nu} z^{-\nu} \right). \quad (\text{F.1})$$

Substituting equation (F.1) into the continuation value function in equation (D.13) with profits from equation (C.35), the ν and a are identical to those from equations (E.3) and (E.4). Evaluating at the threshold,

$$v(1) = a (1 + (N - 1)d^{1-\sigma}) \bar{\pi}_{\min} \left(1 + \frac{\sigma-1}{\nu} \right). \quad (\text{F.2})$$

Solving for the Growth Rate and measure of Varieties when All Firms Export. Using the free entry condition from equation (D.19) with equation (F.2) to find,

$$\frac{x}{\bar{\pi}_{\min}} = a (1 + (N - 1)d^{1-\sigma}) \frac{\chi}{1-\chi} \frac{\sigma + \nu - 1}{\nu}. \quad (\text{F.3})$$

Substitute equations (F.1) and (F.2) into the value matching condition of equation (D.17), and divide by $a\bar{\pi}_{\min}(1 + (N - 1)d^{1-\sigma})$

$$1 + \frac{\sigma-1}{\nu} = \frac{\theta(\nu + \sigma - 1)(\theta + \nu - \sigma + 1)}{\nu(\theta + \nu)(\theta - \sigma + 1)} - \frac{x}{\bar{\pi}_{\min}a (1 + (N - 1)d^{1-\sigma})}. \quad (\text{F.4})$$

Combine equations (F.3) and (F.4), and solve for ν . For any cost function x and minimum profits $\bar{\pi}_{\min}$,

$$\nu = \frac{\chi\theta(\theta + 1 - \sigma)}{\sigma - 1 - \theta\chi}. \quad (\text{F.5})$$

The aggregate growth rate is found by equating equations (E.3) and (F.5) to find

$$g = \underbrace{\mu}_{\text{Drift}} + \underbrace{\frac{(r - \mu)((\sigma - 1)/\chi - \theta)}{\theta^2 - \theta\sigma + (\sigma - 1)/\chi}}_{\text{Deterministic}} + \frac{v^2}{2} \underbrace{\frac{\theta^2(\theta + 1 - \sigma)^2}{(\theta^2 - \theta\sigma + (\sigma - 1)/\chi)(\theta - (\sigma - 1)/\chi)}}_{\text{Stochastic}}. \quad (\text{F.6})$$

In the baseline case of $v = \mu = 0$

$$g = \frac{(\rho + \delta)(\sigma - 1 - \chi\theta)}{\theta\chi(\gamma + \theta - \sigma) - (\gamma - 1)(\sigma - 1)}, \quad (\text{F.7})$$

and,

$$\frac{x}{\bar{\pi}_{\min}} = (1 + (N - 1)d^{1-\sigma}) \frac{\chi}{1-\chi} \frac{1}{r - g}. \quad (\text{F.8})$$

The growth rate is independent of the trade costs, the population, and the number of countries. The intuition—as discussed in the body of the paper—is that the growth rate is driven by the ratio of the minimum to the mean profits, which are proportional and independent of the scale or integration of economies in the absence of any export selection. The constant death rate only enters to increase the discount rate.

Note in the baseline case of $v = \mu = \delta = 0$ and log utility

$$g = \frac{\rho(\sigma - 1 - \chi\theta)}{\theta^2\chi} \bar{\pi}_{\text{rat}}^k, \quad (\text{F.9})$$

where (from eqs. C.40 and C.17 with $\kappa = 0$ and $\hat{z} = 1$) the ratio of average profits to minimum profits is

$$\bar{\pi}_{\text{rat}}^k = \frac{\bar{\pi}_{\text{agg}}^k/\Omega}{\bar{\pi}_{\text{min}}^k} = \frac{(1 + (N - 1)d^{1-\sigma})\bar{\pi}_{\text{min}}E[z^{\sigma-1}]}{(1 + (N - 1)d^{1-\sigma})\bar{\pi}_{\text{min}}} = E[z^{\sigma-1}] = \frac{\theta}{1 + \theta - \sigma}. \quad (\text{F.10})$$

The Measure of Varieties Ω . Here we solve for the measure of varieties Ω in the baseline case of $v = \mu = \delta = 0$.

Substitute into the free entry condition equation (F.8) using the definition of $\bar{\pi}_{\text{min}}$ in terms of \bar{z} and \tilde{L} from equation (C.33), the definition of \bar{z} from equation (C.17), and the definition of \tilde{L} from equation (C.26), to obtain the implicit equation. In the baseline case where $\eta = 0$, an explicit solution is³⁹

$$\Omega = \frac{\chi((\gamma - 1)(\sigma - 1) - \theta\chi(\gamma + \theta - \sigma))}{\zeta(\theta\chi(-\gamma\delta - \sigma(\theta(\delta + \rho) + \rho) + \delta + \theta\rho + \rho) + (\gamma - 1)\delta(\sigma - 1) + \theta^2\sigma\chi^2(\delta + \rho))}. \quad (\text{F.13})$$

Note that the only place that the adoption cost, ζ , has come into the system of Ω and g is in the denominator of F.13. For this reason, the ζ parameter (along with \bar{L}) determines the scale of the economy.

It can be shown that in the Krugman model for all cases with $\eta = 0$, the number of domestic varieties is independent of trade costs d . Thus, both Ω and g are independent of d if $\eta = 0$. This implies through equations (C.26) and (B.14) that the amount of labor dedicated to technology adoption, \tilde{L} , is also independent of d if $\eta = 0$.

