Productivity Overshooting: The Dynamic Impact of Trade Opening with Heterogeneous Firms*

Thomas Chaney†
University of Chicago and NBER

June, 2005

Abstract

In this paper, I build a dynamic model of trade with heterogeneous firms which extends the work of Melitz (2003). As countries open up to trade, they will experience a productivity overshooting. Aggregate productivity increases in the long run, but it increases even more so in the short run. When trade opens up, there are too many firms, inherited from the autarky era. The most productive foreign firms enter the domestic market. Competition is fierce. The least productive firms that are no more profitable are forced to stop production. Not only do the most productive firms increase their size because they export, but the least productive firms stop producing altogether. Aggregate productivity soars. As time goes by, firms start to exit because of age. Competition softens. Some less productive firms resume production. This pulls down aggregate productivity. The slower the exit of firms, the larger this overshooting phenomenon. This model also predicts that the price compression that accompanies trade opening may be dampened in the long run. It also predicts that inequalities should increase at the time when a country opens up to trade, and then gradually recede in the long run.

*I am grateful to Marc Melitz, Xavier Gabaix and Daron Acemoğlu for their constant encouragements and their advice. All remaining errors are mine. First draft: June 2005.

†Contact: Department of Economics, The University of Chicago, Chicago, IL 60637. Email: tchaney@uchicago.edu.

1 Introduction

There are strong empirical evidences that trade opening induces massive reallocations of the factors of production, not only between sectors, but also between individual firms within a given sector. In a series of papers, Bernard and Jensen (1999, 2001a, 2001b, 2002) acknowledge the importance of those reallocations between firms, and the importance of heterogeneity between firms with regard to exports. Bernard, Eaton, Jensen and Kortum (2003) show that differences in productivity among firms may explain the patterns of international trade. Exporters are more productive, larger, and more capital intensive than non exporters. Using a panel of French firms, Eaton, Kortum and Kramarz (2004a, 2004b) document the large differences in size and productivity between exporters and non exporters, and between multiple countries exporters and single country exporters. The reallocation of the factors of production from low productivity firms towards high productivity exporters when trade is opened up accounts for large variations in aggregate volatility.

There is little understanding of the transitional dynamics following trade opening. Empirically, we do observe that the short run impact of trade opening is much larger than the long run impact. The exit of firms in import competing sectors following a trade liberalization is typically large, but short lived. Many adjustments that are typical of trade liberalization episodes seem to be stronger in the short run than in the long run.

I build a dynamic model of international trade with heterogeneous firms that helps explain the dynamics of productivity after a country opens up to trade. As countries open up to trade, they will experience a productivity overshooting. Aggregate productivity increases in the long run, but it increases even more so in the short run. When trade opens up, there are too many firms, inherited from the autarky era. The most productive foreign firms enter the domestic market. Competition is fierce. The least productive firms that are no more profitable are forced to stop production. Not only do the most productive firms increase their size because they export, but the least productive firms stop producing altogether. Aggregate productivity soars. As time goes by, firms start to exit because of age. The competitive situation improves. Some less productive firms resume production. This pulls down aggregate productivity. The slower the exit of firms, and the more competitive the economy, the larger this overshooting phenomenon. This model also predicts that the price compression that accompanies trade opening may be dampened in the long run. It also predicts that inequalities should increase at the time when a country opens up to trade, and then gradually recede in the long run.

This model extends the pioneering work on trade with heterogeneous firms of Marc Melitz (2003). Melitz describes the long run impact of trade opening. He only considers the steady state properties of such a model with heterogeneous firms. By considering the transitional dynamics, I am able to see how differences between the mass of firms in the short run and in the long run may lead to a non monotonic response of aggregate productivity along the transition towards the new steady state. Hopenhaym (1992) also describes the reallocation of production between firms in an dynamic model, but considers only long run predictions. Eaton and Kortum (2002) develop a model of international trade with heterogeneous firms which extends the framework of Dornbusch, Fischer and Samuelson (1977) to a multi country setting. They only describe the steady state properties of this model.

The model most closely related to this is Ghironi and Melitz (2005). They analyze the transitional dynamic in a model of trade with heterogeneous firms. They find that firm heterogeneity may explain systematic departures from purchasing power parity, and provides microfoundations for the Harrod-Balassa-Samuelson effect. In their model, they impose by assumption that no firm will be forced to exit their domestic market when trade is opened up. The only dynamic comes from the entry and exit of firms into the export market, and the exogenous natural death among firms. In addition, Ghironi and Melitz have to rely on numerical simulations to derive dynamics. I adopt a different formalization that allows me to account for the large exit of firms at times of trade liberalizations. I am able to describe analytically the transition towards a steady state where some firms are allowed to resume production.

The mechanism generating productivity overshooting is intimately related to the overshooting model of Rudiger Dornbusch (1976). Technically, it takes more time for the number of firms to adjust than for the relative size of firms. This is what explains the difference between short run and long run adjustments. In the short run, the number of firms cannot adjust discretely beyond a certain point. It only evolves sluggishly, as firms start dying. Therefore, in the short run, the relative size of firms must adjust in order to clear the labor market. The size of less productive firms shrinks whereas that of more productive firms increases. This shift of mass towards the most productive firms explains the large increase of aggregate productivity in the short run. As time goes by, the number of firms gradually adjusts, and the size of less productive firms increases faster than that of more productive ones. Aggregate productivity falls towards its long run steady state.

The remainder of the paper is organized as follows. Section 2 introduces a simple dynamic model of trade with heterogeneous firms. Section 3 describes the transitional dynamics when trade is opened up. Section 4 concludes.

2 A simple model of trade with heterogeneous firms

I build up a simple model of trade with heterogeneous firms based on Melitz (2003). For simplicity, I use similar notations as the ones used by Melitz, and I add new ones only when necessary.

The world is comprised of two identical countries, home and foreign. I will only consider symmetric equilibria.[1] Each country is populated with a mass L of workers. Those workers produce goods, earn wages and dividends, and consume. For simplicity, I assume that all workers own a single share in a mutual fund. The mutual fund owns all domestic firms, collects all their profits, invests in new firms when optimal, and redistributes all remaining profits to the workers. There are no international capital markets.[2] Perfect competition on the labor market, and identical ownership in the mutual fund allow me to consider that everything is as if all decisions were undertaken by a representative consumer.

2.1 Demand

Each worker is endowed with one unit of labor that she supplies inelastically. I normalize wages (equal across countries) to one, and express all prices in terms of wages. Workers share the same intertemporal utility. They consume a CES aggregate of differentiated goods in each period. If they consume a quantity $q_t(\varphi)$ of variety φ in period t, and all varieties in the set Φ_t, they derive a utility, $U_0 \equiv E_0\left[\sum_{t=0}^{+\infty} \beta^t C_t\right]$,with $C_t = \left(\int_{\varphi\in\Phi} q_t(\varphi)^{\frac{\sigma-1}{\sigma}} d\varphi\right)^{\frac{\sigma}{\sigma-1}}$. The elasticity of substitution between any two varieties is constant and equal to σ, β is a subjective discount factor. For simplicity, I will consider the limiting case where $\beta \underset{\beta<1}{\rightarrow} 1$, so that everything is as if $\beta = 1$, but the intertemporal utility is still well defined. The price P_t of one unit of the composite good C depends on the price of each variety: $P_t = \left(\int_{\varphi\in\Phi_t} p_t(\varphi)^{1-\sigma} d\varphi\right)^{\frac{1}{1-\sigma}}$, where Φ_t is the endogenous set of available goods. Given those isoelastic preferences, the representative consumer will spend a fraction of its income on each differentiated variety. How much of each variety she

[1] This assumption, as in Melitz (2003) simplifies the analysis greatly, since it prevents relative wages from adjusting.

