The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity

Marc J. Melitz *
Department of Economics
Harvard University
mmelitz@harvard.edu

First version: May 2, 1999
This version: November 5, 1999

Abstract

This paper builds a dynamic industry model with heterogeneous firms that exhibits the mechanisms by which trade causes reallocations of resources among firms in an industry. The paper shows how the exposure to trade will induce only the more productive firms to enter the export market (while some less productive firms continue to produce only for the domestic market) and will simultaneously force the least productive firms to exit. It then shows how further increases in the industry's exposure to trade (driven by trade liberalization or the addition of new trading partners) lead to additional inter-firm reallocations towards more productive firms. These phenomena have been empirically documented but can not be explained by current general equilibrium trade models, because they rely on a representative firm framework. The paper also shows how the aggregate industry productivity growth generated by the reallocations contributes to a welfare gain, thus highlighting a benefit from trade that has not been examined theoretically before.

The paper adapts Hopenhayn's (1992a) dynamic industry model to monopolistic competition in a general equilibrium setting. In so doing, the paper provides an extension of Krugman's (1980) trade model that incorporates firm level productivity differences. Firms with different productivity levels coexist in an industry because each firm faces initial uncertainty concerning its productivity before making an irreversible investment to enter the industry. Entry into the export market is also costly, but the firm's decision to export occurs after it gains knowledge of its productivity.

*I would like to thank, without implicating, Alan Deardorff and Jim Levinsohn for helpful comments and discussions. Funding from the Alfred P. Sloan Foundation is gratefully acknowledged.

1 Introduction

Recent empirical research using longitudinal plant or firm level data in several countries has overwhelmingly substantiated the existence of large and persistent productivity differences among establishments in the same narrowly defined industries. Foster, Haltiwanger and Krizan (1998) summarize this research by concluding that "... within sector differences dwarf between sector differences in behavior." In related work, Haltiwanger (1997, Table1) reports that 4-digit industry effects explain less than 10 percent of the overall variation in the growth rates of output, employment, capital stocks, and productivity across establishments in the U.S. from 1977 to 1987. Complementing this evidence on the extent of within sector heterogeneity, other studies have shown that the bulk of resource reallocations across firms remains internal to the specific sector. Davis and Haltiwanger (1998) summarize this evidence for the U.S. and report that less than 1 in 10 job reallocations reflect employment shifts across sectors. Levinsohn (1999) reports similar numbers for most industries in Chile following wide-reaching trade liberalization. Evidence reported in Roberts and Tybout, eds (1996) confirms that these patterns are not specific to the U.S. and that substantial within sector reallocations between heterogeneous firms are also prevalent in developing countries.

If these large intra-industry reallocations were unrelated to the heterogeneous characteristics of firms, then their separate existence would not necessarily make them important determining factors of industry performance. On the other hand, if the reallocations are related to firm characteristics, then the nature of the link between the two significantly affects several important aspects of industry performance. Although the analysis of this link between firm characteristics and industry evolution is an ongoing research program, enough evidence has been collected to demonstrate its existence and relevance for industry performance. The main firm characteristic found to be empirically linked to intra-industry reallocations is firm productivity.[1] The strongest evidence of this link pertains to firm entry and exit decisions. Productivity differences between entering and exiting firms significantly contribute to aggregate industry productivity changes over time. Additionally, a large number of studies have documented a strong correlation between firm exit and low productivity (firm age is also correlated with exit: younger firms have disproportionately high failure rates). Finally, some studies have also found evidence that reallocations unrelated to entry and exit contribute to

[1]Firm age and capital vintage are other important explanatory characteristics that have been highlighted in some studies, although their impact may be limited to their effect on productivity.

industry productivity growth by redistributing market shares among incumbent firms.[2] A similar reallocation process has also been studied at a higher level of aggregation: Basu and Fernald (1997) find that U.S. aggregate productivity changes across the business cycle are partly driven by expenditure reallocations across sectors with different average productivity levels. The inherent inability of representative firm industry models to explain the contribution of reallocations to industry performance has prompted the development of a theoretical literature of industry dynamics that emphasizes the role of firm level heterogeneity. This literature, along with the previously mentioned empirical evidence, is reviewed in Foster et al. (1998) and Tybout (1996).

This paper adapts one of these recent industry models with heterogeneous firms in order to analyze the role of international trade as a catalyst for inter-firm reallocations within an industry. It then describes how these reallocations affect both industry performance and welfare. The business press often assumes the existence of this catalyst role of trade when describing how exposure to trade has both enhanced the growth opportunities of some firms while simultaneously contributing to the downfall or "downsizing" of other firms in the same industry. Similarly, protection from trade is reported to shelter inefficient firms. Rigorous empirical work has recently corroborated this anecdotal evidence. Bernard and Jensen (1999) (for the U.S.), Aw, Chen and Roberts (1997) (for Taiwan), and Clerides, Lach and Tybout (1998) (for Colombia, Mexico, and Morocco) all find evidence that the causation of the correlation between firm productivity and export status runs from the former to the latter: more productive firms self-select into the export market. Aw et al. (1997) also find evidence suggesting that exposure to trade forces the least productive firms to exit the industry (firms with higher productivity levels relative to the incumbent average exit after the exposure to trade). Both of these selection effects (into the export market and out of the industry) obviously reallocate market shares from less productive firms (who exit) to more productive ones (who export) and therefore contribute to industry productivity growth.[3]

By relying on a representative firm framework (at least at the level of the industry), general equilibrium trade models have largely ignored these intra-industry reallocations and focused instead on other consequences of trade, such as inter-industry reallocations or phenomena affecting all firms in similar ways.[4] This paper attempts to fill this gap by providing a general equilibrium industry

[2]The importance of this phenomenon varies across studies and is cyclically sensitive (see Foster et al. (1998))

[3]Forces other than trade also affect the reallocation of resources within an industry. Olley and Pakes (1996) find that deregulation in the U.S. telecommunications industry increased productivity predominantly through this channel rather than through intra-firm productivity gains.

[4]This last category includes models that assume a direct link between trade and firm level efficiency. In these models, exposure to trade typically increases the efficiency level of all firms through a variety of channels: learning

model with heterogeneous firms that exhibits the mechanisms by which the exposure to trade affects the selection of firms into the industry and the export markets. This model shows how exposure to trade will induce only the more productive firms to enter the export markets (while some less productive firms continue to produce only for the domestic market) and will simultaneously force the least productive firms exit. The paper then shows how further increases in the industry's exposure to trade (driven either by trade liberalization or the addition of new trading partners) lead to additional inter-firm reallocations towards more productive firms. The model thus explains how trade can generate industry productivity growth without necessarily affecting intra-firm efficiency. It also provides a theoretical foundation for the recent empirical findings described above and rigorously shows how trade can contribute to the Darwinian evolution of industries – forcing the least efficient firms to contract or exit while promoting the growth and success of the more efficient ones.

2 Model Background

Incorporating heterogeneity in a dynamic industry setting, where forward looking firms make entry and export decisions, necessarily increases the technical complexity of this model vis-a-vis its representative firm counterparts. In order to reduce this additional complexity, I abstract from some of the firm level dynamic stochastic processes that are typically modeled in the recent industry dynamics literature, while preserving the necessary components that explain how certain characteristics of industries shape their endogenous composition with heterogeneous firms. The main forces explaining the impact of trade on an industry are nevertheless quite intuitive. The opening of new export markets exclusively benefits the more efficient firms, as entry into these markets is costly and can only be afforded by the more efficient firms (who earn higher profits). The competition generated by the entry of the more efficient firms into the domestic market forces the least efficient domestic firms to exit.

This model builds upon Krugman's (1980) analysis of trade in the presence of product differentiation, increasing returns, and monopolistic competition, by incorporating firm level productivity differences. Given the differentiation of goods, these productivity differences may reflect more than

effects, increased scale of production, increased innovation, higher quality or diversity of intermediate inputs, reduction of agency problems between owners and managers. Clerides et al. (1998) and Bernard and Jensen (1999) specifically test whether new exporting firms become more efficient. Neither of these studies finds evidence supporting this hypothesis. Tybout and Westbrook (1995) test and reject the hypothesis that increased productivity in Mexico's growing export industries was driven by increases in the scale of plant production.

just cost differences among products yielding the same utility. Higher productivity can also be interpreted as producing a better product (generating higher utility) at equal cost. The model draws heavily from Hopenhayn's (1992a, 1992b) work on firm and industry productivity dynamics to explain the endogenous selection of heterogeneous firms in an industry. Instead of assuming some immutable and innate ordering of firms from most to least productive, Hopenhayn derives the equilibrium distribution of firm productivity from the profit maximizing decisions of initially identical firms who are uncertain of their initial and future productivity.[5] This paper adapts his model to a monopolistically competitive industry (Hopenhayn only considers competitive firms) in a general equilibrium setting.[6] Although the current model preserves the initial firm uncertainty over productivity and also the endogenous, forward looking, firm entry decision, it greatly simplifies the ensuing firm productivity dynamics. Hopenhayn shows how these dynamics shape the equilibrium distribution of firm productivity and analyzes the impact of these dynamics on firm value and the performance of cohorts of firms over time. This model foregoes this type of analysis and instead relies on the choice of a suitable distribution of initial firm productivity levels in order to generate a realistic shape for the equilibrium distribution of firm productivity. The benefit of this simplification is a gain in the tractability of the model that permits a detailed analysis of the impact of trade on both aggregate industry performance and the relative performance of different types of firms (indexed by productivity). As in Hopenhayn (1992b), I restrict the equilibrium analysis to equilibria that maintain a stable aggregate industry environment over time. Forward looking firms correctly anticipate this aggregate environment when making all decisions (including entry). As the impact of any firm on the industry is assumed to be negligible, the equilibrium preserves a stable aggregate environment even though the fortunes of individual firms change, generating

[5]As was previously mentioned, one of the robust empirical patterns emerging from recent industry studies is that new entrants are much more likely to have lower productivity and exit than do older incumbents. This evidence conflicts with the notion that firm productivity differences are "innate" and known to firms prior to entry. It rather suggests that uncertainty concerning productivity is an important feature that explains the behavior of prospective and new entrants.

[6]Montagna (1995) also adapts Hopenhayn's model to a monopolistic competition environment (in a partial equilibrium setting), but confines the analysis to a static equilibrium with no entry or exit and further constrains the distribution of firm productivity levels to be uniform. Unfortunately, the relevance of her work is difficult to assess because a questionable formulation of equilibrium is used. Although firms with low productivity draws (below a given cutoff) do not produce in the equilibrium, they do not foresee this while making their entry decisions. The equilibrium entry condition is based on the incorrect assumption made by all prospective entrants that they somehow will be constrained to remain in the industry regardless of their productivity draw and possibly endure a perpetual stream of losses. Instead of weighing the present value of potential positive profits against the entry cost, firms ignore the entry cost and weigh the present value of potential positive profits against the present value of losses that are actually never incurred. The equilibrium described in Montagna (1995) thus does not respond to changes in the entry cost, while being quite sensitive to the magnitude of losses that can never materialize.

simultaneous entry and exit.

3 Setup of the Model

Demand

The preferences of a representative consumer are given by a C.E.S. utility function over a continuum of goods indexed by ω:

$$U = \left[\int_{\omega \in \Omega} q(\omega)^{\rho} \, d\omega \right]^{1/\rho}.$$

The measure of the set Ω will represent the mass (or alternatively, the number) of available goods. These goods will be substitutes, implying $0 < \rho < 1$ and an elasticity of substitution between any two goods of $\sigma = \frac{1}{1-\rho} > 1$. As was originally shown by Dixit and Stiglitz (1977), consumer behavior can be modeled by considering the set of varieties consumed as an aggregate good $Q \equiv U$ associated with an aggregate price

$$P = \left[\int_{\omega \in \Omega} p(\omega)^{1-\sigma} \, d\omega \right]^{\frac{1}{1-\sigma}}. \tag{1}$$

These aggregates can then be used to derive the optimal consumption decisions over individual varieties using

$$q(\omega) = Q \left(\frac{p(\omega)}{P} \right)^{-\sigma},$$

which, in turn, yields an expenditure over varieties of

$$r(\omega) = P\,Q \left(\frac{p(\omega)}{P} \right)^{1-\sigma} = R \left(\frac{p(\omega)}{P} \right)^{1-\sigma}, \tag{2}$$

where $R = \int_{\omega \in \Omega} r(\omega) d\omega$ is aggregate expenditure. The quite special symmetric structure of demand implies that the ratio of quantities consumed or of expenditures on any two varieties is uniquely determined by the price ratio between the two varieties:

$$\frac{q(\omega_1)}{q(\omega_2)} = \left(\frac{p(\omega_1)}{p(\omega_2)} \right)^{-\sigma} \quad \text{and} \quad \frac{r(\omega_1)}{r(\omega_2)} = \left(\frac{p(\omega_1)}{p(\omega_2)} \right)^{1-\sigma},$$

where ω_1 and ω_2 are any two varieties.

Production

There is a continuum of firms, each choosing to produce a different variety ω. Production requires only one factor, labor, which is inelastically supplied at its aggregate level L (which also indexes the size of the economy). Firm technology is represented by a cost function that exhibits constant marginal cost with a fixed overhead cost. Labor used is thus a linear function of output q:

$$l = f + \frac{q}{\varphi}.$$

All firms share the same fixed cost $f > 0$ but have different productivity levels indexed by $\varphi > 0$. Regardless of its productivity, each firm faces a residual demand curve with constant elasticity σ and thus chooses the same profit maximizing markup equal to $\frac{\sigma}{\sigma-1} = \frac{1}{\rho}$. This yields a pricing rule

$$p(\varphi) = \frac{w}{\rho\,\varphi}, \tag{3}$$

where w is the common wage rate hereafter normalized to one. Firm profit is then

$$\pi(\varphi) = r(\varphi) - l(\varphi) = \frac{r(\varphi)}{\sigma} - f,$$

where $r(\varphi)$ is firm revenue and $\frac{r(\varphi)}{\sigma}$ is variable profit. $r(\varphi)$ (and hence $\pi(\varphi)$) also depend on the aggregate price and revenue as shown in (2):

$$r(\varphi) = R(P\,\rho\,\varphi)^{\sigma-1} \tag{4}$$

$$\pi(\varphi) = \frac{R}{\sigma}(P\,\rho\,\varphi)^{\sigma-1} - f. \tag{5}$$

On the other hand, the ratios of any two firms' outputs and revenues only depend on the ratio of the two firms' productivity levels:

$$\frac{q(\varphi_1)}{q(\varphi_2)} = \left(\frac{\varphi_1}{\varphi_2}\right)^{\sigma} \qquad \text{and} \qquad \frac{r(\varphi_1)}{r(\varphi_2)} = \left(\frac{\varphi_1}{\varphi_2}\right)^{\sigma-1}. \tag{6}$$

In summary, a more productive firm (higher φ) will be bigger (larger output and revenues), charge a lower price, and earn higher profits than a less productive firm.

Aggregation

An equilibrium will be characterized by a mass M of firms (and hence M goods) and a distribution $\mu(\varphi)$ of productivity levels over a subset of $(0,\infty)$. In such an equilibrium, the aggregate price P defined in (1) can be written

$$P = \left[\int_0^\infty p(\varphi)^{1-\sigma}\, M\,\mu(\varphi)\,d\varphi\right]^{\frac{1}{1-\sigma}},$$

since $M\,\mu(\varphi)\,d\varphi$ represents the mass of firms in the interval $[\varphi,\varphi+d\varphi]$. Using the pricing rule (3), this can be written

$$P = M^{\frac{1}{1-\sigma}}\,p(\tilde{\varphi}), \qquad \text{where} \qquad \tilde{\varphi} = \left[\int_0^\infty \varphi^{\sigma-1}\,\mu(\varphi)\,d\varphi\right]^{\frac{1}{\sigma-1}}. \tag{7}$$

$\tilde{\varphi}$ is a weighted average of the firm productivity levels φ and is independent of the number of firms M.[7] These weights reflect the relative output shares of firms with different productivity levels.[8] Given a distribution $\mu(\varphi)$, the weight assigned to more productive firms increases with the elasticity of substitution σ ($\tilde{\varphi}$ rises with σ): When the elasticity of substitution is high, consumers spend a disproportionate amount of their income on goods with lower prices (produced by the more productive firms) and generate disproportionately larger market shares for these firms.

