International Trade, Offshoring and Heterogeneous Firms

Richard Baldwin* and Toshihiro Okubo**

Abstract
Recent trade models determine the equilibrium distribution of firm-level efficiency endogenously and show that freer trade shifts the distribution towards higher average productivity due to entry and exit of firms. These models ignore the possibility that freer trade also alters the firm-size distribution via international firm migration (offshoring); firms must, by assumption, produce in their ’birth nation.’ We show that when firms are allowed to switch locations, new productivity effects arise. Freer trade induces the most efficient small-nation firms to move to the large nation. The big country gets an ‘extra helping’ of the most efficient firms while the small nation’s firm-size distribution is truncated on both ends. This reinforces the big-nation productivity gain while reducing or even reversing the small-nation productivity gain. The small nation is nevertheless better off allowing firm migration.

*Richard Baldwin
Graduate Institute, Geneva

**Toshihiro Okubo
Faculty of Economics, Keio University
Recent trade models determine the equilibrium distribution of firm-level efficiency endogenously and show that freer trade shifts the distribution towards higher average productivity due to entry and exit of firms. These models ignore the possibility that freer trade also alters the firm-size distribution via international firm migration (offshoring); firms must, by assumption, produce in their 'birth nation.' We show that when firms are allowed to switch locations, new productivity effects arise. Freer trade induces the most efficient small-nation firms to move to the large nation. The big country gets an ‘extra helping’ of the most efficient firms while the small nation’s firm-size distribution is truncated on both ends. This reinforces the big-nation productivity gain while reducing or even reversing the small-nation productivity gain. The small nation is nevertheless better off allowing firm migration.

JEL H32, P16.

Keywords: heterogeneous firms, trade, economic geography, productivity, offshoring.

1. INTRODUCTION

The flourishing of heterogeneous-firms trade models which started with Eaton and Kortum (2002) and Melitz (2003) focused attention on the way freer trade can boost productivity via changes in the equilibrium distribution of firm-level outputs (the least efficient shuts down and the most efficient expands). Helpman, Melitz and Yeaple (2004) allowed for multinational production, i.e. production by a single firm in both home...
and foreign whose purpose is to avoid trade costs, but the literature to date ignores the issue of international firm migration – what is known as offshoring in North America, ‘delocation’ in Europe and ‘hollowing out’ in Japan – i.e. where a firm ceases production in one nation and sets up production abroad. Since firm migration – the shifting of production location by individual firms – is an important empirical phenomenon, it is striking that heterogeneous-firm trade (HFT) models ignores this spatial relocation.¹

Our paper argues that the assumption of no spatial relocation in the HFT models is not innocent. We work in a framework where nations differ only in size. In this world, freer trade has two distinct effects. The first is the well-known pro-productivity effect discussed above. The second is that it creates “home market effect” pressures that foster the shifting of production to the larger market. In the standard HFT model, all the home-market-effect pressure is alleviate via the entry and exit of firms (see, for example, the asymmetric case in Helpman, Melitz and Yeaple 2004). We argue that since entry/exit is not instantaneous, some of the home-market-effect pressure would be alleviated via relocation of the most efficient firms from the small market to the large market – as long as the possibility of spatial reallocation is not ruled out by assumption. Such firm-migration would have interesting implications for the firm-size distribution in both nations. The big country would get an ‘extra helping’ of the most efficient firms and the small nation’s firm-size distribution would be truncated on both ends. The most efficient firms would relocate and the least efficient firms would exit. In this way, relocation adds a new dimension to the productivity impact of freer trade. The big nation gets an extra-large productivity gain while the small nation’s productivity gain is mitigated or even reversed.

To demonstrate these interactions as cleanly as possible, our paper combines a simple New Economic Geography (NEG) model with a simple heterogeneous-firm trade (HFT) model. Since NEG and HFT models are both on the edge of analytic tractability, it is not surprising that the combined model is impossible to solve analytically when we allow entry/exit and relocation simultaneously. To explore the combined model analytically, we consider two polar cases: One where entry and exit of firms is instantaneous, but relocation is slow, and one where relocation is instantaneous, but entry/exit is slow.

¹ Many empirical studies found the offshoring of production using micro-data, such as Head and Ries (2003), Girma and Görg (2004) and Tomiura (2007).
1.1. Literature review

Our paper has antecedents from two strands of literature. One branch focuses on differences among firms and has been recently reviewed by Redding (2010); the main theoretical papers are Eaton and Kortum (2002), Bernard, Eaton, Jensen and Kortum (2003), Melitz (2003), Helpman, Melitz and Yeaple (2004), Bernard, Redding and Schott (2007), Melitz and Ottaviano (2008), Yeaple (2005), Demidova (2005) and Chaney (2005). These models were motivated by empirical evidence. For example, firm differences within sectors may be more pronounced than differences between sector averages, and most firms – even in traded-goods sectors – do not export at all (Bernard and Jensen 1995, 1999; Clerides, Lach and Tybout 1998, Aw, Chung, and Roberts 2000, Eaton, Kortum and Kramarz 2004, Arnold and Hussinger (2010); see Tybout 2003 for a survey).