Through C.42, since \tilde{L} , Ω , and g are independent of d when $\eta = 0$, in response to a decrease in trade costs d , c increases only due to the $(1 + (N - 1)d^{1-\sigma})^{1/(\sigma-1)}$ term in \bar{z} .

³⁹The solution for the simplest case of $\gamma = 1$ and $\delta = 0$ is,

$$g = \frac{\rho(\sigma - 1 - \theta\chi)}{\theta\chi(1 + \theta - \sigma)} \quad (\text{F.11})$$

$$\Omega = \frac{\chi(1 + \theta - \sigma)}{\zeta\rho((1 + \theta)(\sigma - 1) - \sigma\theta\chi)} \quad (\text{F.12})$$

The key relationship can be summarized by the following elasticities.

$$\frac{d \log \bar{\pi}_{\text{rat}}^k(d)}{d \log(d)} = 0. \quad (\text{F.14})$$

Using equation (C.54) with $\hat{z} = 1$ shows

$$\frac{d \log \lambda_{ii}(d)}{d \log(d)} = (\sigma - 1) \left(1 + \frac{d^{\sigma-1}}{N-1} \right)^{-1} = (\sigma - 1)(1 - \lambda_{ii}) > 0. \quad (\text{F.15})$$

Furthermore, when $\eta = 0$,

$$\frac{d \log(1 - \tilde{L}(d))}{d \log(d)} = \frac{d \log \Omega(d)}{d \log(d)} = \frac{d \log g(d)}{d \log(d)} = 0. \quad (\text{F.16})$$

For the case with $\eta = 0$, the ratio of c for different trade costs d_1 and d_2 that both feature positive trade is

$$\frac{\bar{c}_{d_1}}{\bar{c}_{d_2}} = \frac{\bar{z}_{z_{d_1}}}{\bar{z}_{z_{d_2}}} = \left(\frac{1 + (N-1)d_1^{1-\sigma}}{1 + (N-1)d_2^{1-\sigma}} \right)^{\frac{1}{\sigma-1}}. \quad (\text{F.17})$$

Using the normalized welfare function in equation (D.1), since g is independent of d in the $\kappa = 0$ case, the ratio of welfare for different trade costs d_1 and d_2 that both feature positive trade (for $\gamma > 0$) is,

$$\frac{\bar{U}_{d_1}}{\bar{U}_{d_2}} = \left(\frac{1 + (N-1)d_1^{1-\sigma}}{1 + (N-1)d_2^{1-\sigma}} \right)^{\frac{1-\gamma}{\sigma-1}}. \quad (\text{F.18})$$

Comparing welfare between free trade ($d = 1$) and autarky (sufficiently high d s.t. there are no exporters) gives (for $\gamma \neq 1$)

$$\frac{\bar{U}_{\text{free}}}{\bar{U}_{\text{autarky}}} = N^{\frac{1-\gamma}{\sigma-1}}. \quad (\text{F.19})$$

G. Computing the BGP Equilibrium with Selection into Exporting ($\kappa > 0$)

We solve for g and Ω by reducing the equilibrium conditions to a system of two equations in these two unknowns. First, combining the free entry condition from equation (D.19) with $v(1)$ from equation (E.6) yields

$$\frac{x}{\bar{\pi}_{\text{min}}} = a \frac{\chi}{1-\chi} \frac{\sigma + \nu - 1}{\nu}. \quad (\text{G.1})$$

The second equation is found by evaluating the value matching condition of equation (D.17) by substituting in the domestic and exporter value functions in equations E.1 and E.2, the export threshold \hat{z} in equation (C.36), and the value at the adoption threshold $v(1)$ in equation (E.6) and dividing by $a\bar{\pi}_{\min}$. That is, evaluate

$$\frac{v(1)}{a\bar{\pi}_{\min}} = \frac{\int_1^{\infty} v(z,t)f(z,t)dz}{a\bar{\pi}_{\min}} - \frac{x}{a\bar{\pi}_{\min}} \quad (\text{G.2})$$

to obtain

$$1 + \frac{\sigma - 1}{\nu} = \frac{\frac{\nu(N-1)(\theta-\sigma+1)(d^{1-\sigma}(\theta+\nu)\hat{z}^{-\theta+\sigma-1} - b\theta\hat{z}^{-\theta-\nu})}{a(g-r)} + \theta(\nu(N-1)d^{1-\sigma}(\theta+\nu)\hat{z}^{-\theta+\sigma-1} + (\nu+\sigma-1)(\theta+\nu-\sigma+1))}{\nu(\theta+\nu)(\theta-\sigma+1)} - \frac{\chi}{1-\chi} \frac{\sigma+\nu-1}{\nu}. \quad (\text{G.3})$$

As detailed in Section H, use the definition of $\bar{\pi}_{\min}$ and substitute for x, ν, a, b, \hat{z} and r into equations (G.1) and (G.3) to find a system of 2 equations in Ω and g . Note that the adoption cost x does not appear in equation (G.3), which is why g is independent of the specification of the cost of adoption. Equation (G.1) does explicitly depend on x , which is why the number of varieties is a function of the adoption cost, and ultimately why welfare is also a function of x .