$\tilde{\varphi}$ also represents aggregate productivity because it completely summarizes all the information in the distribution of productivity levels $\mu(\varphi)$ relevant for all aggregate variables (see appendix):

$$P = M^{\frac{1}{1-\sigma}}\,p(\tilde{\varphi}), \qquad R = PQ = M\,r(\tilde{\varphi}),$$
$$Q = M^{1/\rho}\,q(\tilde{\varphi}), \qquad \Pi = M\,\pi(\tilde{\varphi}),$$

where $R = \int_0^\infty r(\varphi)\,M\mu(\varphi)\,d\varphi$ and $\Pi = \int_0^\infty \pi(\varphi)\,M\mu(\varphi)\,d\varphi$ represent aggregate revenue (or expenditure) and profit. Thus, an industry comprised of M firms with any distribution of productivity levels $\mu(\varphi)$ that yields the same average productivity level $\tilde{\varphi}$ will also induce the same aggregate outcome as an industry with M representative firms sharing the same productivity level $\varphi = \tilde{\varphi}$. This variable will be alternatively referred to as aggregate or average productivity. The welfare level associated with this aggregate outcome can be measured using the indirect utility at the aggregate

[7] Subsequent conditions on the equilibrium $\mu(\varphi)$ must of course ensure that $\tilde{\varphi}$ is finite.

[8] Using $\frac{q(\varphi)}{q(\tilde{\varphi})} = \left(\frac{\varphi}{\tilde{\varphi}}\right)^\sigma$ (see (6)), $\tilde{\varphi}$ can be written as $\tilde{\varphi} = \int_0^\infty \varphi\,\frac{q(\varphi)}{q(\tilde{\varphi})}\,\mu(\varphi)\,d\varphi$, where $\frac{q(\varphi)}{q(\tilde{\varphi})}$ indexes the relative output share of a firm with productivity φ.

price and income level (Y):

$$V(P,Y) = \frac{Y}{P} = Y\, M^{\frac{1}{\sigma-1}}\, \rho\, \tilde{\varphi}.$$

Given a level of aggregate income Y, welfare will increase with M (increased product variety), with ρ (lower markups), and with $\tilde{\varphi}$ (higher aggregate productivity).[9] Let $\bar{r} = \frac{R}{M}$ and $\bar{\pi} = \frac{\Pi}{M}$ represent the average revenue and profit per firm. Note that $\bar{r} = r(\tilde{\varphi})$ and $\bar{\pi} = \pi(\tilde{\varphi})$: the firm with productivity equal to the average $\tilde{\varphi}$ will earn revenue and profit equal to the per-firm average for the industry.

4 Firm Entry and Exit

There is a large (unbounded) pool of prospective entrants into the industry. Prior to entry, firms are identical. To enter, firms must first make an initial investment, modeled as a fixed entry cost $f_e > 0$ (measured in units of labor), which is thereafter sunk. Firms then draw their initial productivity parameter φ from a common distribution $g(\varphi)$.[10] $g(\varphi)$ has positive support over $(0, \infty)$ and has a continuous cumulative distribution $G(\varphi)$. The absence of an upper bound on productivity is assumed only for simplicity; an upper bound can be incorporated in the analysis without qualitatively changing any of the main results.

Upon entry with a low productivity draw, a firm may decide to immediately exit and not produce. If the firm does produce, it then faces a constant (across productivity levels) probability δ in every period of a bad shock that would force it to exit. Although there are some realistic examples of severe shocks that would constrain a firm to exit independently of productivity (such as natural disasters, new regulation, product liability, major changes in consumer tastes), it is also likely that exit may be caused by a series of bad shocks affecting the firm's productivity. This type of firm level process is explicitly modeled by Hopenhayn(1992a, 1992b). He then shows how these firm level productivity dynamics give the equilibrium distribution of productivity levels $\mu(\varphi)$ a different shape than the ex-ante distribution $g(\varphi)$, and determine the ex-ante survival probabilities for a firm,

[9] Benassy (1996) calls the elasticity of welfare with respect to M, $\frac{1}{\sigma-1}$, the taste for variety effect. He shows how one can incorporate this directly into the utility function and break the link between this taste for variety and the markup $1/\rho$ in order to model both as separate parameters. This formulation was also independently derived by Brown, Deardorff and Stern (1996) who use a taste for variety parameter smaller than $\frac{1}{\sigma-1}$ to model computationally the effects of trade liberalization in services.

[10] This captures the fact that firms can not know their own productivity with certainty until they start producing and selling their good. (Recall that productivity differences may reflect cost differences as well as differences in consumer valuations of the good.)

conditional on successful entry. The current model foregoes this type of analysis and assumes that the shape of the equilibrium distribution and the ex-ante survival probabilities are exogenously determined by $g(\varphi)$ and δ. On the other hand, the range of productivity levels, and hence the average productivity level, are endogenously determined. The increased tractability afforded by these simplifications permits the detailed analysis of the impact of trade on this endogenous range of productivity levels and on the distribution of market shares and profits across this range. Since the probability δ and especially the shape of $g(\varphi)$ are left unrestricted, it is extremely unlikely that these simplifications will bias the predictions of the model. Importantly, this simplified industry model will nevertheless generate the main empirical patterns described in the introduction: Since a portion of the firms who exit are those who entered with a low productivity draw (and immediately exit), the overall probability of exit will be negatively correlated with both firm productivity and age.[11] Furthermore, the model also generates the empirical pattern that the average productivity level of all entrants and exiting firms (including the firms whose entry is unsuccessful) is lower than the average productivity of incumbents. The assumption of the exogenous probability of exit δ among incumbents does preclude the analysis of the evolution of firm cohorts after the first period following entry. On the other hand, the model preserves the essential features that separates the cohort of new entrants from the incumbents and explain the endogenous selection of heterogeneous firms into the industry.

As previously mentioned, this paper will only consider steady state equilibria, in which the aggregate variables remain constant over time. As a result, since each firm's productivity level does not change over time, its optimal per period profit level (excluding f_e) will also remain constant (see (5)). An entering firm with productivity φ would then immediately exit if this profit level were negative (and hence never produce), or would produce and earn $\pi(\varphi) \geq 0$ in every period until it is hit with the bad shock and is forced to exit. Assuming that there is no time discounting,[12] each firm's value function will be given by

$$v(\varphi) = \max \left\{ 0, \sum_{t=0}^{\infty} (1-\delta)^t \, \pi(\varphi) \right\} = \max \left\{ 0, \frac{1}{\delta} \, \pi(\varphi) \right\},$$

[11] In this stylized model, firms who enter with low productivity draws exit immediately and do not produce. A period in this model would then correspond, in the real world, to a time span of a couple years during which a new firm makes the decision to stay or exit. The output produced by these new firms who subsequently exit during this period is ignored in this model. Passive learning effects, as in Jovanovic (1982), whereby new firms learn about their unknown productivity levels through noisy cost signals, are probably quite important during this early stage.

[12] Again, this is assumed for simplicity. The probability of exit δ introduces an effect similar to time discounting.

where the dependence of $\pi(\varphi)$ on R and P from (5) is understood. Thus, $\varphi^* = \inf\{\varphi : v(\varphi) > 0\}$ identifies the lowest productivity level (hereafter referred to as the cutoff level) of producing firms (recall that $\pi(\varphi)$ is strictly increasing in φ). Since $\pi(0) = -f$ is negative, $\pi(\varphi^*)$ must be equal to zero. This will be referred to as the zero cutoff profit condition.

Any entering firm drawing a productivity level $\varphi < \varphi^*$ will immediately exit and never produce. Since subsequent firm exit is assumed to be uncorrelated with productivity, the exit process will not affect the equilibrium productivity distribution $\mu(\varphi)$. This distribution must then be determined by the initial productivity draw, conditional on successful entry. Hence, $\mu(\varphi)$ will be given by the conditional of $g(\varphi)$ on $[\varphi^*, \infty)$, that is

$$\mu(\varphi) = \begin{cases} \frac{g(\varphi)}{1-G(\varphi^*)} & \text{if } \varphi \geq \varphi^*, \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

where $p_{in} = 1 - G(\varphi^*)$ is the ex-ante probability of successful entry.[13] This defines the aggregate productivity level $\tilde{\varphi}$ as a function of the cutoff level φ^*:[14]

$$\tilde{\varphi}(\varphi^*) = \left[\frac{1}{1 - G(\varphi^*)} \int_{\varphi^*}^{\infty} \varphi^{\sigma-1}\, g(\varphi)\, d\varphi\right]^{\frac{1}{\sigma-1}}. \tag{9}$$

The assumption of a finite $\tilde{\varphi}$ imposes certain restrictions on the size of the upper tail of the distribution $g(\varphi)$: the $(\sigma - 1)^{\text{th}}$ uncentered moment of $g(\varphi)$ must be finite. (8) clearly shows how the shape of the equilibrium distribution of productivity levels is tied to the exogenous ex-ante distribution $g(\varphi)$ while allowing the range of productivity levels (indexed by the cutoff φ^*) to be endogenously determined.[15] (9) then shows how the endogenous range affects the determination of the aggregate productivity level.

[13]The equilibrium distribution $\mu(\varphi)$ can be determined from the distribution of initial productivity with certainty by applying a law of large numbers to $g(\varphi)$: recall that individual firms have zero mass, so a positive mass of entering firms requires an infinite number of draws from the distribution $g(\varphi)$. Although some technical problems may arise when applying a law of large numbers to a continuum of random variables, this will not be the case in the current situation. See Hopenhayn (1992a, Note 5) and the reference to Feldman and Gilles (1985) for further details.

[14]This dependence of $\tilde{\varphi}$ on φ^* is understood when it is subsequently written without its argument.

[15](8) also illustrates the earlier discussion concerning firm cohorts. All cohorts of incumbent firms will have the same distribution of productivity levels $\mu(\varphi)$. The model thus does not differentiate between cohorts of incumbent firms. These incumbents essentially form one cohort. This cohort is then critically differentiated from that formed by new entrants, whose distribution of productivity levels is given by $g(\varphi)$.

Zero Cutoff Profit Condition

As shown in (9), the average productivity level $\tilde{\varphi}$ is completely determined by the cutoff productivity level φ^* (given the exogenous distribution $g(\varphi)$). Thus, the average profit and revenue levels are also tied to the cutoff level φ^* (see (6)):

$$\bar{r} = r(\tilde{\varphi}) = \left(\frac{\tilde{\varphi}(\varphi^*)}{\varphi^*}\right)^{\sigma-1} r(\varphi^*) \quad \text{and} \quad \bar{\pi} = \pi(\tilde{\varphi}) = \left(\frac{\tilde{\varphi}(\varphi^*)}{\varphi^*}\right)^{\sigma-1} \frac{r(\varphi^*)}{\sigma} - f$$

The zero cutoff profit condition, by pinning down the revenue of the cutoff firm, then implies a relationship between the average profit per firm and the cutoff productivity level:

$$\begin{aligned} \pi(\varphi^*) = 0 &\iff \frac{r(\varphi^*)}{\sigma} - f = 0 \\ &\iff r(\varphi^*) = \sigma f \\ &\iff \bar{\pi} = \left(\frac{\tilde{\varphi}(\varphi^*)}{\varphi^*}\right)^{\sigma-1} f - f \\ &\iff \bar{\pi} = f\, k(\varphi^*), \end{aligned} \tag{10}$$

where $k(\varphi^*) = \left(\frac{\tilde{\varphi}(\varphi^*)}{\varphi^*}\right)^{\sigma-1} - 1 = \frac{\bar{r}}{r(\varphi^*)} - 1$ is the percentage difference between the average and cutoff firm revenues, as a function of the cutoff productivity level φ^*.

Free Entry and the Value of Firms

Since all incumbent firms (with $\varphi \geq \varphi^*$) – other than the cutoff firm – earn positive profits, the average profit level $\bar{\pi}$ must be positive. In fact, the expectation of future positive profits is the only reason for firms to consider sinking the investment cost f_e required for entry. Let $\bar{v}$ represent the present value of the average profit flows:

$$\bar{v} = \sum_{t=0}^{\infty} (1-\delta)^t\, \bar{\pi} = \frac{1}{\delta}\bar{\pi}.$$

$\bar{v}$ is also the average value of firms, conditional on successful entry: $\bar{v} = \int_{\varphi^*}^{\infty} v(\varphi)\, \mu(\varphi)\, d\varphi$. Now define v_e to be the net value of entry:

$$v_e = p_{in}\, \bar{v} - f_e = \frac{1 - G(\varphi^*)}{\delta}\bar{\pi} - f_e. \tag{11}$$

If this value were negative, no firm would want to enter. In any equilibrium where entry is unrestricted, this value could further not be positive since the mass of prospective entrants is unbounded.

5 Equilibrium in a Closed Economy

A stationary equilibrium is defined by constant aggregate variables over time and the free entry of firms into the industry. Such an equilibrium is completely referenced by a triplet (φ^*, P, R) satisfying the following conditions (see (10) and (11)):

$$\begin{cases} \pi(\varphi^*) = 0 & \text{(Zero Cutoff Profit)} \\ v_e = 0 & \text{(Free Entry)} \end{cases} \iff \begin{cases} \bar{\pi} = f\,k(\varphi^*) & \text{(ZCP)} \\ \bar{\pi} = \frac{\delta f_e}{1-G(\varphi^*)} & \text{(FE).} \end{cases} \tag{12}$$

Since the aggregate price index (P), revenue (R), and productivity level ($\tilde{\varphi}$) remain constant, all other aggregate variables must also remain constant.[16] The equilibrium mass of firms M is determined from the aggregate price index and aggregate productivity level using the aggregation condition $P = M^{\frac{1}{1-\sigma}}\, p(\tilde{\varphi}) = M^{\frac{1}{1-\sigma}}\, \frac{1}{\rho\tilde{\varphi}}$. This stationary equilibrium also requires a mass M_e of new entrants in every period, such that the mass of successful entrants, $p_{in}M_e$, exactly replaces the mass δM of incumbents who are hit with the bad shock and exit: $p_{in}M_e = \delta M$ (This will be subsequently referred to as the aggregate stability condition.) The equilibrium distribution of productivity $\mu(\varphi)$ is not affected by this simultaneous entry and exit since the successful entrants and failing incumbents have the same distribution over productivity levels.

The labor used by these new entrants for investment purposes must, of course, be reflected in the accounting for aggregate labor L, and affects the aggregate labor available for production. Aggregate income must then reflect payments made to both types of workers:

$$Y = w\,L = L = L_p + L_e,$$

where L_p and L_e represent, respectively, the aggregate labor used for production and investment (by new entrants). The payments to production workers L_p must also equal the difference between aggregate revenue and profit: $L_p = R - \Pi$. This is also the labor market clearing condition for production workers. The market clearing condition for investment workers requires $L_e = M_e\, f_e$. Using the aggregate stability condition ($p_{in}M_e = \delta M$) and the free entry condition $\left(\bar{\pi} = \frac{\delta f_e}{1-G(\varphi^*)}\right)$,

[16] Recall that aggregate productivity $\tilde{\varphi}$ will be a function of the cutoff level φ^* as shown in (9).

L_e can be written:

$$L_e = M_e\, f_e = \frac{\delta M}{p_{in}} f_e = M\,\bar{\pi} = \Pi.$$

The equilibrium conditions thus require that the total payment to investment workers match the aggregate profit derived from production. This, in turn, requires that the aggregate revenue derived from production match the aggregate income for the entire economy:

$$R = L_p + \Pi = L_p + L_e \iff R = Y = L. \tag{13}$$

The aggregate industry revenue is thus fixed by the exogenous aggregate labor parameter (which indexes country size).[17] The aggregate price index P can then be calculated from the aggregation condition:

$$\begin{aligned} P &= M^{\frac{1}{1-\sigma}}\, p(\tilde{\varphi}) \\ &= \left(\frac{R}{\bar{r}}\right)^{\frac{1}{1-\sigma}} p(\tilde{\varphi}) \\ &= \left(\frac{R}{\sigma\,(\bar{\pi}+f)}\right)^{\frac{1}{1-\sigma}} \frac{1}{\rho\,\tilde{\varphi}(\varphi^*)}. \end{aligned} \tag{14}$$

Conditions (12), (13), and (14) completely determine the equilibrium triplet (P, R, φ^*).

