Learning and Knowledge Diffusion in a Global Economy*

Kunal Dasgupta †

First draft : November 2008

This draft : September 2011

Abstract

I develop a dynamic general equilibrium model to understand how multinationals affect host countries through knowledge diffusion. Workers in the model learn from their managers and knowledge diffusion takes place through worker mobility. Unlike in a model without learning, I present a novel mechanism through which an integrated equilibrium represents a Pareto improvement for the host country. I go on to explore other dynamic consequences of integration. The entry of multinationals makes the lifetime earning profiles of host country workers steeper. At the same time, if agents learn fast enough, integration creates unequal opportunities, thereby widening inequality. The ex-workers of foreign multinationals also found new firms which are, on average, larger than the largest firms under autarky.

KEYWORDS : Multinationals, knowledge diffusion, learning, worker mobility, pareto improvement, spin-offs.

JEL Code : F15, F23, F40.

---

*I am indebted to Esteban Rossi-Hansberg, without whose guidance and support this paper would never have been possible, and grateful to Gene Grossman and Marc Melitz for their valuable advice and encouragement. I would also like to thank Francesco Caselli, Thomas Chaney, Arnaud Costinot, Gilles Duranton, Cecilia Fieler, Luis Garicano, Stefania Garetto, Hugo Hopenhayn, Robert McMillan, Jordi Mondria, Peter Morrow, Marc Muendler, Xianwen Shi, Colin Stewart, Daniel Trefler, Vinayak Tripathi, Jonathan Vogel, and especially Alexander Monge-Naranjo and Stephen Redding for their helpful comments. I would finally like to thank the editor, Pol Antràs, and two anonymous referees for suggestions that have greatly improved the paper. Financial support from the International Economics Section, Princeton University is highly appreciated.

†University of Toronto, *E-mail: kunal.dasgupta@utoronto.ca*

1 Introduction

One of the most common arguments in favor of activities of multinational enterprises (MNEs) in developing countries is that knowledge/technology can thereby diffuse from MNEs to domestic firms. This belief is reflected in the widespread use of investment incentives by many host country governments to attract prospective MNE investors (Oman, 2000). Yet, how does this knowledge diffuse to domestic firms? And what are the consequences of such diffusion for the domestic economy? In an era of unprecedented globalization, answers to these questions have taken on a great deal of policy relevance.

The last two decades have seen the development of a large literature that examines whether and how knowledge diffuses from MNEs to domestic firms (See Lipsey and Sjöholm, 2005, for a survey). While the focus of most of these studies has been on finding evidence of knowledge spillovers, the precise mechanisms through which the spillovers occur are less well-understood. In particular, the mobility of workers from MNEs to domestic firms has not received enough attention. Casual observation, however, indicates that this could be an important channel for knowledge diffusion. For example, in the year 1981, four employees working in Patni Computer Systems Limited, a MNE, left to form (along with three others) Infosys Limited, an information and technology (IT) firm. Infosys has gone on to become one of the largest IT firms of not only India, but the entire world. More recently, two former Amazon.com employees founded Flipkart, which has turned into the largest online bookstore in India.

Infosys and Flipkart might be two of the more notable examples, but instances of employees learning from MNEs and then founding their own firms are widely prevalent. For example, Giarratana, Pagano and Torrisi (2004) look at the spin-offs from MNEs that were created in India after the country liberalized in 1991 and find that the founders brought a high level of technological expertise from the MNEs to the new firms. Similarly, many potential entrepreneurs in China perceive the MNEs as schools where they can train themselves; many of them leave to start their own business, once they have the required expertise.[1] Easterly (2001) discusses the Korean company Daewoo's decision to train the workers of a Bangladeshi textile firm in 1979; most of them left the parent firm during the 1980s to start their own garment export firms, laying the foundation for the \$2 billion dollars Bangladeshi garment industry.[2]

Until recently, the evidence on knowledge spillover from MNEs through worker mobility

[1] "China's people problem", *The Economist*, 14th April, 2005.

[2] Evidence on founders of spin-offs inheriting knowledge from their parents in the U.S. economy is provided by Klepper (2002) for the automobile industry, Klepper and Sleeper (2005) for the laser industry, and Filson and Franco (2006) for the rigid disk drive industry.

had been confined to anecdotes and surveys. This, however, is changing slowly. The availability of longitudinal linked employer-employee data sets for a number of countries has made it possible to track workers over time. These data sets allow the researcher to observe workers who switch from one firm to another, together with a host of worker and firm characteristics (See Abowd and Kramarz, 1999, for a discussion of some of these data sets). Taking advantage of such a data set from Brazil, Poole (2011) estimates MNE spillovers through the hiring of workers with MNE experience by domestic firms. Poole shows that the wages of workers in domestic establishments goes up in the presence of MNE-trained workers; at the aggregate level, such spillovers have the potential to generate wage increases to the tune of 0.6% of GDP.[3] Balsvik (2011) uses a similar dataset from Norway to test for knowledge spillovers from MNEs. He finds that workers with prior MNE experience contribute 20% more to productivity of domestic firms, relative to workers without such experience. Other studies to examine this particular channel for knowledge spillovers include Görg and Strobl (2005), Malchow-Møller, Markusen and Schjerning (2007), and Markusen and Trofimenko (2009).[4]

In light of the evidence presented above, in this paper, I develop a model that sheds light on the impact of MNE entry on welfare, wages and occupational choice. At the forefront of my analysis is diffusion of knowledge through learning and worker mobility.[5] Specifically, I develop a dynamic general equilibrium model with three key features: (1) every period, agents with heterogeneous knowledge choose their occupation and sort into production teams, (2) within teams, workers learn from managers, and (3) there is perfect mobility of workers within a country. Complementarity between the worker's and manager's knowledge in the production and learning technologies leads to positive assortative matching (or PAM), whereby more knowledgeable workers team up with more knowledgeable managers to produce and learn.[6]

Globalization or integration allows managers to hire workers from other countries - MNEs

[3] As Poole argues, this is probably a lower bound for spillovers since it ignores the direct increases in establishment productivity (a part of which may be captured by profits, rather than wages).

[4] Looking at a firm survey from Ghana, Görg and Strobl (2005) show that firms which are run by owners who worked for foreign MNEs in the same industry immediately prior to opening their firm, are more productive than other domestic firms. Using a matched employer-employee data set from Denmark, Malchow-Møller et al. (2007) show that both workers, as well as, self-employed agents experience an increase in earnings if they had previous experience in MNEs. Markusen and Trofimenko (2009) provide evidence that the use of foreign experts has a positive effect on both wages and value-added of workers in Colombian plants.

[5] The finding of Keller (2002) that knowledge spillovers are local, combined with the result that speaking the same language facilitates diffusion of knowledge, suggests that one possible channel for diffusion is human interaction. Building on this idea, Keller and Yeaple (2009*a*) build a model where transfer of knowledge requires face-to-face interaction between co-workers.