[2] Note that the symmetry assumption makes international financial markets irrelevant.

consumes depends on the price of this variety relative to the price of others. The information about the price of all other varieties is summed up in the price of the composite consumption good P_t. If total expenditure on differentiated goods is R_t, the representative consumer spends $r_t(\varphi) = R_t (p_t(\varphi)/P_t)^{1-\sigma}$ on variety φ.

2.2 Production and trade

Labor is the only factor of production. Production is done under increasing returns to scale. Each firm must pay an overhead cost each period. This fixed per period cost is identical for all firms. Firms are heterogeneous in terms of productivity. The marginal cost of production is constant for each firm, but differs across firms. Each firm draws a random labor productivity shock φ, meaning that the unit labor requirement is equal to $1/\varphi$. For simplicity, the productivity of a firm is fixed upon entry, and does not evolve over the life-span of the firm. The cost of producing q units of goods for a firm with productivity φ is $c(\varphi) = f + q/\varphi$.

When trade is allowed between the two countries, there are two types of trade barriers: a fixed cost and a variable cost. In order to enter the foreign market, a firm must pay a fixed cost f_x. For simplicity, I assume that this fixed cost is paid each period. Having a sunk entry cost into the foreign market in addition to a fixed per period cost would not change the dynamics fundamentally. It would only slow down the adjustments of trade flows following the opening to trade between the two countries.[3] The variable trade cost is a traditional "iceberg" transportation cost. If 1 unit of good is shipped between the two countries, only a fraction $1/\tau$ arrives. The larger τ, the more expensive transportation.

Each firm is a monopolist for its own variety. If it does export, it is allowed to charge different prices in each market. Given the production technology, and the technology of transportation for international trade, and given that demand is isoelastic, firms will charge a constant mark-up over marginal cost. A firm with productivity φ, if it does survive, will charge a price $p_d(\varphi)$ on the domestic market; if it does export, it will charge a price $p_x(\varphi)$ on the foreign market:

$$p_d(\varphi) = \frac{1}{\rho\varphi} \text{ and } p_x(\varphi) = \frac{\tau}{\rho\varphi} \tag{1}$$

with $\frac{1}{\rho} = \frac{\sigma}{\sigma-1} > 1$ the mark-up charged by each firm.

[3] Outside of steady state, and unlike in Melitz (2003), it does matter whether the fixed cost of exporting is paid once and for all or at the beginning of each period. Those two formulations won't be equivalent anymore. I assume the cost is paid each period, which simplifies greatly the computation of the transitional dynamics. I believe results would not change qualitatively if I opted for the other formulation.

I define the distribution of firm productivity at time t by $\mu_t(\cdot)$, with all firms above productivity φ_t^* selling on the domestic market, and all firms with a productivity above $\varphi_{x,t}^*$ exporting. Both the sequence of μ_t's and of $\{\varphi^*, \varphi_x^*\}_t$'s will be determined in equilibrium. Plugging the prices set by each individual firm from Eq. 1 into the price index, aggregate prices at time t are defined by,

$$(\rho P_t)^{1-\sigma} = \int_{\varphi_t^*}^{+\infty} \varphi^{\sigma-1} \mu_t(\varphi)\, d\varphi + \int_{\varphi_{x,t}^*}^{+\infty} \left(\frac{\varphi}{\tau}\right)^{\sigma-1} \mu_t(\varphi)\, d\varphi \tag{2}$$

I can therefore compute quantities sold by each firm, both at home and abroad, profits earned by firms, and from these profits, I know the set of firms that are able to survive, and the set of firms that are able to export.[4] At time t, a firm with productivity φ, if it does produce at all, produces for the domestic market a quantity $q_{d,t}(\varphi) = \frac{R_t}{P_t}(P_t \rho \varphi)^{\sigma}$, so that its total domestic sales are $r_{d,t}(\varphi) = R_t (P_t \rho \varphi)^{\sigma-1}$, and the total profits it earns from selling domestically are $\pi_{d,t}(\varphi) = \frac{r_{d,t}(\varphi)}{\sigma} - f$. If this firm is able to export, it produces for the foreign market a quantity $q_{x,t}(\varphi)_{f.o.b.} = \frac{R_t}{P_t}\left(\frac{P_t \rho \varphi}{\tau}\right)^{\sigma}$ (or including the shipping cost, $q_{x,t}(\varphi)_{c.i.f.} = \frac{\tau R_t}{P_t}\left(\frac{P_t \rho \varphi}{\tau}\right)^{\sigma}$), so that its total foreign sales are $r_{x,t}(\varphi) = R_t \left(\frac{P_t \rho \varphi}{\tau}\right)^{\sigma-1}$, and the total profits it earns from exporting are $\pi_{x,t}(\varphi) = \frac{r_{x,t}(\varphi)}{\sigma} - f_x$.

Assumption 1 *In any period, a firm may decide not to produce any quantity. In such a case, it does not have to incur the overhead cost. In other words, a firm that wants to survive does not have to earn negative profits in order to stay in business.*[5]

As in Melitz, I can define thresholds for domestic production, and for exports. No firm will produce quantities if it means earning non positive profits, and no firm will export if it means earning non positive profits from exporting. I define φ_t^* as the productivity of the least productive firm earning non negative profits from domestic sales: $\pi_{d,t}(\varphi_t^*) = 0$. By the same token, $\varphi_{x,t}^*$ is the productivity of the least productive firm earning non negative profits from exporting: $\pi_x(\varphi_x^*) = 0$. I solve for the thresholds φ_t^* and $\varphi_{x,t}^*$. The conditions defining those threshold are the zero cutoff

[4] Note that I use the productivity φ as the identity of a firm. Literally, there is zero mass of firms with a productivity exactly equal to φ. In this continuum setting, we can say that the number of firms with a productivity φ is equal to $\mu(\varphi)\, d\varphi$, the density at φ. So potentially, there are "more than one" single firm with such a productivity. Formally, each of these firms has a different identity (each of them produces a unique differentiated variety). However they all behave in exactly the same way. They are indistinguishable from their actions. Hence I can safely abuse language and identify a firm by its productivity, φ.

[5] This assumptions insures that the distribution of productivity among surviving firms is stationary. It greatly simplifies the description of the dynamic adjustments after opening to trade. I discuss relaxing this assumption (and forcing firms to pay overhead costs each period in order to stay in business) in appendix A.

profit conditions (domestic and foreign):

$$\varphi_t^{*\sigma-1} = \frac{\sigma f}{R_t} \times (\rho P_t)^{1-\sigma} \qquad (ZCP_t)$$

$$\varphi_{x,t}^{*\sigma-1} = \frac{\sigma f_x}{R_t} \times \left(\frac{\rho P_t}{\tau}\right)^{1-\sigma} = \left(\tau^{\sigma-1}\frac{f_x}{f}\right) \times \varphi_t^{*\sigma-1} \qquad (ZCP_{x,t})$$

Any firm with a productivity below φ_t^* will not produce for the domestic market, and no firm with a productivity below the threshold $\varphi_{x,t}^*$ will export.[6] From those productivity thresholds, I can define the probability of exporting, conditional on survival: $p_{x,t} = \frac{P(\varphi>\varphi_{x,t}^*)}{P(\varphi>\varphi_t^*)}$.

2.3 Entry and exit of firms

The distribution of firms at any point in time is the result of a history of entry and exit of firms.

Entry is done in the following way. An entrepreneur may decide to start up a firm. In order to do so, she must pay a sunk entry cost f_e. Once this cost is paid, she receives a productivity shock φ, drawn from a random distribution with c.d.f. $g(\cdot)$ and p.d.f. $G(\cdot)$ defined over the support $[\varphi_{\min}, +\infty)$.[7] Any firm that does not expect to earn positive profits in the future exits. In addition, all surviving firms have an exogenous probability δ of dying each period.

