Universität St.Gallen

Globalization and Productivity in the Developing World

Reto Foellmi and Manuel Oechslin

February 2012 Discussion Paper no. 2012-03

School of Economics and Political Science,
Department of Economics

University of St. Gallen

Editor:	Martina Flockerzi University of St. Gallen School of Economics and Political Science Department of Economics Varnbüelstrasse 19 CH-9000 St. Gallen Phone +41 71 224 23 25 Fax +41 71 224 31 35 Email seps@unisg.ch
Publisher:	School of Economics and Political Science Department of Economics University of St. Gallen Varnbüelstrasse 19 CH-9000 St. Gallen Phone +41 71 224 23 25
Electronic Publication:	Fax +41 71 224 31 35 http://www.seps.unisg.ch

Globalization and Productivity in the Developing World

Reto Foellmi and Manuel Oechslin

Author's address:

Prof. Dr. Reto Foellmi
Department of Economics
SIAW-HSG
Bodanstr.8
CH-9000 St. Gallen
Phone +41 71 224 22 69
Fax +41 71 224 22 98
Email reto.foellmi@unisg.ch
Website www.siaw.unisg.ch/foellmi

Prof. Dr. Manuel Oechslin
Tilburg University
Department of Economics
NL-5000 LE Tilburg
Phone +31 13 466 8089
Email m.oechslin@uvt.nl

Abstract

We explore the productivity impact of international trade in a monopolistically competitive economy with endogenous mark-ups due to credit market frictions. We show that reducing trade barriers in such an environment (i) may - but not necessarily must – have a negative impact on productivity and output; (ii) is bound to increase the polarization of the income distribution. The reason is that the pro-competitive effects of trade reduce mark-ups and hence the borrowing capacity of less affluent entrepreneurs. As a result, smaller firms may no longer be able to make the investments required to operate the high-productivity technology. Our findings are consistent with evidence from developing countries which (i) does not suggest a clear-cut impact of trade on economic performance; (ii) hints at an inequality-increasing effect of globalization.

Keywords

International trade, credit market frictions, productivity, inequality.

JEL Classification

O11, F13, O16.

1 Introduction

What is the impact of international trade on productivity? Recent theoretical work (e.g., Melitz, 2003) emphasizes that exposure to trade improves aggregate productivity via a selection effect: By providing access to export markets, lower trade barriers create new profit opportunities for the most productive firms. As a result, these firms increase factor demand so that factor prices rise. Higher factor prices, in turn, drive the least productive firms out of the market – and hence boost aggregate productivity and real output. In this paper, however, we show that greater exposure to international trade may actually have a negative effect on productivity and aggregate output if credit markets are highly imperfect and wealth inequality is substantial, i.e., under the circumstances we encounter throughout the developing world (see, e.g., Banerjee and Duflo, 2010). Our analysis thus provides an explanation for why the empirical literature does not find a clear positive impact of international trade on economic performance in developing countries (see, e.g., DeJong and Ripoll, 2006).[1]

We explore the impact of trade in a monopolistically competitive model (à la Dixit and Stiglitz, 1977) that features an *endogenous distribution of mark-ups* due to credit market frictions (as in Foellmi and Oechslin, 2010). It is assumed that loan repayment is imperfectly enforceable so that an entrepreneur's borrowing capacity depends on her private wealth. As a result, less affluent entrepreneurs are forced to run small firms – and thus charge high prices and mark-ups. Greater exposure to trade, however, is bound to reduce these mark-ups: Competition from abroad reduces the maximum prices smaller firms can charge; moreover, there is a surge in the cost of borrowing since larger firms increase capital demand to take advantage of new export opportunities. Lower mark-ups, in turn, reduce the borrowing capacity of less affluent firm owners – which means that they may no longer be able to make the investments required to operate the high-productivity (i.e., state-of-the-art) technology.

The magnitude and consequences of this reduced access to credit depend on the degree to which the exposure to international trade increases. A strong increase would drive the smaller firms out of the market (as, e.g., in Melitz, 2003) and hence unambiguously improve economic performance. A smaller rise, however, just forces them to resort to the low-productivity (i.e., "traditional") technology – and hence would promote resource misallocation and a fall in aggregate output. It turns out, however, that aggregate economic performance is not only impaired via this productivity channel. Our analysis shows that an ineffective financial system imposes a further and less obvious cost on the economy: The liberalization-induced fall in the

[1] We acknowledge support from the NCCR Trade Regulation.

borrowing capacity – and hence the output – of the smaller firms requires the economy to import larger quantities of goods and hence to spend more resources on transportation. Put differently, opening up to trade leads to a "costly" replacement of domestic output – with potentially adverse overall consequences for the aggregate real income.

Our model's predictions are not just theoretical possibilities but can account for important empirical observations. First, the lack of a clear-cut effect of trade on productivity or output is consistent with the findings of a voluminous empirical literature on trade policy and economic performance which – particularly among developing countries – fails to identify a robust link (see, e.g., Kehoe and Ruhl, 2010, for a recent overview).[2] Second, our theory predicts that trade will amplify the polarization of the income distribution (since small businesses lose while the richest entrepreneurs win due to better export opportunities). Again, this prediction seems to be borne out by the data (see, e.g., Goldberg and Pavcnik, 2007). Finally, our model suggests that opening up to trade may reduce the size of the credit market (as measured by the credit-to-GDP ratio) since those who rely on external finance face tighter credit constraints. This implication is consistent with findings by Bertola and Lo Prete (2010) who show that openness, after controlling for other relevant policies, has a negative impact on financial development.

Although this paper shows that opening up to trade can have negative effects on aggregate output and the income distribution in poor countries, it clearly does not suggest that these countries should refrain from liberalizing trade. Rather, our analysis suggests that a reduction in trade barriers should be implemented in combination with complementary reforms or policies. In particular, if a reduction in trade barriers were accompanied by a corresponding increase in the quality of credit contract enforcement (or, alternatively, by measures which decrease the cost of borrowing for small firms), the effects would be clearly positive. However, to the extent that such complementary reforms or policies take time to implement, our analysis cautions against sudden liberalizations in poor economies (and hence provides a rationale for "special and differential treatment" of developing countries within the WTO).

This paper contributes to the literature on international trade and heterogeneous firms. Yet, focusing on developing countries, our theory deviates from the standard classes of models (i.e., Melitz, 2003; Bernard et al., 2003) and, as a result, suggests an ambiguous relationship between trade and economic performance. In some dimensions, we do arrive at similar conclusions, however. For instance, we find that exporters are a minority and, on average, are bigger

[2] This empirical pattern is also consistent with anecdotal evidence from East Asia. As pointed out by, e.g., Stiglitz and Charlton (2005), many of the East Asian miracle countries did not follow the "orthodox free trade prescriptions" (p. 16) but used to protect selected industries from import competition.

and more productive than firms which exclusively serve the home market. The mechanism behind these implications is an entirely different one, though. If credit markets are imperfect and wealth is unequally distributed, only a small fraction of rich entrepreneurs has the means to operate at a scale which permits them to employ the state-of-the-art technology and to produce more output than the domestic market will absorb. By analyzing how trade affects the distribution of mark-ups in an economy, our paper further connects with recent work by Epifani and Gancia (2011) who show that the pro-competitive effects of international trade can actually reduce welfare when they increases the dispersion of mark-ups. This paper, in contrast, shows that – when there are credit market imperfection – international trade may reduce welfare even if it leads to a more even distribution of mark-ups.

More broadly, our analysis is related to work on how (credit) market frictions lead to resource misallocation and hence compromise total factor productivity in low-income countries (e.g., Banerjee and Duflo, 2005; Hsieh and Klenow, 2009; Song et al., 2011). These papers, however, do not address whether exposure to international trade affects the resource allocation in a positive or a negative way − which is the prime focus here. Earlier papers with a related, but different, focus include Banerjee and Newman (2004) and Matsuyama (2005). These papers elaborate variants of the Ricardo-Viner model and do not address how the pro-competitive effects of trade affect mark-ups and thus access to credit and firms' technology choices.

The rest of this paper is organized as follows. The next section presents and solves the closed-economy model. In Section 3, we describe the effects of opening up to international trade – with special emphasis on an intermediate-openness case. Relying on simulations, the section further gives a systematic overview of the adjustments associated with a continuous decrease in trade barriers from prohibitive levels to zero. Section 4 discusses the analytical and numerical results and links them to the empirical literature. It also includes a brief discussion of policy implications. Section 5, finally, concludes.

2 The Closed Economy

2.1 Endowments, Technologies, and Preferences

Assumptions. The economy is populated by a continuum of (potential) entrepreneurs. The population size is normalized to 1. The entrepreneurs are heterogeneous with respect to their initial capital endowment $\omega_i, i \in [0, 1]$, and their production possibilities. The capital endowments are distributed according to the distribution function $G(\omega)$ which gives the measure of the population with an endowment below ω. We further assume that $g(\omega)$, which refers to the density function, is positive over the entire positive range. The aggregate capital endowment, $\int_0^\infty \omega dG(\omega)$, will be denoted by K.

Each entrepreneur owns a specific skill (or technological know-how) that makes him a monopoly supplier of a single differentiated good. All goods are produced with a simple technology that requires capital as the only input into production. The technology, however, is characterized by a non-convexity. In particular, its productivity is relatively low if the investment falls short of a critical threshold. In formal terms, we impose

$$y_i = \begin{cases} bk_i & : \quad k_i < \kappa \\ ak_i & : \quad k_i \geq \kappa \end{cases}, \; b < a, \tag{1}$$

where y_i and k_i denote, respectively, output and capital and κ refers to the critical scale of investment. In what follows, we say that an entrepreneur operates the "low-productivity technology" if she invests less than the κ-threshold; similarly, we say that an entrepreneur operates the "high-productivity technology" if the investment exceeds this critical threshold.