[6] In their study of variation in management practices across firms and countries, Bloom and Van Reenen (2010) find that that the education of both workers and managers is strongly correlated with management scores; more educated workers are hired by more educated managers.

are formed.[7] I consider a two country model, where the foreign country has relatively more knowledgeable agents (in a sense to be made precise shortly) compared to the home country. This results in lower wages under autarky in the home country, creating incentive for foreign managers to hire home workers. Following integration, new teams are formed as foreign managers try to leverage their superior knowledge with respect to home workers. In this setting, I identify two effects that determine home wages. First, integration increases the competition for workers, which tends to raise wages. This is the labor demand effect. This alone would make some of the incumbent managers worse off. But there is also a new effect: integration creates the possibility for the workers to be matched with very knowledgeable foreign managers. By working for these managers, workers can learn and earn more than under autarky. PAM, however, implies that the less knowledgeable workers can expect to work for the MNEs in the future only if they learn from the less knowledgeable home managers, and this creates a rent. Since learning is fully foreseen by the agents, the managers extract part of this rent by paying lower wages and thereby internalize the knowledge "spillover". This is the learning effect. If agents learn fast enough, this effect dominates and the wage schedule shifts down by enough to make the incumbent managers better off. The workers are better off too, because the increase in their continuation value outweighs the reduction in current wage. The above mechanism through which integration can lead to a Pareto improvement in the host country is new in the literature.[8]

The model is explored further in the numerical section. I present three sets of results concerning lifetime earnings profiles, inequality and spin-offs from MNEs. The learning dynamics of the model imply that the lifetime earnings profiles of agents born as workers are upward sloping. The slope of the earnings profile depends on the matching function. By improving the matches, integration increases the amount of knowledge that agents can acquire in each period, thereby raising the gradient of the lifetime earnings profiles.

Since the distribution of knowledge is endogenous, the model also allows me to talk meaningfully about inequality and how it changes following integration. A change in inequality in this model reflects not only a change in the agents' earnings but also a change in the entire knowledge distribution following integration, a feature that is absent in static models of trade or FDI. I show that if agents learn fast enough, integration can increase aggregate inequality.[9]

[7]In this paper, MNEs are synonymous with international production teams. I abstract from the issue related to the boundaries of international firms. For some recent papers which deal with this issue, see Antràs (2003), Grossman and Helpman (2003) and Antràs and Helpman (2004).

[8]MNEs extracting learning rents from workers also features in the model of Malchow-Møller, Markusen and Schjerning (2007). In their model, MNEs provide workers with learning opportunities while domestic firms do not. Since workers are homogeneous, in equilibrium they must be indifferent between joining both types of firms. This results in the workers in MNEs earning lower wages in the first period; the workers "pay" to acquire knowledge.

[9]See Goldberg and Pavcnik (March 2007) for evidence on episodes of globalization, which include both trade

Another novel prediction of the model concerns the spin-offs from the MNEs. Among the new managers entering the economy every period, a fraction consists of those who were previously working in other firms. I show that the biggest and most productive firms operating at home are, on average, run by foreign managers. Combined with PAM, this implies that founders of spin-offs who have previous MNE experience are, on average, more knowledgeable than those who do not. Furthermore, some of these spin-offs are larger and more productive than the largest firms under autarky.[10]

This paper adds to the large literature on international knowledge diffusion and, in particular, to the literature on knowledge diffusion through MNEs. MNE activity is not the only conduit of knowledge diffusion across countries, but it is widely perceived to be an important one.[11] Griffith et al. (2003) and Haskel et al. (2007) show that small but positive spillovers from MNEs exist in the United Kingdom. Economically and statistically significant positive spillovers have been found for MNEs operating in Lithuania (Javorcik, 2004) and in the United States (Keller and Yeaple, 2009*b*) (For a comprehensive survey, see Keller, 2004). The findings of positive spillovers have renewed interest in the channels through which knowledge diffuses from MNEs to domestic firms.[12]

In developing a model of knowledge diffusion through worker mobility, this paper combines two strands of the existing literature. Kremer and Maskin (2006) and Antràs, Garicano and Rossi-Hansberg (2006) use matching models to analyze how international team formation affects earnings and welfare. In Burstein and Monge-Naranjo (2009), international teams arise due to the interaction of country-embedded productivity (captured by the quality of institutions, infrastructure and productive inputs, among other factors) and firm-embedded productivity (captured by managerial know-how). Those models, however, are static in nature, with the

and FDI liberalization, being accompanied by increasing inequality. IMF (2007) reports that inward FDI has exacerbated inequality in developing countries.

[10]Recent years have witnessed the arrival of a new breed of emerging market multinationals in global business ("The Challengers", *The Economist*, 10th January, 2008).

[11]Other possible channels for knowledge diffusion include trade, ethnic communities and travel. Coe and Helpman (1995) and Coe, Helpman and Hoffmaister (1997) provide evidence of knowledge spillovers through imports, although Keller (1998) casts doubt on their findings. The evidence on knowledge spillovers through exports is mixed too; Clerides, Lach and Tybout (1998), Bernard and Jensen (1999) and Aw, Chung and Roberts (2000) do not find any evidence of learning through exporting, while Biesebroeck (2005) and Loecker (2007) do. Kerr (2008) shows that codified and tacit knowledge flows between U.S. ethnic research and entrepreneurial communities and their home countries. Examining the flow of business travelers between U.S.and seventy-four other countries, Keller and Hovhannisyan (2010) find that international business travel leads to the diffusion of knowledge.

[12]Of course, knowledge diffusion is not the only cause for positive MNE spillovers. Facing competition from MNEs, domestic firms could upgrade technology and thereby raise productivity. Productivity spillovers could also happen through linkages between MNEs and domestic firms (Rodríguez-Clare, 1996; Markusen and Venables, 1999).

distribution of knowledge in a country being given exogenously. I allow workers to learn from their managers, thereby extending the above mentioned models to a dynamic environment. By imposing quite weak restrictions on the production and learning technologies, I am able to characterize the equilibrium of both the closed as well as the integrated economies. In the process, I not only confirm some of their results in a more general setting, but generate new insights as to the nature of welfare gains, inequality and worker mobility.[13]

In related work, Monge-Naranjo (2008) and Beaudry and Francois (2010) develop dynamic general equilibrium models involving multinationals and on-the-job learning.[14] Those models do not involve matching however. A key feature of Monge-Naranjo is costly skill acquisition by individuals. By assuming learning by observing within firms, I instead focus on matching and endogenous occupational choice. In terms of the firm's problem, my paper is closer to Beaudry and Francois. In their model, skills are transferred from managers to workers through learning by observing. As in my model, firms offer a bundle of wage and skill to workers, who trade current wage with continuation value. The authors are primarily concerned, however, with analyzing why managers from developed countries may not relocate to developing countries, despite a shortage of skills in the latter. This is different from my paper in that I focus on the impact of MNEs on domestic skill accumulation and its consequences, assuming that there are sufficient incentives for foreign managers to relocate.[15] Despite these similarities, my model differs from the above mentioned papers in one key respect. In Monge-Naranjo, workers get only one chance to learn (since he has a two period overlapping generations model), while in Beaudry and Francois, workers immediately become managers upon learning. The combination of long-lived agents, PAM and stochastic learning implies that workers in my model can learn throughout their lives and two workers with identical initial conditions can end up with very different life-time earnings profiles. Besides generating richer dynamics at the individual level, this makes it possible for even the less knowledgeable domestic workers to work in MNEs in the future, thereby opening up a new channel for welfare gain.