I assume free entry. If a firm with productivity φ earns total profits $\pi_t(\varphi)$ in period t (domestic profits, which may be equal to zero at some points in time, plus, for some firms, export profits), then the free entry condition at date t states that if there are firms that enter at time t, the value of entering, $v_{e,t}$, must equal the cost of entering, f_e,

$$v_{e,t} = \int_{\varphi_{\min}}^{+\infty} \left(\sum_{s=t}^{+\infty} (1-\delta)^{s-t} \pi_s(\varphi)\right) g(\varphi)\, d\varphi \leq f_e \qquad (FE_t)$$

If $f_e > \int_1^{+\infty} \left(\sum_{s=t}^{+\infty} (\beta(1-\delta))^{s-t} \pi_s(\varphi)\right) \mu_t(\varphi)\, d\varphi$, no firm enters. Free entry prevents the other strict inequality from ever happening. Condition FE_t holds with equality as long as a strictly positive number of firms do enter. Note that it will be crucial to define the profit function $\pi_t(\cdot)$. This function is potentially complex along the transition following opening to trade. Firms

[6] I assume that $\frac{f_x}{f}\tau^{\sigma-1} > 1$. Trade barriers are always sufficiently high so that only a subset of firms export: $\varphi_t^* < \varphi_{x,t}^*$. If this condition were violated, then $\varphi_t^* = \varphi_{x,t}^*$. All active firms would export, and all results would carry through.

[7] The only condition on the distribution of productivity shocks is that the $(\sigma-1)^{\text{th}}$ moment of G is defined, or that the integral $\int_{\varphi_{\min}}^{+\infty} \varphi^{\sigma-1} g(\varphi)\, d\varphi$ converges. This property ensures that the total size of the economy is finite. The choice of $\varphi_{\min}$ is purely arbitrary. The assumption that the support for this distribution is unbounded from above simplifies notations greatly, but it is only a notational assumption. It is perfectly admissible within this model that there is zero mass above a certain threshold $\varphi_{\max}$ $(g(\varphi) = 0,\ \forall \varphi > \varphi_{\max})$.

may not earn any profit over some period of time, and then start earning positive profits. After some point in time, they may even start earning some extra profits from exporting.

The free entry condition and the general equilibrium will determine how many firms enter each period. Call $M_{e,t}$ the number of new entrants at time t. I must now impose that the labor market clears. Labor is used for investment (to cover the sunk entry cost of new entrants), and for production. The labor allocated to production is used both to pay for the fixed costs (fixed overhead cost for domestic production plus fixed trade barrier if the firm exports), and to cover the variable cost of production. The total labor used for investment[8] is $f_e \times M_{e,t}$. If the there is a total mass M_t of firms operating at time t, and a fraction $p_{x,t}$ of those firms are exporters, the total workforce used to cover overhead costs is $(f + p_{x.t} f_x) \times M_t$. Since each firm charges a constant mark-up over marginal cost, it can easily be proven that if total expenditures on differentiated goods is R_t, the total workforce used for producing differentiated goods (variable cost only) is ρR_t. The labor market clearing condition at each point in time during the transition requires,

$$L = f_e \times M_{e,t} + (f + p_{x.t} f_x) \times M_t + \rho R_t \tag{LMC_t}$$

I will see in the next two sections how the free entry condition characterizes the autarky and trade steady states. I will then see how entry of new firms take place along the transition from the autarky steady state to a new trade steady state.

2.4 Autarky steady state

In this section and the next, I recall Melitz (2003) computation of the steady state of this economy, both under autarky, and under trade. I will denote the autarky steady state by the time subscript $t = -\infty$, and the trade steady state by the time subscript $t = +\infty$.

In a steady state, there is as much entry as exit. All firm with a productivity below $\varphi^*_{-\infty}$ exit immediately upon receiving their productivity draw. Labor market clearing insures that total expenditure on differentiated goods is exactly equal to L. All profits are used to invest into starting up new firms, so that the mutual fund's finances are balanced. So in the autarky steady state, $(LMC_{-\infty}) \Leftrightarrow R_{-\infty} = L$.

Following Melitz (2003), I can define a special average productivity among active firms:

$$\tilde{\varphi}(\varphi^*) \equiv \left(\frac{\int_{\varphi^*}^{+\infty} \varphi^{\sigma-1} g(\varphi)\, d\varphi}{\int_{\varphi^*}^{+\infty} g(\varphi)\, d\varphi} \right)^{\frac{1}{\sigma-1}}$$

[8] It is possible that at some point in time, there is no entry, so that $M_{e,t} = 0$.

It will be useful to determine the total mass of firms, $M_{-\infty}$. Using the special average notation, the following accounting identity holds: $R_{-\infty} = M_{-\infty} \times r\left(\tilde{\varphi}_{-\infty}\right)$. Since expected profits are constant over time, I can define average profits conditional on survival, $\bar{\pi}_{-\infty} = \pi_{d,-\infty}\left(\tilde{\varphi}_{-\infty}\right)$. The steady state autarky equilibrium is defined by the zero cutoff profits condition (which defines $\varphi^*_{-\infty}$), the free entry condition, and the labor market clearing condition:

$$\left\{\begin{array}{l} (ZCP_{-\infty}) \\ (FE_{-\infty}) \\ (LMC_{-\infty}) \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} \bar{\pi}_{-\infty} = f\left(\left[\frac{\tilde{\varphi}\left(\varphi^*_{-\infty}\right)}{\varphi^*_{-\infty}}\right]^{\sigma-1} - 1\right) \\ \bar{\pi}_{-\infty} = \frac{\delta f_e}{P\left(\varphi > \varphi^*_{-\infty}\right)} \\ M_{-\infty} = \frac{L}{\sigma\left(\bar{\pi}_{-\infty} + f\right)} \end{array}\right.$$

The zero cutoff profit condition and the free entry condition define two schedules of $\bar{\pi}_{-\infty}$ as a function of $\varphi^*_{-\infty}$, which have a unique intersection.[9] In every period, $\delta M_{-\infty}$ firms die from attrition, and $M_{e,-\infty} = \delta M_{-\infty}/P\left(\varphi > \varphi^*_{-\infty}\right)$ firms are created, among which only those firms with a productivity above $\varphi^*_{-\infty}$ survive, and replace the deceased firms.

In the next section, I describe the steady state that the economy will reach after trade between the two countries is opened up.

2.5 Trade steady state

In the trade steady state, there must also be as much entry as there is exit. Upon receiving their productivity shock, incumbents with a productivity below $\varphi^*_{+\infty}$ exit immediately. Among survivors, all firms with a productivity above $\varphi^*_{x,+\infty}$ export. The average profits that a firms earns, conditional on surviving, are the sum of profits earned domestically, and profits earned from exporting: $\bar{\pi}_{+\infty} = \pi_{d,+\infty}\left(\tilde{\varphi}_{+\infty}\right) + p_{x,+\infty}\pi_x\left(\tilde{\varphi}_x\right)$, where $p_{x,+\infty} = \frac{P\left(\varphi > \varphi^*_{x,+\infty}\right)}{P\left(\varphi > \varphi^*_{+\infty}\right)}$ is the probability of exporting, conditional on survival. The stationary assumption implies that the labor market clearing condition is the same in the trade steady state as in the autarky steady state: $(LMC_{+\infty}) \Leftrightarrow R_{+\infty} = L$. The trade steady state is defined by the two zero cutoff profits conditions (which define $\varphi^*_{+\infty}$ and $\varphi^*_{x,+\infty}$), the free entry condition, and the labor market clearing

[9] See Melitz (2003) for a formal proof of this statement.

condition:

$$\begin{cases} (ZCP_{x,+\infty}) \\ (ZCP_{+\infty}) \\ (FE_{+\infty}) \\ (LMC_{+\infty}) \end{cases} \Leftrightarrow \begin{cases} \varphi^*_{x,+\infty} = \tau \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma-1}} \varphi^*_{+\infty} \\ \bar{\pi}_{+\infty} = f\left(\left[\frac{\tilde{\varphi}(\varphi^*_{+\infty})}{\varphi^*_{+\infty}}\right]^{\sigma-1} - 1\right) + p_{x,+\infty} f_x \left(\left[\frac{\tilde{\varphi}(\varphi^*_{x,+\infty})}{\varphi^*_{x,+\infty}}\right]^{\sigma-1} - 1\right) \\ \bar{\pi}_{+\infty} = \frac{\delta f_e}{P(\varphi > \varphi^*_{+\infty})} \\ M_{+\infty} = \frac{L}{\sigma(\bar{\pi}_{+\infty} + f + p_x f_x)} \end{cases}$$

As is the case in the autarky steady state, $\delta M_{+\infty}$ firms die from attrition each period, and are replaced by new entrants whose productivity is above $\varphi^*_{+\infty}$. One can easily prove that $\bar{\pi}_{+\infty} > \bar{\pi}_{-\infty}$, $\varphi^*_{+\infty} > \varphi^*_{-\infty}$ and $M_{+\infty} < M_{-\infty}$.