The assumptions of both market power and non-convexities play an important role in our model. They will allow us to mirror the idea that opening up exposes firms to more vigorous competition and hence may affect technology choices (especially, as is discussed below, in the presence of credit market imperfections). Of course, the idea that exposure to international trade enhances competition is not restricted to poor economies. Yet, firms in low-income countries might be particularly prone to losing market power because they tend to produce less innovative goods (see, e.g., Acemoglu and Zilibotti, 2001) and since the market structure in these places is often monopolistic (see, e.g., UNCTAD, 2006).

The entrepreneurs' utility function is assumed to be of the familiar CES-form,

$$U = \left(\int_0^1 c_j^{(\sigma-1)/\sigma} dj \right)^{\frac{\sigma}{\sigma-1}}, \tag{2}$$

where c_j denotes consumption of good j and $\sigma > 1$ represents the elasticity of substitution between any two goods. Each entrepreneur i maximizes objective function (2) subject to

$$\int_0^1 p_j c_j dj = m(\omega_i), \tag{3}$$

where p_j is the price of good j and $m(\omega_i)$ refers to entrepreneur i's nominal income (which, in turn, will depend on the initial capital endowment, as will be discussed further below).

Finally, for tractability purposes, we impose a parameter restriction which puts an upper bound on the critical scale of investment:

$$\kappa < K(b/a)^{\sigma-1}. \tag{R1}$$

Implications. Under these conditions, entrepreneur i's demand for good j is given by

$$c_j(y(\omega_i)) = \left(\frac{p_j}{P}\right)^{-\sigma} \frac{m(\omega_i)}{P}, \tag{4}$$

where $P \equiv (\int_0^1 {p_j}^{1-\sigma} dj)^{1/(1-\sigma)}$ denotes the CES price index. In a goods market equilibrium, aggregate demand for good j must be equal to the supply of good j, y_j. Taking this into account, we can express the real price of good j as a function of y_j and Y/P,

$$\frac{p_j}{P} = \frac{p(y_j)}{P} \equiv \left(\frac{Y}{P}\right)^{\frac{1}{\sigma}} y_j^{-1/\sigma}, \tag{5}$$

where $Y \equiv \int_0^1 p(y_j) y_j dj$ denotes the economy-wide nominal output and the ratio Y/P refers to the real output. Notice further that, in a goods market equilibrium, the real price of a good is strictly decreasing in the quantity produced. The reason is simple: Since the marginal utility from consuming any given good falls in the quantity consumed, the only way to make domestic consumers buy larger quantities is to lower the price.

Later on, it will be helpful to have an expression for the aggregate real output (or, equivalently, for the aggregate real income) that depends only on the distribution of firm outputs. Using (5) in the definition of Y, we obtain

$$\frac{Y}{P} = \left(\int_0^1 y_j^{(\sigma-1)/\sigma} dj\right)^{\frac{\sigma}{\sigma-1}} = \frac{M}{P}, \tag{6}$$

where $M \equiv \int_0^1 m(\omega_i) di$ denotes the aggregate nominal income.

2.2 The Credit Market

Assumptions. Entrepreneurs may borrow and lend in an economy-wide credit market. Unlike the goods market, the credit market is competitive in the sense that both lenders and borrowers take the equilibrium borrowing rate as given. However, the credit market is imperfect in the sense that borrowing at the equilibrium rate may be limited. As in Foellmi and Oechslin (2010), such credit-rationing may arise from imperfect enforcement of credit contracts. More specifically, we assume that borrower i can avoid repayment altogether by incurring a cost which is taken to be a fraction $\lambda \in (0, 1]$ of the current firm revenue, $p(y_i)y_i$.

The parameter λ mirrors how well the credit market works. A value close to one represents a near-perfect credit market while a value near zero means that the credit market functions poorly. Intuitively, in the latter case, lenders are not well protected since the borrowers can "cheaply" default on their payment obligations – which invites ex post moral hazard. As a result, lenders are reluctant to provide external finance. Poor creditor protection and the associated problem of moral hazard are in fact important phenomena in many developing economies. It is, for example, well documented that – throughout the developing world – insufficient collateral laws or unreliable judiciaries often make it extremely hard to enforce credit contracts in a court (see, e.g., Banerjee and Duflo, 2005; 2010).

Implications. Taking the possibility of ex post moral hazard into account, a lender will give credit only up to the point where the borrower just has the incentive to pay back. In formal terms, this means that the amount of credit cannot exceed $\lambda p(y_i)y_i/\rho_i$, where ρ_i denotes the interest rate borrower i faces. Note further that – since borrowers always repay and because there are no individual-specific risks associated with entrepreneurship – the borrowing rate must be the same for all agents ($\rho_i = \rho$). Using this information, and accounting for (1), we find that borrower i does not default on the the credit contract ex post if

$$\lambda p(y_i)y_i/\rho \geq \begin{cases} y_i/b - \omega_i & : \quad y_i < a\kappa \\ y_i/a - \omega_i & : \quad y_i \geq a\kappa \end{cases}, \tag{7}$$

where the right-hand side of (7) gives the size of the credit.

We now derive how the maximum amount of borrowing, and hence the maximum output, depends on the initial wealth endowment, ω.[3] To do so, suppose that there is a wealth level $\omega_\kappa < \kappa$ which permits borrowing exactly the amount required to meet the critical investment size κ. Taking (5) and (7) into account, this threshold level is defined by

$$\omega_\kappa + \lambda x \left(a\kappa\right)^{(\sigma-1)/\sigma} = \kappa, \tag{8}$$

where

$$x \equiv P^{(\sigma-1)/\sigma}Y^{1/\sigma}/\rho = (Y/P)^{1/\sigma}/(\rho/P).$$

With these definitions (and expressions 5 and 7) in mind, it is immediately clear that the maximum firm output is implicitly determined by

$$\overline{y} = \begin{cases} b\left(\omega + \lambda x \overline{y}^{(\sigma-1)/\sigma}\right) & : \quad \omega < \omega_\kappa \\ a\left(\omega + \lambda x \overline{y}^{(\sigma-1)/\sigma}\right) & : \quad \omega \geq \omega_\kappa \end{cases} \tag{9}$$

and hence depends on the initial wealth endowment. It is the purpose of the following lemma to clarify the relationship between $\overline{y}$ and ω.

Lemma 1 *A firm's maximum output, $\overline{y}(\omega)$, is a strictly increasing function of the initial capital endowment, ω.*

[3] Since the initial wealth is the only individual-specific factor that determines maximum borrowing, the index for individuals will be dropped in the rest of this section.

Proof. See Appendix. ■

The maximum firm output increases in initial capital for two different reasons. First, and most directly, an increase in ω means that the entrepreneur commands more own resources which can be invested. Second, there is an indirect effect operating through the credit market: An increase in ω allows for higher borrowing since the entrepreneur has more "skin in the game" (Banerjee and Duflo, 2010). Figure 1 shows a graphical illustration of $\overline{y}(\omega)$.

Figure 1 here

Besides the positive slope, the figure highlights two additional properties of the $\overline{y}(\omega)$-function. First, the function is locally concave. This just mirrors the fact that the marginal return on investment falls in the level of investment; thus, the positive impact of an additional endowment unit on the borrowing capacity must decrease. Second, there is a discontinuity at ω_κ since, at that point, an entrepreneur is able to switch to the more productive technology.

2.3 Output Levels

We now discuss how individual firm outputs depend on capital endowments, holding constant the aggregate variables Y/P and ρ/P (and hence x). Our discussion presumes

$$x \geq \frac{1}{a}\frac{\sigma}{\sigma-1}(a\kappa)^{1/\sigma}, \tag{10}$$

which will actually turn out to be true in equilibrium (see Proposition 1).

$\omega \geq \omega_\kappa$. We start by looking at entrepreneurs who are able to use the more productive technology. Resources permitting, these entrepreneurs increase output up to the point where the marginal revenue, $((\sigma-1)/\sigma)P^{(\sigma-1)/\sigma}Y^{1/\sigma}y^{-1/\sigma}$, equals the marginal cost, ρ/a. We denote this profit-maximizing output level by $\widetilde{y}$ and we use $\widetilde{\omega}$ to denote the wealth level which puts an agent exactly in a position to produce $\widetilde{y}$. Using these definitions, we have

$$\widetilde{y} = \left(ax\frac{\sigma-1}{\sigma}\right)^{\sigma} \quad \text{and} \quad \widetilde{\omega} = \left(1-\lambda\frac{\sigma}{\sigma-1}\right)\frac{\widetilde{y}}{a}, \tag{11}$$

where $\widetilde{y}/a \geq \kappa$ due to (10).

Two points should be noted here. First, because of Lemma 1 and $\widetilde{y} \geq a\kappa$, we have $\widetilde{\omega} \geq \omega_\kappa$. Second, as can be seen from the second expression in (11), $\lambda < (\sigma-1)/\sigma$ is sufficient for having a group of credit-constrained entrepreneurs, i.e., entrepreneurs who have too little access to credit to produce at the profit-maximizing output level. On the other hand, if $\lambda \geq (\sigma-1)/\sigma$, even entrepreneurs with a zero wealth endowment can operate at the profit-maximizing scale.

Why? The smaller the elasticity of substitution, the higher is the constant mark-up $\sigma/(\sigma-1)$ over marginal costs. So, if σ is small, even poor agents are able to generate revenues which are large relative to the payment obligation. This means that only a very low λ may induce a borrower to default ex post. Put differently, the credit market imperfection is binding for some entrepreneurs only if it is "more substantial" than the imperfection in the product market.