The remainder of the paper proceeds as follows. Section 2 describes the model while in Section 3, I study the properties of a stationary equilibrium. In Section 4, I analyze how integration affects matches, output and welfare in the host country. I study a numerical example in

[13] Ramondo (2008) and Garetto (2008) develop quantitative models that compute static welfare gains associated with multinational production.

[14] Rodríguez-Clare (2007) develops a model of trade and diffusion where growth is caused by technological progress. Unlike my model, however, diffusion of ideas follows an exogenous process.

[15] Fosfuri, Motta and Rønde (2001) and Glass and Saggi (2002) consider game-theoretic treatment of knowledge diffusion in a partial equilibrium context, where spillovers result from a domestic firm trying to hire a MNE trained worker.

Section 5 and use it to further characterize the equilibrium. Section 6 concludes. All the proofs are in the Appendix.

2 The Model

My model introduces learning and dynamics to a framework that is similar to Antràs, Garicano and Rossi-Hansberg (2006).

2.1 Preferences and Endowments

There is a continuum of heterogeneous agents with different levels of knowledge. Knowledge is embodied in an agent, but can be acquired through interactions, i.e., an agent can learn from others. One can think of knowledge as some composite of different attributes that affects an agent's productive capability.[16] A newborn agent draws his knowledge from an exogenously given distribution $\Phi(k)$ with support $[\underline{k}, \overline{k}]$. I assume that $\Phi(k)$ is continuous, with $\phi(k)$ being the corresponding density. Agents die every period with a constant probability δ and are replaced by newborns such that the population is constant. δ also acts as a discount factor. The actual distribution of knowledge at time t is denoted by $\Psi_t(k)$, with $\psi_t(k)$ being the corresponding density. Agents are risk neutral. Since the size of the population plays no role in the analysis, it is normalized to 1. Time is discrete.

2.2 Production

Firms produce a single, non-storable good. A firm is comprised of a manager and production workers. Workers do routine jobs and each worker combines with the manager to produce $f(y)$ units of output, where y is the knowledge of the manager. Thus, "f(y) captures the indivisibility of management-type decisions and implies a scale economy because it improves productivity of all the workers in the firm, irrespective of their numbers" as in Rosen (1982) p. 314. Notice that the productivity of workers in a firm run by a manager with knowledge y is simply $f(y)$. The manager pays wages to the workers and is the residual claimant on the output.[17]

There is a technological restriction to the number of workers a manager can hire. The span of control of a manager depends only on the knowledge of the workers he hires. This span of

[16] For a trade model where agents have two attributes, see Ohnsorge and Trefler (2007).

[17] Here, as in Monge-Naranjo (2008), I assume that there is no difference between the managers and entrepreneurs. For a model which makes this distinction, see Holmes and Schmitz (1990).

control is given by $\tilde{n}(\min_i[x_i])$, where x_i is the knowledge of the i-th worker hired by a manager. This is similar to the O-ring production function (Kremer, 1993), whereby the workers in a firm are as good as their least skilled counterpart. It is easy to see that given such a technology, the manager will hire only one type of worker in equilibrium. I denote the equilibrium span-of-control by $n(x)$ and use this notation in the rest of the paper. (For a micro-foundation of such a technology, see Garicano, 2000). Total output of a firm is then given by

$$q = f(y)n(x) \tag{1}$$

I make the following assumptions regarding the production technology:

ASSUMPTION 1a : *f is continuous, strictly increasing and weakly convex in y; n is continuous, strictly increasing and strictly concave in x. Furthermore, $\frac{\partial}{\partial y}[\frac{f'(y)}{f(y)}] \leq 0$.*

ASSUMPTION 1b : $\frac{f'(\underline{k})}{f(\bar{k})} > \frac{n'(\underline{k})}{n(\underline{k})}$.

Assumptions 1a and 1b together imply that the output elasticity of the managerial task is greater than that of the production task at any knowledge level. As we shall see, this results in the more knowledgeable agents becoming managers in equilibrium.[18] Assumption 1b also says that for a given knowledge distribution, there should be sufficient asymmetry between the manager and the worker's contribution to output. This is a technical condition required for the existence and uniqueness of equilibrium.[19]

2.3 Learning

Agents also learn in firms. Since the seminal work of Gary Becker (Becker, 1962), economists have been studying on-the-job training. In this paper, I abstract from formal training provided by firms and instead focus on the knowledge that workers acquire while producing. I follow Jovanovic and Rob (1989) in defining the learning technology. Specifically, within each firm,

[18]To see this, note that the output elasticity of the manager's knowledge is $\frac{\partial q}{\partial y}\frac{y}{q} = \frac{f'(y)y}{f(y)}$, while that of the worker is $\frac{n'(x)x}{n(x)}$. Assumption 1a implies that $\frac{f'(y)}{f(y)}$ and $\frac{n'(x)}{n(x)}$ are non-increasing in y and x respectively. From Assumption 1b, we have, $\frac{f'(\bar{k})}{f(\bar{k})} > \frac{f'(\bar{k})}{f(\bar{k})}\frac{f'(\underline{k})}{f'(\bar{k})} > \frac{n'(\underline{k})}{n(\underline{k})}$. This inequality, combined with the previous observation, implies that $\frac{f'(k)k}{f(k)} > \frac{n'(k)k}{n(k)}$ $\forall k$.

[19]The assumption that labor is the only factor of production is without loss of generality. Capital can be easily introduced in the model. The cost of capital has three components - sunk, fixed and variable costs. In the absence of uncertainty in production and credit market imperfections, the first two do not really have any effect; so I just normalize those to zero. As for variable capital requirement, we can think of it as being subsumed in $f(y)$.

a worker learns from the manager.[20] Learning is stochastic and depends both on the knowledge of the manager and the worker. The randomness in learning does not necessarily reflect any randomness in the knowledge transfer process but rather, is a simple way of modeling the heterogeneous capacity to absorb knowledge. A worker with knowledge x at time t has knowledge x' at time $t+1$. The learning distribution is given by $L(x'|x, y)$, $x' \in [x, y]$. For all $h(x')$ increasing in x', I make the following assumptions about the learning technology:

ASSUMPTION 2a : $\frac{\partial}{\partial x} \int h(x') dL(x'|x, y) > 0.$

ASSUMPTION 2b : $\frac{\partial}{\partial y} \int h(x') dL(x'|x, y) > 0.$

ASSUMPTION 2c : $\frac{\partial^2}{\partial x \partial y} \int h(x') dL(x'|x, y) \geq 0.$

The first two conditions are the familiar ones for first-order stochastic dominance. These conditions imply that expected learning is increasing in the knowledge of both the workers and the managers. The third assumption says that there is (weak) complementarity in learning. Although the learning technology is taken as given, it can be derived from a micro-founded model where learning requires effort and workers optimally choose how much effort to allocate (see the online Appendix).

3 Equilibrium

Agents are price-takers. There are two prices in the economy. First, the managers hire workers and pay a price for their marginal product. Second, the workers learn from the managers and pay a price for the acquired knowledge. It is inconsequential whether there are two transactions within the firm or whether the managers simply pay the wage net of the rent (Rosen, 1972). What matters is the net payment to workers $w_t(k)$; let us call this the wage. The absence of aggregate uncertainty in the model, combined with a large number of agents, implies that the evolution of the knowledge distribution is deterministic. Since in equilibrium, $w_t(k)$ is a function of only the knowledge distribution, the evolution of $w_t(k)$ is also deterministic. Therefore, given an initial distribution of knowledge, $\Psi_0(k)$, a competitive equilibrium is characterized by a deterministic sequence $\{\Psi_t(k), w_t(k)\}$.