There are two important properties of the trade steady state to be noted. First, there are fewer firms when trade is allowed than under autarky, $M_{+\infty} < M_{-\infty}$. This is a crucial prediction of the Melitz model of trade with heterogeneous firms. In autarky, firms that have access to a given technology are the only ones to compete for the home market. A given number of those firms can survive. As trade opens up, some highly productive firms are able to export. Those firms increase the toughness of competition. Fewer firms can survive. Those high productivity firms eat up an even larger share of the market. Mechanically, since those firms are more productive, fewer firms are required to service the entire market.

The second prediction, which is the central claim of the Melitz model, is that average productivity increases after trade is opened up. The most productive firms from abroad enter the domestic market, and push the least productive firms out of business. At the same time, the most productive domestic firms, who have access to the export market, benefit disproportionately more from the possibility of exporting. Not only does the productivity threshold for survival go up, but the share of the most productive firms increases compared to that of the least productive ones.

In the next section, I describe the transition towards the new steady state after both countries symmetrically open up to trade.

3 The dynamics of trade opening: productivity overshooting

In this section, I describe the transition of the economy from autarky towards trade. I first describe intuitively the forces driving the transitional dynamics, before turning to the formal derivation of those dynamics.

If firms are heterogeneous in terms of productivity, only a subset of firms, the most productive, are able to overcome trade barriers. The presence of these high productivity exporters, along with the upward shift in average productivity, implies that fewer firms can survive under trade than under autarky. In a sense, in the autarky steady state, domestic firms are alone to satisfy the entire domestic demand, and many firms must operate. When trade is opened up, there are "too many" firms. So during the transition, the mass of firms must shrink.

There is a fundamental asymmetry between the creation and destruction of firms, due to the presence of sunk entry costs. Because I have assumed free entry, as soon as there is some potential for profits, there will always be some firms entering. Exit on the other hand may take time. If the mass of firms at every level of productivity must shrink in order to reach the new steady state, because high productivity firms only die at a slow pace, the transition will be slow. Because existing firms have already paid the sunk entry cost, they are far less vulnerable than potential entrants. They will exit if and only if they do not expect to earn any positive profits at any point in the future.

At the moment when trade is opened up, the most productive foreign firms start exporting. They eat up part of the domestic demand, and push many low productivity firms out of business. The least productive among those firms will never be able to generate positive profits ever again in this globalized world, and exit immediately. Competition is at its fiercest right at the time when trade is opened up. So upon opening to trade, there is a spike of destruction of firms. This implies a large increase in average productivity: only the most productive firms can survive in this new environment.

Because the world has inherited an "overcrowded" economy, there won't be any entry of firms for some period of time. During this transitional period, the natural death eats up the total mass of firms. Competition gradually softens. Firms with a low productivity that had been on hold until then can start producing again. Exporting becomes easier as the mass of firms shrinks. The mass of firms gradually shifts towards less productive firms that can more easily survive now.

Eventually, when the mass of firms has shrunk sufficiently, the new trade steady state is reached, with a lower average productivity than at the time of trade opening. Once this state is reached, entry starts again. One important property of this model is that the dynamics towards the trade steady state only take a finite amount of time. This crucially depends on assumption 1, which guaranties a stationary distribution of firm productivity.

In the next three sections, I derive formally the transitional dynamics, and their properties.

3.1 Transitional dynamics

Before opening to trade, the economy was in an autarky steady state, defined in section 2.4. Trade opens up at time $t = 0$. The opening is unexpected. Because of the dynamic nature of the model, there may be multiple rational expectations equilibria. I will only consider a class of dynamic equilibria, those that converge towards a steady state.

It will be useful to define the following alternative measure of the mass of firms: $\tilde{M}_t = M_t / P(\varphi > \varphi_t^*)$. This measure corresponds to the mass of firms per unit of density at each level of productivity. It also corresponds to an ideal notion of the total number of firms, which includes those firms that cannot survive (with a productivity below φ_t^*). Depending on the value of this alternative measure of mass in the trade steady state, there are two possible scenarii for the transition path. If $\tilde{M}_{+\infty} \geq \tilde{M}_{-\infty}$, the transition will be immediate. The economy jumps to the new trade steady state within one period. If $\tilde{M}_{+\infty} < \tilde{M}_{-\infty}$ on the other hand, the transition towards the new steady state takes a finite time.

The reason for this is simple. If there are more firms per level of productivity in the trade steady state than in the autarky steady state ($\tilde{M}_{+\infty} \geq \tilde{M}_{-\infty}$), competition is softer after trade is opened than it will be in the steady state, average profits are higher than in the steady state. I know that in the trade steady state, the discounted stream of profits is exactly equal to the sunk entry cost. Hence the appeal of extra profits will attract an influx of new firms, in order to restore the free entry condition (FE_t) with equality. Those firms are spread all over the distribution of productivity, so that the mass of firms at each level of productivity jumps immediately to its steady state level.

If on the other hand, there are fewer firms per level of productivity in the trade steady state than in the autarky steady state ($\tilde{M}_{+\infty} < \tilde{M}_{-\infty}$), competition is tougher after trade openes than it will be in the new steady state, average profits are lower. Since in the steady state, profits are just enough to cover the sunk entry cost, lower average profits implies that no firm will enter as long as $\tilde{M}_t \geq \tilde{M}_{+\infty}$. The natural death process gradually erodes the mass of firms, until the mass per level of productivity reaches its steady state level. From that point onward, the economy is in the trade steady state, and entry resumes in order to offset death from attrition.

I now turn to the formal proof of these statements.First, I define the criterion for fast or slow convergence towards the steady state. This criterion depends on whether the economy is overcrowded after trade is opened up or not.

Criterion 2 (Overcrowding)

$$\tilde{M}_{+\infty} < \tilde{M}_{-\infty}$$

Remarks: *If criterion 2 is met, there are more firms at every level of productivity in the autarky steady state than in the trade steady state (the economy is "overcrowded"), then the convergence towards the new steady state will take some finite amount of time, and there is overshooting in productivity. If criterion 2 is not met, then the economy immediately adjusts to its new trade steady state. See appendix B for the full functional form of this criterion.*

Proposition 3 *If criterion 2 is not satisfied, that is if $\tilde{M}_{+\infty} \geq \tilde{M}_{-\infty}$, there is an influx of firms upon trade opening: $M_{e,0} = \left(\tilde{M}_{+\infty} - \tilde{M}_{-\infty}\right) + \delta\tilde{M}_{-\infty}$. From period $t = 1$ onward, the economy is in the trade steady state defined in section 2.5. The productivity thresholds immediately jump to their steady state values, $\varphi_0^* = \varphi_{+\infty}^*$, and $\varphi_{x,0}^* = \varphi_{x,+\infty}^*$.*

Proof. It is sufficient to prove that when trade is allowed and $M_{e,t} = 0$, $\tilde{M}_t < \tilde{M}_{+\infty} \Rightarrow P(\varphi > \varphi_t^*)\,\bar{\pi}_t > P\left(\varphi > \varphi_{+\infty}^*\right)\bar{\pi}_{+\infty}$.