The following lemma is an immediate corollary of the above discussion:

Lemma 2 *Suppose $\lambda < (\sigma-1)/\sigma$. Then, entrepreneurs (i) with $\omega \in [\omega_\kappa, \widetilde{\omega})$ produce $\overline{y}(\omega) < \widetilde{y}$; (ii) with $\omega \in [\widetilde{\omega}, \infty)$ produce $\widetilde{y}$. Otherwise, if $\lambda \geq (\sigma-1)/\sigma$, all entrepreneurs produce $\widetilde{y}$.*

Proof. See Appendix. ■

$\omega < \omega_\kappa$. We now focus on the investment behavior of less affluent entrepreneurs, i.e., agents with a capital endowment below ω_κ (which does not allow for the use of the high-productivity technology). As established above, such entrepreneurs can only exist if $\lambda < (\sigma-1)/\sigma$.

Lemma 3 *Suppose $\lambda < (\sigma-1)/\sigma$. Then, entrepreneurs with a wealth endowment below ω_κ produce $\overline{y}(\omega)$.*

Proof. See Appendix. ■

Putting things together. An immediate implication of Lemmas 2 and 3 is that the equilibrium individual firm outputs are given by

$$y(\omega) = \begin{cases} \overline{y}(\omega) & : \quad \omega < \widetilde{\omega} \\ \widetilde{y} & : \quad \omega \geq \widetilde{\omega} \end{cases}, \tag{12}$$

where $\overline{y}(\omega)$ is implicitly determined by (9) and $\widetilde{y}$ is given in (11). Note that the case $\omega < \widetilde{\omega}$ is only relevant if the parameter restriction $\lambda < (\sigma-1)/\sigma$ holds (and hence $\widetilde{\omega} > 0$). Assuming that the restriction does hold, Figure 2 gives a graphical illustration of (12). The figure shows two possible situations. In panel *a.*, we have $\omega_\kappa > 0$ so that a positive mass of entrepreneurs are forced to use the less productive technology. Panel *b.* shows a situation where $\omega_\kappa \leq 0$ so that all entrepreneurs have access to the more productive technology.

Figure 2 here

The distribution of firm outputs is mirrored in the distribution of output prices. Since each firm faces a downward-sloping demand curve (equation 5), smaller firms charge higher prices – despite the fact that each good enters the utility function symmetrically. Only if there is no credit rationing do output levels across firms fully equalize so that all prices are the same.

2.4 The Equilibrium under Autarky

When characterizing the use of technology and individual firm outputs, we kept constant aggregate real output and the real interest rate (and hence the ratio $x = (Y/P)^{1/\sigma}/(\rho/P)$). We now establish that, in fact, both Y/P and ρ/P are uniquely determined in the macroeconomic equilibrium. To do so, note that we can write aggregate gross capital demand (i.e., the sum of all physical capital investments by firms) as a function of x,

$$K^D(x) = \int_0^{\omega_\kappa} \frac{\overline{y}(\omega;x)}{b} dG(\omega) + \int_{\omega_\kappa}^{\widetilde{\omega}} \frac{\overline{y}(\omega;x)}{a} dG(\omega) + \int_{\widetilde{\omega}}^{\infty} \frac{\widetilde{y}(x)}{a} dG(\omega), \tag{13}$$

where aggregate capital supply, $K = \int_0^\infty \omega dG(\omega)$, is exogenous and inelastic.

Proposition 1 *There exists a unique macroeconomic equilibrium (i.e., real output, Y/P, and the real interest rate, ρ/P, are uniquely pinned down). If $\lambda < (\sigma - 1)/\sigma$, a positive mass of entrepreneurs are credit-constrained (and the poorest among them may be forced to use the low-productivity technology). Otherwise, if $\lambda \geq (\sigma - 1)/\sigma$, no one is credit-constrained.*

Proof. See Appendix. ■

Figure 3 here

Figure 3 shows K^D as a function of x (for the case $\lambda < (\sigma - 1)/\sigma$). The figure also highlights that condition (10), on which both Lemma 2 and 3 rely, is indeed satisfied.[4]

Finally, it is worth emphasizing that the properties of this equilibrium are consistent with firm-level evidence from developing countries (see, e.g., Banerjee and Duflo, 2005). In particular, if the credit market friction is sufficiently severe, we have a coexistence of (i) high and low marginal (revenue) products of capital; (ii) more and less advanced technologies.

3 Integrating into the World Economy

This section explores how a reduction in trade barriers affects the use of technologies, aggregate output, and the income distribution in the home economy. The home economy – which is taken to represent a developing country – will be called the "South". The rest of the world (i.e., the South's trading partner) is referred to as the "North" and represents an advanced economy.

[4] If $\lambda \geq (\sigma - 1)/\sigma$, we have $K^D(x) = (x(\sigma - 1)/\sigma)^\sigma a^{\sigma - 1}$, and it can be easily checked that $K^D(x) = K$ defines a unique x (with $Y/P = aK$ and $\rho/P = a(\sigma - 1)/\sigma$).

3.1 Assumptions

Trade barriers. So far, the trade barriers have been assumed to be sufficiently high to prevent trade between South and North. This section focuses on a situation in which trade between the two regions may occur. Yet, North and South are less than perfectly integrated due to the existence of per-unit trade costs (which may be composed of tariffs and transport costs). In particular, we rely on the usual "iceberg" formulation and assume that $\tau \geq 1$ units of a good have to be shipped in order for one unit to arrive at the destination.

The North. The North differs from the South in that its markets function perfectly. In particular, the northern credit market is frictionless so that there are no credit constraints. Moreover, in the North, each variety is produced by a large number of firms so that the northern goods market is perfectly competitive. Regarding access to technology and preferences, there are no differences between the two regions (i.e., technology and preferences are also represented by equations 1 and 2, respectively). Moreover, for the sake of simplicity, the North produces the same spectrum of goods as the South does.[5] Thus, following Banerjee and Newman (2004) and Foellmi and Oechslin (2010), poor and rich countries are not distinguished in terms of technology or endowments but according to how well important markets work.

Given our assumptions regarding markets and technologies, it is immediately clear that all northern firms operate the high-productivity technology and charge a uniform price – which, in turn, is equal to the marginal cost. In what follows, it is convenient to normalize the northern price level to one. Obviously, this normalization implies that all goods prices in the North (as well as the northern marginal cost) are also equal to one.

3.2 An Equilibrium with Intermediate Trade Costs

Under the assumptions made above, it is clear that τ gives the (marginal) cost of producing one unit of a good in the North and selling it in the South. As a result, since the northern firms operate under perfect competition, the price of any good produced in the North and exported to the South is given by τ. This, in turn, implies that all southern producers face a northern competitive fringe and cannot set a price above τ (in terms of the numéraire).

[5] It may be more natural to assume that the North produces a larger number of varieties than the South. Yet, doing so would increase the gains from trade but not change the qualitative implications otherwise.

3.2.1 Characterizing the Equilibrium

Intermediate per-unit trade costs. In what follows, we focus on an "intermediate" τ which makes a positive fraction of entrepreneurs – but not all of them – unable to set the price that would make domestic demand equal to the output produced by the firm. More specifically, we discuss an equilibrium where τ is such that (i) the price that would imply a domestic demand of $a\kappa$ units exceeds the upper bound τ; (ii) the profit-maximizing price charged by unconstrained entrepreneurs lies below the upper bound. In formal terms,

$$p(a\kappa) > \tau > p(\widetilde{y}), \tag{14}$$

where $p(y)$ and $\widetilde{y}$ are defined in (5) and (11), respectively.

Changes relative to the closed economy. Allowing for international trade leads to two formal adjustments (relative to the closed-economy variant of the model). First, the fact that there is a binding upper bound on prices changes the relationship between the endowment and the maximum firm output. For price-constrained firms, the relationship is now given by

$$\overline{y}^I = \begin{cases} b\left(\omega + \lambda\tau\rho^{-1}\overline{y}^I\right) & : \quad 0 \leq \omega < \omega_\kappa^I \\ a\left(\omega + \lambda\tau\rho^{-1}\overline{y}^I\right) & : \quad \omega_\kappa^I \leq \omega < \omega_\tau^I \end{cases}, \tag{9'}$$

where ω_κ^I denotes the level which permits borrowing of exactly the amount required to meet the critical investment size κ; ω_τ^I refers to the threshold which allows an entrepreneur to produce a quantity of output that goes exactly together with an equilibrium price of τ.[6] A straightforward derivation of the two thresholds in (9') gives

$$\omega_\kappa^I = \left(1 - \frac{\lambda a\tau}{\rho}\right)\kappa \text{ and } \omega_\tau^I = \left(1 - \frac{\lambda a\tau}{\rho}\right)(Y/P)(\tau/P)^{-\sigma}/a. \tag{15}$$

The second formal change concerns the determination of the borrowing rate. Since we are looking at an equilibrium in which a positive mass of entrepreneurs is price-constrained, the economy imports goods from abroad. This, in turn, implies that there must be positive aggregate exports (because our framework is static, trade needs to be balanced). The fact that the equilibrium involves exports allows us to explicitly pin down the borrowing rate. Since exporting one unit of an arbitrary good (which requires $1/a$ units of capital) generates an income of $1/\tau$, the domestic borrowing rate must be a/τ. If the equilibrium borrowing rate were higher, nobody would export since lending would generate a higher return; on the other hand, if the borrowing rate were lower, demand for capital would exceed supply since even the richest agents in the economy would seek credit in order to export as much as possible.

[6] For capital endowments equal to or bigger than ω_τ^I, the maximum output a firm can produce continues to be implicitly determined by $\overline{y}^I = a(\omega + \lambda x\left(\overline{y}^I\right)^{(\sigma-1)/\sigma})$.