At this point, I impose more structure on the form of the equilibrium. As long as δ is not

[20] Unlike Jovanovic and Rob (1989), learning is one-sided. Assuming that managers also learn from workers could be an interesting extension and would be one channel through which growth can be introduced in this model.

too small (as I show in the next section, this implies that agents do not learn fast enough), this is without loss of generality as I prove later in Proposition 1. In particular, I assume that the equilibrium exhibits positive assortative matching or PAM, whereby more knowledgeable workers match with more knowledgeable managers; if $m_t(k)$ represents the equilibrium matching function (i.e., $m_t(k)$ is the knowledge of the manager who hires a worker with knowledge k), then $m_t'(k) > 0$. Intuitively, positive assortative matching is a consequence of there being strong complementarity in production and weak complementarity in learning. I also assume that at every period t, the equilibrium is characterized by a threshold level of knowledge k_t^* such that all agents with knowledge less than k_t^* are workers, while those with knowledge above k_t^* are managers.[21]

A convenient property of this equilibrium is its block recursive structure - matches can be determined completely once we know the knowledge distribution. One does not need to know the wage schedule in order to determine the matches; rather, once the matches are determined, wages adjust so as to support the matches that emerge. Of course, this does not mean that the way agents match does not depend on wages. In this economy, wages (and profits) not only determine the remuneration of the agents but they also play an allocative role (Sattinger (1993)). But for the purpose of solving the model, the matching function can be derived without any information on the wage function.

Given a sequence of wage functions $\{w_t\}_{t=0}^{\infty}$, the manager's problem is then defined recursively as

$$V_M(y, w_t) = \max_x f(y)n(x) - w_t(x)n(x) + (1-\delta)\max[V_W(y, w_{t+1}), V_M(y, w_{t+1})]. \tag{2}$$

where $V_W(y, w_t)$ is the value function of an agent with knowledge y, if he chooses to be a worker while $V_M(y, w_t)$ is the value function if, instead, he chooses to be a manager. The value of a manager depends on the current distribution $\Psi_t(k)$ through the net wage schedule w_t, where the dependence of wage on k is understood. The second term on the right allows for the possibility that an agent, who is a manager at time t, might choose to be a worker at time $t+1$.

[21]The key assumption driving this result is 1b. There are two opposing forces that determine matching in this economy - complementarity, which encourages agents with similar levels of knowledge to match together and asymmetry, which goes in the opposite direction (Kremer and Maskin, 1996). Assumption 1b, combined with a large enough δ (as I formally show in Proposition 1) implies that the latter effect dominates, whereby managers and workers with different levels of knowledge match.

For a worker with knowledge x, $V_W(x, w_t)$ is given by

$$V_W(x, w_t) = w_t(x) + (1-\delta) \int_{x}^{m_t(x)} \max[V_W(x', w_{t+1}), V_M(x', w_{t+1})] dL(x'|x, m_t(x)). \quad (3)$$

The term within the integral denotes the expected value of the worker if he works for a manager with knowledge $m_t(x)$. Depending on how much he learns, the worker might become a manager or continue as a worker at $t+1$. As before, this decision depends not only on the worker's own knowledge but also on the wage function at $t+1$. Facing $V_M(.)$ and $V_W(.)$, an agent picks the maximum.

In each period, the supply and demand for workers (or managers) must be equal. Labor market clearing requires that

$$\int_{\underline{k}}^{k} d\Psi_t(s) = \int_{k_t^*}^{m_t(k)} n(m_t^{-1}(s)) d\Psi_t(s) \quad \forall k \le k_t^* \quad (4)$$

The left hand side denotes the supply of workers in the interval $[\underline{k}, k]$, while the right hand side denotes the demand for workers coming from managers in the interval $[k_t^*, m_t(k)]$. Measure consistency requires that these two values be equal *for every* k.[22] This follows from PAM, since the workers hired by managers with knowledge in $[k_t^*, m_t(k)]$, must have knowledge in $[\underline{k}, k]$. Differentiating equation (4) with respect to k yields

$$m_t'(k) = \frac{\psi_t(k)}{n(k)\psi_t(m_t(k))} \quad (5)$$

The above differential equation, along with the boundary conditions $m_t(\underline{k}) = k_t^*$ and $m_t(k_t^*) = \overline{k}$, allows us to solve for the matching function. I show in the Appendix that given $\Psi_t(k)$, the threshold k_t^* and consequently the matching function is uniquely determined.

Over time, workers' knowledge increases due to learning. But agents also die every period with probability δ and are replaced by newborns who draw knowledge from the exogenous distribution $\Phi(k)$. Birth, learning and death implies a rule for the evolution of the knowledge distribution $\Psi_t(k)$:

[22] Note that the labor market-clearing condition is not standard. It is not enough for the above equation to be satisfied only for $k = k_t^*$, as would be the case in a standard Lucas span-of-control model without matching.

$$\Psi_{t+1}(k) = \delta\Phi(k) + (1-\delta)\int_{\underline{k}}^{k}\int_{s}^{k} dL(s'|s, m_t(s))d\Psi_t(s) \text{for all } k \in [\underline{k}, \overline{k}] \tag{6}$$

The first term on the right-hand side denotes the fraction of agents who are born in period $t+1$ with knowledge less than k. The second term denotes the agents who remain below k in period $t+1$, despite learning from their managers in period t.[23] Equation (6) implies that Ψ_{t+1} is determined by how individuals acquire knowledge in period t, and the acquisition of knowledge by individuals is determined only by who they match with at time t, which in turn depends only on Ψ_t. Therefore, Ψ_{t+1} is a function of Ψ_t. We seek a fixed point of Ψ_t, i.e., an invariant knowledge distribution Ψ^*.

Let us know turn towards the determination of wage. It might seem natural to write the wage as a function of the knowledge of both the worker and the manager, given that (a) the same manager can produce different levels of output by hiring different types of workers and (b) the same worker can acquire different levels of knowledge by working for different managers. Condition (a) suggests that the price of labor should be specific to a worker-manager pair while (b) suggests that the same should be true for the price of knowledge. In order to understand why $w_t(x)$ only has the knowledge of the worker as its argument, let us look at the underlying mechanism that determines the wage function.

In this economy, a manager with knowledge y offers a wage schedule $w_t(x, y)$ such that y *is indifferent between hiring any* x. But along with the wage, a manager also offers a continuation value (from learning) to each worker. Let us denote this by $C(x, y)$. Therefore, a manager offers a bundle of wage and continuation value, and a worker with knowledge x picks the manager with knowledge $m_t(x)$ where $m_t(x) = arg \max_y [w_t(x, y) + C(x, y)]$.[24] In equilibrium, the wage earned by x is the wage offered to him by $m_t(x)$, i.e., $w_t(x, m_t(x)) = w_t(x)$.[25]

This completes the characterization of the competitive equilibrium. The following proposi-

[23]Note that agents with knowledge $s > k_t^*$ are managers and hence, $\int_s^k dL(s'|s, m_t(s)) = 1 \forall k$. The managers match with workers who have less knowledge and accordingly, they do not learn.