Assume, as proven in appendix B, that this property holds.

If $\tilde{M}_0 = \tilde{M}_{+\infty}$, the economy has already reached its new steady state, and no further adjustment occurs.[10]

If the mass of firms per level of productivity were strictly lower than its trade steady state value, $\tilde{M}_0 < \tilde{M}_{+\infty}$, total expected profits would be larger than the sunk entry cost:

$$\begin{aligned} v_{e,0} &= \sum_{t\geq 0} (1-\delta)^t \times P(\varphi > \varphi_0^*)\,\bar{\pi}_t \\ &> v_{e,+\infty} = \sum_{t\geq 0} (1-\delta)^t \times P\left(\varphi > \varphi_{+\infty}^*\right)\bar{\pi}_{+\infty} = f_e \end{aligned}$$

Free entry prevents the occurence of such an imbalance. So there will be a net entry of new firms that increases the mass of firms to its steady state level. In order to reach the trade steady state mass of firms, those firms destroyed by attrition must be replaced by new entrants: $\delta\tilde{M}_{-\infty}$ must enter. In addition, the ideal mass of firms $\left(\tilde{M}\right)$ must be increased to $\tilde{M}_{+\infty}$, from $\tilde{M}_{-\infty}$. So the total number of entrants is $M_{e,0} = \left(\tilde{M}_{+\infty} - \tilde{M}_{-\infty}\right) + \delta\tilde{M}_{-\infty}$. From $t = 1$ onward, we are in the trade steady state, and new firms enter only to replace attrition deaths: $M_{e,t} = M_{e,+\infty} = \delta\tilde{M}_{+\infty}$.

Along this path, expected profits and survival thresholds are constant, $P(\varphi > \varphi_0^*)\,\bar{\pi}_0 = ... = P\left(\varphi > \varphi_{+\infty}^*\right)\bar{\pi}_{+\infty}$, so that the free entry condition (FE_t) is satisfied with equality for all $t \geq 0$. ■

[10] This will be the case with Pareto distributed productivity shocks.

Proposition 4 *If criterion 2 is satisfied, that is if* $\tilde{M}_{+\infty} < \tilde{M}_{-\infty}$*, there a finite length of time* $T \geq 0$ *during which no firm enters. At time* $t = T$*, the trade steady state defined in section 2.5 is reached.* T *is uniquely defined as the minimum integer such that* $\tilde{M}_T \leq \tilde{M}_{+\infty}$*.*

Proof. Assume, as proven in appendix B, that when $M_{e,t} = 0$, $\tilde{M}_t > \tilde{M}_{+\infty} \Rightarrow P(\varphi > \varphi_t^*)\,\bar{\pi}_t < P\left(\varphi > \varphi_{+\infty}^*\right)\bar{\pi}_{+\infty}$.

I need to prove that

$$\left\{\begin{array}{c} M_{e,0} = M_{e,1} = ... = M_{e,T-1} = 0 \\ M_{e,T} = \left(\tilde{M}_{+\infty} - \tilde{M}_{T-1}\right) + \delta \tilde{M}_{T-1} \\ M_{e,T+1} = ... = M_{e,+\infty} = \delta \tilde{M}_{+\infty} \end{array}\right\}$$

is a rational expectations equilibrium. To do so, I prove that if firms expect that the economy will follow this path, indeed their entry decisions will follow these patterns.

As long as no entry takes place, the mass of firms evolves according to the following law of motion,

$$\tilde{M}_t = (1-\delta)\,\tilde{M}_{t-1} \tag{M_t}$$

or $\tilde{M}_t = (1-\delta)^t\,\tilde{M}_{-\infty}$. The mass of firms steadily declines. Define T as the minimum integer such that $\tilde{M}_T \leq \tilde{M}_{+\infty}$. Since $\tilde{M}_{-\infty} > \tilde{M}_{+\infty} > 0$, since $\tilde{M}_t$ is strictly decreasing in t and converges towards zero, T is uniquely defined.[11] By definition of T, I know that $\tilde{M}_0 > ... > \tilde{M}_{T-1} > \tilde{M}_{+\infty} \geq \tilde{M}_T$.

I now prove recursively that:

"at $t = T-1$, no firm enters if agents expect the equilibrium path to be followed after $t = T-1$"

"at $t = T-2$, no firm enters if agents expect the equilibrium path to be followed after $t = T-2$"

"..."

"at $t = 0$, no firm enters if agents expect the equilibrium path to be followed after $t = 0$"

From $\tilde{M}_{T-1} > \tilde{M}_{+\infty} \Rightarrow P(\varphi > \varphi_t^*)\,\bar{\pi}_t < P\left(\varphi > \varphi_{+\infty}^*\right)\bar{\pi}_{+\infty}$, I know that $P\left(\varphi > \varphi_{T-1}^*\right)\bar{\pi}_{T-1} < P\left(\varphi > \varphi_{+\infty}^*\right)\bar{\pi}_{+\infty}$. If agents expect the equilibrium path to be followed after $t = T-1$, from the free entry condition in the trade steady state ($v_{e,+\infty} = f_e$), I know that,

$$\begin{aligned} v_{e,T-1} &= \sum_{t \geq T-1} (1-\delta)^t\, P(\varphi > \varphi_t^*)\,\bar{\pi}_t \\ &= P\left(\varphi > \varphi_{T-1}^*\right)\bar{\pi}_{T-1} + (1-\delta)\,v_{e,+\infty} < f_e \end{aligned}$$

[11] I consider the interesting case where $T > 0$. If $T = 0$, the adjustment to the new steady state is immediate. It is always possible to change the unit of observation for time, and hence to reduce δ sufficiently so that $T > 0$.

So if agents expect the equilibrium path to be followed after $t = T-1$, no firm enters at $t = T-1$.

By the same reasoning, if agents expect the equilibrium path to be followed after $t = T-2$, $P\left(\varphi > \varphi^*_{T-2}\right) \bar{\pi}_{T-2} < P\left(\varphi > \varphi^*_{+\infty}\right) \bar{\pi}_{+\infty}$, and no firm enters at $t = T-2$.

No firm enters until $t = T$, and from that point onward, the economy is in steady state.

Along the transition, the equilibrium will be determined by the zero cutoff profit conditions, the labor market clearing condition, and the law of motion for the mass of firms. Once the steady state is reached, the law of motion of the mass of firms is replaced by the free entry condition:

$$
\begin{aligned}
0 \leq\ & t < T \begin{cases} (ZCP_t) \\ (ZCP_{x,t}) \\ (M_t) \\ (LMC_t) \end{cases} \\
t \geq\ & T \begin{cases} (ZCP_{x,+\infty}) \\ (ZCP_{+\infty}) \\ (FE_{+\infty}) \\ (LMC_{+\infty}) \end{cases}
\end{aligned}
$$

■

As can be guessed from the description of the dynamic evolution of average profits and the mass of firms, the productivity threshold jumps upon opening to trade, and then gradually falls towards its steady state level. I describe this phenomenon of productivity overshooting in the next section.

3.2 Productivity overshooting

In this section, I will only consider the interesting case where transitional dynamics are not collapsed into one single period.

As trade opens up, there are too many firms. Competition is fierce, profits are low. Many firms are forced to stop producing, since they could not even cover their overhead costs. Only the most productive firms are still active, and the most productive among them are exporters. As time goes by, since no firm enters during the transition, firms at every level of productivity start dying. For the survivors, the situation gradually improves. This means that some firms can start exporting. Those new exporters experience a discrete increase in the volume of their sales, and their employment. As the productivity threshold for exports falls, mass is shifted towards those less productive exporters. At the same time, some firms that had stopped producing altogether resume their production. Some of the varieties that had disappeared when the most productive foreign exporters had entered start being produced again. The threshold for domestic sales falls.