Parameters. We now work towards a description of the parameter constellations under which this equilibrium can occur. The first step is to note that using $\rho = a/\tau$ in (15) yields

$$\omega_\kappa^I = \left(1 - \lambda\tau^2\right)\kappa \ \text{ and } \ \omega_\tau^I = \left(1 - \lambda\tau^2\right)(Y/P)(\tau/P)^{-\sigma}/a.$$

Thus, for a positive mass of price-constrained entrepreneurs to exist, we need $\tau^2 < 1/\lambda$. Secondly, observe that imposing condition (14) leads to a lower bound on τ. Using both $\rho = a/\tau$ and the definition of $\widetilde{y}$ in expression (5) gives $p(\widetilde{y}) = (1/\tau)(\sigma/(\sigma-1))$. As a result, $\tau > p(\widetilde{y})$ implies $\tau^2 > (\sigma/(\sigma-1))$. Thus, in sum, what we necessarily must have is

$$\frac{\sigma}{\sigma-1} < \tau^2 < \frac{1}{\lambda}. \tag{R2}$$

Finally, we want to make sure that entrepreneurs with $\omega < \omega_\kappa^I$ do indeed run a firm (instead of becoming lenders). To get a condition, note that each capital unit invested in a low-productivity firm generates a return of τb. On the other hand, lending is associated with a return of a/τ. We assume that the former exceeds the latter:

$$a/b < \tau^2. \tag{R3}$$

3.2.2 Establishing the Equilibrium

We now establish the existence of the equilibrium described above, assuming that the two additional parameter restrictions hold. We proceed in two steps. First, we derive an expression for aggregate imports. Second, we establish that the real income is uniquely pinned down.

Aggregate exports. Total consumption expenditures on an arbitrary good supplied by an entrepreneur with $\omega < \omega_\tau^I$ are $\tau c(\tau) = Y P^{\sigma-1}\tau^{1-\sigma}$. To get the value of imports, one has to deduct the value of the domestic production. Moreover, in a balanced trade equilibrium, the total value of all imports must be equal to the value of all exports, EXP. As a result, we have

$$\begin{aligned} EXP &= YP^{\sigma-1}\tau^{1-\sigma}G(\omega_\tau^I) \\ &\quad -\tau\int_0^{\omega_\kappa^I} \frac{b}{1-\lambda\tau^2 b/a}\omega dG(\omega) - \tau\int_{\omega_\kappa^I}^{\omega_\tau^I} \frac{a}{1-\lambda\tau^2}\omega dG(\omega), \end{aligned}$$

where the expression on the right-hand side of the first line gives total expenditures on all goods that are imported (i.e., goods produced by entrepreneurs with $\omega < \omega_\tau^I$); the first expression of the second line is the total value of the goods produced by domestic entrepreneurs with $\omega < \omega_\kappa^I$ (i.e., by low-productivity firms); the second expression of the second line gives the total value of the goods produced by domestic entrepreneurs with $\omega_\kappa^I \leq \omega < \omega_\tau^I$ (i.e., by high-productivity firms with an output that is too small to meet the demand at price τ).

Resource constraint. To find an expression for (gross-)capital demand, note first that from (11) and $\rho = a/\tau$ we have $\widetilde{y} = (Y/P)P^{\sigma}\tau^{\sigma}\left((\sigma-1)/\sigma\right)^{\sigma}$ and $\widetilde{\omega} = (1-\lambda(\sigma/(\sigma-1))\left(\widetilde{y}/a\right)$. With these expressions in mind, the credit market equilibrium condition reads

$$\begin{aligned} K &= \int_0^{\omega_\kappa^I} \frac{1}{1-\lambda\tau^2 b/a}\omega dG(\omega) + \int_{\omega_\kappa^I}^{\omega_\tau^I} \frac{1}{1-\lambda\tau^2}\omega dG(\omega) + \int_{\omega_\tau^I}^{\widetilde{\omega}} \frac{\overline{y}^I(\omega)}{a} dG(\omega) \\ &+ \int_{\widetilde{\omega}}^{\infty} \frac{\widetilde{y}}{a} dG(\omega) + \tau\frac{EXP}{a}, \end{aligned}$$

where $\overline{y}^I(\omega)$ is implicitly determined by (9'). Using the expression for total exports, EXP, derived above, the equilibrium condition can be rewritten as

$$\begin{aligned} K &= \int_0^{\omega_\kappa^I} \frac{1-\tau^2 b/a}{1-\lambda\tau^2 b/a}\omega dG(\omega) + \int_{\omega_\kappa^I}^{\omega_\tau^I} \frac{1-\tau^2}{1-\lambda\tau^2}\omega dG(\omega) + \int_{\omega_\tau^I}^{\widetilde{\omega}} \frac{\overline{y}^I(\omega)}{a} dG(\omega) \\ &+ \frac{1}{a} Y P^{\sigma-1}\tau^{\sigma}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}\left[1-G(\widetilde{\omega})\right] + \frac{1}{a} Y P^{\sigma-1}\tau^{2-\sigma} G(\omega_\tau^I). \end{aligned}$$

The following proposition shows that this equilibrium condition pins down a unique real income.

Proposition 2 *Suppose that conditions (R2) and (R3) hold and that κ is sufficiently low (in a sense made clear in the proof). Then, there exists a unique macroeconomic equilibrium (i.e., an equilibrium with the values of Y/P and ρ/P uniquely pinned down) where (i) the poorest entrepreneurs use the low-productivity technology; (ii) all poorer (and middle-class) entrepreneurs are price-constrained and face import competition; (iii) all richer entrepreneurs set the profit-maximizing price; (iv) the richest entrepreneurs export parts of their output.*

Proof. See Appendix. ■

The properties of this equilibrium are – in addition to the evidence discussed after Proposition 1 – consistent with stylized facts about the relative performance of exporting firms (see, e.g., Bernard et al., 2003). In particular, the firms that export parts of their production tend to be the biggest ones and they are also more productive than the average firm in the economy (since some import-competing small firms use the low-productivity technology). Moreover, to the extent that the set of richest entrepreneurs is relatively small, exporting firms are in a minority. Obviously, though, the mechanism behind these implications is entirely different from the one in the standard models of trade and heterogeneous firms (i.e., Melitz, 2003; Bernard et al., 2003). Here, in an environment characterized by credit market frictions and inequality, it is the wealth endowment that determines whether an entrepreneur can access the resources required to operate the high-productivity technology and to enter export markets.

3.2.3 Comparative-Static Properties

We now discuss a number of comparative-static properties of the equilibrium.

Trade barriers and average productivity. The first interesting comparative-static result is that reducing trade barriers (i.e., a fall in τ) increases the number of firms which use the low-productivity technology: Since the critical threshold in this regard, $\omega_\kappa^I = \left(1 - \lambda\tau^2\right)\kappa$, increases as τ falls, it must be the case that $G(\omega_\kappa^I)$ is higher when τ is lower. This is quite an intuitive finding. As τ shrinks, the maximum price that can be demanded (by the price-constrained firms) decreases while the cost of borrowing ($\rho = a/\tau$) increases. As a result, the profit margins shrink – which means that these firms face a reduction in the collateral they can put up. Less collateral, in turn, implies a lower borrowing capacity so that some additional firms become unable to meet the κ minimum investment threshold.

A higher number of low-productivity firms, however, does not necessarily mean that a larger fraction of the aggregate capital stock is allocated to less efficient technologies (which would imply a decline in capital-weighted average firm productivity). Since the already existing low-productivity firms invest less, the impact on average productivity is ambiguous. To see this, note that the share of total capital invested in low-productivity firms is given by

$$\frac{1}{1 - \lambda\tau^2 b/a} \int_0^{\omega_\kappa^I} \omega dG(\omega)/K = \frac{1}{1 - \lambda\tau^2 b/a} \int_0^{(1-\lambda\tau^2)\kappa} \omega dG(\omega)/K.$$

Obviously, the impact of lower trade barriers on the above expression depends on the parameters of the model and on the mass of entrepreneurs at ω_κ^I. If the latter is sufficiently big, a gradual reduction in trade barriers implies that a larger fraction of the capital stock is used less productively so that capital-weighted average firm productivity (and, as the discussion below shows, potentially aggregate real output) falls.

Trade barriers and the distribution. The second interesting finding relates to the income distribution. Given that the equilibrium discussed in Proposition 2 prevails, a reduction in trade barriers amplifies the polarization of the income distribution. To see this, note that the nominal rate of return in unconstrained firms is a/τ whereas the rate in price-constrained firms is given by $(1-\lambda)\tau b/(1-\lambda\tau^2(b/a))$ if the the low-productivity technology is operated; and by $(1-\lambda)\tau a/(1-\lambda\tau^2)$ if the high-productivity technology is used. Thus, lowering τ increases nominal incomes in the higher parts of the distribution and diminishes those at the bottom.

Figure 4 illustrates the impact of a fall in τ on the income distribution in qualitative terms.

Figure 4 here

To summarize, the nominal incomes of the poor fall because of lower output prices and a higher cost of borrowing. The rich gain because the return from exporting goods (or, alternatively, from lending to less affluent entrepreneurs) increases as τ declines.

It is noteworthy that the above discussion just illustrates through which channels a reduction in trade barriers might impair aggregate real output and individual real incomes. The discussion does not, however, imply that these variables necessarily fall. The reason is that reducing trade barriers goes hand in hand with a lower price level – which could overcompensate a less efficient use of factors or a decline in nominal incomes. We address this issue in detail by means of an example in the following subsection.