Existence and Uniqueness of the Equilibrium

The Free Entry FE and Zero Cutoff Profit ZCP conditions represent two different relationships between the average profit level $\bar{\pi}$ and the cutoff productivity level φ^* (see (12)). These two conditions are discussed in further detail below. I first summarize their important properties for the determination of the equilibrium values of φ^* and $\bar{\pi}$: In (φ, π) space the FE curve is increasing and is cut by the ZCP curve only once from above. This ensures the existence and uniqueness of the equilibrium φ^* and $\bar{\pi}$. Furthermore, upward (downward) shifts of the ZCP curve or downward (upward) shifts of the FE curve must lead to an increase (decrease) in the equilibrium φ^*. The determination of the equilibrium φ^* and $\bar{\pi}$ is graphically shown in Figure 1.[18]

[17] It is important to emphasize that this result is not a direct consequence of aggregation and market clearing conditions: it is a property of the model's stationary equilibrium.

[18] Although the ZCP curve must cut the FE curve from above, it is not necessarily downward sloping as represented in the graph. The following discussion provides some mild additional assumptions on the shape of $g(\varphi)$ that ensure

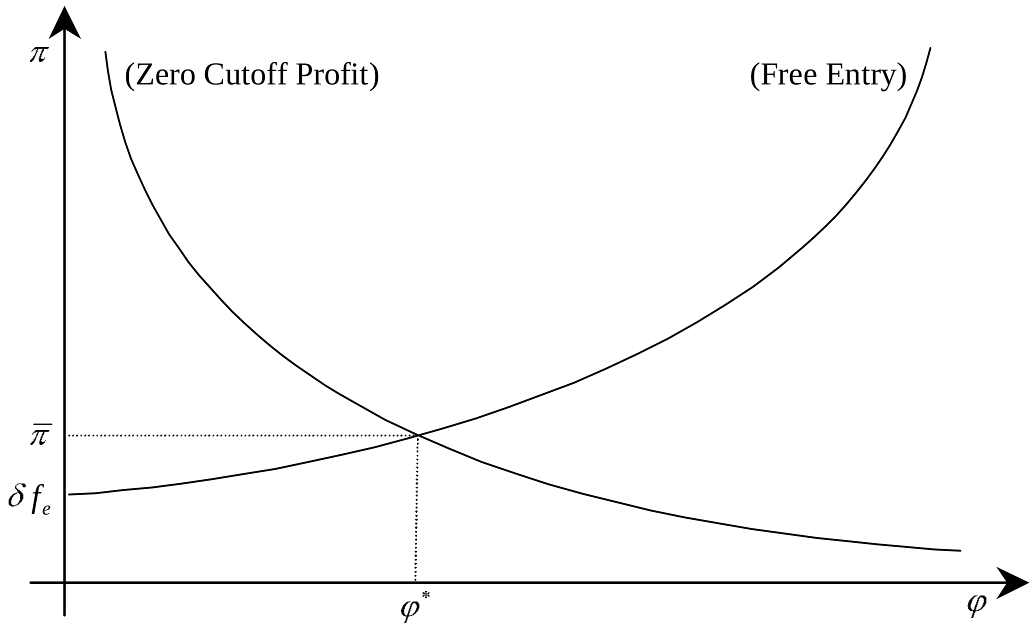

Figure 1: Determination of the Equilibrium Cutoff φ^* and Average Profit $\bar{\pi}$

Inspection of the FE condition $\left(\bar{\pi} = \frac{\delta f_e}{1-G(\varphi^*)}\right)$ reveals that it represents an increasing relationship between these two variables. Along the FE curve, $\bar{\pi}$ increases from δf_e to infinity for $\varphi^* \in (0, \infty)$:[19] As φ^* increases, the probability of successful entry $\left(p_{in} = 1 - G(\varphi^*)\right)$ decreases – average profits must therefore increase for firms to remain indifferent about entry.

The relationship between $\bar{\pi}$ and φ^* implied by the zero cutoff profit condition $\left(\bar{\pi} = f\,k(\varphi^*)\right)$ will depend on the properties of $k(\varphi)$, which are in turn determined by the properties of the distribution $g(\varphi)$ and the elasticity σ. Recall that $k(\varphi^*) = \left(\frac{\tilde{\varphi}(\varphi^*)}{\varphi^*}\right)^{\sigma-1} - 1 = \frac{\bar{r}}{r(\varphi^*)} - 1$ represents the percentage difference between the average and cutoff firm revenues. As the cutoff level φ^* goes to zero, the revenue of the cutoff firm also goes to zero (as its price then becomes infinite – see (3) and (4)). Since the average revenue level is always positive ($\tilde{\varphi} > 0$, and hence $r(\tilde{\varphi}) > 0$, even when $\varphi^* \to 0$), the ratio of the average to cutoff firm revenue becomes infinite as φ^*goes to zero: Along the ZCP curve, $\bar{\pi} \to \infty$ as $\varphi^* \to 0$. Further properties of $k(\varphi)$ require some extra regularity assumptions on the distribution $g(\varphi)$. If $g(\varphi)$ belongs to most of the common families of distributions (including the lognormal, exponential, Gamma, or Weibull distributions or truncations on $(0, \infty)$ of the normal, logistic, extreme value, or Laplace distributions), then $k(\varphi)$ will monotonically decrease to zero on

that the ZCP curve monotonically decreases to zero as shown in the graph.

[19]Since the cumulative distribution $G(\varphi)$ must be increasing from 0 to 1 on $(0, \infty)$.

$(0,\infty)$.[20] In these cases, $\bar{\pi}$ decreases from infinity to zero for $\varphi^* \in (0,\infty)$ along the ZCP curve, as shown in Figure 1 . The regularity conditions on $g(\varphi)$ ensure that an increase in the cutoff level φ^* redistributes the mass of incumbent firms towards the cutoff level. This pushes the average productivity level $\tilde{\varphi}$ closer to the cutoff φ^*, and hence reduces the percentage difference between the revenues of the average and cutoff firms.

As was previously mentioned, the ZCP curve must cut the FE curve only once from above. This property, which ensures the existence and uniqueness of the equilibrium and determines the direction of some comparative statics, does not depend on the additional restrictions on $g(\varphi)$ that were just discussed (see the appendix for a proof).

Analysis of the Equilibrium

As was just shown, the equilibrium productivity cutoff level, φ^*, and average firm profit, $\bar{\pi}$, do not depend on the country size L. Hence, the equilibrium distribution of productivity levels $\mu(\varphi)$ and average productivity level $\tilde{\varphi}$ will also be independent of country size. Average firm revenue $\bar{r} = \sigma(\bar{\pi} + f)$ will also be independent of L and can be used to determine the number of firms:

$$M = \frac{R}{\bar{r}} = \frac{L}{\sigma(\bar{\pi} + f)}. \tag{15}$$

Hence, a large and small country will share the same firm level variables (same φ^*, $\tilde{\varphi}$, $\bar{r}$, $\bar{\pi}$). The large country will just have more firms in an amount proportional to its country size. This larger number of firms will nevertheless be identically distributed over the same productivity range $[\varphi^*, \infty)$ as will be the firms from the small country. Welfare per worker W, which is given by

$$W = \frac{V(P,Y)}{L} = P^{-1} = M^{\frac{1}{\sigma-1}}\rho\tilde{\varphi}, \tag{16}$$

will be higher in the large country due only to increased product variety. This influence of country size on the determination of aggregate variables is identical to that derived by Krugman (1980) with representative firms. Once $\tilde{\varphi}$ and $\bar{\pi}$ are determined, the aggregate outcome predicted by this model is identical to one generated by an economy with representative firms who share the same productivity level $\tilde{\varphi}$ and profit level $\bar{\pi}$. On the other hand, this model with heterogeneous firms explains how the aggregate productivity level $\tilde{\varphi}$ and the average firm profit level $\bar{\pi}$ are endogenously

[20] Sufficient conditions for this property are that $\frac{g(\varphi)\varphi}{1-G(\varphi)}$ be increasing and unbounded from above on $(0,\infty)$.

determined and how both can change in response to various shocks. In particular, a country's production technology (referenced by the distribution $g(\varphi)$) need not change in order to induce changes in aggregate productivity. In the following sections, I argue that the exposure of a country to trade creates precisely the type of shock that induces reallocations between firms and generates increases in aggregate productivity. These results can not be explained by representative firm models where the aggregate productivity level is exogenously given as the productivity level common to all firms. Changes in aggregate productivity can then only result from changes in firm level technology and not from reallocations.

Comparative Statics

Before using this model to analyze the industry and firm level responses to changes in trade regime, two of the comparative statics of the closed economy model are briefly described. An increase in the sunk entry cost f_e will shift up the FE curve and lower the equilibrium cutoff level φ^*. Aggregate productivity must then decrease with the cutoff level ($\tilde{\varphi}$ is an increasing function of φ^*). The change in $\bar{\pi}$ depends on the direction of the ZCP curve. If it is downward sloping (the most likely case) then $\bar{\pi}$ will rise. In this case, both the change in $\bar{\pi}$ and φ^* contribute to defray the higher investment cost f_e: the lower cutoff φ^* increases the ex-ante probability of successful entry (p_{in}) while the higher $\bar{\pi}$ raises the expected gain from successful entry. The higher level of average profit is associated with a reduction in product variety (see (15)). This lower product variety and the lower aggregate productivity then both contribute to a welfare loss (see (16)). The sign of the welfare change does not depend on the additional assumption that the ZCP curve be downward sloping: a rise in the investment cost will always generate a welfare loss (see appendix). Also, note that if an upper bound on firm productivity is incorporated into this model, then the possibility of representative firms is obtained as a limiting case when the entry cost goes to zero. In this case, the cutoff level φ^* and average level $\tilde{\varphi}$ are both pushed towards the productivity upper bound and the average profit level is driven to zero, as shown in Figure 2.

An increase in the fixed production cost f will shift up the ZCP curve and therefore raise the equilibrium cutoff level φ^*, along with aggregate productivity: firms previously producing with low productivity levels can no longer earn positive profits, and they exit. The average profit level $\bar{\pi}$ must increase, as the FE curve is upward sloping . The higher profit level will, in turn, imply a decrease in product variety (see (15)). The product variety decrease and the aggregate productivity increase have opposite effects on welfare. In the appendix, it is shown that the sign of the welfare

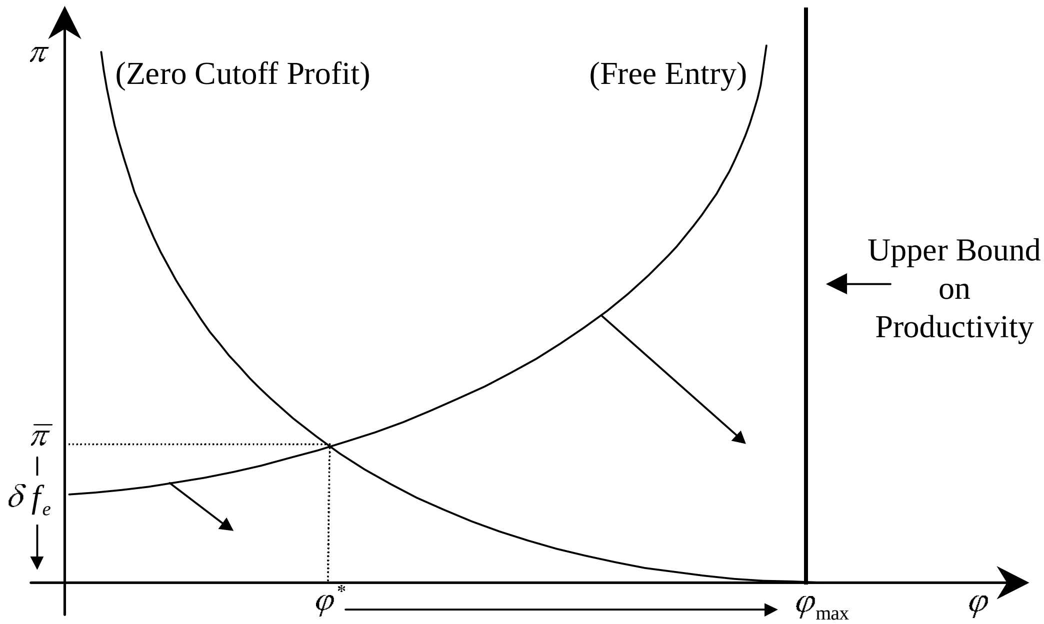

Figure 2: Representative Firms as a Limiting Equilibrium When $f_e \to 0$

change is negative and hence that the lower product variety effect dominates that of the higher aggregate productivity. This yields the reasonable property that higher production costs have an overall negative impact on welfare.

6 Overview and Assumptions of the Open Economy Model

I now examine the impact of trade in a world (or trading group) that is composed of countries whose economies are of the type that was previously described.[21] When there are no additional costs associated with trade, then trade allows the individual countries to replicate the outcome of the integrated world economy.[22] Trade then provides the same opportunities to an open economy as would an increase in country size to a closed economy. As was previously discussed, an increase in country size has no effect on firm level outcomes. The transition to trade will thus not affect any of the firm level variables $(\varphi^*, \tilde{\varphi}, \bar{r}, \bar{\pi})$: The same number of firms in each country produce at the same output levels and earn the same profits as they did in the closed economy. All firms in a

[21] Although I will refer to these trading economies as countries, one could also consider these to be regions within a country or groups of closely integrated countries. The appropriateness and relevance of this choice of economic entity will become clear once trading costs between economies are introduced.

[22] Consumers in every country have access to the same bundle of goods at the same aggregate price index. Firms behave as if they were selling their product in the integrated world market. The FE and ZCP conditions will be identical across countries and will not be affected by the transition to trade.

given country divide their sales between domestic and foreign consumers, based on the size of their country relative to the integrated world economy. Thus, in the absence of any costs to trade, the existence of firm heterogeneity does not affect the impact of trade. This impact is identical to the one described by Krugman (1980) with representative firms: Although firms are not affected by the transition to trade, consumers enjoy welfare gains driven by the increase in product variety.[23]

On the other hand, there are strong reasons to believe that the fixed costs associated with exporting are significant. The existence of such costs provides the most consistent explanation for the widely observed pattern that, across countries and industries, certain firms do not export while others in the same narrowly defined industry do. Interviews with managers making export decisions confirm that firms in differentiated product industries face significant fixed costs associated with the entry into export markets (see Roberts and Tybout (1997b)): A firm must find and inform foreign buyers about its product and learn about the foreign market. It must then research the foreign regulatory environment and adapt its product to ensure that it conforms to foreign standards (which include testing, packaging, and labeling requirements). An exporting firm must also set up new distribution channels in the foreign country and conform to all the shipping rules specified by the foreign customs agency. Although some of these costs can not be avoided, others are often manipulated by governments in order to erect non-tariff barriers (NTBs) to trade. Regardless of their origin, these costs are most appropriately modeled as independent of the firm's export volume decision.[24]

When there is uncertainty concerning the export market, the timing and sunk nature of the costs become quite relevant for the export decision (most of the previously mentioned costs must be sunk prior to entry into the export market).[25] The strong and robust empirical correlations at

[23]The irrelevance of firm heterogeneity for the impact of trade is not just a consequence of negligible trade costs. The assumption of an exogenously fixed elasticity of substitution between varieties also plays a significant role in this result. The presence of heterogeneity (even in the absence of trade costs) plays a significant role in determining the impact of trade once this assumption is dropped. In a separate appendix (available upon request to the author), the current model is modified by allowing the elasticity of substitution to endogenously increase with product variety. This link between trade and the elasticity of substitution was studied by Krugman (1979) with representative firms. In the context of the current model, the appendix shows how the size of the economy then affects the aggregate productivity level and the skewness of market shares and profits across firms with different productivity levels. Larger economies have higher aggregate productivity levels – even though they have the same firm level technology index by $g(\varphi)$. Therefore, even in the absence of trade costs, trade increases the size of the "world" economy and induces reallocations of market shares and profits towards more productive firms and generates an aggregate productivity gain.

[24]The modeling of a fixed export cost is not new. Bernard and Jensen (1999), Clerides et al. (1998), Roberts and Tybout (1997a), and Roberts, Sullivan and Tybout (1995) all introduce a fixed export cost into the theoretical sections of their work in order to explain the self-selection of firms into the export market. However, these analyses are restricted to a partial equilibrium setting in which the distribution of firm productivity levels is fixed.