G.1. Case with $v = \mu = \delta = 0$

As the GBM does not qualitatively impact the solution, we concentrate our analytical theory on the simple baseline case. The cost of adoption is a function of minimum profits, parameters, and $r - g$:

$$x = \bar{\pi}_{\min} \frac{\chi}{1-\chi} \frac{1}{r-g}. \quad (\text{G.4})$$

Evaluating the general equation (G.3) with the substitutions for x, ν, a, b, \hat{z} and r that correspond to the baseline case of $v = \mu = \delta = 0$ yields a unique g that satisfies the value function and the value matching equations, given by the implicit equation,

$$g = \frac{(\sigma - 1 + (N - 1)\theta d^{1-\sigma} \hat{z}^{-\theta+\sigma-1}) \bar{\pi}_{\min} + (N - 1)(-\theta + \sigma - 1) \hat{z}^{-\theta} \kappa}{x(\gamma + \theta - 1)(\theta - \sigma + 1)} - \frac{\rho}{\gamma + \theta - 1}. \quad (\text{G.5})$$

Using $\kappa/\bar{\pi}_{\min} = \hat{z}^{\sigma-1} d^{1-\sigma}$ from equation (C.36), simplify to

$$g = \frac{(\sigma - 1) [\bar{\pi}_{\min} + (N - 1)\kappa\hat{z}^{-\theta}]}{x(\gamma + \theta - 1)(\theta - \sigma + 1)} - \frac{\rho}{\gamma + \theta - 1}. \quad (\text{G.6})$$

Using equation (C.47), realize this relates the growth rate to the difference between average and minimum profits

$$g = \left(\frac{\bar{\pi}_{\text{agg}}/\Omega - \bar{\pi}_{\min}}{x(\gamma + \theta - 1)} \right) - \frac{\rho}{\gamma + \theta - 1}. \quad (\text{G.7})$$

Substitute for x using the free entry condition in equation (G.4), use the definition of the average to minimum profit ratio ($\bar{\pi}_{\text{rat}} := \frac{\bar{\pi}/\Omega}{\bar{\pi}_{\text{min}}}$), and substitute for r using equation (D.3) to obtain

$$g = (\rho + (\gamma - 1)g) \frac{1 - \chi}{\chi(\gamma + \theta - 1)} (\bar{\pi}_{\text{rat}} - 1) - \frac{\rho}{\gamma + \theta - 1}. \quad (\text{G.8})$$

Solving for g gives an equation for g as a function exclusively of parameters and the ratio of average to minimum profits

$$g = \frac{\rho}{\chi\theta((1 - \chi)\bar{\pi}_{\text{rat}} - 1)^{-1} + 1 - \gamma}. \quad (\text{G.9})$$

Furthermore, with log utility $\gamma = 1$ and

$$g = \frac{\rho(1 - \chi)}{\chi\theta} \bar{\pi}_{\text{rat}} - \frac{\rho}{\chi\theta}. \quad (\text{G.10})$$

The relationship between growth and trade costs. First, note that the growth rate is increasing in the profit ratio, (since $\chi \in (0, 1)$):

$$\frac{dg(\bar{\pi}_{\text{rat}})}{d\bar{\pi}_{\text{rat}}} = \frac{\rho\theta\chi(1 - \chi)}{((\gamma - 1)(1 - (1 - \chi)\bar{\pi}_{\text{rat}} + \theta\chi))^2} > 0. \quad (\text{G.11})$$

To determine if whether growth is increasing in d , use the chain rule

$$\frac{dg(d)}{dd} = \frac{dg(\bar{\pi}_{\text{rat}})}{d\bar{\pi}_{\text{rat}}} \frac{d\bar{\pi}_{\text{rat}}(d)}{dd}. \quad (\text{G.12})$$

Given equations (G.11) and (G.12), a sufficient condition to conclude that $\frac{dg(d)}{dd} < 0$ is $\frac{d\bar{\pi}_{\text{rat}}(d)}{dd} < 0$. To show this differentiate C.48 w.r.t. d to find,

$$\frac{d\bar{\pi}_{\text{rat}}(d)}{dd} \propto - \left((\sigma - 1)\hat{z}(d) + d(1 + \theta - \sigma) \frac{d\hat{z}(d)}{dd} \right). \quad (\text{G.13})$$

Since $d > 0$, $\hat{z}(d) > 0$, $\sigma > 1$, and $1 + \theta - \sigma > 0$, a sufficient condition for $\frac{d\bar{\pi}_{\text{rat}}(d)}{dd} < 0$ is if $\frac{d\hat{z}(d)}{dd} > 0$. Differentiate equation (C.36) to find

$$\frac{d\hat{z}(d)}{dd} \propto (\sigma - 1) - \frac{d}{\bar{\pi}_{\text{min}}(d)} \frac{d\bar{\pi}_{\text{min}}(d)}{dd}. \quad (\text{G.14})$$

Therefore, a sufficient condition to conclude that $\frac{d\hat{z}(d)}{dd} > 0$ is

$$\frac{d \log \bar{\pi}_{\text{min}}(d)}{dd} < \frac{\sigma - 1}{d}. \quad (\text{G.15})$$

Summarizing,

$$\text{sign} \frac{dg(d)}{dd} = \text{sign} \frac{d\bar{\pi}_{\text{rat}}}{dd} = -\text{sign} \frac{d\hat{z}(d)}{dd}. \quad (\text{G.16})$$

G.2. Baseline Case with $v = \mu = \delta = \eta = 0$ and log utility.