3.3 From Autarky to Full Integration: A Two-Groups Example

Using a two-groups example we now take a broader look at the relationship between trade barriers and macroeconomic outcomes and analyze all possible equilibrium constellations. In particular, we explore the behavior of the aggregate real output (or, equivalently, the aggregate real income), individual real incomes, and the income distribution as the trade costs continuously decrease from very high levels to zero. To do so, we impose a two-group distribution as this will allow us to obtain convenient closed-form solutions. More specifically, we assume that a fraction β of entrepreneurs are "poor" (P). The capital endowment of these entrepreneurs is given by $\omega_P = \theta K$, where $\theta < 1$. The remaining entrepreneurs, the "rich" (R), are endowed with $\omega_R = (1 - \beta\theta)K/(1 - \beta)$ capital units. It is easy to check that this specification implies an aggregate capital endowment of K.

To understand the exposition below, note that there are 2×2 possible equilibrium constellations under which international trade occurs: (i) either only the poor (τP) or all entrepreneurs (τE) are price-constrained; (ii) either only the rich (aR) or all entrepreneurs (aE) use the high-productivity technology. No trade occurs if the poor entrepreneurs are not price-constrained. Then, the economy is in an autarky equilibrium.

3.3.1 Analytical Characterization

Only poor agents price-constrained (τP). We now describe the two trade equilibria in which only the poor entrepreneurs are price-constrained. Suppose first that all agents use the high-productivity technology. Then, the output by the poor entrepreneurs is $a\theta K/(1-\lambda\tau^2)$. As a result, without facing a competitive fringe, they would charge $P^{(\sigma-1)/\sigma}Y^{1/\sigma}\left(1-\lambda\tau^2\right)^{1/\sigma}$ $(\theta aK)^{-1/\sigma}$. This expression must be larger than τ for the competitive fringe to be binding. To determine $P^{(\sigma-1)/\sigma}Y^{1/\sigma}$, we use the credit market equilibrium condition,

$$K=\beta\frac{1-\tau^2}{1-\lambda\tau^2}\theta K+\frac{1}{a}(1-\beta)YP^{\sigma-1}\tau^{\sigma}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}+\frac{1}{a}\beta YP^{\sigma-1}\tau^{2-\sigma},$$

which can be rearranged to obtain

$$YP^{\sigma-1}=aK\left(1-\beta\theta\frac{1-\tau^2}{1-\lambda\tau^2}\right)\left((1-\beta)\tau^{\sigma}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}+\beta\tau^{2-\sigma}\right)^{-1}. \tag{16}$$

This result allows us to express the condition for the competitive fringe to be binding in terms of exogenous variables only. In particular, we obtain

$$\theta<\left(1-\lambda\tau^2\right)\left(\beta+(1-\beta)\left(\tau^2\frac{\sigma-1}{\sigma}\right)^{\sigma}\right)^{-1}. \tag{17}$$

We proceed to explicitly calculate aggregate real output, Y/P, which can be interpreted as the welfare level of the average entrepreneur. To do so, we first have to determine P. Note that a share β of goods is priced at τ whereas the price of the remaining goods is $p(\widetilde{y})=\sigma/((\sigma-1)\tau)$. As a result, we have $P^{1-\sigma}=\beta\tau^{1-\sigma}+(1-\beta)\left(\sigma/((\sigma-1)\tau)\right)^{1-\sigma}$. We use this latter expression in (16) and obtain (recall $U=Y/P$)

$$U^{aE,\tau P}=aK\left(1-\beta\theta\frac{1-\tau^2}{1-\lambda\tau^2}\right)\frac{\left(\beta\tau^{2(1-\sigma)}+(1-\beta)\left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1}\right)^{\sigma/(\sigma-1)}}{\beta\tau^{2(1-\sigma)}+(1-\beta)\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}}. \tag{18}$$

Suppose now that the the poor entrepreneurs use the low-productivity technology. This happens if $\omega_\kappa^I=\left(1-\lambda\tau^2\right)\kappa>\theta K$ and $\tau^2>a/b$. After going through a similar series of steps, we find that aggregate real output in this case is given by

$$U^{aR,\tau P}=aK\left(1-\beta\theta\frac{1-\tau^2 b/a}{1-\lambda\tau^2 b/a}\right)\frac{\left(\beta\tau^{2(1-\sigma)}+(1-\beta)\left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1}\right)^{\sigma/(\sigma-1)}}{\beta\tau^{2(1-\sigma)}+(1-\beta)\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}},$$

which is obviously smaller than the expression in (18). The condition for the competitive fringe to be binding is $\theta<\left(1-\lambda\tau^2 b/a\right)\left(\beta+(1-\beta)\left(\tau^2(\sigma-1)/\sigma\right)^{\sigma}\right)^{-1}$.

All agents price-constrained (τE). We now turn to the equilibria in which all entrepreneurs are price-constrained and hence set their prices equal to τ (so that $P=\tau$). This happens if $\tau<p(\widetilde{y})$ or, equivalently, $\tau<(\sigma/(\sigma-1))^{1/2}$. As in the two cases above, $YP^{\sigma-1}$ can be determined by looking at the credit market equilibrium condition. In the constellation where all entrepreneurs use the high-productivity technology, this condition reads $K=$

$\beta(1-\tau^2)(1-\lambda\tau^2)^{-1}\theta K + a^{-1}YP^{\sigma-1}\tau^{-\sigma}(1-\beta) + a^{-1}\beta YP^{\sigma-1}\tau^{2-\sigma}$ (and there is a related condition if the poor entrepreneurs use the low-productivity technology). Real output is then given by

$$U^{aE,\tau E} = aK\left(1-\beta\theta\frac{1-\tau^2}{1-\lambda\tau^2}\right)\frac{1}{1-\beta+\beta\tau^2}$$

if all entrepreneurs operate the high-productivity technology and by

$$U^{aR,\tau E} = aK\left(1-\beta\theta\frac{1-\tau^2 b/a}{1-\lambda\tau^2 b/a}\right)\frac{1}{1-\beta+\beta\tau^2}$$

if the poor are forced to rely on the low-productivity technology. Finally, condition (17) is replaced by a condition stating that the poor agents run smaller firms than the rich ones do. If all entrepreneurs use the high-productivity technology, this holds if

$$\theta < 1-\lambda\tau^2.$$

In the case where the poor use the low-productivity technology, the condition is $\theta < 1-\lambda\tau^2 b/a$. Note that $\theta < 1-\lambda\tau^2$ implies that (17) holds (and a fortiori for the conditions relevant in the constellations where the poor use the low-productivity technology).

Group-specific real incomes. To see how individual welfare depends on trade costs, we derive the group-specific real incomes. The nominal income (revenue minus cost of borrowing) of the poor entrepreneurs, m_P, is given by $(1-\lambda)a\tau\theta K/(1-\lambda\tau^2)$ if they use the high-productivity technology; by $(1-\lambda)b\tau\theta K/(1-\lambda\tau^2 b/a)$ if they operate the low-productivity technology. Thus, the welfare level incurred by the representative poor agent, $U_P = m_P/P$, is given by

$$U_P = \begin{cases} \max\left\{\frac{(1-\lambda)b\tau}{1-\lambda\tau^2 b/a}, \frac{a}{\tau}\right\}\theta K/P & : \quad \theta K < \omega_\kappa^I = \kappa(1-\lambda\tau^2) \\ \frac{(1-\lambda)a\tau}{1-\lambda\tau^2}\theta K/P & : \quad \theta K \geq \omega_\kappa^I = \kappa(1-\lambda\tau^2) \end{cases}.$$

The nominal income of the rich entrepreneurs, m_R, reads $(p(\widetilde{y})-\rho)\,\widetilde{y} + \rho(1-\beta\theta)(1-\beta)^{-1}K$. Taking into account that $\widetilde{y} = YP^{\sigma-1}\tau^{\sigma}\left((\sigma-1)/\sigma\right)^{\sigma}$, we find that

$$U_R = \begin{cases} \left(\tau^2-1\right)YP^{\sigma-2}\tau^{\sigma-1}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma} + \frac{a}{\tau}\frac{1-\beta\theta}{1-\beta}K/P & : \quad \tau^2 < \sigma/(\sigma-1) \\ \frac{1}{\sigma-1}YP^{\sigma-2}\tau^{\sigma-1}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma} + \frac{a}{\tau}\frac{1-\beta\theta}{1-\beta}K/P & : \quad \tau^2 \geq \sigma/(\sigma-1) \end{cases}.$$

3.3.2 Numerical Example

Trade barriers and aggregate real output. To understand how a continuous fall in trade costs affects the domestic economy, it is most convenient to focus on a numerical example. The parameter values chosen for this exercise are given in Figure 5 which shows aggregate real output, $U = Y/P$, as a function of the trade costs, τ. The figure highlights that lower trade costs might affect real output through two different channels. First, a change in trade costs affects the domestic price structure and the current account. Second, a change in trade costs affects firms' profits and hence their ability to operate the high-productivity technology.

Figure 5 here

To identify the different channels in a systematic way, suppose first that $\tau > 1.48$. Then, according to condition (17), even the poor entrepreneurs are not price-constrained (moreover, they are able to use the high-productivity technology). Thus, as long as $\tau > 1.48$, we are in an autarky equilibrium and a reduction in trade costs has no impact on the domestic economy. This is no longer true, however, as soon as τ reaches 1.48. Then, a further decrease in trade costs makes the poor entrepreneurs price-constrained and forces them to charge lower prices. Let us now focus first on the consequences for aggregate real output via the first channel. It turns out that the change in the domestic price structure has two opposing effects on real output. On the one hand, there is a positive *pro-competitive effect*: Ceteris paribus, the decrease in prices for some goods boosts aggregate output since these goods are now consumed in higher quantities. On the other hand, however, the lower prices force the price-constrained firms to produce less output, and this fall in domestic production must be compensated through higher imports. As a result, the economy spends more resources on transportation. In other words, a fall in τ leads to a partial replacement of domestic output with costly imports from the North. Figure 5 illustrates that this negative *replacement effect* dominates at higher values of τ so that the real output decreases (even if there is no impact on the use of technologies).[7] Only if τ is sufficiently low, is the pro-competitive effect the dominant one – and real output increases as the trade costs go down. Finally, with fully integrated markets ($\tau = 1$), all monopolistic distortions vanish and the first-best utility level $U = aK$ is achieved.