[24]This is similar to Boyd and Prescott (1987), where the old workers in a firm offer the young a package of current consumption and future expertise.

[25]To see why this is an equilibrium, consider an arbitrary worker with knowledge x^*. By construction, $w_t(x^*)$ maximizes the present discounted value of x^*. Therefore, deviation is not profitable. Let the knowledge of the manager who offers this wage be denoted by y^*. To see why y^* is also maximizing his profit $\pi(y^*)$, observe that

$$\begin{aligned}
\pi_t(y^*) &= f(y^*)n(x^*) - w_t(x^*)n(x^*) \\
&= f(y^*)n(x^*) - w_t(x^*, y^*)n(x^*) \\
&= f(y^*)n(x) - w_t(x, y^*)n(x) \quad \forall x
\end{aligned}$$

tion provides for the existence and uniqueness of the equilibrium.

Proposition 1. *There exists a δ^*, such that $\forall\delta > \delta^*$, a threshold equilibrium exists for every t and it is unique. The equilibrium exhibits positive assortative matching. Finally, a unique, invariant knowledge distribution Ψ^* exists and any initial distribution Ψ_0 weakly converges to Ψ^*.*

The proof of the existence and uniqueness of the threshold equilibrium consists of a number of steps. First, I derive properties of the equilibrium wage function from the assumptions about the production and learning technologies. Recall that the equilibrium wage is net of the learning rents that workers pay to the managers. The derived properties are then used to show that there is (a) positive assortative matching and that (b) the value functions are increasing in knowledge. In the next step, I prove the existence of a threshold equilibrium under no-learning ($\delta = 1$). By continuity arguments, it is shown that the equilibrium exists even under some learning ($\delta < 1$). Finally, the uniqueness of the threshold equilibrium is established using the same conditions that guarantee existence. The readers are referred to the online Appendix for the proof.

In the long run, the knowledge distribution converges to Ψ^*, with threshold k^*. Agents who are born with knowledge above k^* become managers instantaneously. Since managers do not learn, these agents are stuck with the level of knowledge they are born with. Of course, during the transition, we could observe an agent switching from being a manager to a worker, and back to a manager. On the other hand, agents who are born with knowledge below k^* start their lives as workers. These agents learn every period and move up, until they eventually cross the threshold and become managers themselves. For the remainder of the paper, I shall restrict my attention to the stationary equilibrium of the model.

4 Analytical Results

In this section, I present some analytical results of my model. To simplify exposition, I make the following assumption about the learning technology:

ASSUMPTION 2d : *If, in period t, a worker with knowledge x works for a manager with knowledge y, then in period $t+1$, the worker's knowledge x' could be x with probability θ and y with probability $1-\theta$.*

where the last line follows from the definition of $w_t(x, y)$. Now pick any $\hat{x} \neq x^*$. Since $m_t'(x) > 0$, $y^* \neq m_t(\hat{x})$, i.e., we must have $w_t(\hat{x}, y^*) + C(\hat{x}, y^*) < w_t(\hat{x}) + C(\hat{x}, m_t(\hat{x}))$. Hence, in order to hire $\hat{x}$, y^* has to offer a wage $\bar{w}$, where $\bar{w} > w_t(x, y^*)$. But then $\pi_t(y^*)$ would decrease. Therefore, y^* maximizes his profit by hiring x^*. Since, x^* was chosen arbitrarily, the result follows.

Thus, learning is an all-or-nothing proposition for the worker. In the next section, I relax this assumption and work with a more general learning technology. Notice that, in spite of the learning distribution having just two points, it still satisfies Assumptions 2a, 2b and 2c.[26]

4.1 Autarky

I begin by examining the equilibrium under autarky. Recall that the density function for the newborn distribution is given by $\phi(k)$. The learning technology, along with the newborn distribution, implicitly defines the invariant distribution and allows me to solve for the threshold k^*.

Proposition 2. *k^* is defined implicitly by the following equation*

$$\int_{\underline{k}}^{k^*} \frac{\phi(k)}{n(k)} dk = \int_{k^*}^{\overline{k}} \phi(k) dk + (1-\theta)(\frac{1-\delta}{\delta}).$$

Moreover, k^* has the following properties :

$$\frac{\partial k^*}{\partial \theta} < 0; \frac{\partial k^*}{\partial \delta} < 0.$$

Proposition 2 sheds light on how the distribution changes as the rate of learning increases. δ, being the probability of death in a period, proxies for the length of a time period. A lower δ, holding θ unchanged, implies that agents are acquiring the same expected knowledge over a smaller interval of time. On the other hand, a lower θ, holding δ unchanged, implies that agents are acquiring more expected knowledge over the same interval of time. Both these cases translate into faster learning for the agents. An increase in the rate of learning makes the knowledge distribution negatively-skewed, as more and more mass shifts to the upper tail. Consequently, labor market-clearing requires that the threshold shift to the right.

Recall that a worker with knowledge k produces $f(m(k))$ units of output. Hence, total output produced in this economy is given by

$$Y = \int_{\underline{k}}^{k^*} f(m(k)) d\Psi(k) \tag{7}$$

[26] To see this, note that for any $h(.), h' > 0$, the expected value of $h(x')$ is $\theta h(x) + (1-\theta)h(y)$. Therefore, $\frac{\partial E[h(x')]}{\partial x} = \theta h'(x)$ and $\frac{\partial E[h(x')]}{\partial y} = (1-\theta)h'(y)$. Furthermore, $\frac{\partial^2 E[h(x')]}{\partial x \partial y} = 0$.

Total welfare is given by

$$W = \int_{\underline{k}}^{k^*} V_W(k)d\Psi(k) + \int_{k^*}^{\overline{k}} V_M(k)d\Psi(k) \tag{8}$$

In this model, individual welfare equals the present value of consumption (or income, since the good is non-storable) because agents are risk-neutral.

4.2 Integration

Integration, in the context of my model, means that managers from one country can hire workers in another country, i.e., integration leads to the creation of MNEs. The managerial input is rival and as a result, managers cannot operate plants in both countries. Whether managers travel from the source-country to the host-country or not, however, is irrelevant. The motive behind the formation of MNEs is exploiting differences in factor prices.[27] In this paper, I focus on full integration, i.e., I assume that MNEs are formed costlessly. In particular, I assume away any cost that might be associated with opening a plant in another country. I do acknowledge that these costs are important, but the introduction of such costs introduces complexities into the model thereby diverting our attention from the main issue at hand, knowledge diffusion.

Let us introduce some notation. Define the subscripts $i = \{A, I\}$, $j = \{H, F\}$, where A and I stand for autarky and integration respectively, while H and F stand for home and foreign respectively. The home newborn distribution is denoted by $\Phi_H(k)$ with support $[\underline{k}, \overline{k}_H]$, while the foreign newborn distribution is $\Phi_F(k)$, with $[\underline{k}, \overline{k}_F]$ being the corresponding support. I assume that $\overline{k}_F > \overline{k}_H$ and that $\Phi_F(k)$ first-order stochastic dominates $\Phi_H(k)$. The latter assumption, which reflects the relative abundance of more knowledgeable agents in the foreign country, is key for the welfare results shown later.[28] The steady-state knowledge distributions are indexed by i and j. So, for example, $\Psi_{A,H}(k)$ is the home steady-state knowledge distribution under autarky. I also assume that the two countries have the same population. With integration, the fundamental change is in the distribution of newborns, which is given by

[27] This motive for establishing subsidiaries in other countries is the same as in Helpman (1984).