[25]Roberts and Tybout (1997a) find that the sunk nature of these costs and the foreign market uncertainty play a

the firm level between export status and productivity suggest that the export market entry decision occurs after the firm gains knowledge of its productivity, and hence that uncertainty concerning the export markets is not predominantly about productivity (as is the uncertainty prior to entry into the industry). I therefore assume that a firm who wishes to export must make an initial fixed investment, but that this investment decision occurs after the firm's productivity is revealed. For simplicity, I do not model any additional uncertainty concerning the export markets.

Although the size of a country relative to the rest of the world (which constitutes its trading partners) is left unrestricted, I do assume that the world (or trading group) is comprised of some number of identical countries. In other words, a representative country framework is assumed. This assumption is made in order to ensure factor price equalization across countries and hence focus the analysis on firm selection effects that are independent of wage differences. In this model with fixed export costs, countries who differ only in country size can exhibit different wage rates in the equilibrium with trade. These wage differences then induce further firm selection effects and aggregate productivity differences across countries.[26] I thus assume that the economy under study can trade with $n \geq 1$ other countries (the world is then comprised of $n+1 \geq 2$ countries). Firms can export their products to any country, although entry into each of these export markets requires a fixed investment cost of $f_{ex} > 0$ (measured in units of labor).[27] Regardless of export status, a firm still incurs the same overhead production cost f. For simplicity, the variable costs associated with trade (such as transportation and tariffs) are assumed to be negligible and are not modeled.

7 Equilibrium in the Open Economy

The symmetry assumption ensures that all countries share the same wage, which is still normalized to one, and also share the same aggregate variables. Given the absence of variable trade costs, a firm who exports sells its output at the same price on domestic and foreign markets: $p(\varphi) = \frac{w}{\rho\varphi} = \frac{1}{\rho\varphi}$ (see (3)). This firm then also earns the same revenue from its sales to any country – including its

significant role in explaining hysteresis effects associated with firm level export decisions in Colombia.

[26]In these asymmetric equilibria with fixed export costs, large countries enjoys higher aggregate productivity, welfare, and wages relative to smaller countries. The analysis of these equilibria is left for future work.

[27]The restriction that the export entry costs are equal across countries can be relaxed. The effects of such costs differences are discussed at the end of the following section.

domestic sales (see (4)):[28]

$$r_c(\varphi) = R\,(P\rho\varphi)^{\sigma-1}, \tag{17}$$

where R and P denote the aggregate expenditure and price index in every country. The balance of payments condition implies that R also represents the aggregate revenue of firms in any country. The combined revenue of a firm, $r(\varphi)$, thus depends on its export status:

$$r(\varphi) = \begin{cases} r_c(\varphi) & \text{if the firm does not export,} \\ (n+1)\,r_c(\varphi) & \text{if the firm exports to all countries.} \end{cases} \tag{18}$$

If some firms do not export, then there no longer exists an integrated world market for all goods. Even though the symmetry assumption ensures that all the characteristics of the goods available in every country are similar, the actual bundle of goods available will be different across countries: consumers in each country have access to goods (produced by the non-exporting firms) that are not available to consumers in any other country.

Firm Entry, Exit, and Export Status

All the exogenous factors affecting firm entry, exit, and productivity levels remain unchanged by trade. Prior to entry, firms face the same ex-ante distribution of productivity levels $g(\varphi)$. Firms whose entry is successful produce with the same productivity level φ in every period. They all face the same probability δ of a bad shock that would force them to exit. Upon entry with a low productivity draw, a firm may decide to immediately exit and not produce. In a stationary equilibrium, any incumbent firm with productivity φ earns variable profits $\frac{r_c(\varphi)}{\sigma}$ in every period from its sales to any given country. Since the export cost is assumed equal across countries, a firm will either export to all countries in every period or never export.[29] Given that the export decision occurs after firms know their productivity φ, and since there is no additional export market uncertainty, firms are indifferent between paying the one time investment cost f_{ex}, or paying the amortized per period portion of this cost $f_x = \delta f_{ex}$ in every period (as before, there is no additional time discounting other than the probability of the exit inducing shock δ). This per-period representation of the export cost is henceforth adopted for notational simplicity. In the stationary equilibrium,

[28] The c subscript on this revenue function indicates that the revenue is specific to sales *in* a given country.
[29] As was previously mentioned, this restriction can be relaxed. See the discussion at the end of this section.

the aggregate labor resources used in every period to cover the export costs do not depend on this choice of representation.[30] The per-period profit flow of any exporting firm then reflects the per-period fixed cost f_x, which is incurred per export country.

Since no firm will ever export and not also produce for its domestic market,[31] each firm's profit can be separated into portions earned from domestic sales ($\pi_d(\varphi)$) and export sales per country ($\pi_x(\varphi)$) by accounting for the entire overhead production cost in domestic profit:

$$\pi_d(\varphi) = \frac{r_c(\varphi)}{\sigma} - f \qquad \text{and} \qquad \pi_x(\varphi) = \frac{r_c(\varphi)}{\sigma} - f_x. \tag{19}$$

A firm who produces for its domestic market exports to all n countries if $\pi_x(\varphi) \geq 0$. Each firm's combined profit can then be written:

$$\pi(\varphi) = \pi_d(\varphi) + \max\left\{0, n\, \pi_x(\varphi)\right\}.$$

As before, a firm's value is the present value (discounted by the probability of the bad shock) of its profit flows, $v(\varphi) = \max\left\{0, \frac{1}{\delta}\pi(\varphi)\right\}$, and $\varphi^* = \inf\{\varphi : v(\varphi) > 0\}$ still identifies the cutoff productivity level for successful entry into the industry. Additionally, $\varphi_x^* = \inf\{\varphi : \varphi \geq \varphi^* \text{ and } \pi_x(\varphi) > 0\}$ now represents the cutoff productivity level for exporting firms. If $\varphi_x^* > \varphi^*$, then some firms (with productivity levels between φ^* and φ_x^*) produce exclusively for their domestic market. These firms earn $\pi(\varphi) = \pi_d(\varphi) \geq 0$ since $\pi_x(\varphi) < 0$. This can only be possible when $f_x > f$ (see (19)). If $\varphi_x^* = \varphi^*$, then all firms in the industry export. In this case, the cutoff firm (with productivity level $\varphi^* = \varphi_x^*$) earns zero total profit ($\pi(\varphi^*) = \pi_d(\varphi^*) + n\,\pi_x(\varphi^*) = 0$) and non-negative export profit ($\pi_x(\varphi^*) \geq 0$). This will only be possible when $f_x \leq f$. Thus, the size of the fixed export cost relative to the overhead production cost will determine whether firms in an industry are partitioned by export status. When this export cost attains a certain level ($f_x > f$) then partitioning will occur based on the firms' productivity levels. The relative level of the export cost also determines the

[30] In one case, only the new entrants who export expend resources to cover the full investment cost f_{ex}. In the other case, all exporting firms expend resources to cover the smaller amortized portion of the cost $f_x = \delta f_{ex}$. In equilibrium, the ratio of new exporters to all exporters is δ (see appendix), so the same aggregate labor resources are expended in either case.

[31] A firm would earn strictly higher profits by also producing for its domestic market since the associated variable profit $\frac{r_c(\varphi)}{\sigma}$ is always positive and the overhead production cost f is already incurred.

form of the zero cutoff profit condition as follows:

$$
\begin{array}{cc}
\underline{\text{Case 1: } f_x > f} & \underline{\text{Case 2: } f_x \leq f} \\
\text{(Some firms do not export)} & \text{(All firms export)} \\
\varphi_x^* > \varphi^* & \varphi_x^* = \varphi^* \\
\pi_d(\varphi^*) = 0 \quad \text{and} \quad \pi_x(\varphi_x^*) = 0, & \pi(\varphi^*) = \pi_d(\varphi^*) + n\,\pi_x(\varphi^*) = 0.
\end{array}
\tag{20}
$$

The partitioning in Case 1 can obviously not occur in a representative firm model, where a fixed export cost would either preclude all firms from exporting (if $f_x > f$) or otherwise permit all of them to export (if $f_x \leq f$).[32]

As before, the equilibrium distribution of productivity levels for incumbent firms, $\mu(\varphi)$, is determined by the ex-ante distribution of productivity levels, conditional on successful entry: $\mu(\varphi) = \frac{g(\varphi)}{1-G(\varphi^*)}$, $\forall\varphi \geq \varphi^*$. $p_{in} = 1 - G(\varphi^*)$ still identifies the ex-ante probability of successful entry. Furthermore, $p_x = \frac{1-G(\varphi_x^*)}{1-G(\varphi^*)}$ now represents the ex-ante probability that one of these successful firms will export. p_x must then also represent the ex-post fraction of firms that export. Let M denote the equilibrium mass of incumbent firms in any country. $M_x = p_x\,M$ then represents the mass of exporting firms while $M_t = M + n\,M_x$ represents the total mass of varieties available to consumers in any country (or alternatively, the total mass of firms competing in any country).

Aggregation

Using the same weighted average function defined in (9), let $\tilde{\varphi} = \tilde{\varphi}(\varphi^*)$ and $\tilde{\varphi}_x = \tilde{\varphi}(\varphi_x^*)$ denote the average productivity levels of, respectively, all firms and exporting firms only. The average productivity across all firms, $\tilde{\varphi}$, is based only on domestic market share differences between firms (as reflected by differences in the firms' productivity levels). If some firms do not export, then this average will not reflect the additional export shares of the more productive firms. Let $\tilde{\varphi}_t$ be the weighted average of productivity levels that reflects the combined domestic and export shares of all firms. This average is obtained by multiplying by $n + 1$ the weights of exporting firms relative to non-exporting firms (recall from (18) that a firm multiplies its output and revenues by $n + 1$ when

[32] If the fixed export cost were exactly equal to the overhead production cost ($f_x = f$) then the representative firms would be indifferent between exporting and only producing for the domestic market. There is no reason to assume that this precise equality would hold.

it exports to all countries):

$$\tilde{\varphi}_t = \left[\frac{1}{[G(\varphi_x^*) - G(\varphi^*)] + (n+1)[1 - G(\varphi_x^*)]} \int_{\varphi^*}^{\infty} \varphi^{\sigma-1} \, \chi(\varphi) \, g(\varphi) \, d\varphi \right]^{\frac{1}{\sigma-1}}, \tag{21}$$

$$\text{where } \chi(\varphi) = \begin{cases} 1 & \text{if } \varphi \le \varphi_x^*, \\ n+1 & \text{if } \varphi > \varphi_x^*. \end{cases}$$

By symmetry across countries, $\tilde{\varphi}_t$ also represents the average productivity level of all firms (domestic and foreign) competing in any single country (weighted by each firm's share of that country's marekt). That is, $\tilde{\varphi}_t$ is the weighted average of $\tilde{\varphi}$ (the average productivity of domestic firms selling in their home market) and $\tilde{\varphi}_x$ (the average productivity of foreign firms – from n different countries – exporting to the home market), and can also be written:[33]

$$\tilde{\varphi}_t = \left[\frac{1}{1 + p_x \, n} \left(\tilde{\varphi}^{\sigma-1} + p_x \, n \, \tilde{\varphi}_x^{\sigma-1} \right) \right]^{\frac{1}{\sigma-1}}. \tag{22}$$

The aggregate outcome in any country is thus equivalent to one in a closed economy with M_t firms and aggregate productivity level $\tilde{\varphi}_t$. The aggregation properties derived earlier for the closed economy can then be applied. The price index of the bundle of goods available in any country is thus

$$P = M_t^{\frac{1}{1-\sigma}} p(\tilde{\varphi}_t) = M_t^{\frac{1}{1-\sigma}} \frac{1}{\rho \, \tilde{\varphi}_t}, \tag{23}$$

while $r_c(\tilde{\varphi}_t) = \frac{R}{M_t}$ represents the average expenditure that consumers spend on any single variety.

As was previously mentioned, the balance of payments condition implies that R also represents the aggregate revenue of all M domestic firms in any given country. As before, let Π denote the

[33]The two types of averages are best illustrated using a simple example: Assume there are only two countries (A and B) and two firms in each country with productivity levels φ_1 and φ_2. Let φ_i^c denote the firm from country c with productivity level φ_i. Further assume that $\varphi_1 < \varphi_2$ and that only firms φ_2^A and φ_2^B export. Let $r_1 = r_c(\varphi_1)$ and $r_2 = r_c(\varphi_2)$ denote the domestic revenues of the firms in either country. Firms φ_2^A and φ_2^B also earn export sales equal to r_2, so their combined revenues are $2\,r_2$. Aggregate expenditure or revenue in either country is $R = r_1 + 2\,r_2$. There are then two types of productivity averages for country A: the first type averages the productivity levels of firms from country A (φ_1^A and φ_2^A), weighted by each firm's combined market share in the world (r_1/R and $2\,r_2/R$). The second type averages the productivity levels of all firms competing in country A (φ_1^A, φ_2^A, and φ_2^B), weighted by each firm's market share *in* country A (r_1/R, r_2/R, and r_2/R). The two averages must be equal since φ_2^A's export sales to country B, r_2, exactly matches φ_2^B's export sales to country A. The equality of the productivity averages remains true with $n+1$ countries and a continuum of firms: For any given firm from country A (φ^A) who exports to the other n countries, there are n other identical firms in each country whose combined export sales to country A exactly match φ^A's total exports to the n countries.

aggregate profit earned by these firms. Then,

$$R = \int_{\varphi^*}^{\infty} r_c(\varphi)\, M\mu(\varphi)\; d\varphi + n \int_{\varphi_x^*}^{\infty} r_c(\varphi)\, M\mu(\varphi)\; d\varphi$$
$$= R_d + n\, R_x,$$

and

$$\Pi = \int_{\varphi^*}^{\infty} \pi_d(\varphi)\, M\mu(\varphi)\; d\varphi + n \int_{\varphi_x^*}^{\infty} \pi_x(\varphi)\, M\mu(\varphi)\; d\varphi$$
$$= \Pi_d + n\, \Pi_x,$$

where R_d and Π_d represent the aggregate revenue and profit derived from domestic sales while R_x and Π_x represent those derived from export sales to any given country. The construction of the average productivity levels $\tilde{\varphi}$ and $\tilde{\varphi}_x$ further implies (see appendix):

$$r_c(\tilde{\varphi}) = \frac{R_d}{M} \qquad \pi_d(\tilde{\varphi}) = \frac{\Pi_d}{M},$$
$$r_c(\tilde{\varphi}_x) = \frac{R_x}{M_x} \qquad \pi_x(\tilde{\varphi}_x) = \frac{\Pi_x}{M_x}.$$

$r_c(\tilde{\varphi})$ and $\pi_d(\tilde{\varphi})$ thus also represent the average domestic revenue and profit across all domestic firms. Similarly, $r_c(\tilde{\varphi}_x)$ and $\pi_x(\tilde{\varphi}_x)$ represent the average export revenue and profit (to any given country) across all domestic firms who export. The combined average revenue ($\bar{r}$) and profit ($\bar{\pi}$) – derived from both domestic and export sales – can then be written in terms of these averages and the export probability p_x:

$$\begin{aligned} \bar{r} &= \frac{R}{M} \\ &= \frac{R_d}{M} + n\,\frac{R_x}{M} \\ &= r_c(\tilde{\varphi}) + p_x\, n\, r_c(\tilde{\varphi}_x), \end{aligned} \qquad \text{and} \qquad \begin{aligned} \bar{\pi} &= \frac{\Pi}{M} \\ &= \frac{\Pi_d}{M} + n\,\frac{\Pi_x}{M} \\ &= \pi_d(\tilde{\varphi}) + p_x\, n\, \pi_x(\tilde{\varphi}_x). \end{aligned} \tag{24}$$

Zero Cutoff Profit Condition

As in the closed economy equilibrium, the zero cutoff profit condition will imply a relationship between the average profit per firm $\bar{\pi}$, and the cutoff productivity level φ^*. The steps needed to show this are broken down into two cases, based on the form of the zero cutoff profit condition (20):

Case 1: $f_x > f$ (some firms do not export)

The zero cutoff profit condition (20) yields

$$\begin{cases} \pi_d(\varphi^*) = 0 \\ \pi_x(\varphi_x^*) = 0 \end{cases} \iff \begin{cases} r_c(\varphi^*) = \sigma f \\ r_c(\varphi_x^*) = \sigma f_x \end{cases} \tag{25}$$

$$\iff \begin{cases} r_c(\tilde{\varphi}) = \left(\frac{\tilde{\varphi}}{\varphi^*}\right)^{\sigma-1} \sigma f \\ r_c(\tilde{\varphi}_x) = \left(\frac{\tilde{\varphi}_x}{\varphi_x^*}\right)^{\sigma-1} \sigma f_x \end{cases}$$

$$\iff \begin{cases} \pi_d(\tilde{\varphi}) = f\, k(\varphi^*) \\ \pi_x(\tilde{\varphi}_x) = f_x\, k(\varphi_x^*), \end{cases}$$

where $k(\varphi) = \left[\left(\frac{\tilde{\varphi}(\varphi)}{\varphi}\right)^{\sigma-1} - 1\right]$ as was previously defined. Note that (25) also yields an expression of φ_x^* as a function for φ^*:

$$\frac{r_c(\varphi_x^*)}{r_c(\varphi^*)} = \left(\frac{\varphi_x^*}{\varphi^*}\right)^{\sigma-1} = \frac{f_x}{f} \iff \varphi_x^* = \varphi^* \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma-1}}. \tag{26}$$

Case 2: $f_x \leq f$ (all firms export)

The zero cutoff profit condition (20) now yields

$$\begin{aligned} \pi_d(\varphi^*) + n\,\pi_x(\varphi^*) = 0 &\iff (n+1)\, r_c(\varphi^*) = \sigma(f + n\, f_x) \\ &\iff (n+1)\, r_c(\tilde{\varphi}) = \left(\frac{\tilde{\varphi}}{\varphi^*}\right)^{\sigma-1} \sigma(f + n\, f_x) \\ &\iff \pi_d(\tilde{\varphi}) + n\,\pi_x(\tilde{\varphi}) = (f + n\, f_x)\, k(\varphi^*), \end{aligned} \tag{27}$$

where $\varphi_x^* = \varphi^*$, leading to $\tilde{\varphi}_x = \tilde{\varphi}$ and $p_x = 1$.