Adding the restriction that $\eta = 0$ and $\gamma = 1$ to the adoption cost equation (C.43) and the firms' discount rate equation (D.4) delivers the key simplifications that permit solving for the BGP equilibrium in closed form:

Calculating the Growth Rate and the Measure of Varieties.

$$x = \zeta, \quad (\text{G.17})$$

$$r - g = \rho. \quad (\text{G.18})$$

Using equation (G.4) gives an expression for $\bar{\pi}_{\text{min}}$ in terms of model parameters

$$\bar{\pi}_{\text{min}} = \frac{(1 - \chi)\zeta\rho}{\chi}. \quad (\text{G.19})$$

Substitute equation (G.19) into equation (C.36) to find the export threshold in terms of parameters,

$$\hat{z} = d \left(\frac{1}{\zeta} \frac{\kappa\chi}{\rho(1-\chi)} \right)^{\frac{1}{\sigma-1}}. \quad (\text{G.20})$$

Substitute $\gamma = 1$ and equations (G.17), (G.19), and (G.20) into equation (G.6) and simplify to obtain g in closed form:

$$g = \frac{\rho(1-\chi)}{\chi\theta} \frac{\sigma-1}{(\theta-\sigma+1)} \left(1 + (N-1)d^{-\theta} \left(\frac{\kappa}{\zeta} \frac{\chi}{\rho(1-\chi)} \right)^{1-\frac{\theta}{\sigma-1}} \right) - \frac{\rho}{\theta}, \quad (\text{G.21})$$

$$= \frac{\rho(1-\chi)}{\chi\theta} \frac{\left(\theta + (N-1)(\sigma-1)d^{-\theta} \left(\frac{1}{\zeta} \frac{\kappa\chi}{\rho(1-\chi)} \right)^{1-\frac{\theta}{\sigma-1}} \right)}{(\theta-\sigma+1)} - \frac{\rho}{\chi\theta}. \quad (\text{G.22})$$

For ease of comparison to g as a function of $\bar{\pi}_{\text{rat}}$ use equation (G.8) evaluated at $\gamma = 1$ with equation (G.21) to realize

$$\bar{\pi}_{\text{rat}} - 1 = \frac{\left(1 + (N-1)d^{-\theta} \left(\frac{1}{\zeta} \frac{\kappa\chi}{\rho(1-\chi)} \right)^{1-\frac{\theta}{\sigma-1}} \right)}{\left(\frac{\theta}{\sigma-1} - 1 \right)}, \quad (\text{G.23})$$

or equation (G.10) with equation (G.22) to see

$$\bar{\pi}_{\text{rat}} = \frac{\left(\theta + (N-1)(\sigma-1)d^{-\theta} \left(\frac{1}{\zeta} \frac{\kappa\chi}{\rho(1-\chi)} \right)^{1-\frac{\theta}{\sigma-1}} \right)}{(\theta - \sigma + 1)}. \quad (\text{G.24})$$

Note, g is decreasing in d and κ (since $\theta > 0$, $\sigma > 1$, and $1 + \theta - \sigma > 0$). Thus, from equation (G.16) or direct differentiation of equation (G.21),

$$\frac{dg}{dd} < 0; \quad \frac{d\bar{\pi}_{\text{rat}}}{dd} < 0; \quad \frac{d\hat{z}}{dd} > 0; \quad \frac{dg}{d\kappa} < 0. \quad (\text{G.25})$$

See by comparing to equation (F.7) that the limit of g as $d \rightarrow \infty$ in equation (G.22) equals the autarky and all-export economy growth rates, so there is no discontinuity in the economy in this direction.

Note that the parameter ζ (previously interpreted as the scale in equation F.13) and κ only enter the growth rate multiplicatively. This is because the fixed costs of adoption, entry, and export in levels are proportional to the scale of the economy. Since the calibration strategy targets relative moments (i.e., proportion of exporters, relative size of exporters to domestic firms, growth rates, trade shares), κ is not separately identifiable from ζ without some moment that targets the level of the economy.

To find the number of varieties, maintain $\gamma = 1$, $\eta = 0$. To solve for Ω , start with the definition of $\bar{\pi}_{\text{min}}$ from equation (C.33):

$$\bar{\pi}_{\text{min}} = \frac{1-\tilde{L}}{(\sigma-1)\tilde{z}^{\sigma-1}}. \quad (\text{G.26})$$

For $\bar{\pi}_{\text{min}}$, substitute from equation (G.19). For the right hand side, substitute for \tilde{L} , \tilde{z} and S with equations (C.26), (C.17), and (B.14). Then, use g and \hat{z} from equations (G.21), (G.20), and solve for Ω in terms of model parameters:

$$\Omega = \frac{1}{\zeta} \frac{\chi(1+\theta-\sigma)}{(1-\chi)\theta\sigma\rho} \left(1 + (N-1)d^{-\theta} \left(\frac{\kappa\chi}{\zeta\rho(1-\chi)} \right)^{1-\frac{\theta}{\sigma-1}} - \frac{1+\theta-\sigma}{\theta\sigma(1-\chi)} \right)^{-1}. \quad (\text{G.27})$$

Note, from equations (C.54) and (G.20), the home trade share is,

$$\lambda_{ii} = \frac{1}{1 + (N-1)d^{-\theta} \left(\frac{\kappa\chi}{\zeta\rho(1-\chi)} \right)^{1-\frac{\theta}{\sigma-1}}}. \quad (\text{G.28})$$