The other strand is the NEG literature (see Fujita, Krugman and Venables 1999, and Fujita and Thisse 2002). In this literature, increasing returns to scale and trade costs create forces that foster spatial agglomeration. Since the well-known core-periphery model was introduced by Krugman (1991), various models have been provided by the literature. The simplest is Martin and Rogers (1995); this is the NEG model we embrace.

The paper that is closest to ours in combining heterogeneous firms and NEG considerations is our earlier paper, Baldwin and Okubo (2006b). That paper ignores two key elements that are crucial to the heterogeneous-firm literature, namely entry and exit of firms, and fixed market entry costs. The main findings in that paper are that the most efficient firms are likely to first relocate to the bigger market. This suggests that standard econometric tests of agglomeration economies are biased. Okubo (2009), Okubo et al. (2010) and Forslid and Okubo (2012) are other papers in this line.2

The rest of the paper is organised into three sections. The next section introduces the model and solves it for the closed economy case. The following section works out the trading equilibrium and the welfare impact on the small nation. The final section presents our concluding remarks.

2. THE MODEL

Our model, which combines the Martin and Rogers (1995) NEG model with Melitz (2003) HFT model, is well known, so we present it briefly (see Baldwin and Okubo 2006a for details). The two nations (North and South) are identical except North is larger. One sector is numeraire, Walrasian with costless trade and uses only labour; the other (M-sector) is a Dixit-Stiglitz with iceberg trade costs that uses labour and capital.

A typical M-sector firm \( j \) faces constant marginal costs \( (a_j \text{ units of labour}) \) and a fixed cost (one unit of capital). Capital is internationally mobile but labour is not. Since there is one unit of capital per firm, international migration by firms (i.e. relocation) is synonymous with capital mobility. Capital is owned by labourers who are immobile between nations and capital income is repatriated costlessly, so capital/firms move in search of the highest reward irrespective of national cost-of-living considerations. Following Melitz (2003), a potential firm pays a start-up cost of \( F_I \) units of labour to create a unit of capital and intertemporal discounting is ignored but firms die according to a Poisson process with the hazard rate \( \delta \).

An M-sector firm faces per-unit and per-market ‘selling’ costs in each market; per-unit costs are zero in the local market but ‘iceberg’ in the export market (firms must ship \( 1+t \geq 1 \) units to sell one unit abroad). The per-market costs – what we call ‘beachhead costs’ and denote as \( F \) – reflect the fixed cost of establishing a foothold for each new variety in each market; \( F \) involves only labour inputs.

Trade in the numeraire equalises wage rates and these are normalised to unity (assuming nations are different but similar in size so both goods are produced in both nations). By the usual Dixit-Stiglitz properties, local and export market prices of a variety are \( a_j/(1-1/\sigma) \) and \( ta_j/(1-1/\sigma) \), where \( \sigma \) is the constant elasticity of substitution in CES utility function. Tastes are logarithmic quasi-linear, so demand for a North variety sold locally is:

\[
(1) \quad c_j = \frac{(p_j)^\sigma}{\bar{p}^{1-\sigma}} E; \quad \bar{p} = \left( \int_{\Theta^{*}} p_i^{1-\sigma} di + \int_{\Theta^{*}} \phi \int_{\Theta^{*}} p_h^{1-\sigma} dh \right)^{1/(1-\sigma)}, \quad \sigma > 1 \geq \phi \equiv (1+t)^{1-\sigma} \geq 0
\]

where \( p_j \) is variety-\( j \)’s producer price, \( \bar{p} \) is the price index and \( E \) is North expenditure on M; \( \Theta \) and \( \Theta^{*} \) are sets of goods produced in the North and the South, respectively and \( \phi \) indicates the ‘free-ness’ of trade.
Southern demand functions are isomorphic. With logarithmic quasi-linear preferences, expenditure on manufactured goods is invariant to income, trade costs and firm location.