Aggregate productivity, measured as a size-weighted average of productivities across firms, falls down for two reasons. First, those firms that resume production have a low productivity, and they pull down average productivity. At the same time, the mass of sales is shifted gradually away from the high productivity exporters towards the new lower productivity exporters. So as more and more firms die, aggregate productivity falls as well. Eventually, the economy reaches its steady state when productivity stays constant.

Following trade opening, aggregate productivity increases in the long run. This is the main prediction of the Melitz model. I predict that in the short run, productivity will increase more than in the long run. This rapid increase in productivity followed by a gradual deterioration of productivity is what I call productivity overshooting. The forces driving this overshooting in productivity are very similar in spirit to the forces driving exchange rate overshooting in Dornbusch (1976). In that model, the reason why exchange rate overshoots is because exchange rates can adjust much faster than domestic prices. In this model, productivity overshoots because it takes time for firms that are already here to die. Because those existing firms will not die right away, something else must adjust in the meantime to offset the imbalances created by the sudden entry of foreign exporters. This is done through a temporary reduction in the share of low productivity firms. Some of these less productive firms stop producing altogether for some time (until sufficiently many high productivity firms have died, and the situation has sufficiently improved). Some other firms do not stop producing, but they reduce their production relative to higher productivity firms (which start exporting). This adjustment, the reduction in the share of less productive firms, can happen much faster than forcing existing firms out. So in the short run, this will be the only variable of adjustment. In a sense, one could say that the extensive margin of productivity (how many levels of productivity can be active) can adjust much faster than the intensive margin of productivity (how many firms there are at each level of productivity). This reduction in the share of less productive firms in the short run will cause productivity to increase substantially. In the longer run, as firms die and the mass of firms shrinks, this increase in aggregate productivity is dampened. The death of firms allows for the share of less productive firms to increase again, partially offsetting the short run productivity gains.

Note that some firms will exit definitively at the time when trade is opened up. Those are the firms that will not survive, even once the new steady state is reached. If criterion 2 does not hold, there will actually be no overshooting, and productivity directly jumps to its steady state level. This is because I do allow for immediate exit of firms when trade is opened up. Firms have to pay a fixed overhead production cost each period, and therefore some firms, after trade is opened up, know that they will never be profitable again, and exit immediately. If I remove the assumption of a fixed overhead cost, as is done in a similar setting by Ghironi and Melitz (2005), adjustments will be even more sluggish, and productivity overshooting will be larger. In such a setting, the prediction that the total mass of firms must be reduced still holds. There is no possibility of discretely adjusting this mass, so the adjustment will take much more time than in the current model. I believe however that the prediction that opening up to trade will have a sudden negative impact on the least productive domestic firms is a plausible feature. It is interesting to see that even when some firms exit, transition towards the new steady state may still take some time, and we may observe a phenomenon of productivity overshooting.

I will not go into the details of computing the average productivity of firms (weighted by the size of sales, or by employment) in the economy along the transition. I will only prove that along the transition, as the mass of firms shrinks, the productivity threshold for survival falls. It is intuitive to see that as the productivity threshold falls towards its new steady state level, aggregate productivity also falls. The following proposition proves formally the productivity overshooting triggered by trade opening.

Proposition 5 (Overshooting) *When the transition is not instantanuous, there is overshooting in productivity,*

$$\varphi_0^* > \varphi_1^* > ... > \varphi_T^* = ... = \varphi_{+\infty}^* > \varphi_{-\infty}^*$$

Proof. We already know from Melitz (2003) that productivity is higher in the trade steady state than in the autarky steady state: $\varphi_{+\infty}^* > \varphi_{-\infty}^*$

If criterion 2 is met, I know that the mass of firms per level of productivity, $\tilde{M}_t$, will gradually fall from its autarky level, towards its trade steady state level (at a constant rate $(1-\delta)$):

$$\tilde{M}_{-\infty} > \tilde{M}_0 > \tilde{M}_1 > ... > \tilde{M}_T = ... = \tilde{M}_{+\infty}$$

See appendix B for a proof that $\tilde{M}_t > \tilde{M}_s \Rightarrow \varphi_t^* > \varphi_s^*$. ■

It is important for this overshooting phenomenon to happen that the opening to trade is unexpected. Looking at the transitional dynamics from an ex ante point of view, we can see further justifications for this overshooting in productivity. There are too many firms at the time when trade is opened up. The profitability conditions worsen in such a way that no firm has any incentive to enter during the transition towards the new steady state. This implies that entrepreneurs who started up their company before the opening to trade had formed the wrong expectations regarding their future expected stream of profits. Average realized profits will fall below their expected level for some time. Had those entrepreneurs known in advance about this opening to trade, they would not have started up their company.[12] This means that there was overinvestment in autarky.

In other words, the "technology" of production has improved, in an ex ante expectation sense, because there has been overinvestment in improving the available technology (creation of too many firms). This is why less investment (payments to start up new firms) is required each period in the trade steady state than in the autarky steady state. Since the economy had overinvested while in autarky, it is natural that the opening to trade creates a temporary spike in productivity. This is only temporary. As the economy converges towards its new steady state, there is disinvestment (no creation of new firms while old ones die). Thanks to the available "better technology" offered by trade, the aggregate productivity in the new steady state is higher than in autarky. But it is lower than during the transitional period, which was a time of abundance of "capital".

3.3 Price compression and inequality

As trade opens up, the dynamics of productivity are parallel to movements in prices and inequalities. Such a model with heterogeneous firms is perfectly fit for describing the impact of trade opening on aggregate prices as well as inequalities (between firms, in the absence of any formalization of wage bargaining).

As trade opens up, prices are driven down by the entry of high productivity foreign exporters. Literally speaking, since the model imposes that mark-ups are constant, no single firm changes the price it sets. However, low productivity/ high price firms exit, and since the share of high productivity/ low price firms increases, aggregate prices fall when trade is opened up. This fall in aggregate prices is exactly symmetrical to average productivity increases. In this model, productivity is directly a mirror image of prices. This fall in aggregate prices happens in the long run.

[12] I describe qualitatively what happens when the opening to trade is announced in advance in appendix C.

But exactly in the same way as there is overshooting in productivity, there will be "overcompression" of prices in the short run. In the short run, only the most productive firms are active. So prices are pushed down substantially. As firms exit, and as competition softens along the transition path towards the new steady state, less productive firms resume production. Those firms charge a higher price than their competitors. So average prices will "undershoot", in the sense that there is a sudden drop of prices in the short run, and a gradual price increase in the medium run.

Melitz (2003) points out that in the presence of firm heterogeneity, inequalities between firms increase. This is due to the fact that more productive firms benefit from trade, whereas the situation of less productive firms worsens. So inequalities increase in the long run. But once again, inequalities increase more in the short run than in the long run. There is overshooting in inequalities. In the short run, only the very best firms benefit from the access to foreign markets, whereas many less productive firms have to incur a temporary loss in profits (which goes all the way to zero profits for some time for firms with a productivity $\varphi^*_{+\infty} < \varphi < \varphi^*_0$). In the short run, inequalities soar. In the medium run, as some firms exit, inequalities among survivors recede, until the new steady state is reached. Inequalities will always be higher in a trade regime than in autarky, but less so in the long run than in the short run. It takes some time for the economy to adapt to the new regime.

4 Conclusion

In this paper, I have shown that when a country opens up to trade, it will experience an sudden increase in productivity, followed by a gradual reduction towards a steady state. Opening up to trade does lead to a permanent increase of productivity, but there is an even larger increase in the short run. I call this phenomenon productivity overshooting. The reason for this productivity overshooting is simple. Because trade allows are more efficient allocation of factors of production, there will be fewer firms in a globalized world than in a world composed of countries in autarky. After countries open up to trade, the number of firms will have to fall. However, in the presence of sunk entry costs, such a reduction in the number of firms takes time. In the short run, competition is fierce, and many firms cannot profitably produce. Only the most productive firms are active. Aggregate productivity soars. As the number of firms gradually falls (as existing firms age), competition softens, and low productivity firms are able to resume production. Aggregate productivity gradually falls as the world economy converges towards its new steady state. Along with this productivity overshooting, I expect to observe price compression in the short run, followed by a gradual increase in aggregate prices, as well as an overshooting in inequalities.