Note, by substituting eq. G.28 into eqs. G.21 and G.27, the growth rate and Ω can be written as a function of the home trade share:

$$g = \frac{\rho(1-\chi)}{\chi\theta} \frac{\sigma-1}{(\theta-\sigma+1)} \lambda_{ii}^{-1} - \frac{\rho}{\theta}. \quad (\text{G.29})$$

$$\Omega = \frac{\chi}{\zeta\rho} \left(\frac{(1-\chi)\theta\sigma}{1+\theta-\sigma} \lambda_{ii}^{-1} - 1 \right)^{-1}. \quad (\text{G.30})$$

Calculating Consumption. By the resource constraint, $c = y$ when $\eta = 0$ (equation C.41). Thus, using equation (C.30), consumption is given by

$$c = y = (1 - \tilde{L})\bar{z}. \quad (\text{G.31})$$

Equations (C.26), (B.14), (C.36), and (G.29) combine to yield the amount of labor dedicated to variable goods production in terms of the home trade share:

$$1 - \tilde{L} = (\sigma - 1) \left(\sigma - \frac{1 + \theta - \sigma}{\theta(1 - \chi)} \lambda_{ii} \right)^{-1}. \quad (\text{G.32})$$

Equation (C.50) gives

$$\bar{z} = \Omega^{\frac{1}{\sigma-1}} \lambda_{ii}^{\frac{1}{1-\sigma}} (\mathbb{E} [z^{\sigma-1}])^{\frac{1}{\sigma-1}}. \quad (\text{G.33})$$

Substituting equations (G.32), (G.33), and (G.30) into equation (G.31) yields consumption as a function of parameters and the home trade share:

$$c = \frac{(\sigma - 1)\theta\sigma(1 - \chi)}{\sigma(1 + \theta - \sigma)} \left(\frac{\chi}{\rho\zeta} \right)^{\frac{1}{\sigma-1}} \left(\frac{(1 - \chi)\theta\sigma}{1 + \theta - \sigma} - \lambda_{ii} \right)^{\frac{\sigma}{1-\sigma}} (\mathbb{E} [z^{\sigma-1}])^{\frac{1}{\sigma-1}}. \quad (\text{G.34})$$

Trade Cost Elasticities. Comparative statics are analyzed by calculating elasticities with respect to trade costs using equations (G.20), (G.21), (G.24), (G.27), and (G.28):

$$\frac{d \log \hat{z}(d)}{d \log(d)} = 1, \quad (\text{G.35})$$

$$\frac{d \log g(d)}{d \log(d)} = -\theta \left(1 + \frac{\rho d^\theta (-\theta\chi + \sigma - 1) \left(\frac{\kappa\chi}{\zeta\rho(1-\chi)} \right)^{\frac{\theta}{\sigma-1}}}{\kappa(N-1)(\sigma-1)\chi} \right)^{-1} < 0, \quad (\text{G.36})$$

$$\frac{d \log \bar{\pi}_{\text{rat}}(d)}{d \log(d)} = -\theta \left(1 + \frac{\theta d^\theta \left(\frac{\kappa\chi}{\zeta\rho(1-\chi)} \right)^{\frac{\theta}{\sigma-1}-1}}{(N-1)(\sigma-1)} \right)^{-1} < 0, \quad (\text{G.37})$$

$$\frac{d \log \Omega(d)}{d \log(d)} = \theta \left(1 + \frac{\rho d^\theta ((\theta + 1)(\sigma - 1) - \theta\sigma\chi) \left(\frac{\kappa\chi}{\zeta\rho(1-\chi)} \right)^{\frac{\theta}{\sigma-1}}}{\theta\kappa(N-1)\sigma\chi} \right)^{-1} > 0, \quad (\text{G.38})$$

$$\frac{d \log \lambda_{ii}(d)}{d \log(d)} = \theta \left(1 + \frac{d^\theta}{N-1} \left(\frac{\kappa\chi}{\zeta\rho(1-\chi)} \right)^{\frac{\theta}{\sigma-1}-1} \right)^{-1} > 0. \quad (\text{G.39})$$

Let $\varepsilon_{f,x}$ be the elasticity of any $f(x)$ w.r.t. x . These elasticities can be rearranged to highlight their relationship to trade volume. To see this, first define the ratio of the home trade share to the share of

goods purchased away from home:

$$\frac{\lambda_{ii}}{1 - \lambda_{ii}} = \frac{d^\theta \left(\frac{\kappa\chi}{\zeta\rho(1-\chi)} \right)^{\frac{\theta}{\sigma-1} - 1}}{N - 1}. \quad (\text{G.40})$$

Substituting this into the result above, yields

$$\frac{d \log \lambda_{ii}(d)}{d \log(d)} = \theta \left(1 + \frac{\lambda_{ii}}{1 - \lambda_{ii}} \right)^{-1} = \theta(1 - \lambda_{ii}), \quad (\text{G.41})$$

$$\frac{d \log g(d)}{d \log(d)} = -\theta \left[1 + \left(\frac{-\theta\chi + \sigma - 1}{(\sigma - 1)(1 - \chi)} \right) \frac{\lambda_{ii}}{1 - \lambda_{ii}} \right]^{-1} \quad (\text{G.42})$$

$$= \left(\frac{\chi(1 + \theta - \sigma)}{(\sigma - 1)(1 - \chi)} \lambda_{ii} - 1 \right)^{-1} \varepsilon_{\lambda_{ii},d}, \quad (\text{G.43})$$

$$\frac{d \log \Omega_{ii}(d)}{d \log(d)} = \left(1 - \frac{1 + \theta - \sigma}{\theta\sigma(1 - \chi)} \lambda_{ii} \right)^{-1} \varepsilon_{\lambda_{ii},d}. \quad (\text{G.44})$$