2.1. Solution difficulties

The model is marked by a complex matrix of interactions involving firm heterogeneity, export behaviour, free entry-exit and international relocation. The spatial allocation of firms poses the greatest difficulties. The location choice depends upon the degree of competition in each market but this depends in turn upon the mass and efficiency of the firms that choose to locate and enter in each market. The crux of the difficulty lies in defining $\bar{p}$ and $\bar{p}^*$ when the production location of varieties is not determined a priori. To determine the equilibrium location of firms, we need to know the degree of competition in each market, but this in turn depends upon the distribution of the marginal costs in each market. In general, a firm with any marginal cost could locate in either nation, so it is not obvious how one would apply the density function for the $a$’s in forming the integrals behind the $\bar{p}$’s.

The literature has leaned on two extreme assumptions to get past this difficulty. The NEG literature ignores heterogeneity; the HFT literature ignores firm mobility. Determining the equilibrium when firms are heterogeneous and internationally mobile is the central technical difficulty of our model.

3. Equilibrium with trade

We start with autarky, i.e. zero mobility of goods and firms, where the solutions are well known. Assuming the distribution of $a_j$’s is Pareto, whose cumulative density function is denoted as $G[a]$ with a shape parameter, $\rho$, the mass of active firms and the threshold $a$ below which firms produce ($a_D$) are:

$$a_D = \left( \frac{f_j (\beta - 1)}{f} \right)^{1/\rho}, \quad n = \frac{E(\beta - 1)}{f \beta} \quad \text{where} \quad \beta \equiv \frac{\rho}{\sigma - 1} > 1, \quad f \equiv \sigma \delta F \quad \text{and} \quad f_i \equiv \sigma \delta F_i$$

As trade opens up, firms with $a$’s below the threshold $a_X$ export, and the large market becomes more attractive to firms in the increasing-returns M-sector (home market effect, henceforth HME). There are two adjustment margins for the HME pressure: (1) Firms can physically migrate from South to North, or (2) firm births can exceed firm deaths in the North, while the opposite occurs in the South.
The existence of two adjustment margins makes characterisation of the equilibrium difficult. We know that the distribution of firm-births and firm-deaths in equilibrium must be $\delta n G[a]$ and $\delta n^* G[a]$ ("*" denotes Southern one), but we do not know, a priori, where these varieties will be manufactured. Determining production location requires knowledge of the degree of local competition, but these depend upon the location of production.

To tackle this problem, it is useful to work out the equilibrium assuming that entry/exit and relocation occur in sequence. First entry/exit is instantaneous, and then relocation becomes possible. Second, relocation is instantaneous and then entry/exit becomes possible.

### 3.1. Instantaneous entry/exit followed by relocation

When entry/exit is instantaneous but relocation is slow, all adjustment to the HME pressure occurs via entry/exit. The solution is thus similar Helpman, Melitz and Yeaple (2004), namely:

$$a_D = \left( \frac{f_I (\beta - 1)}{f(1 + \phi^\beta)} \right)^{1/n}, \quad a_X = \left( \frac{f_I \phi^\beta (\beta - 1)}{f(1 + \phi^\beta)} \right)^{1/\rho}, \quad n^w = \frac{1 - 1/\beta}{f(1 + \phi^\beta)}, \quad s_n = \frac{1}{2} + \frac{1 + \phi^\beta}{1 - \phi^\beta} \left( s_E - \frac{1}{2} \right)$$

where $s_E \equiv \frac{E}{E+E^*}, \quad s_n \equiv \frac{n}{n+n^*}, \quad n^w \equiv n+n^*$

Note that $s_n$ and $s_E (>0.5)$ are the North’s share of the worldwide mass of firms $n+n^*$ and worldwide expenditure on all M-sector varieties $E+E^*$.

When firms are allowed to change locations (i.e. relocate), each existing firm compares the operating profit it would earn in the two markets. For exporting firms, which are already selling in both markets, the question boils down to the minimization of trade costs. Operating profit when located in the North is $a_j^{1-\sigma}(B+\phi B^*)/\sigma$ for a typical firm; when in the South it is $a_j^{1-\sigma}(B+\phi B^*)/\sigma$, where $B \equiv \frac{E}{\bar{m}}, \quad \bar{m} \equiv n \int_0^{a_n} a^{1-\sigma} dG[a|a_D] \cdot B^*$ are isomorphic. Since $B=B^*$ in equilibrium, no firm would gain from relocation. To summarise:

**Result 1:** Instantaneous entry/exit relieves all pressure to relocate, i.e. instantaneous entry/exit is a perfect substitute for relocation, so the instant-entry-then-relocation case is identical to the standard HFT model.
The productivity effects of greater openness in this case are well known from Helpman, Melitz and Yeaple (2004), so we do not repeat them. However, productivity effects are substantially more complex when relocation occurs. To set the stage, we replicate the well-known productivity results using a flexible analytic technique. The basic idea is to focus on how freer trade changes the weighted average of the a’s (i.e. average productivity measure) – denoted \( \bar{a} \) – of the underlying distribution of a’s, denoted as \( G[a] \). The Northern average productivity measure is defined as
\[
\bar{a} = \int_{a_x}^{a_0} \left( \frac{a^{-\sigma} B(1 + \phi)}{TP} \right) n dG[a|a_D] + \int_{a_x}^{a_0} \left( \frac{a^{-\sigma} B}{TP} \right) n dG[a|a_B],
\]
where \( TP \), total production in the North, is given by
\[
TP = \int_{0}^{a_x} a^{-\sigma} B(1 + \phi) n dG[a|a_D] + \int_{a_x}^{a_0} a^{-\sigma} B n dG[a|a_B].
\]
The weights represent the importance of firms that export (X-types) and firms that sell only locally (D-types). Solving the integrals, the equilibrium expression for \( a_D \) in (3) and \( G[a]=a^{1-\rho} \), productivity measure in the North is written as:
\[
\bar{a} = \int_{0}^{a_x} W_X(a^{1-\rho}) da + \int_{a_x}^{a_0} W_D(a^{1-\rho}) da
\]
where \( W_X \) and \( W_D \) are the weights on X and D type firms:
\[
W_X = \frac{(1 + \phi)}{(1 + \phi^\beta)^{1/\beta}} \left( \frac{\rho(1-1/\beta)}{f(x(\beta - 1)/f)^{1-1/\beta}} \right), \quad W_D = \frac{1}{(1 + \phi^\beta)^{1/\beta}} \left( \frac{\rho(1-1/\beta)}{f(x(\beta - 1)/f)^{1-1/\beta}} \right)
\]
Plainly, \( W_X \) rises and \( W_D \) falls as trade gets freer \((d\phi>0)\). Thus:
\[
\frac{d\bar{a}}{d\phi} = \int_{0}^{a_x} \frac{dW_X}{d\phi} a^{1-\rho} da + \int_{a_x}^{a_0} \frac{dW_D}{d\phi} a^{1-\rho} da + \int_{0}^{a_x} \frac{dW_X}{d\phi} a^{1-\rho} da + \int_{a_x}^{a_0} \frac{dW_D}{d\phi} a^{1-\rho} da + \frac{da_X}{d\phi} (W_Xa_X^{1-\rho} - W_Da_D^{1-\rho}) + \frac{da_D}{d\phi} W_D a_D^{1-\rho}
\]
When \( d\bar{a}/d\phi \) is written in this way, we see that the productivity impact of freer trade depends upon the impact of \( d\phi \) on the ‘weighting’ terms – the \( dW/d\phi \)’s – under the integrals – and the impact on the limits of integration – i.e. the last two terms. We know that the derivative (4) is negative (i.e. there is a pro-

productivity effect) since freer trade uniformly shifts weight from high values of \( a^{1-\rho} \) to low values.

\[3\] See Baldwin and Okubo (2006a) for details.
To be more clearly, Figure 1 shows this graphically. The term that is being weighted, namely $G[a]=a^{1-\sigma+\rho}$, is shown with the rising solid line and we note that it is invariant to the freeness of trade. The weighting function (i.e. $W_X$ and $W_D$) before the increase in trade freeness is shown with the solid horizontal line, “Weights (pre)”. After the opening, the weighting function shifts to “Weights (post)”, so the weight on $a$’s below $a_X$ rises and the weight on $a$’s above $a_X$ falls. The impact of the change in the limits of integration is fourfold and corresponds to the third and fourth terms in (4). The first and second terms in (4) represent the switch of the weighting lines holding $a_D$ and $a_X$ at the initial level. The third term shows the impact of shift in weight for the range of $a$’s between the pre and post export cut-off. The final term shows the shift in weight (from positive to zero) for the range of $a$’s between the pre and post D-type cut-off.

Plainly the post-integration weighting curve unambiguously shifts mass to lower values of $a$, so we know that $\bar{a}$ falls, i.e. manufacturing productivity rises since the average unit labour input coefficient falls. In a sense, freer trade “pivots” the density clockwise around $a_X'$, lowering the weight on $a$’s above $a_X'$ and raising it on $a$’s below $a_X'$. Freer trade is pro-productivity in the South for the same reasons.

![Figure 1: Trade liberalisation’s productivity impact HFT model, free entry case.](image)

### 3.2. Instantaneous relocation followed by entry/exit

The assumption of instantaneous birth-and-death of firms is clearly unrealistic. If entry/exit takes time, some firms may adjust to the HME pressure by moving to the big market. To explore this possibility, we reverse the ordering and assume that relocation is instantaneous and then considers possibility of entry/exit
(starting from autarky as before). With entry/exit suspended initially, the two free entry conditions and the
domestic cut-off condition are moot, so all firms that were active in autarky remain active initially when
trade and firm-migration become possible.