References

[1] Bernard, Andrew B.; Jonathan Eaton; J. Bradford Jensen and Sam Kortum, 2003, "Plants and Productivity in International Trade," *American Economic Review, v. 93, pp. 1268-1290.*

[2] Bernard, Andrew B. and J. Bradford Jensen, 1999, "Exceptional exporter performance: cause, effect, or both?", *Journal of International Economics, v. 47, pp. 1–25.*

[3] Bernard, Andrew B. and J. Bradford Jensen, 2001a, "Why Some Firms Export," *The Review of Economics and Statistics, v. 86 no. 2.*

[4] Bernard, Andrew B. and J. Bradford Jensen, 2001b, "Who Dies? International Trade, Market Structure, and Industrial Restructuring," *NBER Working Paper # 8327* .

[5] Bernard, Andrew B. and J. Bradford Jensen, 2002, "Entry, Expansion, and Intensity in the US Export Boom, 1987-1992," *Review of International Economics*, forthcoming.

[6] Dornbusch, Rudiger, 1976, "Expectations and Exchange Rate Dynamics," *Journal of Political Economy, v. 84, pp. 1161-76.*

[7] Dornbusch, Rudiger; S. Fischer, and Paul A. Samuelson, "Comparative Advantage, Trade, and Payments in a Ricardian Model with a continuum of Goods," *American Economic Review, vol. 67, pp. 823-839.*

[8] Eaton, Jonathan; Sam Kortum, 2002, "Technology, Geography, and Trade," *Econometrica, vol. 70 no. 5 (September 2002), pp. 1741-1779.*

[9] Eaton, Jonathan; Sam Kortum, and Francis Kramarz, 2004, "An Anatomy of International Trade: Evidence from French Firms," *American Economics Review, Papers and Proceedings*, forthcoming.

[10] Eaton, Jonathan; Sam Kortum, and Francis Kramarz, 2004, "Dissecting Trade: Firms, Industries, and Export Destinations," *NBER Working Paper 10344.*

[11] Ghironi, Fabio, and Marc J. Melitz, 2005, "International Trade and Macroeconomic Dynamics with Heterogeneous Firms." *Working Paper.*

[12] Helpman, Elhanan; Marc Melitz and Yona Rubinstein, 2004, "Trading Partners and Trading Volumes," *Working Paper.*

[13] Hopenhaym, Hugo, 1992, "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica, v.60, pp. 621-653.*

[14] Krugman, Paul. 1980. "Scale Economies, Product Differentiation, and the Pattern of Trade." *American Economic Review, vol. 70, pp. 950—959.*

[15] Melitz, Marc, 2003, "The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations," *Econometrica, v. 71 no. 6 (November 2003), pp. 1695-1725.*

Appendix

A Non stationary productivity distribution and dynamic adjustments

Absent assumption 1, the distribution of productivity among firms during the transition may be non stationary.

The reason is the following. When trade opens up, the productivity threshold for domestic sales soars. In this paper, I have described what happens when existing firms have the option of not producing at all during some period, without going out of business (that is without having to pay the sunk entry cost, get a productivity draw from a lottery, all over again). Therefore, the exit decision was simple. All firms with a productivity above $\varphi^*_{+\infty}$ stay in business. because they know that in the long run (once the trade steady state is reached), they will be able to generate some strictly positive profits. However, some of these firms have to stay idle for some time. If those firms had to pay a fixed cost each period in order to stay in business, they would have to gauge the cost of staying in business (a given number of periods with negative profits), against the expected benefits from staying in business (expected positive profits starting up in the future). The least productive among those firms will exit. It is easy to prove that there is a range of productivity $\left[\varphi^*_{+\infty}; \bar{\varphi}_0\right]$ over which firms exit immediately.

Let us now move along the transition. Firms die (there were too many firms to start with), the productivity threshold for domestic sales goes down. Eventually, this threshold goes down below $\bar{\varphi}_0$. Then, with each new cohort of entrants, some firms with a productivity below $\bar{\varphi}_0$ will survive. But until that point, there were no firms with such a low productivity, whereas there is a strictly positive number of firms with a productivity above $\bar{\varphi}_0$ that have survived. Because of the nature of the technology, there will be as many new firms with $\varphi > \bar{\varphi}_0$ as firms with $\varphi < \bar{\varphi}_0$ created (up to some relative probability). But since to start with, the stock of firms with $\varphi > \bar{\varphi}_0$ is strictly positive, and there are no firms with $\varphi < \bar{\varphi}_0$, there is an imbalance. The distribution of productivity among existing firms is no longer given by $g(\cdot) \times (Constant)$. The distribution is tilted towards high productivity firms. There are disproportionately many firms with a high productivity. Those are the firms that survived the opening to trade.

In such a configuration, the steady state cannot be reached in a finite number of periods. The disequilibrium of the distribution towards high productivity firms gradually fades away, as those old firms who survived since the time when trade was opened for the first time die, and are replace with new cohorts.

However, if the distribution of firms is non stationary, it won't be possible to determine even one equilibrium path analytically. And iterative numerical methods will be highly computationally intensive. In order to know how many firms enter each period from say t onward, I must guess a path for the future beyond t. But what equilibrium is reached at time t depends on the current distribution of firms, which is inherited from the past. So I must also guess what has happened in the past. The equilibrium solution is both forward looking and backward looking. The only possibility is to assume that the steady state is approximatively reached at a given point in time, say $t = T$ large (and assume that adjustments beyond that point are negligible), guess an entire path for the entry of firms $\left\{M_{e,0}^{(1)}; ...; M_{e,T}^{(1)}\right\}$, solve for the equilibrium in each period given this path of entry has been followed in the past and is expected to be followed in the future, and extract what would be the optimal path entry of forward looking agents, $\left\{M_{e,0}^{(2)}; ...; M_{e,T}^{(2)}\right\}$, and so on until this algorithm converges. The problem of such an algorithm is that it is computationally demanding, since each iteration (from $\left\{M_{e,0}^{(n)}; ...; M_{e,T}^{(n)}\right\}$ to $\left\{M_{e,0}^{(n+1)}; ...; M_{e,T}^{(n+1)}\right\}$) requires to solve the entire sequence of T equilibria simultanuously. Moreover, in such a model, the strategy of a firm may potentially be complex: some firms may decide to incur negative profits for some time because they expect to earn positive profits in the future. The exit decision of firms is forward looking and depends on the option value of staying in business.

B Criterion 2 and proofs of propositions 3, 4 and 5

Criterion 2 (reminded) The full functional form of this criterion[13] is:

$$\frac{\int_{\varphi_{-\infty}^{*}}^{+\infty}\left(\varphi^{\sigma-1}-\varphi_{-\infty}^{*\sigma-1}\right)dG\left(\varphi\right)}{\int_{\varphi_{-\infty}^{*}}^{+\infty}\varphi_{-\infty}^{*\sigma-1}dG\left(\varphi\right)} > \frac{f\frac{\int_{\varphi_{+\infty}^{*}}^{+\infty}\left(\varphi^{\sigma-1}-\varphi_{+\infty}^{*\sigma-1}\right)dG(\varphi)}{\int_{\varphi_{+\infty}^{*}}^{+\infty}\varphi_{+\infty}^{*\sigma-1}dG(\varphi)}+p_{x,+\infty}f_{x}\frac{\int_{\varphi_{x,+\infty}^{*}}^{+\infty}\left(\varphi^{\sigma-1}-\varphi_{x,+\infty}^{*\sigma-1}\right)dG(\varphi)}{\int_{\varphi_{x,+\infty}^{*}}^{+\infty}\varphi_{x,+\infty}^{*\sigma-1}dG(\varphi)}}{f\frac{\int_{\varphi_{+\infty}^{*}}^{+\infty}\varphi^{\sigma-1}dG(\varphi)}{\int_{\varphi_{+\infty}^{*}}^{+\infty}\varphi_{+\infty}^{*\sigma-1}dG(\varphi)}+p_{x,+\infty}f_{x}\frac{\int_{\varphi_{x,+\infty}^{*}}^{+\infty}\varphi^{\sigma-1}dG(\varphi)}{\int_{\varphi_{x,+\infty}^{*}}^{+\infty}\varphi_{x,+\infty}^{*\sigma-1}dG(\varphi)}}$$