[28] In this paper, I want to highlight the dynamic gains to workers in developing countries arising from the access to better managers from developed countries. As Bloom and Van Reenen (2010) point out, better managers (or better management practices, to be precise) are relatively more abundant in the U.S. compared to developing countries like India or China.

$$\Phi_I(k) = \begin{cases} \frac{1}{2}\Phi_H(k) + \frac{1}{2}\Phi_F(k) & \text{for } k \in [\underline{k}, \overline{k}_H] \\ \frac{1}{2} + \frac{1}{2}\Phi_F(k) & \text{for } k \in [\overline{k}_H, \overline{k}_F] \end{cases} \tag{9}$$

$\Phi_I(k)$, combined with the learning technology, determines the integrated knowledge distribution $\Psi_I(k)$. The new threshold, k_I^*, would typically be different from $k_{A,j}^*$, the autarky threshold in country j. Equation (9), along with the assumption that $\Phi_F(k)$ first-order stochastic dominates $\Phi_H(k)$, implies that $\Phi_I(k)$ first-order stochastic dominates $\Phi_H(k)$. In the benchmark case of no-learning, the knowledge distributions in both the countries coincide with the newborn distributions. Consequently, under no-learning, $k_I^* > k_{A,H}^*$ (See Lemma 1 in the Appendix for a proof). With learning, however, the knowledge distributions are no longer exogenous. Still, one can derive a relation between $k_{A,H}^*$ and k_I^*, as shown in the following proposition.

Proposition 3. *$k_I^* > k_{A,H}^*$, where k_I^* is defined implicitly by the following equation*

$$\int_{\underline{k}}^{k_I^*} \frac{\phi_H(k) + \phi_F(k)}{2n(k)} dk = \frac{1}{2} \int_{k_I^*}^{\overline{k}_F} (\phi_H(k) + \phi_F(k))dk + (1-\theta)(\frac{1-\delta}{\delta}).$$

The range of knowledge for the home workers expands under integration. The agents with knowledge in $[k_{A,H}^*, k_I^*]$ switch from being managers to workers. The entry of highly knowledgeable foreign managers raises the opportunity cost of being a manager for a home agent. This is not only due to (possibly) higher wages paid by the MNEs but also due to the better learning opportunities provided by the MNEs. An incumbent home manager weighs the cost of becoming a worker for a MNE (forgone current profits) against the benefit (higher expected profits in the future). For the managers in $[k_{A,H}^*, k_I^*]$, benefits outweigh costs and consequently they switch.

Although, in the new steady state, there are home agents with knowledge in $[\overline{k}_H, \overline{k}_F]$, but at the time of birth, these agents still draw their knowledge from the home newborn distribution, which does not change with integration. Home agents attain knowledge in $[\overline{k}_H, \overline{k}_F]$ *through learning, not through birth.* Thus globalization creates a class of firms which are bigger and more productive than the "best" firms under autarky. This result relies on the assumption that the upper bounds of the home and foreign newborn distributions are different.

Proposition 3 indicates the direction of change for the threshold, but it says nothing about its magnitude. In particular, the following two scenarios are possible:

Case I ($k_I^* > \overline{k}_H$) : In this case, every agent born in the home country starts his life as a worker.

The support of $\Psi_{I,H}(k)$ is $[\underline{k}, m(\overline{k}_H)]$, despite the fact that, the home newborn distribution $\Phi_H(k)$ still has the smaller support.[29] Though theoretically an interesting case, this situation is quite extreme because it implies that integration results in the destruction of all incumbent firms (managers), who are replaced by a new class of bigger and more productive firms.

Case II ($k_I^* < \overline{k}_H$) : In this case, the support of $\Psi_{I,H}(k)$ is $[\underline{k}, \overline{k}_F]$. This case is characterized by the birth of a new class of home firms (with knowledge in $[\overline{k}_H, \overline{k}_F]$), who are on par with the foreign MNEs in terms of size and productivity. But unlike Case I, a set of incumbent home managers with knowledge in $[k_I^*, \overline{k}_H]$ continues to operate in the integrated economy.

Whether we are in Case I or Case II depends on the parameters of the model. As long as $\overline{k}_F$ is not too different from $\overline{k}_H$, there will be some incumbent managers operating in the home country under integration.[30] Intuitively, a large gap between $\overline{k}_F$ and $\overline{k}_H$ implies that following integration, the home agents have an opportunity to work for very knowledgeable foreign managers. This is also true for every incumbent home manager, who would rather work in foreign MNEs, learn and become much better managers in the future than remain managers with low levels of knowledge. Irrespective of which case we are in, integration affects the matching of agents.

Proposition 4. *A positive measure of home workers have a better match in the integrated equilibrium compared to autarky.*

An immediate implication of Proposition 3 is that $m_I(\underline{k}) > m_{A,H}(\underline{k})$, where $m_{A,H}(.)$ and $m_I(.)$ are the matching functions under autarky and integration respectively. To see this, note that $m_{A,H}(\underline{k}) = k_{A,H}^*$ and $m_I(\underline{k}) = k_I^*$. Proposition 3 then gives the result. Hence, the least knowledgeable worker in the home country, and by continuity, a set of less knowledgeable workers, is matched with better managers. On the other hand, some of the home managers are now matched with less able workers. Since the output of firms depends positively on workers' knowledge, the output of some of the home firms under integration are necessarily lower than in autarky.

Corollary 1. *Under integration, the output of a positive measure of home firms goes down.*

[29] It is not the case that every home agent is a worker. There are home managers in $[k_I^*, m(\overline{k}_H)]$. This, however, means that the home managers in the integrated economy have knowledge greater than $\overline{k}_H$.

[30] Note that for a given $\overline{k}_H$, there exists a k' such that $\overline{k}_F < k'$ implies that $k_I^* < \overline{k}_H$. This follows from the result that k_I^* is monotone increasing in $\overline{k}_F$, and $k_I^* < \overline{k}_H$ when $\overline{k}_F = \overline{k}_H$.

According to Aitken and Harrison (1999), FDI was accompanied by a decline in the productivity of domestically owned firms in Venezuela. This decline, the authors report, is due to a contraction in output of domestic firms due to the "market stealing effects" of foreign firms. In this paper, the output reduction is a natural consequence of complementarity in production and learning.[31] Under integration, the most knowledgeable workers are hired by the MNEs leaving less knowledgeable workers to work for incumbent home firms. I call this the "worker stealing effect". Despite this effect, the home managers could actually be better off if learning is fast enough, as we show later.