When relocation is allowed, the order of firm migration matters. To see which firms move first we
calculate the incipient gain and assume firms move in order their then gain ranking. The present-value gain
from relocation is:

\[ v_X[a] = \frac{a^{1-\sigma}}{\delta \sigma} (1-\phi)(B-B^*) > 0; \quad 0 \leq a \leq a_X \]

This is highest for efficient firms and it is always positive since \( B > B^* \) in the initial situation. For Southern
D-type firms migrating to the North, the gain would be:

\[ v_D[a] = \frac{a^{1-\sigma}}{\delta \sigma} (B-B^*) > 0; \quad a_X \leq a \leq a_D \]

**Result 2:** The most efficient export firms in the small nation move to the large nation first. The local
small firms have less incentive to move to the large nation.

Relocation raises the degree of competition in big North market and lowers it in the South, so the value to
any given firms of moving declines as the range of firms that have moved expands. The process continues
until the gain from moving is zero, i.e. the location condition, \( B = B^* \), holds.

Formally, when the range of Southern firms that have moved to the North is \( a \in [0,a_R] \) the \( m \)'s are:

\[ m = n^{\alpha} \int_0^{a_R} a^{1-\sigma} dG[a] na_d[a] + n^{*} \int_0^{a_R} a^{1-\sigma} dG[a] na_D[a] + n^{*} \int_{a_D}^a a^{1-\sigma} dG[a] na_D[a] \]

\[ m^* = \phi n^{\alpha} \int_0^{a_R} a^{1-\sigma} dG[a] na_D[a] + \phi n^{*} \int_0^{a_R} a^{1-\sigma} dG[a] na_D[a] + n^{*} \int_{a_D}^a a^{1-\sigma} dG[a] na_D[a] \]

Solving these and using the solutions to define the B’s, we have:

\[ B = \frac{s_E (1-1/\beta) a_D^{\alpha} / n^w}{s_a a_D^{\alpha} + (1-s_a)(1-\phi) a_R^{\alpha} + (1-s_a)\phi a_X^{\alpha}}; \quad B^* = \frac{(1-s_E)(1-1/\beta) a_D^{\alpha} / n^w}{s_a a_D^{\alpha} - (1-s_a)(1-\phi) a_R^{\alpha} + (1-s_a) a_D^{\alpha}} \]

\[ \alpha = 1 - \sigma + \rho > 0 \]
Solving the location condition, \( B = B^* \), yields the ‘relocation threshold’ \( a_R \) as a function of the export threshold \( a_X \):

\[
(8) \quad a_R \alpha = \frac{s_n + s_E - 1}{(1 - s_n)(1 - \phi)} \phi a_X \alpha
\]

Employing the initial condition \( s_n = s_E \), \( a_R \) and the export cut-off condition are written as:

\[
(9) \quad a_R = \left( \frac{\phi(2s_E - 1)}{1 - \phi(1 - s_E)} \right)^{\frac{1}{\alpha}} a_X, \quad \frac{\phi a_X \alpha - (\beta - 1) f}{1 - \phi(1 - s_E)} \phi a_X \alpha = f
\]

The second expression cannot be solved for \( a_X \), but it has a unique and positive solution since the left-hand side is monotonically falling as \( a_X \) rises and it starts at zero when \( a_X = 0 \).

### 3.2.1 Two phases of migration-- from XX to DD relocation phases

Inspection of (9) reveals that \( a_R \) starts out at zero but gets progressively closer to \( a_X \) as the level of trade freeness (i.e. \( \phi \)) rises. This implies that the migrating firms are initially X-types as discussed above. When trade is sufficiently free, \( a_R \) equals \( a_X \). At this level of trade freeness, all Southern X-types will have moved to the North. Beyond this level of trade freeness, a second phase of relocation begins where all the migrating firms are D-types. We refer to these two phases of migration as XX and DD migration phases. The exact level of \( \phi \) where DD migration phase begins is simple to calculate. Setting \( a_R = a_X \) and solving for \( \phi \) yields:

\[
(10) \quad \phi^p = \frac{1 - s_E}{s_E}
\]

where \( \phi^p \) is the ‘partitioning’ level of trade freeness.