I now turn to the missing parts in the proofs leading to the productivity overshooting proposition. I will first prove proposition 5: a reduction in the mass of firms leads to a fall in the

[13] With Pareto distributed shocks, that is if $P(\varphi > \varphi^{*}) = (\varphi^{*}/\varphi_{\min})^{-\gamma}$, with $\gamma > 1$ some scaling parameter, this criterion is not met, and the is an exact equality. This implies that with Pareto distributed shocks, the adjustment towards the new steady state is immediate. There are as many high productivity firms ($\varphi > \varphi_{+\infty}^{*}$) in autarky and in the steady state. There are as many firms that enter (and die from attrition plus immediate exit) each period in both regime. The only difference is that under the trade regime, fewer firms survive. But, with Pareto distributed productivity shocks, the distribution among survivors is exactly identical under trade and under autarky.

productivity threshold for domestic sales. This will allow me to prove the missing part in the proof of propositions 3 and 4: a reduction in the mass of firms per level of productivity is equivalent to an increase in expected per period profits.

Proposition 6 *(i) Under the trade regime, and when* $M_{e,t} = 0$*,*

$$\tilde{M}_t > \tilde{M}_s \Rightarrow \varphi_t^* > \varphi_s^*$$

(ii) Under the trade regime, and when $M_{e,t} = 0$*,*

$$\tilde{M}_t < \tilde{M}_s \Leftrightarrow P(\varphi > \varphi_t^*)\,\bar{\pi}_t > P(\varphi > \varphi_s^*)\,\bar{\pi}_s$$

Proof. (i) The equilibrium along the transition is determined by four conditions: the zero cutoff profit conditions, (ZCP_t) and $(ZCP_{x,t})$, the labor market clearing condition, (LMC_t), and the law of motion for the mass of firms which determins the value of $\tilde{M}_t$ at each point along the transition, (M_t).

Since the distribution of firms is stationary, I know that $\mu_t(\varphi) = \tilde{M}_t g(\varphi)$. I can therefore compute the price index along the transition path. Plugging the zero cutoff profits condition for exporters, $(ZCP_{x,t})$, into the price index in Eq. (2), and using the definition of μ_t along the transition, and rearranging, I get,

$$(\rho P_t)^{1-\sigma} = \frac{\varphi_t^{*\sigma-1} M_t}{f}\left(f\left(\frac{\tilde{\varphi}(\varphi_t^*)}{\varphi_t^*}\right)^{\sigma-1} + p_{x,t} f_x \left(\frac{\tilde{\varphi}(\varphi_{x,t}^*)}{\varphi_{x,t}^*}\right)^{\sigma-1}\right)$$

Plugging this price index into the zero cutoff profit condition for domestic sales, (ZCP_t), I get,

$$R_t = \sigma M_t \left(f\left(\frac{\tilde{\varphi}(\varphi_t^*)}{\varphi_t^*}\right)^{\sigma-1} + p_{x,t} f_x \left(\frac{\tilde{\varphi}(\varphi_{x,t}^*)}{\varphi_{x,t}^*}\right)^{\sigma-1}\right)$$

Plugging this into the labor market clearing condition, (LMC_t), I can solve for the aggregate demand, R_t, and rearranging, I get an implicit relationship between the mass $\tilde{M}_t$ and the productivity threshold φ_t^*,

$$\begin{aligned} L &= M_t\left[(\sigma-1)\,\bar{\pi}_t + \sigma\,(f + p_{x,t} f_x)\right] \\ \text{or } L &= \tilde{M}_t\left[f\int_{\varphi_t^*}^{+\infty}\left((\sigma-1)\left(\frac{\varphi}{\varphi_t^*}\right)^{\sigma-1} + 1\right) dG(\varphi) + \int_{\varphi_{x,t}^*}^{+\infty}\left((\sigma-1)\, f\left(\frac{\varphi}{\tau\varphi_t^*}\right)^{\sigma-1} + f_x\right) dG(\varphi)\right] \end{aligned}$$

$\left((\sigma-1)\left(\frac{\varphi}{\varphi_t^*}\right)^{\sigma-1} + 1\right)$ and $\left((\sigma-1)\, f\left(\frac{\varphi}{\tau\varphi_t^*}\right)^{\sigma-1} + f_x\right)$ are positive and increasing in φ_t^* and $\varphi_{x,t}^*$ respectively. $\varphi_{x,t}^*$ is increasing in φ_t^*. So as $\tilde{M}_t$ falls along the transition towards the trade steady state, φ_t^* must fall in order to preserve the equilibrium on the labor market.

So I have proven that along the transition, φ_t^* increases as the mass of firms $\tilde{M}_t$ falls.

(ii) I will now prove that as φ_t^* increases, $P(\varphi > \varphi_t^*)\,\bar{\pi}_t$ decreases.

Plugging the zero cutoff profit conditions (ZCP_t) and $(ZCP_{x,t})$ into the formula for average profits, I get:

$$\bar{\pi}_t = f\left(\left(\frac{\tilde{\varphi}_t}{\varphi_t^*}\right)^{\sigma-1} 1\right) + p_{x,t} f_x \left(\left(\frac{\tilde{\varphi}_{x,t}}{\varphi_{x,t}^*}\right)^{\sigma-1} - 1\right)$$

Multiplying by the probability of survival, I get,

$$P(\varphi > \varphi_t^*)\,\bar{\pi}_t = f \int_{\varphi_t^*}^{+\infty} \left[\left(\frac{\varphi}{\varphi_t^*}\right)^{\sigma-1} - 1\right] dG(\varphi) + f_x \int_{\varphi_{x,t}^*}^{+\infty} \left[\left(\frac{\varphi}{\varphi_{x,t}^*}\right)^{\sigma-1} - 1\right] dG(\varphi)$$

Since $\left[\left(\frac{\varphi}{\varphi_t^*}\right)^{\sigma-1} - 1\right]$ and $\left[\left(\frac{\varphi}{\varphi_{x,t}^*}\right)^{\sigma-1} - 1\right]$ are non negative and decreasing in φ_t^* and $\varphi_{x,t}^*$ respectively, and since $\varphi_{x,t}^*$ is increasing in φ_t^*, $P(\varphi > \varphi_t^*)\,\bar{\pi}_t$ is decreasing in φ_t^*. ■

C Expected opening to trade

If the opening to trade is announced in advance, productivity overshooting will be dampened. Indeed, the very reason why there is overshooting in productivity is because the economy has overaccumulated capital while in autarky, in the sense that too many firms have been started. The number of firms must go down to its steady state level, and the absence of entry during the transition explains why some less productive firms have to temporarily stop their production, and therefore why aggregate productivity overshoots in the short run. If the opening to trade is expected, the economy will start disinveting in advance. Fewer firms are created each period, so that new firms are not sufficient to replace exit each period. There is net exit of firms even before trade is opened up. This is because rational expectation agents do know that their investment (creation of a new firm) would not be worthwhile at least for some time. The speed at which the economy disinvests obviously depends on the rate at which existing firms die (the coefficient δ), the impatience of entrepreneurs (the discount factor β), as well as the horizon at which trade will be opened up. However, unless trade opening has been sufficiently early, there will still be a period of productivity overshooting. There are too many of those firms that had been created before trade opening were announced, and those firms will have to disappear.

We do observe that even those trade agreements that have been have been negotiated long in advance, still have a large impact at the time when they are implemented.