Total Home output (GDP) produced in the integrated equilibrium is given by

$$Y = \int_{\underline{k}}^{k_I^*} f(m_I(k)) d\Psi_{I,H}(k) P_H \tag{10}$$

This is different from Gross National Income (GNI), which is given by

$$GNI = \int_{\underline{k}}^{k_I^*} w_I(k) d\Psi_{I,H}(k) P_H + \int_{k_I^*}^{\overline{k}_F} \pi_I(k) d\Psi_{I,H}(k) P_H \tag{11}$$

The difference between the two arises because in an integrated equilibrium, a part of the home output goes to the foreign country as profits of foreign MNEs, while some of the home firms may become multinationals and earn profits from their operations in the foreign country. Finally, aggregate welfare in the home country is given by

$$W = \int_{\underline{k}}^{k^*} V_W(k) d\Psi_{I,H}(k) P_H + \int_{k^*}^{\overline{k}_F} V_M(k) d\Psi_{I,H}(k) P_H \tag{12}$$

To sum up, with integration, the threshold of the knowledge distribution shifts to the right. This necessarily means that some of the home workers are hired by more knowledgeable managers. These workers also learn more compared to autarky. At the same time, some of the incumbent firms suffer a decline in output.

[31] In this model, since the output produced by a worker depends only on the knowledge of the manager he is matched with, the productivity of a firm, as measured by the value-added per worker, does not change.

4.3 Change in Welfare

Integration changes individual, as well as, aggregate welfare in the home country. I focus my attention on the case where there are surviving home managers, i.e., Case II.[32] In order to understand how learning affects welfare, first let us look at the benchmark case of no learning. This is similar to the static framework presented in Antràs et al. (2006). A key result that emerges from Antràs et al. (2006) is that integration raises aggregate consumption, and with risk-neutral agents, the aggregate welfare of the home economy. What about individual welfare? In the previous section, we showed that the output produced by the less knowledgeable home managers goes down under integration (This is true for both the learning and the no-learning case). The actual change in profits and welfare, however, depends on the wages they pay, which would be different from those under autarky. Of course, as wages change, the welfare of the workers change too.

Proposition 5. *In the absence of learning, an integrated steady-state equilibrium with incumbent home firms can never represent a Pareto improvement relative to the autarky steady-state equilibrium in the home country.*

In the absence of learning, integration creates winners and losers. The identity of the winners and losers, though, depends on the specific parameter values. If we think of workers and managers as separate factors of production, Proposition 5 essentially gives us a Heckscher-Ohlin like result. But, does Proposition 5 continue to hold when we introduce learning? In order to prove otherwise, I have to show that every agent in the home economy is strictly better off under integration. Since output equals revenue in this model, Corollary 1 implies that some of the home managers earn lower revenue compared to autarky. Hence, for these managers to be better-off under integration, the wage bill has to go down more than revenue.

In this model, there are two forces that determine wages. First, there is a *labor demand effect*. The entry of foreign managers increases the demand for home workers. At the same time, integration increases competition faced by the home workers from their foreign counterparts. As shown by Antràs et al. (2006), (1) if the two countries are not too similar and, (2) if the span of control is not too small, the labor demand effect raises the wages of all home workers.

Second, there is a *learning effect*. A worker, in this model, can be hired by any manager with a positive probability. Working for a more knowledgeable manager means higher expected

[32] The reason for this is the following : If $\overline{k}_F$ is very different from $\overline{k}_H$, then irrespective of whether agents learn or not, every home agent is better off working for the more knowledgeable foreign managers. Thus we get Pareto improvement, but the home firms disappear completely. Apart from being an extreme case, this is also the only scenario where home agents experience a Pareto improvement in Antràs et al. (2006).

learning and consequently, higher earnings. Hence, the entry of highly knowledgeable foreign managers raises the continuation value of the home workers. PAM, however, implies that the most knowledgeable managers hire only the most knowledgeable workers. Therefore, the less knowledgeable workers can work for the MNEs only if they learn and acquire enough knowledge from their current managers, some of whom are the incumbent home managers. A positive value of learning implies that workers are willing to pay in order to learn. Thus learning creates a rent. This allows the managers to compress the wage. The workers accept this wage reduction because they expect to be compensated in the future. So the learning effect tends to lower the wage schedule.

The final impact on wages depends on the relative strengths of the two effects. The following Proposition shows that under certain conditions, it is possible for the home economy to realize Pareto gains.

Proposition 6. *The integrated steady-state equilibrium is a Pareto improvement for the Home country if*

$$(1-\theta)(\frac{1-\delta}{\delta}) > \chi.$$

χ is a function of technology and the bounds of the home newborn distribution.[33] Let us denote by Ω, the set of pairs of θ and δ that satisfy the above condition. The left-hand side of the above expression is positive by definition. For Ω to be non-empty, the value of χ must be positive but small enough. It can be shown that χ is an decreasing function of $\frac{f'(k)}{f(k)} - \frac{n'(k)}{n(k)}$, i.e., the degree of asymmetry between the manager's and the worker's contribution to output. Intuitively, the greater is this asymmetry, the greater is the increase in the worker's earning when he becomes a manager; and the greater is the wage cut that the worker is willing to accept in order to learn.

Proposition 6 also sheds light on how welfare changes as the rate of learning changes. Assuming that the right-hand side of the expression in Proposition 6 is positive, the inequality is not satisfied for high enough δ (or θ). In the limiting case of no-learning, $\delta = 1$ (or $\theta = 1$), the left-hand side is equal to zero. As δ (or θ) falls, the left-hand side starts to increase and at some point, exceeds the right-hand side. According to Corollary 1, some of the incumbent firms produce less under integration relative to autarky. For these firms to be better off, they must be paying lower wages to the workers. Corollary 2 follows naturally.

[33] $\chi = \frac{2f(\overline{k}_H)[n(\overline{k}_H)-n(\underline{k})]+n(\underline{k})[\mu(\overline{k}_H)f(\overline{k}_H)-\mu(\underline{k})f(\underline{k})]}{\mu(\underline{k})[f(\overline{k}_F)+f(\underline{k})]n(\underline{k})-\mu(\overline{k}_H)[n(\underline{k})+n(\overline{k}_H)]f(\overline{k}_H)}$, where $\mu(\underline{k}) = \frac{n(\underline{k})}{1+n(\underline{k})}$ and $\mu(\overline{k}_H) = \frac{n(\overline{k}_H)}{1+n(\overline{k}_H)}$.

Corollary 2. *If all home agents gain from integration, some of the home workers must earn a lower wage compared to autarky.*

Pareto gains are important for political economy reasons. Gaining support for trade or FDI liberalization becomes much easier if every agent gains from integration.[34] Notice, however, that proposition 6 compares welfare across two steady-states. It is possible that even though the new steady-state generates a Pareto improvement, the policy lowers the welfare of some agents relative to autarky. Numerically, however, it can be shown that for a wide range of parameter values, integration generates a Pareto improvement not only across steady-states, but also at the instant the policy is enacted.

4.4 Remarks

Few remarks are in order. First, the rent-sharing in the model takes place without any bargaining between the workers and the managers. Rather, the division is determined by labor market clearing. Two assumptions of the model are key for this result - fixed supply of agents and endogenous occupational choice. To see this, first consider a model without learning. In equilibrium, the agent with the threshold knowledge k^* earns the same wage and profit and accordingly, has the same value from both occupations. Now assume that agents can learn. Can we have a situation where the workers extract the entire rent from learning? No, because then the value from being a worker will be higher than the value from being a manager for the threshold agent; although the current wage and profit are the same, the continuation value of a worker will be higher than the continuation value of a manager because the worker, almost surely, will be a manager with knowledge greater than k^* in the next period. This, however, cannot be an equilibrium since the agent with knowledge k^*, and by continuity, agents with knowledge just above k^*, will strictly prefer to be workers. Given the fixed supply of agents, this will create an excess supply of workers. For the labor market to clear, wages would have to go down. The amount by which wage falls for each worker is the rent that he pays to the manager. A similar argument shows that managers cannot extract the entire learning rent either.