When \( \phi \) rises above \( \phi^p \), migration affects the B’s differently because the South no longer exports manufactures to the North and the migrating firms no longer sell back into the Southern market. The \( \bar{m} \)’s in DD relocation phase are now:

\[
\bar{m} = (n + n^*) \int_0^{a_x} a^{-1 - \sigma} dG[a|a_D] + n \int_{a_x}^{a_D} a^{-1 - \sigma} dG[a|a_D], \quad \bar{m}^* = (n + n^*) \int_0^{a_x} a^{-1 - \sigma} dG[a|a_D] + n^* \int_{a_x}^{a_D} a^{-1 - \sigma} dG[a|a_D]
\]

which solve to:
International Trade, Offshoring and Heterogeneous Firms

Using $s_n = s_E$, and solving the location condition ($B=B^*$) for $a_R$ in the DD relocation phase yields:

$$a_R = a_D \left( \frac{\phi s_E}{1-s_E} \right)^{1/\beta}$$

We see that as before, $a_R$ continues to rise gradually as trade gets freer.

Using the ratio of the cut-off conditions (12) can be written in terms of $a_D$. Solving for the $\phi$ where $a_R = a_D$ tells us the level of openness at which all firms have left the small nation – what is known as the ‘sustain point’ in economic geography, $\phi^S$. It is $\phi^S = \left((1-s_E)/s_E\right)^{1/\beta}$.

Note that since $\beta > 1$, and $s_E > 1/2$ the $\phi^P < \phi^S < 1$ and thus as $\phi$ limits to 1, it eventually passes the sustain point $\phi^S$ and so full agglomeration occurs. However, as $\beta$ increases, full agglomeration is less likely. At extreme, when $\beta = \infty$, full agglomeration never happens except for costless trade in which case any spatial location of production is an equilibrium, including full agglomeration. For $\beta$ to approach infinity, $\rho$ must approach infinity (this means that the firm marginal cost distribution is heavily skewed to the left, i.e. towards the highest marginal cost), or $\sigma$ must approach unity. To summarise:

**Result 3: (Two phases of firm relocation).** The gradually freer trade produces a gradual relocation of firms from the small nation to the big nation with the most efficient firms leaving first. This process is marked by two phases: migration of small-nation X-types (who remain X-types after their relocation), and – once all X-type firms have left the small nation – migration of small-nation D-type firms (who remain D-types in their new location).

The ‘partitioning’ threshold between the two phases of migration is defined by the level of trade freeness, $\phi^P$, which equals $(1-s_E)/s_E$; all firms are in the big nation when trade freeness $\phi$ surpasses the sustain point $\left((1-s_E)/s_E\right)^{1/\beta}$. The spatial distribution of firms according to their level of efficiency is depicted in Figure 2 (when relocation is still in the XX phase). The big market has an extra allocation of the most efficient firms while the Southern market has none.
We turn now to consider entry/exit by adding back the two free entry conditions and the domestic cut-off condition. To be concrete, we assume that \( \phi \) is such that marginal migration is of the XX phase, i.e. \( 0 < \phi < \phi^p \), although the analysis would be qualitatively identical for \( \phi > \phi^p \). As always, the B’s are critical. Using (8) to eliminate \( a_R \) and the ratio of cut-offs to eliminate \( a_X \), the domestic cut-off condition is:

\[
a_D^{1-\alpha} B = f, \quad B = \frac{s_E(1-1/\beta)a_D^{\alpha}}{s_xa_D^{\alpha} + s_E\phi^a a_D^{\alpha}}
\]

Solving this for \( n^w \):

\[
(13) \quad n^w = \frac{(\beta - 1)/f^2}{s_x/s_E - \phi^a}
\]

Since \( s_n = s_E \) in the case at hand, the comparison of (13) and (3) makes it clear that allowing entry and exit after instantaneous relocation results in a rise in the number of active firms. In other words, relocation does not fully relieve the HME pressure; global mass of firms must rise to restore expected pure profits to zero. This of course, requires the firm-birth rate to exceed the firm-death rate. Since the location condition assures the equality of profitability in the two nations, we assume that this affects each potential innovator in the world in the same way. This is natural to assume that the overall increase in \( n^w \) will rise pro rata on Northern and Southern varieties, so the share of firms owned by Northern citizens, \( s_n \), will not change. In short, we assume that \( s_n = s_E \) even after the global mass of firms adjusts.

Using the ratio of cut-offs and the free entry conditions we have:

\[
\frac{1}{n^w} = f + f(a_X)^{\rho} + f_l (\frac{1}{a_D})^{\rho}
\]
Using (13) to solve this for \( a_D \) and using the ratio of cut-off conditions to get \( a_X \), we have:

\[
 a_D = \left( \frac{f_1(\beta - 1) / f}{\beta (s_n / s_E - 1) + (1 + \phi^\rho)} \right)^{\rho}, \quad a_X = \left( \frac{\phi^\rho f_1(\beta - 1) / f}{\beta (s_n / s_E - 1) + (1 + \phi^\rho)} \right)^{\rho}
\]

(13), (14) and (12) characterise the equilibrium.