These two cases both yield the same expression for the average profit per firm derived in (24):

$$\begin{aligned} \bar{\pi} &= \pi_d(\tilde{\varphi}) + p_x\, n\, \pi_x(\tilde{\varphi}_x) \\ &= f\, k(\varphi^*) + p_x\, n\, f_x\, k(\varphi_x^*) \qquad \text{(ZCP)} \end{aligned} \tag{28}$$

since $\varphi_x^* = \varphi^*$, $\tilde{\varphi}_x = \tilde{\varphi}$, and $p_x = 1$ when $f_x \leq f$. Note that the zero cutoff profit condition (28) implies that $\bar{\pi}$ can be written as a function of only the cutoff level φ^* since φ_x^* and hence $p_x\, k(\varphi_x^*)$, can also be written as functions of only φ^* using (26):

$$\varphi_x^* = \max\left\{\varphi^*, \varphi^* \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma-1}}\right\}. \tag{29}$$

Free Entry and the Value of Firms

As before, let $\bar{v} = \sum_{t=0}^{\infty}(1-\delta)^t \bar{\pi} = \frac{1}{\delta}\,\bar{\pi}$ represent the present value of the average profit flows. The net value of entry $v_e = p_{in}\,\bar{v} - f_e$, and hence the free entry condition $v_e = 0 \iff \bar{\pi} = \frac{\delta f_e}{p_{in}}$ will be identical to those derived for the closed economy: regardless of profit differences across firms (based on export status), the expected value of future profits must equal the fixed investment cost.

Determination of the Equilibrium

As in the closed economy case, a stationary equilibrium is uniquely determined by the triplet (φ^*, P, R) satisfying the free entry and zero cutoff profit conditions. It is shown in the appendix that the free entry condition and the new zero cutoff profit condition (summarized in (28) and (29)) identify a unique φ^* and $\bar{\pi}$ (the new ZCP curve still cuts the FE curve from above). The equilibrium φ^*, in turn, determines the export productivity cutoff φ_x^* as well as the average productivity levels $\tilde{\varphi}$, $\tilde{\varphi}_x$, $\tilde{\varphi}_t$, and the ex-ante successful entry and export probabilities p_{in} and p_x. As was the case in the closed economy equilibrium, the free entry condition and the aggregate stability condition[34] $(p_{in}\,M_e = \delta\,M)$ ensure that the aggregate payment to the investment workers L_e equals the aggregate profit level Π. Thus, once again, the aggregate revenue R is determined by the aggregate income Y, which just reflects the payments to the aggregate labor force of the country: $R = Y = L$. As was shown in (23), the aggregate price index is determined by the aggregate number of goods available in each country (M_t) and the average productivity level across all firms selling these goods $(\tilde{\varphi}_t)$. It therefore remains to be shown that the number of firms M in either country (and hence, the number of varieties available $M_t = (1 + p_x)\,M$) is uniquely determined by the equilibrium conditions.

Following a line of argument similar to the one used in the closed economy case, the number of firms is obtained from the equilibrium conditions by using the property that these conditions

[34] Recall that this condition ensures that the mass of successful entrants matches the mass of incumbent firms who exit.

identify the average revenue per firm ($\bar{r}$) independently of M (see (24)):

$$\begin{aligned}\bar{r} &= r_c(\tilde{\varphi}) + p_x\, n\, r_c(\tilde{\varphi}_x)\\ &= \sigma[\pi_d(\tilde{\varphi}) + f] + p_x\, n\, \sigma[\pi_x(\tilde{\varphi}_x) + f_x]\\ &= \sigma(\bar{\pi} + f + p_x\, n\, f_x),\end{aligned}$$

which thus determines the number of firms:

$$M = \frac{R}{\bar{r}} = \frac{L}{\sigma(\bar{\pi} + f + p_x\, n\, f_x)}. \tag{30}$$

Analysis of the Equilibrium

The preceding analysis has shown that the existence of fixed export costs – which would typically represent the amortized portion of an export market entry cost – can induce a partitioning of firms by export status, based on firm productivity. The analysis also revealed that this partitioning is not a necessary consequence of fixed export costs and firm heterogeneity: when the export cost is low enough (relative to the overhead production cost), the model predicts that all firms would enter the export market. Note that, in these cases, the distribution of the fixed costs between the overhead production cost and the export cost is irrelevant (since both are incurred, only the sum of the two matters). In particular, the equilibrium will be identical to one where firms incur an overhead production cost of $f + n\, f_x$ and no export cost. Thus, when $f_x \leq f$, the trade equilibrium replicates that of an integrated world economy with $(n + 1)$ times the labor endowment of each country and an overhead production cost of $f + n\, f_x$. All the firm level variables, as well as the welfare level per worker, will be identical in the country's open economy equilibrium and this integrated closed world economy.

Although there is no direct empirical evidence on the relative magnitudes of export entry costs and overhead production costs, the indirect evidence on the widespread partitioning of firms by export status and productivity level suggests – if the basic framework of the current model is accepted – that the case where $f_x > f$ is the most relevant for drawing inferences about the effects of trade. Of course, even when $f_x > f$, the predictions of the model are still highly stylized. One of the most glaring simplifications, that firms either do not export or export to all countries, can be addressed without compromising the tractability of the model.

Suppose that the countries are all evenly divided into a certain number of identical-sized trade blocs such that the costs of exporting within the bloc are lower than those of exporting across trade blocs.[35] Since the symmetry across countries is preserved, the preceding equilibrium analysis can be easily extended to cover this additional feature. If the export costs within trade blocs is high enough (such that the flow cost per period is greater than the overhead production cost), then firms will be partitioned into three categories based on their productivity: as before, the most productive firms will export to all countries and the least productive firms will not export. There will now also be a third category of firms with productivity levels in an intermediate range. These firms will export to countries within their trade bloc but will not export to countries outside their trade bloc. The volume of trade between countries in the same trade bloc will then obviously be higher than the volume of trade between countries in different trade blocs. Thus, the incorporation of export cost differences across countries enables the current model to explain why firms export to an increasing number of countries as their productivity rises and to explain how the pattern of trade between countries is influenced by differences in export market entry costs.

8 The Impact of Trade

The result that the modeling of fixed export costs explains the partition of firms by export status and productivity level is not exactly earth-shattering. This can be explained quite easily within a simple partial equilibrium model with a fixed distribution of firm productivity levels. On the other hand, such a model would be ill-suited to address several important questions concerning the impact of trade in the presence of export market entry costs and firm heterogeneity: What happens to the range of firm productivity levels? Do all firms benefit from trade or does the impact depend on a firm's productivity? How is aggregate productivity and welfare affected? The current model is much better suited to address these questions,[36] which are answered in the following sections. The current section examines the effects of a transition from a closed economy (described in section 5) to the open economy described in the previous section. The following section then studies the impact of trade liberalization.

I now analyze the aggregate and firm level responses within a country that transitions from autarky to the open economy equilibrium with n other countries. Let φ_a^* and $\tilde{\varphi}_a$ denote the cutoff

[35] These lower costs could be driven either by trade agreements within blocs to reduce non-tariff-barriers to trade or by other factors reducing export costs within a bloc (such as common languages or similar legal systems, or exchange rate stability).

[36] In order to plausibly address these questions, a model should allow for the endogenous selection of the heterogeneous firms into the industry and incorporate the general equilibrium feedback link between wages and productivity.

and average productivity levels in autarky. I use the notation of the previous section for all variables and functions pertaining to the new open economy equilibrium. As was previously mentioned, the FE condition is identical in both the closed and open economy. Inspection of the new ZCP condition in the open economy (28) relative to the one in the closed economy (12) immediately reveals that the ZCP curve shifts up: the exposure to trade induces an increase in the cutoff productivity level ($\varphi^* > \varphi_a^*$) and in the average profit per firm.[37] The least productive firms with productivity levels between φ_a^* and φ^* can no longer earn positive profits in the new trade equilibrium and therefore exit. The cause of this selection effect depends on the level of the export cost. If $f_x \leq f$, then all incumbent firms (including the cutoff firm) export and incur the additional export cost f_x. The rise in the equilibrium cutoff level is then driven by the increase in the overall fixed costs (recall that this outcome also occurs when the overhead production cost f increases in the closed economy). If $f_x > f$, then the cutoff firm does not export and hence does not incur any additional fixed costs. On the other hand, the cutoff firm now competes in its domestic market with the new foreign exporters who are all more productive ($\varphi_x^* > \varphi^*$). This additional competition forces the cutoff firm to exit. Another selection process also occurs when $f_x > f$, as only the firms with productivity levels above φ_x^* enter the export markets. This export market selection effect and the domestic market selection effect (of firms out of the industry) both reallocate market shares towards more efficient firms and contribute to an aggregate productivity gain ($\tilde{\varphi}_t > \tilde{\varphi}_a$).[38]

Inspection of the equations for the equilibrium number of firms ((15) and (30)) reveals that $M < M_a$ where M_a represents the number of firms in autarky.[39] Although the number of firms in a country decreases after the transition to trade, consumers in the country still typically enjoy greater product variety ($M_t = (1 + p_x)\, M > M_a$). That is, the decrease in the number of domestic firms following the transition to trade is typically dominated by the number of new foreign exporters. It is nevertheless possible, when the export cost is high, that these foreign firms replace a larger number of domestic firms (if the latter are sufficiently less productive).[40] Although the impact of product variety on welfare is ambiguous, the contribution of the aggregate productivity gain to welfare is always enough to guarantee a welfare gain (see appendix). Trade – even though it is

[37] Recall that the FE curve must be upward sloping and cuts the ZCP curve from below.

[38] $\tilde{\varphi} - \tilde{\varphi}_a$ represents the contribution of the domestic market selection effect to the productivity gain, while $\tilde{\varphi}_t - \tilde{\varphi}$ represents the contribution of the export market selection effect. (Of course, this second contribution is zero when $f_x \leq f$ and all firms export.)

[39] Recall that the average profit $\bar{\pi}$ must be higher in the open economy equilibrium.

[40] Given the previously mentioned regularity conditions on $g(\varphi)$, which ensure that $k(\varphi)$ is decreasing, it can further be shown that the possibility of a product variety loss can not occur when $f_x \leq f$ and must occur when f_x is above a certain threshold level.

costly – necessarily generates a welfare gain.

The Reallocation of Market Shares and Profits Across Firms

I now examine how the impact of trade on individual firms changes with the firm's productivity level. To do this, I track the performance of a firm with productivity $\varphi \geq \varphi_a^*$ during the transition from autarky to trade. Let $r_a(\varphi) > 0$ and $\pi_a(\varphi) \geq 0$ denote the firm's revenue and profit in autarky. Recall that, in both the closed and open economy equilibria, the aggregate revenue of domestic firms is exogenously given by the country's size ($R = Y = L$). Hence, $\frac{r_a(\varphi)}{L}$ and $\frac{r(\varphi)}{L}$ represent the firm's market share (within the domestic industry) in autarky and in the equilibrium with trade. Additionally, in this equilibrium with trade, $\frac{r_c(\varphi)}{L}$ represents the firm's share of its domestic market (since $r_c(\varphi)$ is the firm's domestic revenue and $R = L$ also represents aggregate consumer expenditure in the country). The impact of trade on this firm's market share can be evaluated using the following inequalities (see appendix):

$$r_c(\varphi) < r_a(\varphi) < (n+1)\, r_c(\varphi), \qquad \forall\, \varphi \geq \varphi^*.$$

The first part of the inequality indicates that all firms incur a loss in their share of their domestic market with the transition to trade. A firm who does not export then also incurs a total revenue loss. The second part of the inequality indicates that a firm who exports more than makes up for its loss of domestic sales with export sales and increases its total revenues.[41] Thus, a firm who exports increases its share of industry revenues while a firm who does not export loses market share. φ_x^* is then also the cutoff level that partitions the firms between those who gain and those who lose market share. Note that, when $f_x \leq f$, $\varphi_x^* = \varphi^*$ and all firms export – so firms either remain in the industry and gain market share, or they exit.

Now consider the change in profit earned by a firm with productivity level φ during the transition to trade. If the firm does not export in the new trade equilibrium, then it must incur a profit loss, since its revenue, and hence variable profit, is now lower (the overhead production cost f remains unchanged). The direction of the profit change for an exporting firm is not immediately clear since it involves a tradeoff between the increase in total revenue (and hence variable profit) and the increase in fixed cost due to the additional export cost. For such a firm ($\varphi \geq \varphi_x^*$), this profit change

[41] Recall that $r(\varphi) = (n+1)\, r_c(\varphi)$ for all exporting firms.

can be written:[42]

$$\Delta\pi(\varphi) = \pi(\varphi) - \pi_a(\varphi) = \frac{1}{\sigma}\left[(n+1)\,r_c(\varphi) - r_a(\varphi)\right] - f_x$$
$$= \frac{\varphi^{\sigma-1}}{\sigma}\left[\frac{(n+1)\,r_c(\varphi^*)}{(\varphi^*)^{\sigma-1}} - \frac{\sigma f}{(\varphi_a^*)^{\sigma-1}}\right] - f_x,$$

where $\left[\frac{(n+1)\,r_c(\varphi^*)}{(\varphi^*)^{\sigma-1}} - \frac{\sigma f}{(\varphi_a^*)^{\sigma-1}}\right] = \frac{1}{\varphi^{\sigma-1}}\left[(n+1)\,r_c(\varphi) - r_a(\varphi)\right] > 0$. The profit change for an exporting firm, $\Delta\pi(\varphi)$, will thus be an increasing function of the firm's productivity level φ. In addition, this change must be negative for the exporting firm with the cutoff productivity level φ_x^*:

$$\Delta\pi(\varphi_x^*) = [\pi_d(\varphi_x^*) + n\,\pi_x(\varphi_x^*)] - \pi_a(\varphi_x^*) = \pi_d(\varphi_x^*) - \pi_a(\varphi_x^*) = \frac{1}{\sigma}[r(\varphi_x^*) - r_a(\varphi_x^*)] < 0.$$

Therefore, there must exist another cutoff level $\varphi^\dagger > \varphi_x^*$ that partitions the firms between those who gain and those who lose profits. Within this group of efficient firms who both export and increase their profits ($\varphi > \varphi^\dagger$), the profit gain increases with the firm's productivity level.