The elasticity of $(1 - \tilde{L})$ w.r.t. d is

$$\frac{d \log(1 - \tilde{L}(d))}{d \log(d)} = \left(\frac{\theta\sigma(1 - \chi)}{1 + \theta - \sigma} \lambda_{ii}^{-1} - 1 \right)^{-1} \varepsilon_{\lambda_{ii},d} > 0. \quad (\text{G.45})$$

The elasticity of \bar{z} w.r.t. d is

$$\frac{d \log(\bar{z}(d))}{d \log(d)} = \frac{\varepsilon_{\Omega,d} - \varepsilon_{\lambda_{ii},d}}{\sigma - 1} > 0. \quad (\text{G.46})$$

From equation (G.31)

$$\varepsilon_{c,d} = \varepsilon_{1-\tilde{L},d} + \varepsilon_{\bar{z},d}. \quad (\text{G.47})$$

Using equations (G.34) and (G.41) yields

$$\varepsilon_{c,d} = -\frac{\sigma}{\sigma - 1} \left(1 - \frac{\theta\sigma(1 - \chi)}{(1 + \theta - \sigma)} \lambda_{ii}^{-1} \right)^{-1} \varepsilon_{\lambda_{ii},d}. \quad (\text{G.48})$$

Finally, from D.2

$$\varepsilon_{\bar{U},d} = \frac{\rho \varepsilon_{c,d} + g \varepsilon_{g,d}}{g + \rho \log(c\bar{L}M(0))} \quad (\text{G.49})$$

$$= \frac{\rho^2}{\bar{U}} (\rho \varepsilon_{c,d} + g \varepsilon_{g,d}). \quad (\text{G.50})$$

This can be further organized by substitution for $\varepsilon_{c,d}$ and $\varepsilon_{g,d}$ from equations (G.48) and (G.43) into equation (G.50).⁴⁰

$$\varepsilon_{\bar{U},d} = -\varepsilon_{\lambda_{ii},d} \frac{\rho^3}{\bar{U}} \left(\frac{\sigma}{(\sigma-1) \left(1 - \frac{\theta\sigma(1-\chi)}{(\theta-\sigma+1)\lambda_{ii}}\right)} + \frac{(\sigma-1)(1-\chi)}{\theta\chi(\theta-\sigma+1)\lambda_{ii}} \right). \quad (\text{G.51})$$

Therefore, if the term in brackets is positive, then the elasticity of utility is of the opposite sign of the home trade share, and hence always decreasing in trade costs. The first term in the brackets is always negative and the second term is always positive. There exist parameter values such that sum is negative, such that economies with lower trade costs have lower welfare in a comparison of steady states. This occurs when $\sigma-1$ is close to its lower bound of $\theta\chi$. This unintuitive result is possible because this analysis ignores transition dynamics and because this is an inefficient economy, so economics of the second best applies.

Firm Adoption Timing. A firm adopts when normalized productivity equals 1 by definition. On the BGP, firms drift backwards towards $z = 1$ at constant rate g . Thus, the time until adoption $\tau(z)$ is given by

$$e^{-g\tau} z = 1$$

$$\tau(z) = \frac{\log(z)}{g} \quad (\text{G.52})$$

The expected time until adoption for a firm that is about to draw a new productivity, $\bar{\tau}$, is just the expected adoption time integrated over the distribution of the new z :

$$\bar{\tau} = \int_1^\infty \frac{\log(z)}{g} dF(z) = \frac{1}{g} \int_1^\infty \log(z) \theta z^{-\theta-1} = \frac{1}{\theta g} \quad (\text{G.53})$$

Since firms draw a new z from the unconditional distribution, the expected time to adoption for a newly adopting firm is the same as the average time to adoption.

H. Computing the BGP Equilibrium in General

In the general case of the $\kappa = 0$, the equilibrium g can be calculated through an explicit equation, and the Ω found separately. If $\kappa > 0$, then a system of 2 non-linear algebraic equations in g and Ω are solved. Summarizing equations for easy reference against the code:

General Substitutions. The following substitutions are used in reducing the equilibrium conditions into a simple system of equations that can be solved for g and Ω . Given g and Ω all other equilibrium

⁴⁰For determining the direction of the change, as an elasticity is $d\bar{U}'(d)/\bar{U}(d)$, and since $d > 1$, if $\bar{U}(d) > 0$ then the sign of this elasticity calculation matches the sign of the derivative. Otherwise, the sign of the derivative is negative of the elasticity. As equation (G.50) divided by the utility in the calculation, this sign cancels, and ensures that negative utility does not affect the direction of the changes (as expected with a monotone function with the possibility of arbitrarily small initial conditions).