### 3.2.2 Trade, production and productivity effects

The effects of integration are much richer in the instantaneous-relocation-then-entry case. As usual, we get share-shifting and selection effects via the change in \( a_D \). However, we get a new effect from relocation, namely the change in \( a_R \). This means that freer trade has an extra-large productivity effect on the big nation since the firms that migrate to the big nation in response to freer trade are systematically more efficient than the least efficient of the existing firms. Correspondingly, the small nation experiences less of a pro-productivity effect. Indeed as we shall see, the small nation’s productivity may actually fall as trade gets freer.

Figure 3: Productivity effects, free relocation case.

Figure 3 shows the impact on the North’s and South’s average productivity when the economy is in the XX-phase of firm migration (relocating firms are X-types). For the North, freer trade pivots the density around \( a_X \) and so unambiguously shifts weight to lower values of \( a^{1-\sigma+\rho} \). The North’s weighted-average ‘a’ thus falls, i.e. average productivity rises. The right panel shows the impact on the South. Here, we get the usual share-shifting and selection effects, but since the South loses its most efficient firms (the weight on \( a^{1-\sigma+\rho} \)
\[ \alpha^{+p} \text{ between } a_R \text{ and } a_R' \text{ falls to zero}, \text{ the overall productivity impact on the sales-weighted ‘a’ is dampened.} \]

Indeed if the \[ a^{1-\alpha/p} \] curve is rising steeply enough, the average can actually rise, as we shall see below.

In the XX phase, our productivity measure is:

\[
\bar{\alpha}^* = \int_{a_X}^{a_D} \left( a^{1-\sigma} \frac{B(1 + \phi)}{TP^*} \right) n^* dG[a|a_D; a_R] + \int_{a_X}^{a_R} a^{1-\sigma} B(1 + \phi) n^* dG[a|a_D; a_R];
\]

where TP* denotes Southern total production specified as

\[
TP^* = \int_{a_X}^{a_D} a^{1-\sigma} Bn^* dG[a|a_D; a_R] + \int_{a_X}^{a_R} a^{1-\sigma} B(1 + \phi)n^* dG[a|a_D; a_R], \quad dG[a|a_D; a_R] = \frac{\rho a^{\rho-1}}{a_D^\rho - a_R^\rho} da
\]

Solving the integral, we get:

\[
\bar{\alpha}^* = \left( \frac{\alpha}{1 + \alpha} \right) \frac{(1 + \phi) a_X^{1+\alpha} - (1 + \phi) a_R^{1+\alpha} + a_D^{1+\alpha} - a_R^{1+\alpha}}{(1 + \phi) a_X^\alpha - (1 + \phi) a_R^\alpha + a_D^\alpha - a_R^\alpha}
\]

In the case of DD-phase relocation (firms that move are D-types), we have explicit solutions and the formal expression for the weight average ‘a’ can be written as:

\[
\bar{\alpha}^* = \int_{a_D}^{a_R} \left( a^{1-\sigma} \frac{Bn^*}{TP^*} \right) n^* dG[a|a_D; a_R], \quad TP^* = \int_{a_D}^{a_R} a^{1-\sigma} Bn^* dG[a|a_D; a_R]
\]

Solving the integral, we get:

\[
\bar{\alpha}^* = \left( \frac{\alpha}{1 + \alpha} \right) \frac{a_D^{1+\alpha} - a_R^{1+\alpha}}{a_D^\alpha - a_R^\alpha}. \quad \text{Since } a_D \text{ falls with freer trade and } a_R \text{ rises, the productivity effect is ambiguous. In particular, using (12), the ratio of cut-offs and (14), we have}
\]

\[(15) \quad \bar{\alpha}^* = \left( 1 + \phi^\beta \right)^{-\frac{1}{\rho}} \left( \frac{1}{1 - \phi^\beta s/(1 - s)} \right)^{1-\frac{1}{1-\sigma+p}} \left( \frac{1}{2 - \sigma + \rho \left( f_s (1 - \sigma + \rho) \right)} \right)
\]

where \( s = s_n = s_E \) is given by the initial conditions. Note that the first right-hand term is decreasing in \( \phi \) and so tends to imply that freer trade is pro-productivity; the second term however is increasing in \( \phi \) and thus anti-productivity. We can vary the relative importance to the two conflicting terms by varying \( \rho \). When \( \rho \) is small (the regularity condition only requires \( \rho \) to be bigger than \( \sigma-1 \), the pro-productivity term gets stronger relative to the anti-productivity term, so the overall derivative is negative (freer trade lowers the average \( a \)).
Conversely, when $\rho$ is large, freer trade tends to be anti-productivity. Heuristically speaking, when $\rho$ is large, then the line $a^{1+\rho-\sigma}$ is very steep and this amplifies the productivity impact of the loss of probability weight on low values of $a$'s that occurs due to a rise in $a_R$. For example, in the special case of $\sigma = \rho = 2$, the productivity impact of freer trade, $d\bar{\pi}^*/d\phi$, equals $\phi(s(2+\phi^2)/(1-s) - 1)/2(1+\phi^2)^{3/2}$, which is positive since $s > \frac{1}{2}$. In other words, freer trade lowers productivity in the small nation in this case. To summarise, we write