The impact through the second channel becomes visible in the two discontinuous jumps. The reason for these jumps is that the poor agents use the low-productivity technology if $1.12 < \tau < 1.29$. As soon as τ falls below the upper bound, $((1 - \theta K/\kappa)/\lambda)^{1/2} \simeq 1.29$, the capital endowment of the poor agents, θK, falls short of ω_κ^I so that they are forced to operate the low-productivity technology.[8] Yet, the poor entrepreneurs stop using the low-productivity technology if τ reaches $(a/b)^{1/2} \simeq 1.12$. At this point, they decide to give up their business

[7] This replacement effect is reminiscent of a mechanism discussed in a paper by Brander and Krugman (1983). Brander and Krugman show that the rivalry of oligopolistic firms can lead to "reciprocal dumping" (i.e., two-way trade in the same product) and hence to "wasteful" spending on transportation.

[8] Note that the jump is an artefact of the discrete two-group distribution, with a continuous distribution there would be a gradual increase in entrepreneurs relying on the inefficient technology as τ shrinks.

and become lenders instead. Thus, to sum up, aggregate real output may fall substantially below the autarky level for "intermediate" levels of trade costs. A full integration, however, necessarily lifts the real output above the autarky level.

Figure 6 here

Trade barriers and the distribution. It is interesting to explore further how a decline in trade costs affects group-specific real incomes and the income distribution (as measured by the income ratio $m_R/m_P = U_R/U_P$). Figure 6 illustrates the relationship between trade costs and the income ratio for the same parameter values as above. We see that lower trade costs go together with a higher income ratio (and hence higher polarization and inequality). On the one hand, inequality goes up because the poor entrepreneurs are affected by some negative effects. Most notably, the poor entrepreneurs face a higher cost of borrowing since the borrowing rate $\rho = a/\tau$ is a negative function of τ (the poor are borrowers unless $\tau \leq (a/b)^{1/2}$). Moreover, this direct negative effect is amplified by the fact that higher borrowing costs (and lower prices) lead to a weaker borrowing capacity. Panel a. of Figure 7 illustrates that these negative effects may be sufficiently strong to make the poor worse off in absolute terms.

Figure 7 here

On the other hand, as illustrated in Panel b. of Figure 7, inequality goes up because the rich entrepreneurs uniformly gain from lower trade costs (with the exception of levels of τ which make the poor use the low-productivity technology). Overall, this means that there is strong increase in inequality as the trade costs fall from very high levels to zero. For instance, under the parameter values chosen, the income ratio under free trade ($\tau = 1$) is $(1-\beta\theta)/\left((1-\beta)\theta\right) = 8.5$ whereas it is only 4.1 under autarky.

Figure 8 here

Different degrees of credit market frictions. Figure 8, finally, illustrates the trade cost-output relationship for two different degrees of credit market frictions, $\lambda = 0.2$ (as above) and $\lambda = 0.15$. As predicted by the model, we see that the range of trade costs which leads to the use of the less-productive technology is broader if the friction is stronger (i.e., if $\lambda = 0.15$). Yet, the direct effects of λ on the real output (controlling for technology use) are quantitatively small. Any substantial effects are due to the higher prevalence of the use of the inefficient technology.

4 Discussion

This section relates our findings to the empirical macro literature on the consequences of international trade in developing countries and also discusses some policy implications.

4.1 The Model and the Evidence

Trade barriers and economic performance. The main implication of our theory is that – in an environment with credit market imperfections and wealth inequality – opening up to trade has a non-monotonic impact on real output. A partial liberalization may be detrimental since smaller firms – while not being driven out of the market – are forced to use a less productive technology so that aggregate productivity is impaired. In addition, a partial liberalization forces smaller firms to downsize their production – which means that the resulting shortfall in domestic output has to be imported at "unfavorable" terms of trade. A full integration, however, is clearly beneficial from an aggregate perspective: It brings a more even and diverse supply of goods and drives the inefficient firms out of the market (while the negative terms-of-trade effect disappears). Note that this non-monotonicity result stands in contrast to the predictions by models which are now standard in the literature on trade and heterogeneous firms (e.g., Bernard et al., 2003; Melitz, 2003; Melitz and Ottaviano, 2008). Relying on perfect credit markets (and abstracting from wealth inequality), these papers suggest that a reduction in trade barriers has a clear positive impact on productivity and real output.

It is undisputed that credit market imperfections and high wealth inequality are important phenomena throughout the developing world (see, e.g., Banerjee and Duflo, 2005). Thus, an obvious implication of our theory is that the empirical literature on trade barriers and economic performance in developing countries may not find robust results. A brief overview confirms this prediction (see Kehoe and Ruhl, 2010, for a more comprehensive overview). In particular, there are a number of studies (Dorwick and Golley; 2004; DeJong and Ripoll, 2006) that identify a positive impact of openness on growth in more advanced economies but no effect whatsoever in developing countries. Moreover, there are also papers which report that – in developing countries – more openness is actually harmful for growth (Yanikkaya, 2003); yet, other contributions suggest exactly the opposite effect (e.g., Warner, 2003).

The fact that the empirical literature on openness and growth is rather inconclusive also matches well with a number of country studies. On the one hand, it appears that two of East Asia's most successful "miracle countries" – South Korea and Taiwan – used to rely on non-orthodox policies, at least at the early stages of their development (see, e.g., Stiglitz and

Charlton, 2005). On the other hand, there is evidence that India's sudden and comprehensive import liberalization in the early 1990s was accompanied by a significant increase in the extent of resource misallocation across manufacturing firms (Hsieh and Klenow, 2009). Similarly, there is hardly any evidence that the sweeping liberalizations of the 1980s or 1990s in Mexico, Brazil, and Turkey promoted economic performance (see, e.g., Rodrik, 2010). Moreover, according to Lall (1999), it seems that African countries like Kenya, Tanzania, or Zimbabwe did not respond well to the significant liberalization steps taken in the early 1990s.

Trade barriers and the distribution. A further strong implication of our theory is that opening up to international trade leads to higher incomes at the top end of the entrepreneurial income distribution but reduces those at the bottom. As a consequence, trade fosters the polarization of the income distribution (in the sense of Esteban and Ray, 1994) and, most likely, increases income inequality.[9]

The result that the gains from lower trade barriers are concentrated at the top is consistent with recent evidence on the evolution of top incomes in several developing economies. For instance, in the aftermath of significant liberalization steps in the early 1990s, the top-1% income shares in Argentina and India surged (Atkinson et al., 2011, Figure 11; Banerjee and Piketty, 2005, Figure 4). Similarly, there is evidence of surging top-income shares in Mexico after the country comprehensively liberalized trade in the mid-1980s (Foellmi and Oechslin, 2010). More generally, in developing countries, it seems that globalization goes hand in hand with increases in various measures of overall inequality (Goldberg and Pavcnik, 2007, Table 1). Again, our theory is consistent with this observation.

An example: India in the 1990s. Figure 9 summarizes the Indian experience in the 1990s. The figure highlights that the country liberalized international trade substantially in the first half of the decade: Between 1990 and 1995, the average weighted tariffs on manufactured goods fell by more than 50%. In Panel *a.*, the evolution of the average tariff rate is combined with a measure for economic performance, actual output relative to efficient output. The latter measure comes from Hsieh and Klenow (2009) and gives an indication of the loss in output due to misallocation of input factors (i.e., capital and labor) across firms. Obviously, in the early 1990s, opening up to trade went hand in hand with a significant deterioration in the allocation

[9] Note that, in principle, the impact of τ on measures of overall inequality (such as the Gini index) is ambiguous. The reason is that lowering τ may push entrepreneurs in the middle closer to the poorer ones. Yet, given the strong effects at the bottom and the top, lower trade barriers are likely to increase overall inequality.

of input factors (unfortunately, there is no data for the post-1994 period).[10]

Figure 9 here

Yet, as Panel *b.* of Figure 9 shows, the deterioration in the allocation of resources in the 1990-94 period did not have a negative impact on top incomes. During the same period, the top-1% income share rose from 7.4% to 8.1% (and was 9% at the end of the decade). Moreover, as can be seen from Figures 2 and 3 in Banerjee and Piketty (2005), the top-0.01% and the top-0.1% income shares experienced even steeper relative increases.

4.2 Policy Implications

Although we show that – in places with financial market frictions and high wealth inequality – globalization may cause deterioration in the factor allocation and impair aggregate output, our analysis does not suggest that poor countries should stay away from trade liberalization. Such a conclusion would be inappropriate for two different reasons. First, our model implies that only an incomplete liberalization of trade, i.e., a reform that falls short of fully integrating the country into the world economy, may have detrimental effects on macroeconomic aggregates. As highlighted above, a reform that brings the cost of trading with the outside world reasonably close to zero will always have a positive effect on the aggregate variables.

Second, even a modest reduction in trade barriers may boost aggregate output if it were implemented together with complementary reform measures.[11] Since the potentially negative effect of a partial liberalization comes from tighter credit constraints, the complementary measures should concentrate on the credit market. One option in this regard would be to improve credit contract enforcement (i.e., to increase λ). Obviously, if the improvement were sufficiently high, the borrowing constraints would ease or disappear even though mark-ups shrink. As a result, smaller firms would no longer have to cut production or even switch to the less productive technology. Moreover, as a positive side effect, strengthening contract enforcement would ensure that the smaller firms face less steep decreases in income.