values are determined. We use equations (B.13), (B.14), (E.3), (E.4), (E.5), (D.3), (C.26), (C.17), (C.36), (C.19), and (C.43):

$$F(z) = 1 - z^{-\theta} \quad (\text{H.1})$$

$$S = \theta \left(g - \mu - \theta \frac{v^2}{2} \right) \quad (\text{H.2})$$

$$\nu = \frac{\mu - g}{v^2} + \sqrt{\left(\frac{g - \mu}{v^2} \right)^2 + \frac{r - g}{v^2/2}} \quad (\text{H.3})$$

$$a = \frac{1}{r - g - (\sigma - 1)(\mu - g + (\sigma - 1)v^2/2)} \quad (\text{H.4})$$

$$b = (1 - a(r - g)) d^{1-\sigma} \hat{z}^{\nu+\sigma-1} \quad (\text{H.5})$$

$$r = \rho + \gamma g + \delta \quad (\text{H.6})$$

$$\tilde{L} = \Omega [(N - 1)(1 - F(\hat{z}))\kappa + (1 - \eta)\zeta(S + \delta/\chi)] \quad (\text{H.7})$$

$$\bar{z} = [\Omega (\mathbb{E}[z^{\sigma-1}] + (N - 1)(1 - F(\hat{z}))d^{1-\sigma}\mathbb{E}[z^{\sigma-1} | z > \hat{z}])]^{1/(\sigma-1)} \quad (\text{H.8})$$

$$\hat{z} = d \left(\frac{\kappa}{\bar{\pi}_{\min}} \right)^{\frac{1}{\sigma-1}} \quad (\text{H.9})$$

$$w = \frac{1}{\sigma} \bar{z} \quad (\text{H.10})$$

$$x = \zeta (1 - \eta + \eta\Theta/w) \quad (\text{H.11})$$

(Note, since $\bar{\pi}_{\min}$ is an implicit function through the \hat{z} in \bar{z} , it is easiest to add it to the system of equations instead of substituting it out).

All Firms Export Case. For any $v \geq 0$, the growth rate is given by equation (F.6), substituting for r from D.3.

$$g = \mu + \frac{(r - \mu)((\sigma - 1)/\chi - \theta)}{\theta^2 - \theta\sigma + (\sigma - 1)/\chi} + \frac{v^2}{2} \frac{\theta^2(\theta + 1 - \sigma)^2}{(\theta^2 - \theta\sigma + (\sigma - 1)/\chi)(\theta - (\sigma - 1)/\chi)}. \quad (\text{H.12})$$

Using the equilibrium g above, Ω can be found by solving the following system of equations in Ω and $\bar{\pi}_{\min}$ from equations (F.3) and (C.33) (where Ω is implicitly in the \bar{z} and \tilde{L} terms):

$$\frac{x}{\bar{\pi}_{\min}} = a (1 + (N - 1)d^{1-\sigma}) \frac{\chi}{1 - \chi} \frac{\sigma + \nu - 1}{\nu}. \quad (\text{H.13})$$

$$\bar{\pi}_{\min} = \frac{1 - \tilde{L}}{(\sigma - 1)\bar{z}^{\sigma-1}}. \quad (\text{H.14})$$

Selection into Exporting Case. Equations (G.1), (G.3), and (C.33) provide a system of 3 equations in g , Ω , and $\bar{\pi}_{\min}$. To solve this non-linear system, substitute for $\bar{\pi}_{\min}$, ν , a , b , x , r , S , \tilde{L} , \bar{z} and \hat{z} using the

general substitutions listed above to eliminate dependence on all other endogenous variables.

$$\frac{x}{\bar{\pi}_{\min}} = a \frac{\chi}{1-\chi} \frac{\sigma + \nu - 1}{\nu}, \quad (\text{H.15})$$

$$1 + \frac{\sigma - 1}{\nu} = \frac{\nu(N-1)(\theta-\sigma+1)(d^{1-\sigma}(\theta+\nu)\hat{z}^{-\theta+\sigma-1} - b\theta\hat{z}^{-\theta-\nu})}{a(g-r)} + \theta \frac{\nu(N-1)d^{1-\sigma}(\theta+\nu)\hat{z}^{-\theta+\sigma-1} + (\nu+\sigma-1)(\theta+\nu-\sigma+1)}{\nu(\theta+\nu)(\theta-\sigma+1)} - \frac{\chi}{1-\chi} \frac{\sigma+\nu-1}{\nu}, \quad (\text{H.16})$$

$$\bar{\pi}_{\min} = \frac{1-\tilde{L}}{(\sigma-1)\tilde{z}^{\sigma-1}}. \quad (\text{H.17})$$

This system of equations holds for both the general case and the baseline case of $\nu = \mu = 0$ (if using Mathematica, make sure to substitute ν to reorganize the formulas to avoid any singularity). Alternatively, for the baseline case, together with the same definition of $\bar{\pi}_{\min}$ use equations (G.4) and G.9 as the system of equations, using the same substitutions as in the more general case.

Post Solution Calculations. In either case, given the equilibrium g and Ω , the following equilibrium values can be calculated through equations (C.40), (C.30), (D.1), (D.2), (C.54), and (C.42).

$$\bar{\pi}_{\text{agg}} = \bar{\pi}_{\min} \tilde{z}^{\sigma-1} - \Omega(N-1)(1-F(\hat{z}))\kappa, \quad (\text{H.18})$$

$$y = (1 - \tilde{L}) \tilde{z}, \quad (\text{H.19})$$

$$\bar{U} = \begin{cases} \frac{1}{1-\gamma} \frac{(c\bar{L}M(0))^{1-\gamma}}{\rho+(\gamma-1)g} & \gamma \neq 1 \\ \frac{\rho \log(c\bar{L}M(0))+g}{\rho^2} & \gamma = 1 \end{cases}, \quad (\text{H.20})$$

$$\lambda_{ii} = \frac{1}{1 + (N-1)\hat{z}^{\sigma-1-\theta}d^{1-\sigma}}, \quad (\text{H.21})$$

$$c = (1 - \tilde{L}) \tilde{z} - \eta\zeta\Omega\Theta(S + \delta/\chi). \quad (\text{H.22})$$