Summarizing the previous results on the firm level impact of trade by stating that the benefits of trade are not equally spread across firms would be quite an understatement! It was just shown how the exposure to trade generates the type of Darwinian evolution described in the introduction: the most efficient firms thrive and grow – they export and increase both their market share and profits. Some less efficient firms still export and increase their market share but incur a profit loss. Some even less efficient firms remain in the industry but do not export and incur losses of both market share and profit. Finally, the least efficient firms are driven out of the industry.

9 The Impact of Trade Liberalization

The preceding analysis compared the equilibrium outcomes of an economy undergoing a massive change in trade regime from autarky to trade. Very few, if any, of the world's current economies can be considered to operate in an autarky environment. It is therefore reasonable to ask whether an *increase* in the exposure of an economy to trade will induce the same effects as were previously described for the transition of an economy from autarky. The current model is well-suited to address two different types of increases in trade exposure, which correspond to specific (and plausible) policy experiments. The first policy experiment involves the incorporation of additional countries into a

[42]Using $r_c(\varphi) = \left(\frac{\varphi}{\varphi^*}\right)^{\sigma-1} r_c(\varphi^*)$ and $r_a(\varphi) = \left(\frac{\varphi}{\varphi_a^*}\right)^{\sigma-1} \sigma f$.

trade bloc. The second experiment involves a multi-lateral agreement within a trade bloc – holding the number of countries fixed – to reduce non-tariff-barriers to trade. The main impact of the transition from autarky to trade was an increase in aggregate productivity and welfare generated by a reallocation of market shares towards more productive firms. I will show that an increase in the exposure to trade generated by the first type of policy experiment (the incorporation of additional countries into a trade bloc) has an impact that is qualitatively almost identical to the one previously described for the transition from autarky. I then show that an increase in the exposure to trade generated by the second type of policy experiment (a reduction in NTBs) typically induces similar reallocations and an aggregate productivity gain, and it unequivocally engenders a welfare gain.

Increase in the Number of Trading Partners

In the previous section, I analyzed the impact of trade on a country initially in autarky who joins a trade bloc with n other countries. I now examine how these n other countries, assuming they previously traded among themselves, are affected by the addition of this country (and possibly more) to their trade bloc. To keep things simple, I assume that the n countries in the original trade bloc did not trade at all with the additional countries before their incorporation into the trade bloc. After this, I assume that the export costs between any two countries are identical and equal to their level prior to the expansion of the trade bloc. In terms of the current model, this leads to a comparison of the open economy equilibrium with n and $n' > n$ countries. For the sake of brevity (which, admittedly, has already taken a beating in the current paper), I only derive the results for the most plausible case where some firms do not export ($f_x > f$). For simplicity, I further assume that the regularity conditions on $g(\varphi)$ are satisfied, so that the $k(\varphi)$ is decreasing. Throughout this comparative static exercise, I use the notation of the open economy equilibrium in section 7 to describe the old equilibrium with n countries. I then add primes ($'$) to all variables and functions when they pertain to the new equilibrium with n' countries.

Inspection of equation (29) relating φ_x^* to φ^* reveals that the relationship between these two variables is not affected by the number of countries, and that the ratio of the two must remain constant: $\frac{\varphi_x^*}{\varphi^*} = \frac{\varphi_x^{*\prime}}{\varphi^{*\prime}} = \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma-1}}$. It is therefore clear from the zero cutoff profit condition (28) that the ZCP curve will shift up with an increase in n. Hence, the industry cutoff productivity level, and therefore the export cutoff productivity level, both increase with n: $\varphi^{*\prime} > \varphi^*$ and $\varphi_x^{*\prime} > \varphi_x^*$. The increase in the number of trading partners thus forces the least productive firms to exit. As was the case with the transition from autarky, the increased exposure to trade forces all firms to relinquish a portion of their share of their domestic market (see appendix): ${r_c}'(\varphi) < r_c(\varphi),\ \forall\, \varphi \geq \varphi^*$. The less productive firms who do not export thus incur a revenue and profit loss – and the least productive among them therefore exit. Again, as was the case with the transition from autarky, the firms who export more than make up for their loss of domestic sales with their export sales and increase their combined revenues (see appendix): $(n'+1)\, {r_c}'(\varphi) > (n+1)\, r_c(\varphi)$. The most productive of these firms also increase their profits (see appendix): $\Delta\pi(\varphi) = \pi'(\varphi) - \pi(\varphi) > 0$ for $\varphi > \varphi^\dagger$ where $\varphi^\dagger > {\varphi_x^*}'$.[43] Both market shares and profits are thus reallocated towards the more efficient firms.

The domestic market selection effect (of firms out of the industry) clearly contributes to an aggregate productivity gain. On the other hand, the impact of the export market selection effect on aggregate productivity now potentially involves a tradeoff: The market shares of the exporting firms (with $\varphi \geq {\varphi_x^*}'$) increase, and this contributes to an aggregate productivity gain; but new firms with productivity levels between φ_x^* and ${\varphi_x^*}'$ do not export in the new equilibrium with n' countries (whereas they would have exported in the old equilibrium with n countries).[44] If these firms have productivity levels above the aggregate average, then the reduction in their market share contributes to an aggregate productivity loss. Given the assumption that $k(\varphi)$ is decreasing, it is shown in the appendix that the aggregate productivity must nevertheless increase.[45] This assumption also permits one to sign the comparative static effect on the number of firms (see appendix): $M' < M$. The number of firms in each country decreases with the exposure to trade. Recall that this does not imply a decrease in product variety $M_t = (1 + p_x)M$ (as was the case with the transition from autarky, the direction of the change in M_t is ambiguous and depends on the level of f_x). Finally, regardless of the level of the export cost f_x and regardless of the slope of $k(\varphi)$, an increase in the number of trading partners must unequivocally generate a welfare gain (once again, see appendix for proof).

[43] As was the case with the transition from autarky, the profit gain of these firms (with $\varphi > \varphi^\dagger$) increases with the firm's productivity.

[44] There is a transitional issue associated with the exporting status of firms with productivity levels between φ_x^* and ${\varphi_x^*}'$. The loss of export sales to any given country (from $r_c(\varphi)$ down to ${r_c}'(\varphi)$) is such that firms entering with productivity levels between φ_x^* and ${\varphi_x^*}'$ will not export as the lower variable profit $\frac{{r_c}'(\varphi)}{\sigma}$ no longer covers the amortized portion of the entry cost f_x. On the other hand, incumbent firms with productivity levels in this range have already incurred the sunk export entry cost and have no reason to exit the export markets until they are hit with the bad shock and exit the industry. Eventually, all these incumbent firms exit and no firm with a productivity level in that range will export once the new steady state equilibrium is attained.

[45] If $f_x \leq f$ (all firms export), then aggregate productivity must rise with the number of countries, regardless of this additional assumption.

Reduction in Non-Tariff-Barriers

I now examine the effects of a different kind of increase in the exposure to trade, associated with a particular type of trade liberalization: a reduction in NTBs that reduces the export market entry cost for all firms. This reduction is modeled as a decrease in f_x, holding the number of countries fixed at n. As was previously mentioned, the fixed trade costs faced by potential exporters include NTBs but also some other "real" costs. It would thus be unrealistic to think of trade liberalization as reducing the fixed trade cost to zero. As in the previous section, I focus on the case where f_x remains greater than f, so firms are still partitioned by export status after the liberalization occurs.[46]

When $f_x > f$, the comparative statics on the FE and ZCP equilibrium conditions yield the following inequalities (see appendix):

$$\frac{\partial \varphi^*}{\partial f_x} < 0 \quad \text{and} \quad \frac{\partial \varphi_x^*}{\partial f_x} > 0 \qquad \text{when} \quad f_x > f.$$

Trade liberalization thus induces an increase in the industry cutoff productivity level and a decrease in the export cutoff level. The drop in f_x induces the entry of previously non-exporting firms into the export markets. The least productive firms with productivity levels near the cutoff φ^* have nothing to gain from trade liberalization. They do not export, and so are not directly affected by the decrease in the export cost. On the other hand, they now face the additional competition from the new foreign exporters (who are more productive than they, since $\varphi_x^* > \varphi^*$), which forces them to exit. Trade liberalization therefore induces reallocations that are similar to those previously described for the transition from autarky: market shares are reallocated away from the least productive firms (who exit) towards more productive firms (who enter the export markets).

The change in aggregate productivity ($\tilde{\varphi}_t$) induced by these reallocations depends on the contribution of the export market selection effect. The domestic market selection effect necessarily has a positive impact on aggregate productivity, whereas the impact of the export market selection effect depends on the relative productivity of the new exporters. If these exporters have higher than average productivity ($\varphi_x^* > \tilde{\varphi}_t$), then their entry into the export markets has a positive impact on aggregate productivity. If they have lower than average productivity, then the impact of their

[46] Recall that when $f_x \leq f$, the open economy equilibrium is identical to a closed world economy where the overhead production cost is given by $f + n\, f_x$. A reduction in f_x therefore has identical effects to a reduction in f. One can thus invoke the comparative statics of the closed economy equilibrium for changes in f.

entry on aggregate productivity is reversed. The occurrence of either case depends on the size of f_x. When this cost is high, φ_x^* will be also be high (φ_x^* is an increasing function of f_x), and the proportion of exporting firms p_x will be low. There necessarily exists a threshold export cost level such that $\varphi_x^* > \tilde{\varphi}_t$ when f_x is above this threshold.[47] Both firm selection effects induced by trade liberalization would then contribute to an aggregate productivity gain. As trade is further liberalized, the new exporters will have progressively lower productivity levels until this productivity level drops below the aggregate level $\tilde{\varphi}_t$. At this point, the selection effects induced by trade liberalization have opposite impacts on aggregate productivity. At first, the impact of the domestic market selection effect will dominate, and aggregate productivity will continue to rise as trade is liberalized. It can further be shown that the negative impact of the export market selection effect on aggregate productivity must dominate when f_x drops to a level close enough to f. Aggregate productivity then decreases with trade liberalization.[48] If further trade liberalization that would push the export cost f_x below f were possible, then aggregate productivity would necessarily decrease with this further liberalization.[49] Figure 3 shows a graph of the cutoff and aggregate productivity levels as a function of the export cost f_x.

Again, regardless of the level of the export cost f_x, trade liberalization will unequivocally generate a welfare gain (see appendix). In the preceding analysis of the transition from autarky and the increase in the number of trading partners, the accompanying welfare gain was always partially driven by an increase in aggregate productivity. Although trade liberalization also unambiguously increases welfare, the relative contributions of product variety and aggregate productivity changes to this welfare increase now depend on the level of f_x. When f_x is high, the welfare increase is still driven (at least in part) by an aggregate productivity increase. On the other hand, when f_x is low, trade liberalization now induces a decrease in aggregate productivity. Product variety must then not only be increasing, its impact on welfare must furthermore dominate the negative impact of lower aggregate productivity. Finally, there also exists a middle range for f_x where aggregate productivity and product variety are both increasing with trade liberalization.

[47]It can be shown that the limiting case as f_x goes to infinity is in fact identical to the autarky case: φ_x^* goes to infinity and p_x goes to zero (no firms export) while φ^* and $\tilde{\varphi}_t$ go to their autarky levels of φ_a^* and $\tilde{\varphi}_a$.

[48]The differential change in aggregate productivity can be written as

$$d(\tilde{\varphi}_t)^{\sigma-1} = \alpha[(\tilde{\varphi}_t)^{\sigma-1} - (\varphi^*)^{\sigma-1}]\, d\varphi^* + \alpha_x[(\tilde{\varphi}_t)^{\sigma-1} - (\varphi_x^*)^{\sigma-1}]\, d\varphi_x^*,$$

where α and α_x are both positive. The first term captures the impact of the domestic market selection effect while the second one captures the impact of the export market selection effect.

[49]See the comparative static for f in the closed economy and note 46.

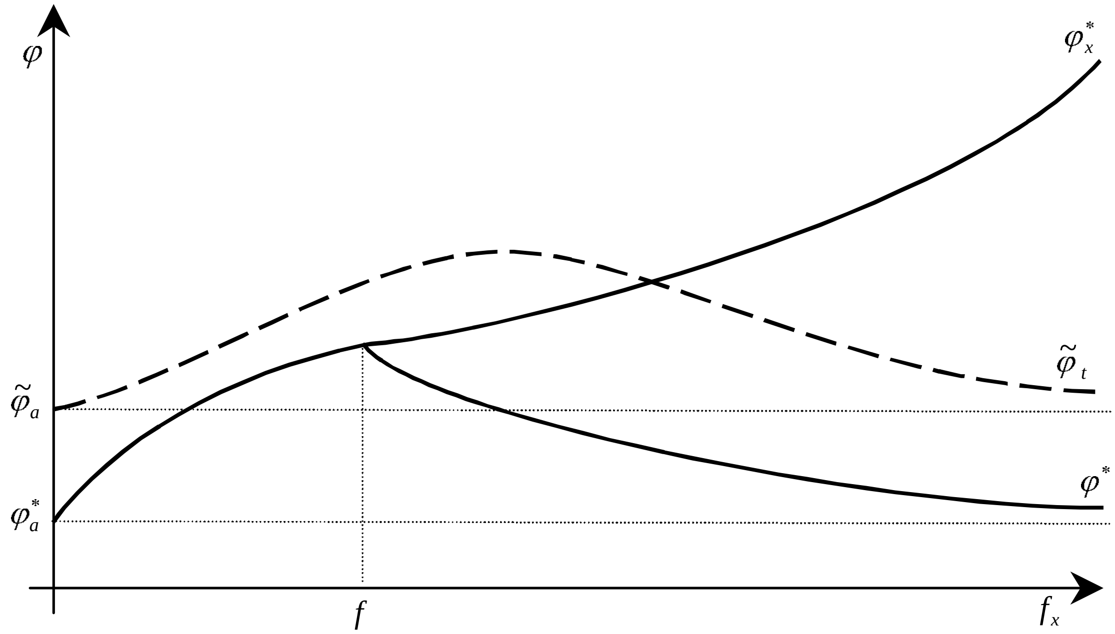

Figure 3: Cutoff and Aggregate Productivity Levels as a Function of the Export Cost

10 Conclusion

This paper has described and analyzed a new transmission channel for the impact of trade on industry structure and performance. Since this channel works through intra-industry reallocations among firms, it can only be studied within an industry model that incorporates firm level heterogeneity. Recent empirical work has highlighted the importance of this channel for understanding and explaining the effects of trade on firm and industry performance.

The paper investigates how the existence of export market entry costs affects the impact of trade both on a country's aggregate environment and on the performance of individual firms. The paper shows that the existence of such costs to trade does not affect the welfare-enhancing properties of trade: one of the most robust results of this paper is that increases in a country's exposure to trade lead to welfare gains. On the other hand, the paper shows how the export costs significantly alter the distribution of the gains from trade across firms. In fact, only a portion of the firms – the more efficient ones – reap benefits from trade in the form of gains in market share and profit. Less efficient firms lose both. The exposure to trade, or increases in this exposure, force the least efficient firms out of the industry. These trade-induced reallocations towards more efficient firms explain why trade may generate aggregate productivity gains without necessarily improving the productive efficiency of individual firms.

This model also explains the patterns reported by Roberts et al. (1995) that some export booms are mainly driven by the entry of new firms into the export market (over half of the substantial export growth in Colombia and Morocco was driven by new exporters). This assumes that a reduction in the fixed export cost was a contributing factor to the export boom – something that was not directly tested by the authors. It is also possible that the causality between the export cost and the number of exporting firms runs in the other direction. Clerides et al. (1998) and Aitken, Hanson and Harrison (1997) find evidence suggesting that the number of exporters in a region creates positive spillover effects that reduce the export costs for other potential exporters. This possibility is not incorporated in the current model but suggests an extension for future work: as exporting firms do not internalize these cost reducing benefits to other firms, a *potential* role is then created for a government export subsidy program.

Although this model mainly highlights the long-run benefits associated with the trade-induced reallocations within an industry, the reallocation of these resources also obviously entails some short-run costs. It is therefore important to have a model that can predict the impact of trade policy on inter-firm reallocations in order to design accompanying policies that would address issues related to the transition towards a new regime. These policies could help palliate the transitional costs while taking care not to hinder the reallocation process. Of course, the model also clearly indicates that policies that hinder the reallocation process or otherwise interfere with the flexibility of the factor markets may prevent a country from reaping the full benefits from trade.