[10] As documented in Hsieh and Klenow's Table 1, this decline in efficiency is mirrored in a rise of the dispersion of the marginal physical productivities across firms. Such a rise may also occur in our set-up as a fall in τ increases the share of firms using the less productive technology. If this share rises from a low level, the standard deviation of the marginal products increases first and then decreases. On the other hand, Table 2 in Hsieh and Klenow suggests a decline in the dispersion of the marginal revenue productivities. This is again consistent with our model which implies that the dispersion of mark-ups falls due to the pro-competitive effect of trade.

[11] A sizeable reduction might be infeasible because, e.g., the remoteness of the place implies high trade costs even if tariffs are negligible; the lack of a tax bureaucracy means that the state is forced to rely on trade taxes.

Yet, a significant improvement in the quality of credit contract enforcement may be difficult to achieve. Arguably, it would require substantial institutional reform (such as the introduction of India-style Debt Recovery Tribunals) and hence be very time-consuming or infeasible. There is, however, a less ambitious alternative. Since a firm's borrowing capacity is negatively related to the borrowing rate (equation 9'), introducing a subsidized-credit scheme for constrained firms would have a very similar effect. The subsidy could be financed through an income tax (i.e., a tax on m_i) which has upon introduction only welfare costs of second order (in the present framework it would not lead to any further distortions at all). Alternatively, if the reduction in τ increased tariff revenues, these additional resources could be used as a source of finance.

5 Summary and Conclusions

We study the macroeconomic implications of trade liberalization in a monopolistically competitive economy that features significant credit market frictions and wealth inequality. Our analysis generates two main findings. First, in contrast to much of the recent literature which emphasizes a positive selection effect of trade, we find that a partial integration into world markets may actually worsen the allocation of production factors and reduce aggregate real output. The reason is that a partial integration lowers mark-ups and hence the borrowing capacity of the less affluent entrepreneurs – who then have to switch to a less-productive technology but are not driven out of the market. In the present framework, only a full integration ensures a positive impact on productivity and output. The second finding relates to the income distribution. We show that integrating into world markets amplifies the pre-existing income inequality. The reason is that the less-affluent entrepreneurs are forced to downsize their firms and to charge lower mark-ups whereas the richer entrepreneurs profit from the access to new markets abroad. Note that both implications are consistent with empirical evidence from the developing world, i.e., from places where strong credit market frictions and high wealth inequality abound. One conclusion from our analysis is that developing countries should liberalize trade as part of a broader reform agenda that also addresses credit market frictions. In particular, developing countries should implement complementary measures that avoid a tightening of credit constraints in response to opening up. According to our analysis, there are numerous possible measures, some more modest (like the introduction of subsidized-credit schemes) and some more ambitious (like improving the quality of credit contract enforcement).

References

[1] Acemoglu, Daron and Fabrizio Zilibotti (2001); "Productivity Differences," *Quarterly Journal of Economics*, 116(2), 563-606.

[2] Atkinson, Anthony, Thomas Piketty, and Emmanuel Saez (2011); "Top Incomes in the Long Run of History", *Journal of Economic Literature*, 49(1), 3-71.

[3] Banerjee, Abhijit and Esther Duflo (2005); "Growth Theory Through the Lens of Development Economics," in: Aghion, P. and S. Durlauf, eds., *Handbook of Economic Growth*, Vol. 1A. Amsterdam: Elsevier.

[4] Banerjee, Abhijit and Esther Duflo (2010); "Giving Credit Where It is Due", *Journal of Economic Perspectives*, 24(3), 61-80.

[5] Banerjee, Abhijit and Andrew Newman (2004); "Notes for Credit, Growth, and Trade Policy", unpublished manuscript.

[6] Banerjee, Abhijit and Thomas Piketty (2005); "Top Indian Incomes, 1922-2000," *World Bank Economic Review* 19(1), 1-20.

[7] Bernard, Andrew, Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum (2003); "Plants and Productivity in International Trade", *American Economic Review*, 93(4), 1268-1290.

[8] Bertola, Giuseppe and Anna Lo Prete (2010); "Whence Policy? Government Policies, Finance, and Economic Integration", C.E.P.R. Discussion Paper 7820.

[9] Brander, James and Paul Krugman (1983); "A 'Reciprocal Dumping' Model of International Trade", *Journal of International Economics*, 15(3-4), 313-321.

[10] DeJong, David and Marla Ripoll (2006); "Tariffs and Growth: An Empirical Exploration of Contingent Relationships", *Review of Economics and Statistics*, 88(4), 625-640.

[11] Dixit, Avinash and Joseph Stiglitz (1977); "Monopolistic Competition and Optimum Product Diversity", *American Economic Review*, 67(3), 297-308.

[12] Dowrick, Steve and Jane Golley (2004); "Trade Openness and Growth: Who Benefits?", *Oxford Review of Economic Policy*, 20(1), 38-56.

[13] Epifani, Paolo and Gino Gancia (2011); "Trade, Markup Heterogeneity and Misallocations", *Journal of International Economics*, 83(1), 1-13.

[14] Esteban, Joan-María and Debraj Ray (1994); "On the Measurement of Polarization", *Econometrica*, 62(4), 819-851.

[15] Foellmi, Reto and Manuel Oechslin (2010); "Market Imperfections, Wealth Inequality, and the Distribution of Trade Gains", *Journal of International Economics*, 81(1), 15-25.

[16] Goldberg, Pinelopi and Nina Pavcnik (2007); "Distributional Effects of Globalization in Developing Countries", *Journal of Economic Literature*, 45(1), 39-82.

[17] Hsieh, Chang-Tai and Peter Klenow (2009); "Misallocation and Manufacturing TFP in China and India", *Quarterly Journal of Economics*, 124(4), 1403-1448.

[18] Kehoe, Timothy and Kim Ruhl (2010); "Why Have Economic Reforms in Mexico Not Generated Growth?", *Journal of Economic Literature*, 48(4), 1005-10273.

[19] Lall, Sanjaya, ed., (1999); *The Technological Response to Import Liberalization in Sub-Saharan Africa*. London: Palgrave MacMillan.

[20] Matsuyama, Kiminori (2005); "Credit Market Imperfections and Patterns of International Trade and Capital Flows", *Journal of the European Economic Association*, 3(2-3), 714-723.

[21] Melitz, Marc (2003); "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, 71(6), 1695–1725.

[22] Melitz, Marc and Giancarlo Ottaviano (2008); "Market Size, Trade, and Productivity", *Review of Economic Studies*, 75(1), 295-316.

[23] Rodrik, Dani (2010); "Diagnostics before Prescription", *Journal of Economic Perspectives*, 24(3), 33-44.

[24] Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti (2011); "Growing Like China", *American Economic Review*, 101(1), 202-241.

[25] Stiglitz, Joseph and Andrew Charlton (2005); *Fair Trade For All. How Trade Can Promote Development*. New York: Oxford University Press.

[26] UNCTAD (2006); *The Least Developed Countries Report 2006. Developing Productive Capacities*. New York and Geneva: United Nations Publication.

[27] Warner, Andrew (2003); "Once More Into the Breach: Economic Growth and Global Integration", Center for Global Development Working Paper Number 34.

[28] Yanikkaya, Halit (2003); "Trade Openness and Economic Growth: A Cross-Country Empirical Investigation", *Journal of Development Economics*, 72(1), 57-89.

APPENDIX: PROOFS

Proof of Proposition 1. (i) We first focus on the case $\lambda < (\sigma-1)/\sigma$ (credit rationing). In order to establish that there is a unique macroeconomic equilibrium, we proceed in two steps. We first show the existence of a unique equilibrium value of x. The second step is then to prove also that Y/P and ρ/P are uniquely pinned down.

To achieve the first step, observe that the equilibrium value of x must solve $K^D(x) = K$, where $K^D(x)$ is given by (13). Suppose now that x is exactly equal to the threshold given in (10). Then, $\widetilde{y}(x)/a$ is equal to κ whereas both $\overline{y}(\omega; x)/a$ (with $\omega \in [\omega_\kappa, \widetilde{\omega})$) and $\overline{y}(\omega; x)/b$ (with $\omega < \omega_\kappa$) are strictly smaller than κ. As a result, K^D must also be strictly smaller than κ. Moreover, since $\kappa < K$ due to (R1), we have $K^D < K$. Assume now that $x \to \infty$. Obviously, under these circumstances, we have $K^D \to \infty > K$. Finally, to show that there is a unique value that solves the equilibrium condition $K^D(x) = K$, we now establish that K^D increases monotonically as x rises from the threshold in (10) to infinity. Expressions (9) and (11) imply that both $\overline{y}(\omega; x)$ and $\widetilde{y}(x)$ are monotonically increasing in x. Moreover, the threshold ω_κ falls in x which reinforces the increase in capital demand since

$$\left[\frac{\overline{y}(\omega_\kappa^-)}{b} - \frac{\overline{y}(\omega_\kappa^+)}{a}\right] g(\omega_\kappa)\frac{d\omega_\kappa}{dx} \geq 0.$$

Thus, we have $K^D(x)/dx > 0$, and the proof of the first step is complete.