Consumption Equivalents. Here we focus on the log utility case. Let subscript t represent variables associated with the initial equilibrium and subscript T denote variables in the new equilibrium. From equation (46), the present discounted value of utility in steady state is

$$\bar{U}(c, g) = \frac{\rho \log(c) + g}{\rho^2}. \quad (\text{H.23})$$

Consumption equivalent welfare is the permanent change in the level of consumption c needed to make the representative consumer in the initial equilibrium indifferent to living in the new equilibrium. That is, the consumption equivalent α is such that

$$\bar{U}((1 + \alpha)c_t, g_t) = \bar{U}(c_T, g_T) \quad (\text{H.24})$$

$$\frac{\rho \log((1 + \alpha)c_t) + g_t}{\rho^2} = \frac{\rho \log(c_T) + g_T}{\rho^2} \quad (\text{H.25})$$

$$1 + \alpha = \frac{c_T}{c_t} \exp\left(\frac{g_T - g_t}{\rho}\right) = \exp(\rho(\bar{U}_T - \bar{U}_t)) \quad (\text{H.26})$$

I. Notation

General notation principle for normalization: move to lowercase after normalizing to the scale of the economy, from nominal to real, per-capita, and relative wages (all where appropriate). For symmetric countries, denote variables related to the trade sector with an x subscript. An overbar denotes an aggregation of the underlying variable. Drop the t subscript where possible for clarity in the static equilibrium conditions.

Notation Summary

Parameters	
Multiplicative Adoption Cost	$\zeta > 0$
Labor/Goods Proportion of Adoption Cost	$0 \leq \eta \leq 1$
Goods Adoption Cost	$\Theta \geq 0$
Iceberg Trade Cost	$d \geq 1$
Fixed Export Cost	$\kappa \geq 0$
Consumer CRRA	$\gamma \geq 0$
Elasticity of Substitution between Products	$\sigma > 1$
Markup	$\bar{\sigma} := \sigma/(\sigma - 1)$
Tail Parameter of the Pareto Distribution	$\theta > 1$
Consumer's Discount Rate	$\rho > 0$
Population per Country	$\bar{L} > 0$
Adoption to Entry Cost Ratio	$\chi < 1$
Firm Death Rate	$\delta \geq 0$
Drift of Productivity Process	$\mu \geq 0$
Standard Deviation of Productivity Process	$v \geq 0$

Notation Summary

Equilibrium Variables	
Productivity	Z
CDF of the Productivity Distribution	$\Phi(Z, t)$
PDF of the Productivity Distribution	$\phi(Z, t)$
Representative Consumers Flow Utility	$U(t)$
Representative Consumers Welfare	$\bar{U}(t)$
Real Firm Value	$V(Z, t)$
Optimal Search Threshold	$M(t)$
Optimal Export Threshold	$\hat{Z}(t)$
Aggregate Nominal Expenditures on Final Goods	$Y(t)$
Aggregate Real Consumption	$C(t)$
Domestic Labor demand	$\ell_d(Z, t)$
Export Labor demand	$\ell_x(Z, t)$
Domestic Quantity	$Q_d(Z, t)$
Export Quantity	$Q_x(Z, t)$
Real Search Cost	$X(t)$
Domestic idiosyncratic prices	$p_d(Z, t)$
Export idiosyncratic prices	$p_x(Z, t)$
Nominal Wages	$W(t)$
Real Domestic Profits	$\Pi_d(Z, t)$
Real Per-market Export Profits	$\Pi_x(Z, t)$
Firm Effective Discount Rate	$r(t)$
Trade Share	$\lambda(t)$
Price level	$P(t)$
Number of Varieties	$\Omega(t)$

Normalization Notation Summary (implicit t where appropriate)

Real, Normalized, and Per-Capita Variables

Per-capita Labor Demand/Supply	$L := \bar{L}/\bar{L}$
Normalized Productivity	$z := Z/M$
Normalized Optimal Export Threshold	$\hat{z} := \hat{Z}/M$
Normalized CDF of the Productivity Distribution	$F(z, t) := \Phi(zM(t), t)$
Normalized PDF of the Productivity Distribution	$f(z, t) := M(t)\phi(zM(t), t)$
Expectation of the Normalized Productivity Distribution	$\mathbb{E}[\Psi(z)] := \int_1^\infty \Psi(z)f(z)dz$
Conditional Expectation of the Normalized Productivity Distribution	$\mathbb{E}[\Psi(z) z > \hat{z}] := \int_{\hat{z}}^\infty \Psi(z)\frac{f(z)}{1-F(\hat{z})}dz$
Normalized, Per-capita Real Firm Value Normalized by Real Wages	$v(z, t) := \frac{1}{LMw}V(Z, t)$
Normalized, Per-capita, Real Expenditures on Final Goods (i.e., Output)	$y := \frac{1}{LM}Y$
Normalized, Per-capita Real consumption	$c := \frac{1}{LM}C$
Normalized, Per-capita, Real Adoption Cost Relative to Real Wages	$x := \frac{1}{LMw}X$
Normalized Real Wages	$w := \frac{1}{PM}W$
Normalized, Per-capita, Real, Aggregate Domestic Profits	$\bar{\pi}_d$
Normalized, Per-capita, Real, Aggregate Per-market Export Profits	$\bar{\pi}_x$