Appendix

A Aggregation Conditions in the Closed Economy

Using the definition of $\tilde{\varphi}$ in (7), the aggregation conditions relating the aggregate variables to the number of firms M and aggregate productivity level $\tilde{\varphi}$ are derived:

$$\begin{aligned}
Q &= \left[\int_0^\infty q(\varphi)^\rho \, M\mu(\varphi)\, d\varphi\right]^{1/\rho} && \text{by definition of } Q \equiv U \\
&= \left[\int_0^\infty q(\tilde{\varphi})^\rho \left(\frac{\varphi}{\tilde{\varphi}}\right)^{\sigma\rho} M\mu(\varphi)\, d\varphi\right]^{1/\rho} \\
&= M^{1/\rho}\, q(\tilde{\varphi}) \left[\frac{1}{\tilde{\varphi}^{\sigma-1}} \int_0^\infty \varphi^{\sigma-1}\, \mu(\varphi)\, d\varphi\right]^{1/\rho} \\
&= M^{1/\rho}\, q(\tilde{\varphi}),
\end{aligned}$$

and using the definition of R and Π as aggregate revenue and profit,

$$\begin{aligned}
R &= \int_0^\infty r(\varphi)\, M\mu(\varphi)\, d\varphi \\
&= \int_0^\infty r(\tilde{\varphi}) \left(\frac{\varphi}{\tilde{\varphi}}\right)^{\sigma-1} M\mu(\varphi)\, d\varphi \\
&= M\, r(\tilde{\varphi})\, \frac{1}{\tilde{\varphi}^{\sigma-1}} \int_0^\infty \varphi^{\sigma-1}\, \mu(\varphi)\, d\varphi \\
&= M\, r(\tilde{\varphi}),
\end{aligned}$$

$$\begin{aligned}
\Pi &= \int_0^\infty \pi(\varphi)\, M\mu(\varphi)\, d\varphi \\
&= \int_0^\infty \left[\frac{r(\varphi)}{\sigma} - f\right] M\mu(\varphi)\, d\varphi \\
&= \frac{1}{\sigma} \int_0^\infty r(\varphi)\, M\, \mu(\varphi)\, d\varphi - Mf \\
&= M \left[\frac{r(\tilde{\varphi})}{\sigma} - f\right] \\
&= M\, \pi(\tilde{\varphi}).
\end{aligned}$$

B Closed Economy Equilibrium

Existence and Uniqueness of the Equilibrium Cutoff Level φ^*

Following is a proof that the FE condition $\left(\bar{\pi} = \frac{\delta f_e}{1-G(\varphi^*)}\right)$ and ZCP condition $\left(\bar{\pi} = f\, k(\varphi^*)\right)$ in (12) identify a unique cutoff level φ^* and that the ZCP curve cuts the FE curve from above in (φ, π) space. I do this by showing that $[1 - G(\varphi)]\, k(\varphi)$ is monotonically decreasing from infinity to zero on $(0, \infty)$. (This is a sufficient condition for both properties.) Recall that $k(\varphi) = \left(\frac{\tilde{\varphi}(\varphi)}{\varphi}\right)^{\sigma-1} - 1$ where

$$\tilde{\varphi}(\varphi)^{\sigma-1} = \frac{1}{1 - G(\varphi)} \int_\varphi^\infty \xi^{\sigma-1}\, g(\xi)\, d\xi$$

as defined in (9). Differentiating with respect to φ yields

$$\frac{\partial \tilde{\varphi}(\varphi)^{\sigma-1}}{\partial \varphi} = \frac{g(\varphi)}{1-G(\varphi)} \left[\tilde{\varphi}(\varphi)^{\sigma-1} - \varphi^{\sigma-1}\right]$$

and hence

$$\begin{aligned} k'(\varphi) &= \frac{g(\varphi)}{1-G(\varphi)} \left[\left(\frac{\tilde{\varphi}(\varphi)}{\varphi} \right)^{\sigma-1} - 1 \right] - \left(\frac{\tilde{\varphi}(\varphi)}{\varphi} \right)^{\sigma-1} \frac{\sigma-1}{\varphi} \\ &= \frac{k(\varphi) g(\varphi)}{1-G(\varphi)} - \frac{(\sigma-1)[k(\varphi)+1]}{\varphi}. \end{aligned}$$

Define

$$j(\varphi) = [1-G(\varphi)]\, k(\varphi). \tag{B.1}$$

Then its derivative and elasticity are given by:

$$j'(\varphi) = -\frac{1}{\varphi}(\sigma-1)[1-G(\varphi)][k(\varphi)+1] < 0, \tag{B.2}$$

$$\frac{j'(\varphi)\, \varphi}{j(\varphi)} = -(\sigma-1)\left(1 + \frac{1}{k(\varphi)}\right) < -(\sigma-1). \tag{B.3}$$

Since $j(\varphi)$ is non-negative and its elasticity with respect to φ is negative and bounded away from zero, $j(\varphi)$ must be decreasing to zero as φ goes to infinity. Furthermore, $\lim_{\varphi \to 0} j(\varphi) = \infty$ since $\lim_{\varphi \to 0} k(\varphi) = \infty$. Therefore, $j(\varphi) = [1-G(\varphi)]\, k(\varphi)$ decreases from infinity to zero on $(0, \infty)$.

Comparative Statics

Several of the comparative statics will use the property that welfare per worker can be written as a function of only the cutoff level φ^*:[50]

$$W = M^{\frac{1}{\sigma-1}} \rho\, \tilde{\varphi} = L^{\frac{1}{\sigma-1}} \rho \left(\frac{1}{\sigma f}\right)^{\frac{1}{\sigma-1}} \varphi^*. \tag{B.4}$$

Note that the property that welfare decreases with a rise in the entry cost f_e is then immediately obtained as it was shown that φ^* decreases in that situation. The direction of the welfare change induced by a rise in the overhead production cost is not immediately obvious as f enters into the welfare equation in (B.4). (Recall that a rise in f induces an increase in φ^*.) The direction of the

[50]using the relationship $\left(\frac{\tilde{\varphi}}{\varphi^*}\right)^{\sigma-1} = \frac{r(\tilde{\varphi})}{r(\varphi^*)} = \frac{R/M}{\sigma f} = \frac{L}{M}\frac{1}{\sigma f}$.

welfare change therefore depends on the direction of the change in $\left(\frac{1}{f}\right)^{\frac{1}{\sigma-1}} \varphi^*$, or alternatively, on the direction of the change in $\frac{(\varphi^*)^{\sigma-1}}{f}$.

Proof that $\frac{(\varphi^*)^{\sigma-1}}{f}$ *Decreases when* f *Increases*

The FE and ZCP equilibrium conditions for φ^* imply $\bar{\pi} = f\,k(\varphi^*) = \frac{\delta f_e}{1-G(\varphi^*)}$, and thus

$$f\,j(\varphi^*) = \delta\, f_e, \tag{B.5}$$

using the definition for $j(\varphi)$ in (B.1). Differentiating (B.5) with respect to f yields:

$$\begin{aligned} j(\varphi^*) + f\,j'(\varphi^*)\,\frac{\partial \varphi^*}{\partial f} = 0 \iff \frac{\partial \varphi^*}{\partial f} &= -\frac{j(\varphi^*)}{f\,j'(\varphi^*)} \\ \iff \frac{\partial \varphi^*}{\partial f} &= \frac{\varphi^*}{f(\sigma-1)[1+1/k(\varphi^*)]}, \end{aligned}$$

where the last step is obtained using (B.3). The differential change in $\frac{(\varphi^*)^{\sigma-1}}{f}$ is then given by

$$\begin{aligned} \frac{\partial\left(\frac{(\varphi^*)^{\sigma-1}}{f}\right)}{\partial f} &= \frac{1}{f^2}\left[f(\sigma-1)(\varphi^*)^{\sigma-2}\,\frac{\partial \varphi^*}{\partial f} - (\varphi^*)^{\sigma-1}\right] \\ &= \frac{(\varphi^*)^{\sigma-1}}{f^2}\left[\frac{1}{1+1/k(\varphi^*)} - 1\right] \\ &= \frac{(\varphi^*)^{\sigma-1}}{f^2}\,\frac{-1}{1+k(\varphi^*)} \\ &< 0. \end{aligned} \tag{B.6}$$

Hence, $\frac{(\varphi^*)^{\sigma-1}}{f}$ decreases when f increases. An increase in f therefore generates a welfare loss.

C Aggregation Conditions in the Open Economy

Since R_d and R_x represent the aggregate revenues derived from domestic and export sales, they can be written:

$$
\begin{aligned}
R_d &= \int_{\varphi^*}^{\infty} r(\varphi)\, M\mu(\varphi)\, d\varphi \\
&= M \frac{1}{1-G(\varphi^*)} \int_{\varphi^*}^{\infty} r(\varphi) g(\varphi)\, d\varphi \\
&= M \frac{1}{1-G(\varphi^*)} \int_{\varphi^*}^{\infty} r(\tilde{\varphi}) \left(\frac{\varphi}{\tilde{\varphi}}\right)^{\sigma-1} g(\varphi)\, d\varphi \\
&= M\, r(\tilde{\varphi}) \frac{1}{\tilde{\varphi}^{\sigma-1}} \left[\frac{1}{1-G(\varphi^*)} \int_{\varphi^*}^{\infty} \varphi^{\sigma-1} g(\varphi)\, d\varphi\right] \\
&= M\, r(\tilde{\varphi}) \qquad \text{by definition of } \tilde{\varphi} = \tilde{\varphi}(\varphi^*), \text{ and} \\
R_x &= \int_{\varphi_x^*}^{\infty} r(\varphi)\, M\mu(\varphi)\, d\varphi \\
&= M \frac{1}{1-G(\varphi^*)} \int_{\varphi_x^*}^{\infty} r(\varphi) g(\varphi)\, d\varphi \\
&= M \frac{1-G(\varphi_x^*)}{1-G(\varphi^*)} \frac{1}{1-G(\varphi_x^*)} \int_{\varphi_x^*}^{\infty} r(\tilde{\varphi}_x) \left(\frac{\varphi}{\tilde{\varphi}_x}\right)^{\sigma-1} g(\varphi)\, d\varphi \\
&= M_x\, r(\tilde{\varphi}_x) \frac{1}{\tilde{\varphi}_x^{\sigma-1}} \left[\frac{1}{1-G(\varphi_x^*)} \int_{\varphi_x^*}^{\infty} \varphi^{\sigma-1} g(\varphi)\, d\varphi\right] \\
&= M_x\, r(\tilde{\varphi}_x) \qquad \text{by definition of } \tilde{\varphi}_x = \tilde{\varphi}(\varphi_x^*).
\end{aligned}
$$

Similarly, the definitions of the aggregate profits Π_d and Π_x imply

$$
\Pi_d = \int_{\varphi^*}^{\infty} \pi_d(\varphi)\, M\mu(\varphi)\, d\varphi \quad \text{and} \quad \Pi_x = \int_{\varphi_x^*}^{\infty} \pi_x(\varphi)\, M\mu(\varphi)\, d\varphi.
$$

Writing $\pi_d(\varphi) = \frac{r(\varphi)}{\sigma} - f$ and $\pi_x(\varphi) = \frac{r(\varphi)}{\sigma} - f_x$ as defined in (19), one obtains $\Pi_d = M\,\pi_d(\varphi)$ and $\Pi_x = M_x\,\pi_x(\varphi)$ using the same derivations shown for the aggregate profit Π in a closed economy.

D Open Economy Equilibrium

Aggregate Labor Resources Used to Cover the Export Costs

It was asserted in note 30 that the ratio of new exporters to all exporters was δ, and hence that the aggregate labor resources used to cover the export cost did not depend on its representation as either a one time sunk entry cost or a per-period fixed cost. As before, let M_e denote the mass of all new entrants. The ratio of new exporters to all exporters is then $\frac{p_x p_{in} M_e}{p_x M} = \frac{p_x p_{in} M_e}{p_x M}$. This ratio must be equal to δ as the aggregate stability condition for the equilibrium ensures that $p_{in} M_e = \delta M$.

Existence and Uniqueness of the Equilibrium Cutoff Level φ^*

Following is a proof that the FE condition, $\left(\bar{\pi} = \frac{\delta f_e}{1-G(\varphi^*)}\right)$ and the new ZCP condition $\left(\bar{\pi} = f\, k(\varphi^*) + p_x\, n\, f_x\, k(\varphi_x^*)\right)$ in (28) identify a unique cutoff level φ^* and that the new ZCP curve cuts the FE curve from above in (φ, π) space. These conditions imply:

$$\frac{\delta f_e}{1 - G(\varphi^*)} = f\, k(\varphi^*) + p_x\, n\, f_x\, k(\varphi_x^*) \iff f\, j(\varphi^*) + n\, f_x\, j(\varphi_x^*) = \delta\, f_e \tag{D.1}$$

using the definition for $j(\varphi)$ in (B.1). Using the relationship between φ^* and φ_x^* in (29), (D.1) can be rewritten:

$$\delta\, f_e = \begin{cases} f\, j(\varphi^*) + n\, f_x\, j\big((f_x/f)^{\frac{1}{\sigma-1}}\varphi^*\big) & \text{if } f_x > f, \\ (f + n\, f_x)\, j(\varphi^*) & \text{if } f_x \le f. \end{cases} \tag{D.2}$$

As $j(\varphi)$ is decreasing from infinity to zero on $(0, \infty)$, both expressions for the right hand side in (D.2) must also monotonically decrease from infinity to zero on $(0, \infty)$. In both cases, (D.2) identifies a unique cutoff level φ^*. Furthermore, since $\big[1 - G(\varphi^*)\big]\big[f\, j(\varphi^*) + n\, f_x\, j(\varphi_x^*)\big]$ is decreasing in φ^*, the new ZCP curve must cut the FE curve from above.

E The Impact of Trade

Welfare

Using (B.4), welfare per worker in autarky can be written:

$$W_a = M_a^{\frac{1}{\sigma-1}}\, \rho\, \tilde{\varphi}_a = L^{\frac{1}{\sigma-1}} \rho \left(\frac{1}{\sigma f}\right)^{\frac{1}{\sigma-1}} \varphi_a^*.$$

Similarly, welfare in the open economy can be written as a function of only the cutoff productivity level:[51]

$$W = M_t^{\frac{1}{\sigma-1}}\, \rho\, \tilde{\varphi}_t = L^{\frac{1}{\sigma-1}} \rho \left(\frac{1}{r_c(\varphi^*)}\right)^{\frac{1}{\sigma-1}} \varphi^*, \tag{E.1}$$

$$\text{where} \quad r_c(\varphi^*) = \begin{cases} \sigma f & \text{if } f_x > f, \\ \frac{\sigma}{n+1}(f + n\, f_x) & \text{if } f_x \le f. \end{cases} \tag{E.2}$$

Since $r_c(\varphi^*) \le \sigma f \ \ (\forall\, f_x)$ and $\varphi^* > \varphi_a^*$, welfare in the open economy must be higher than in autarky: $W > W_a$.