To show also that ρ/P (and hence Y/P) is uniquely pinned down, we make use of the CES price index. The first step is to find an expression for the price associated with an output level $\widetilde{y}$. To do so, we apply the expressions for x and $\widetilde{y}$ in (5) and get $p(\widetilde{y}) = (\rho/a)(\sigma/(\sigma-1))$. With this expression in mind, the definition of the CES price index implies

$$P^{1-\sigma} = \int_0^{\widetilde{\omega}} [p(\overline{y}(\omega))]^{1-\sigma}\, dG(\omega) + \left[\frac{\sigma}{\sigma-1}\frac{\rho}{a}\right]^{1-\sigma} [1 - G(\widetilde{\omega})]. \tag{19}$$

Then, relying again on (5) to substitute for $p(\overline{y}(\omega))$, we eventually obtain

$$\left(\frac{\rho}{P}\right)^{\sigma-1} = \int_0^{\widetilde{\omega}(x)} x^{1-\sigma}\, [\overline{y}(\omega; x)]^{(\sigma-1)/\sigma}\, dG(\omega) + \left[\frac{\sigma}{\sigma-1}\frac{1}{a}\right]^{1-\sigma} [1 - G(\widetilde{\omega}(x))],$$

which pins down the real interest rate ρ/P as a function of x (note that we can choose P as the numéraire and normalize to 1).

(ii) Assume now that $\lambda \geq (\sigma-1)/\sigma$ (no credit rationing). In this situation, all firms produce $\widetilde{y}$ and hence invest $\widetilde{y}/a$ capital units. As a result, (gross-)capital demand is given by $\int_0^\infty (\widetilde{y}/a)\, dG(\omega) = (Y/P)a^{\sigma-1}(\rho/P)^{-\sigma}((\sigma-1)/\sigma)^\sigma$. Moreover, since all firms invest $\widetilde{y}/a$, we must have that $K = \widetilde{y}/a$ – which implies $Y/P = aK$ (equation 6). Hence, the equilibrium interest rate is determined by

$$aKa^{\sigma-1}(\rho/P)^{-\sigma}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma} = K,$$

which results in $\rho/P = a(\sigma-1)/\sigma$.

Proof of Proposition 2. To start the proof, we introduce a number of definitions. First, we have $z \equiv P^{\sigma-1}Y$ so that (i) $p(y)$ defined in (5) reads $p(y) = z^{1/\sigma}y^{-1/\sigma}$; (ii) we have $x = (\tau/a)z^{1/\sigma}$. Second, it is convenient to introduce $\underline{z}$ which is the value of z that makes $p(a\kappa)$ equal to τ. Hence, we have $\underline{z} = (a\kappa)\tau^{\sigma}$. Thirdly, we write capital demand as a function of z:

$$\begin{aligned} K^D(z) \quad = \quad & \int_0^{\omega_\kappa^I} \frac{1-\tau^2 b/a}{1-\lambda\tau^2 b/a}\omega dG(\omega) + \int_{\omega_\kappa^I}^{\omega_\tau^I} \frac{1-\tau^2}{1-\lambda\tau^2}\omega dG(\omega) + \int_{\omega_\tau^I}^{\widetilde{\omega}} \frac{\overline{y}^I(\omega;z)}{a}dG(\omega) \\ & + \frac{1}{a}z\tau^{\sigma}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}[1-G(\widetilde{\omega}_I)] + \frac{1}{a}z\tau^{2-\sigma}G(\omega_\tau^I). \end{aligned}$$

Finally, note that $\overline{y}^I(\omega;z)$ is increasing in z and that $\omega_\kappa^I = \omega_\tau^I$ if $z = \underline{z}$.

We now show that – if κ is sufficiently low – $K^D(z) = K$ uniquely pins down z. The first step is to observe that, as z rises from $\underline{z}$ to infinity, $K^D(z)$ monotonically increases (as marginal changes in ω_τ^I and $\widetilde{\omega}$ leave K^D unaffected), where $\lim_{z\to\infty} K^D(z) = \infty$. The second step is to establish that $K^D(\underline{z}) < K$ if κ is sufficiently low. Since the first term in the above expression is negative and – at $z = \underline{z}$ – the second one is zero, we have

$$K^D(\underline{z}) < \int_{\omega_\tau^I}^{\widetilde{\omega}} \frac{\overline{y}^I(\omega;\underline{z})}{a}dG(\omega) + \frac{1}{a}\underline{z}\tau^{\sigma}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}[1-G(\widetilde{\omega})] + \frac{1}{a}\underline{z}\tau^{2-\sigma}G(\omega_\tau^I).$$

Moreover, using $\underline{z} = (a\kappa)\tau^{\sigma}$ and taking into account that $\overline{y}^I(\omega;z) \leq \widetilde{y} = \underline{z}\tau^{\sigma}\left((\sigma-1)/\sigma\right)^{\sigma}$ gives us

$$K^D(\underline{z}) < \kappa\left(\frac{\tau^2}{\sigma/(\sigma-1)}\right)^{\sigma}[1-G(\omega_\tau^I)] + \kappa\tau^2 G(\omega_\tau^I).$$

Note that the right-hand side (RHS) of the above expression depends only on exogenous parameters (and the distribution of ω). Thus, if κ is sufficiently low, we have $K^D(\underline{z}) < K$. Moreover, since $K^D(z)$ monotonically increases in z (and is unbounded), there exists a unique z which satisfies $K^D(z) = K$.

As in the proof of Proposition 1, the final step is to show that Y/P is uniquely pinned down (given that there is a unique z). To do so, we exploit again the CES price index which – in this case – can be written as

$$P^{(1-\sigma)} = \tau^{1-\sigma} G(\omega_\tau^I) + \int_{\omega_\tau^I}^{\widetilde{\omega}} \left[p(\overline{y}^I(\omega; z))\right]^{1-\sigma} dG(\omega) + \left[\frac{\sigma}{\sigma-1}\frac{1}{\tau}\right]^{1-\sigma} [1 - G(\widetilde{\omega})].$$

Note that $\overline{y}^I(\omega; z)$ as well as the thresholds ω_τ^I and $\widetilde{\omega}$ are functions of z (and the exogenous parameters of the model). As a result, P – and hence $Y/P = zP^{-\sigma}$ – are uniquely determined.

Proof of Lemma 1. The proof is most easily provided by a graphical argument. Consider the case $\omega < \omega_\kappa$. Whereas the left-hand side (LHS) of equation (9) is linear in $\overline{y}$ starting from zero, the RHS starts at ω and its slope reaches zero as $\overline{y}$ grows very large. Thus, $\overline{y}$ is uniquely determined. An increase in ω shifts up the RHS such that the new intersection of the LHS and the RHS lies to the right of the old one. The analogous argument holds true for $\omega \geq \omega_\kappa$. Finally, the definition of ω_κ implies that $\overline{y}(\omega_\kappa) = a\kappa > b\kappa > \lim_{\omega \to \omega_\kappa^-} \overline{y}(\omega)$. Hence, $\overline{y}(\omega)$ is strictly monotonic in ω.

Proof of Lemma 2. Suppose first $\lambda < (\sigma - 1)/\sigma$ so that $\widetilde{\omega} > 0$. Under these circumstances, entrepreneurs with $\omega \in [\omega_\kappa, \widetilde{\omega})$ have access to the efficient technology but their maximum output, $\overline{y}(\omega)$, falls short of $\widetilde{y}$. But this means that, when producing $\overline{y}(\omega)$, the marginal revenue still exceeds marginal costs. Thus, producing the maximum quantity is indeed optimal. On the other hand, entrepreneurs with $\omega \geq \widetilde{\omega}$ will not go beyond $\widetilde{y}$ because, if they chose a higher level, the marginal revenue would be lower than the cost of borrowing (if $\omega < \widetilde{y}/a$) or the income from lending (if $\omega \geq \widetilde{y}/a$). The second part of the claim is obvious and does not require further elaboration.

Proof of Lemma 3. To establish the claim, we show that the marginal revenue at the output level $b\kappa$ is not smaller than the marginal cost associated with the less efficient technology, ρ/b. This implies that for all $y < b\kappa$ marginal revenues strictly exceed marginal costs so that all entrepreneurs with $\omega < \omega_\kappa$ strictly prefer the maximum firm output. The marginal revenue at $y = b\kappa$ is given by $((\sigma - 1)/\sigma)P^{(\sigma-1)/\sigma}Y^{1/\sigma}(b\kappa)^{-1/\sigma}$, and so what we have to prove is

$$\begin{aligned} \frac{\sigma-1}{\sigma} P^{(\sigma-1)/\sigma} Y^{1/\sigma} (b\kappa)^{-1/\sigma} \quad &\geq \quad \frac{\rho}{b} \\ \frac{P^{(\sigma-1)/\sigma} Y^{1/\sigma}}{\rho} \quad &\geq \quad \frac{\sigma}{\sigma-1}\frac{1}{b}(b\kappa)^{1/\sigma}. \end{aligned} \tag{20}$$

In order to do so, we will establish a lower bound for the LHS of the second line in the above expression. Note that $((\sigma - 1)/\sigma)P^{(\sigma-1)/\sigma}Y^{1/\sigma}\widetilde{y}^{-1/\sigma} = \rho/a$. Notice further that, in an equilibrium, we must have that $\widetilde{y}/a \geq K$ since there are no firms operating at a higher scale of investment. Thus, we have $((\sigma-1)/\sigma)P^{(\sigma-1)/\sigma}Y^{1/\sigma}(aK)^{-1/\sigma} \geq \rho/a$ or, equivalently,

$$\frac{P^{(\sigma-1)/\sigma}Y^{1/\sigma}}{\rho} \geq \frac{\sigma}{\sigma-1}\frac{1}{a}(aK)^{1/\sigma}.$$

It is now straightforward to check that, due to the parameter restriction (R1), $(1/a)(aK)^{1/\sigma} > (1/b)(b\kappa)^{1/\sigma}$. But this means that (20) must be satisfied.