Result 4: (New productivity effects). Since the most efficient firms migrate first and freer trade encourages such relocation, freer trade leads to an extra-large productivity gain in the large nation which receives the migrating firms, but dampens the productivity gains in the small nation. During the second phase of migration (see Result 3), freer trade always boosts productivity in the large nation, but may raise or lower productivity in the small nation. If $\rho$ is large enough, freer trade lowers the small nation’s productivity.

3.3. General model

Modelling the entry/exit and relocation process more generally than we have with our two polar cases runs into the fundamental indeterminacy of the system. We have five equilibrium conditions – the two cut-off conditions, the two free-entry conditions and the location condition – and five equilibrating variables, $a_D$, $a_X$, $n$, $n^*$ and $a_R$. However, the two free-entry conditions and the location condition are not independent; any two imply the third. For example, if the two free-entry conditions hold, the location condition $B = B^*$ holds automatically. If the location condition and one free entry condition hold, the other free entry condition is automatically satisfied.

This means that we cannot determine $s_n$, $n^w$ and $a_R$ in general; the most we can say is that:

$$a_R = \frac{s_n + s_E - 1}{(1-s_n)(1-\phi)} (1-\phi) a_X$$
$$n^w = \frac{(1-1/\beta)/f}{s_n/s_E + \phi^\beta}$$

where the cut-offs are defined as in (14). The standard HFT model resolves the indeterminacy by ignoring relocation, i.e. by assuming $a_R = 0$. The standard NEG model resolves it by ignoring firm heterogeneity. The relocation-then-entry/exit case fixes $s_n$ with initial conditions and is thus able to find $a_R$. 
To pin down $s_n$, $n^w$ and $a_R$ in general would require us to introduce quadratic adjustment costs for both relocation and entry/exit. The result would be a system of six differential equations for the six state variables (the three state variables $n^w$, $s_n$ and $a_R$, and their corresponding co-state variables). Working with more than two differential equations is difficult and rarely rewarding. In this case, the final result would be that some of the HME adjustment would occur via relocation and some would occur via entry/exit. Thus, $s_n$ would be higher than $s_E$ in the general model as some of the HME pressure would have been relieved via relocation and some via entry/exit.

The productivity effects on the small nation raises the question of whether the small nation might, under certain parameter conditions, benefit from forbidding relocation while trade costs fall. In the working paper version we show that:

**Result 5: Allowing relocation in the face of trade liberalisation improves the small nation’s welfare more when neither free entry/exit nor relocation is instantaneous.**

4. **CONCLUDING REMARKS**

Eaton and Kortum (2002) and Melitz (2003) opened a new line of research in international trade by providing tractable models where the entry and exit of firms determines the distribution of firm-level efficiency endogenously. One of the most notable results in this HFT literature concerns the impact of trade integration on productivity; freer trade raises average productivity by forcing out the weakest firms and shifting production shares to the most productivity firms. These models, however, ignore the possibility that freer trade may alter the firm-size distribution via international firm migration, i.e. offshoring; firms in the standard HFT models are assumed to produce in the nation in which they are ‘born.’

Our paper relaxes this assuming in presenting a model that allows the equilibrium firm-size distribution to be influenced by relocation as well as entry and exit. We show that this additional channel of adjustment implies that freer trade affects average productivity and the firm-size distribution in new ways. In particular, the HME pressure created by freer trade results in the most efficient firms moving from the small nation to the large nation. This in turn implies that freer trade has an extra-large productivity impact
in the large nation. In the small nation, the relocation always mitigates the positive productivity effect, and may, in some circumstances, result in an anti-productivity effect.

The paper also shows that relocation is a complex phenomenon in the presence of heterogeneous firms and fixed market-entry costs. When trade costs are high, the firm-migration (relocation) involves small-nation exporting firms (i.e. firms that sell locally and export both before and after moving). When trade costs get low enough, all exporting firms will have already moved to the big nation so subsequent relocation involves firms that sell only locally. In this phase, relocation reduces the number of varieties available in the small market. Regardless of the anti-productivity anti-variety effects of relocation, we show that the small nation always benefits by allowing relocation.

REFERENCES