Reallocations

Proof that $r_c(\varphi) < r_a(\varphi) < (n+1)\, r_c(\varphi)$

Recall that $r_a(\varphi) = \left(\frac{\varphi}{\varphi_a^*}\right)^{\sigma-1} \sigma f \ \ (\forall\, \varphi \ge \varphi_a^*)$ and that $r_c(\varphi) = \left(\frac{\varphi}{\varphi^*}\right)^{\sigma-1} r_c(\varphi^*) \ \ (\forall\, \varphi \ge \varphi^*)$. It was just shown in the previous section that $W > W_a$ because $r_c(\varphi^*) \le \sigma f$ and $\varphi^* > \varphi_a^*$. These conditions also necessarily imply that $r_c(\varphi) < r_a(\varphi), \ \forall\, \varphi > \varphi^*$. The second part of the inequality $r_a(\varphi) < (n+1)\, r_c(\varphi)$, is equivalent to (see (E.2)):

$$\frac{(\varphi_a^*)^{\sigma-1}}{f} > \begin{cases} \frac{(\varphi^*)^{\sigma-1}}{(n+1)f} & \text{if } f_x > f \\ \frac{(\varphi^*)^{\sigma-1}}{f+nf_x} & \text{if } f_x \le f \end{cases}$$

When $f_x \le f$, recall that the equilibrium cutoff φ^* is identical to one derived for a closed economy with an overhead production cost equal to $f + nf_x$. The difference between $\frac{(\varphi^*)^{\sigma-1}}{f+nf_x}$ and $\frac{(\varphi_a^*)^{\sigma-1}}{f}$ is thus identical to the change in $\frac{(\varphi^*)^{\sigma-1}}{f}$ generated by an increase of the overhead production

[51] using $\left(\frac{\tilde{\varphi}_t}{\varphi^*}\right)^{\sigma-1} = \frac{r_c(\tilde{\varphi}_t)}{r_c(\varphi^*)} = \frac{R/M_t}{r_c(\varphi^*)} = \frac{L}{M_t}\frac{1}{r_c(\varphi^*)}$, (25), and (27).

cost from f to $f + nf_x$ in a closed economy. This change must be negative as it was previously shown that $\frac{(\varphi^*)^{\sigma-1}}{f}$ decreases when the overhead production cost f rises (see (B.6)). Therefore, $\frac{(\varphi_a^*)^{\sigma-1}}{f} > \frac{(\varphi^*)^{\sigma-1}}{f+nf_x}$ when $f_x \leq f$. The first part of the inequality (when $f_x > f$), is a direct consequence of the comparative static for φ^* when f_x changes. When $f_x = f$, the second part of the inequality yields $\frac{(\varphi_a^*)^{\sigma-1}}{f} > \frac{(\varphi^*)^{\sigma-1}}{(n+1)f}$. As f_x rises above f, we know that the cutoff level φ^* falls $\left(\frac{\partial\varphi^*}{\partial f_x} < 0\right)$. Hence $\frac{(\varphi_a^*)^{\sigma-1}}{f} > \frac{(\varphi^*)^{\sigma-1}}{(n+1)f}$ will remain true for any level of $f_x > f$.

F The Impact of Trade Liberalization

Increase in the Number of Trading Partners

Recall that an increase in the number of trading partners from n to n' shifts up the ZCP curve and causes an increase in the equilibrium cutoff level from φ^* to $\varphi^{*\prime}$. Since the FE curve is upward sloping and cuts the ZCP curve from below, the average profit level must also increase: $\bar{\pi}' > \bar{\pi}$. The relationship between the two cutoff levels, φ_x^* and φ^* in (29) implies that the export cutoff level φ_x^* increases by the same proportion as the industry cutoff level φ^*: $\frac{\varphi_x^{*\prime}}{\varphi^{*\prime}} = \frac{\varphi_x^*}{\varphi^*}$, and therefore $\varphi_x^{*\prime} > \varphi_x^*$. The equilibrium condition (D.1), $f\, j(\varphi^*) + n\, f_x\, j(\varphi_x^*) = \delta\, f_e$, then implies:[52]

$$\begin{aligned} n'\, j(\varphi_x^{*\prime}) &> n\, j(\varphi_x^*), \text{and} \\ n'\, [1 - G(\varphi_x^{*\prime})] &> n\, [1 - G(\varphi_x^*)], \text{and} \\ n'\, p_x' &> n\, p_x. \end{aligned} \tag{F.1}$$

Number of Firms and Aggregate Productivity

(F.1) can be used to determine the direction of the change in both the number of firms and in the aggregate productivity level: Inspection of the equation for the number of firms (30) reveals that $M' < M$ since $\bar{\pi}' > \bar{\pi}$ and $n'\, p_x' > n\, p_x$. The comparison of the aggregate productivity levels is

[52] These derivations use the inequalities $\varphi^{*\prime} > \varphi^*$ and $\varphi_x^{*\prime} > \varphi_x^*$. In addition, the first step uses the property that $j(\varphi)$ is strictly decreasing; the second step uses the definition of $j(\varphi) = [1 - G(\varphi)]\, k(\varphi)$ and the additional assumption that $k(\varphi)$ is decreasing; and the third step uses the definition of $p_x = \frac{1-G(\varphi_x^*)}{1-G(\varphi^*)}$ and the property that $G(\varphi)$ is increasing.

obtained using:

$$
\begin{aligned}
(\tilde{\varphi}_t')^{\sigma-1} &= \frac{1}{1+p_x'n'}\left((\tilde{\varphi}')^{\sigma-1} + p_x'n'(\tilde{\varphi}_x')^{\sigma-1}\right) && \text{(from (22))} \\
&> \frac{1}{1+p_xn}\left((\tilde{\varphi}')^{\sigma-1} + p_xn(\tilde{\varphi}_x')^{\sigma-1}\right) && \text{(since } n'p_x' > np_x \text{ and } \tilde{\varphi}_x' > \tilde{\varphi}') \\
&> \frac{1}{1+p_xn}\left((\tilde{\varphi})^{\sigma-1} + p_xn(\tilde{\varphi}_x)^{\sigma-1}\right) = (\tilde{\varphi}_t)^{\sigma-1}. && \text{(since } \tilde{\varphi}' > \tilde{\varphi} \text{ and } \tilde{\varphi}_x' > \tilde{\varphi}_x)
\end{aligned}
$$

Reallocations

Recall that $r_c{}'(\varphi) = \left(\frac{\varphi}{\varphi^{*\prime}}\right)^{\sigma-1}\sigma f$ and $r_c(\varphi) = \left(\frac{\varphi}{\varphi^{*}}\right)^{\sigma-1}\sigma f$ when $f_x > f$ since $r_c{}'(\varphi^{*\prime}) = r_c(\varphi^*) = \sigma f$. Thus, $r_c{}'(\varphi) < r_c(\varphi),\ \ \forall \varphi \geq \varphi^{*\prime}$ since $\varphi^{*\prime} > \varphi^*$: All firms who produce in the new equilibrium with n' countries lose a portion of their domestic market. Differentiation of the equilibrium condition (D.1) with respect to n, along with the inequality (B.3), yields $\frac{\partial \varphi^*}{\partial n} < \frac{\varphi^*}{(n+1)(\sigma-1)}$, which further leads to $\frac{\partial \frac{(\varphi^*)^{\sigma-1}}{n+1}}{\partial n} < 0$.[53] This implies that $\frac{(\varphi^*)^{\sigma-1}}{n+1}$ decreases when n rises and hence that $\frac{(\varphi^{*\prime})^{\sigma-1}}{n'+1} < \frac{(\varphi^*)^{\sigma-1}}{n+1}$. This last inequality implies that

$$
\begin{aligned}
(n'+1)\, r_c{}'(\varphi) &= (n'+1)\left(\frac{\varphi}{\varphi^{*\prime}}\right)^{\sigma-1}\sigma f \\
&> (n+1)\left(\frac{\varphi}{\varphi^{*}}\right)^{\sigma-1}\sigma f = (n+1)\, r_c(\varphi).
\end{aligned}
$$

A firm who exports in the new equilibrium with n' countries thus increases its total revenues.

The change in profit earned by a firm who exports in the new equilibrium ($\varphi \geq \varphi^{*\prime}$) can be written:

$$
\begin{aligned}
\Delta\pi(\varphi) &= \pi'(\varphi) - \pi(\varphi) \\
&= \frac{1}{\sigma}\left[(n'+1)\, r_c{}'(\varphi) - (n+1)\, r_c(\varphi)\right] - (n'-n)\, f_x \\
&= \varphi^{\sigma-1} f\left[\frac{n'+1}{(\varphi^{*\prime})^{\sigma-1}} - \frac{n+1}{(\varphi^{*})^{\sigma-1}}\right] - (n'-n)\, f_x.
\end{aligned}
$$

$\Delta\pi(\varphi)$ is thus an increasing function of φ since $\left[\frac{n'+1}{(\varphi^{*\prime})^{\sigma-1}} - \frac{n+1}{(\varphi^{*})^{\sigma-1}}\right] > 0$. Furthermore, $\Delta\pi(\varphi_x^{*\prime}) < 0$ since the firm with productivity $\varphi = \varphi_x^{*\prime}$ earns lower profits from both export and domestic sales in the new equilibrium.[54] Therefore, there exists a new cutoff $\varphi^\dagger > \varphi_x^{*\prime}$ such that $\Delta\pi(\varphi) > 0,\ \ \forall\varphi >$

[53]This derivation is similar to the one showing that $\frac{\partial \frac{(\varphi^*)^{\sigma-1}}{f}}{\partial f} < 0$ in the closed economy equilibrium.

[54]Its profits from export sales decrease from a positive level to zero. Its profits from domestic sales decrease since

$\varphi^\dagger$.

Welfare

Inspection of conditions (E.1) and (E.2) that determine welfare in the open economy equilibrium indicates that welfare must always be higher in the equilibrium with n' countries since $\varphi^{*\prime} > \varphi^*$ and $r_c{}'(\varphi^{*\prime}) \leq r_c(\varphi^*), \;\; \forall f_x > 0$. Note that this inequality remains valid for any level of the export cost f_x and does not depend on the additional assumption on the slope of $k(\varphi)$.

Decrease in Non-Tariff-Barriers

Proof that $\frac{\partial \varphi^*}{\partial f_x} < 0$ *and* $\frac{\partial \varphi_x^*}{\partial f_x} > 0$ *when* $f_x > f$

Differentiating the equilibrium condition (D.1) with respect to f_x yields:

$$f\, j'(\varphi^*)\, \frac{\partial \varphi^*}{\partial f_x} + n\, j(\varphi_x^*) + n\, f_x\, j'(\varphi_x^*)\, \frac{\partial \varphi_x^*}{\partial f_x} = 0, \tag{F.2}$$

where $\varphi_x^* = \varphi^* \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma-1}}$ (since $f_x > f$). (F.2) can be rewritten as

$$\begin{aligned} \frac{\partial \varphi^*}{\partial f_x} \left[f\, j'(\varphi^*) + n\, f_x\, j'(\varphi_x^*)\, \frac{\varphi_x^*}{\varphi^*} \right] + n\, j(\varphi_x^*) + n\, j'(\varphi_x^*) \frac{\varphi_x^*}{\sigma - 1} &= 0, \\ \frac{\partial \varphi^*}{\partial f_x} \left[f\, j'(\varphi^*) + n\, f_x\, j'(\varphi_x^*)\, \frac{\varphi_x^*}{\varphi^*} \right] - n\,[1 - G(\varphi_x^*)] &= 0, \text{ or} \\ \frac{\partial \varphi^*}{\partial f_x} &= \frac{n\,[1 - G(\varphi_x^*)]}{f\, j'(\varphi^*) + n\, f_x\, j'(\varphi_x^*)\,(\varphi_x^*/\varphi^*)}, \end{aligned} \tag{F.3}$$

using $\frac{\partial \varphi_x^*}{\partial f_x} = \frac{\partial \varphi^*}{\partial f_x} \frac{\varphi_x^*}{\varphi^*} + \frac{1}{\sigma-1} \frac{\varphi_x^*}{f_x}$, (B.1), and (B.2). Hence, $\frac{\partial \varphi^*}{\partial f_x} < 0$ since the denominator in (F.3) must be negative ($j'(\varphi) < 0, \forall \varphi > 0$). Furthermore, as $j'(\varphi^*)\, \frac{\partial \varphi^*}{\partial f_x}$ is positive, inspection of (F.2) reveals that $n\, f_x\, j'(\varphi_x^*)\, \frac{\partial \varphi_x^*}{\partial f_x}$ must be negative and hence that $\frac{\partial \varphi_x^*}{\partial f_x} > 0$.

Welfare

Recall from (E.1) and (E.2) that welfare per worker is given by $W = L^{\frac{1}{\sigma-1}} \rho \left(\frac{1}{\sigma f}\right)^{\frac{1}{\sigma-1}} \varphi^*$ when $f_x > f$. A decrease in f_x induces an increase in the cutoff level ($\frac{\partial \varphi^*}{\partial f_x} < 0$ when $f_x > f$) and hence generates a welfare gain. As f_x decreases below f, the direction of the welfare change is identical to the one derived for a decrease in f in the closed economy. Welfare thus continues to rise as

---

$r_c{}'(\varphi_x^{*\prime}) < r_c(\varphi_x^{*\prime})$.

f_x decreases below f. Trade liberalization – interpreted as a reduction in f_x– always generates a welfare gain.

References

Aitken, Brian, Gordon Hanson, and Ann Harrison (1997) 'Spillovers, foreign investment, and export behavior.' *Journal of International Economics* 43, 103–32

Aw, Bee Yan, Xiaomin Chen, and Mark Roberts (1997) 'Firm-level evidence on productivity differentials.' *NBER Working Paper No. 6235*

Basu, Susanto, and John Fernald (1997) 'Aggregate productivity and aggregate technology.' *International Finance Discussion Paper No. 593*

Benassy, Jean-Pascal (1996) 'Taste for variety and optimum production patterns in monopolistic competition.' *Economic Letters* 52, 41–47

Bernard, Andrew, and Bradford Jensen (1999) 'Exceptional exporter performance: Cause, effect, or both?' *Journal of International Economics* 47, 1–25

Brown, Drusilla, Alan Deardorff, and Robert Stern (1996) 'Modelling multilateral trade liberalization in services.' *Asia Pacific Economic Review* 2, 21–34

Clerides, Sofronis, Saul Lach, and James Tybout (1998) 'Is learning by exporting important? micro-dynamic evidence from colombia, mexico, and morocco.' *Quarterly Journal of Economics* 113, 903–947

Davis, Stephen, and John Haltiwanger (1998) 'Gross job flows.' In *Handbook of Labor Economics,* ed. Ashenfelter and Card (North Holland). (forthcoming)

Dixit, Avinash, and Joseph Stiglitz (1977) 'Monopolistic competition and optimum product diversity.' *American Economic Review* 67, 297–308

Feldman, Mark, and Chirstian Gilles (1985) 'An expository note on individual risk without aggregate uncertainty.' *Journal of Economic Theory* 35, 26–32

Foster, Lucia, John Haltiwanger, and C.J. Krizan (1998) 'Aggregate productivity growth: Lessons from microeconomic evidence.' *NBER Working Paper No. 6803*

Haltiwanger, John (1997) 'Measuring and analyzing aggregate fluctuations: The importance of building from microeconomic evidence.' *Federal Reserve Bank of St. Louis Economic Review* pp. 55–77

Hopenhayn, Hugo (1992a) 'Entry, exit, and firm dynamics in long run equilibrium.' *Econometrica* 60, 1127–1150

—— (1992b) 'Exit, selection, and the value of firms.' *Journal of Economic Dynamics and Control* 16, 621–653

Jovanovic, Boyan (1982) 'Selection and the evolution of industry.' *Econometrica* 50, 649–70

Krugman, Paul (1979) 'Increasing returns, monopolistic competition, and international trade.' *Journal of International Economics* 9, 469–479

—— (1980) 'Scale economies, product differentiation, and the pattern of trade.' *American Economic Review* 70, 950–959

Levinsohn, James (1999) 'Employment responses to international liberalization in chile.' *Journal of International Economics* 47, 321–44

Montagna, Catia (1995) 'Monopolistic competition with firm-specific costs.' *Oxford Economic Papers* 47, 318–28

Olley, Steven, and Ariel Pakes (1996) 'The dynamics of productivity in the telecommunications equipment industry.' *Econometrica* 64, 1263–97

Roberts, Mark, and James Tybout (1997a) 'The decision to export in colombia: An empirical model of entry with sunk costs.' *American Economic Review* 87, 545–64

—— (1997b) *What Makes Exports Boom?* (The World Bank)

Roberts, Mark, and James Tybout, eds (1996) *Industrial Evolution in Developing Countries* (Oxford University Press)

Roberts, Mark, Theresa Sullivan, and James Tybout (1995) 'Micro-foundations of export booms.' *World Bank, mimeo*

Tybout, James (1996) 'Heterogeneity and productivity growth: Assessing the evidence.' In *Industrial Evolution in Developing Countries,* ed. Mark Roberts and James Tybout (Oxfor University Press) pp. 43–72

Tybout, James, and Daniel Westbrook (1995) 'Trade liberalization and the dimensions of efficiency change in mexican manufacturing industries.' *Journal of International Economics* 39, 53–78