Second, if a MNE has more knowledge than an incumbent home firm, PAM implies that the workers in the MNE are more knowledgeable than the ones in the home firm. PAM also implies that after working in the (more knowledgeable) MNE, a worker never works for the (less knowledgeable) home manager. Therefore, there is no flow of knowledge from the MNE to the

[34] It is not possible to derive policy prescription from such a stylized model. The model does, however, suggest a channel that could affect welfare calculations and hence calls for more analysis.

home firm. Despite this, the incumbent firm could be better off if learning is fast enough.[35] Of course, some of the former MNE workers set up their own firms and these managers directly benefit from the superior knowledge of MNEs. This effect is similar to Monge-Naranjo (2008) where the transfer of skills from MNEs materializes in a new sector of firms, not in the pre-existing sector of firms.

Third, knowledge in this model has only one dimension, i.e., it is completely general. I abstract from firm-specific knowledge along the lines of Hashimoto (1981) and Carmichael (1983) among others. Unlike Beaudry and Francois (2010), I also assume that the same knowledge can be used for both production and management. This assumption is necessary for a key feature of the model - workers can smoothly move from one firm to another, until they become managers themselves. This assumption, however, can be relaxed without sacrificing the tractability of the model. One can introduce a parameter σ, where $\sigma \leq 1$ measures the fraction of the knowledge that a worker acquires in a firm, that would be useful in other firms. In this case, all the qualitative results would go through, but with lower welfare gains. The scenario where all knowledge is firm-specific would correspond to $\sigma = 0$. In this case, there would be no movement of workers across firms just as in a static model.

Fourth, the assumption of risk neutrality is not necessary for the above results. In the absence of credit market frictions, the Separation Theorem applies - an agent's decision to maximize present value of income would be independent of his risk preference. In other words, the equilibrium matching and earnings will be unaffected by risk-aversion; only the absolute level of welfare will be altered.

And finally, Corollary 1 suggests that some of the incumbent home firms produce less following integration. If agents learn fast enough, the current earnings of some of the home workers is also lower (Corollary 2). Hence, simply looking at current wages or output might give the impression that workers and firms in the home country are worse off following integration when, in fact, they could all be better-off. My model therefore suggests that one should interpret lower wages or output with caution, especially when drawing conclusions about welfare.

[35] The traditional view regarding knowledge spillover is that workers with experience in MNEs are hired by domestic managers. These workers bring with them knowledge regarding better technology and management practices and this raises the productivity of the domestic firms. See, for example, Barba Navaretti and Venables (2004). In fact, this is the channel for knowledge spillovers that has been examined by both Poole (2011) and Balsvik (2011).

5 Numerical Results

In this section, I parameterize the model and solve it numerically. The exercise allows me to derive additional dynamic results which cannot be obtained analytically. I present three sets of results concerning lifetime earnings profiles, inequality and spin-offs from MNEs. Throughout, the focus is on the home economy.

The only change from the last section is in the learning technology. I assume that a worker with knowledge x, and working for a manager with knowledge y, draws his knowledge in the next period from a distribution which is uniform on $[x, y]$. The production function is given by $f(y) = y^{\alpha}$, $n(x;\beta) = x^{\beta}$. Finally, I assume that the distribution of newborns is a truncated exponential in $[1, \overline{k}]$ with parameter λ. By setting $\underline{k} = 1$, the size of the smallest firm is implicitly set to two (one manager and one worker). The following figures are drawn for $\alpha = 1$, $\beta = 0.5$, $\lambda = 1$, $\overline{k}_H = 1.5$, $\overline{k}_F = 1.75$.

5.1 Earnings Dynamics

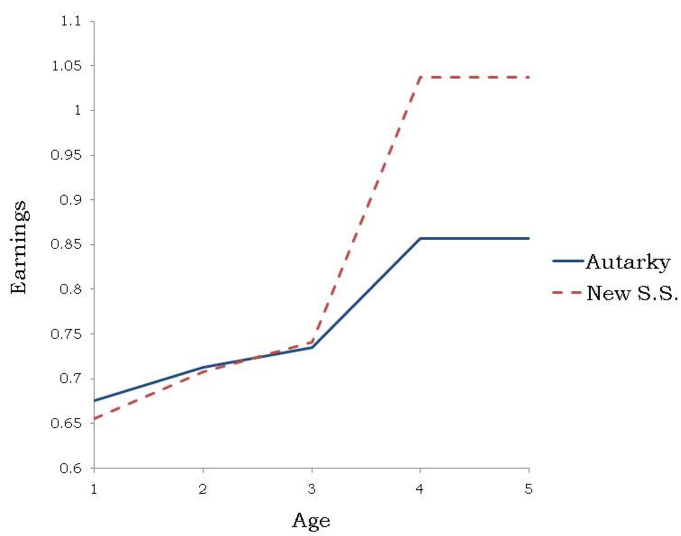

Figure 1: *Evolution of individual earnings*

Although the economy as a whole does not grow in the steady-state, individual earning grows over the lifetime. Figure 1 plots the earnings path of the median worker for $\delta = 0.5$. The figure is drawn under the assumption that the actual knowledge he acquires every period is the expected knowledge that an agent with his level of knowledge would acquire. In the figure,

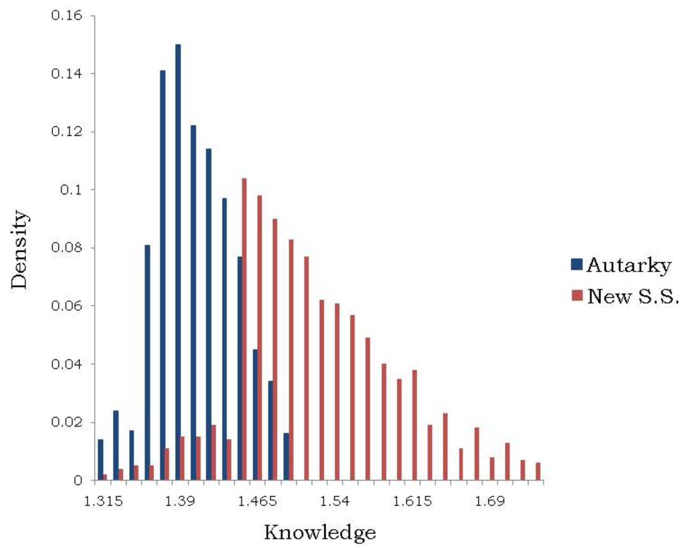

Figure 2: *Distribution of knowledge of an agent after 5 periods*

the agent works for the first three periods and manages from the fourth period onwards. Under integration, a lower wage in the first two periods is more than compensated by the increase in future profits. The lifetime earnings schedule under integration is also steeper than that under autarky.

With integration, the matches of the workers improve. Accordingly, at each period, the worker learns more which causes the earnings schedule to become steeper. Figure 2 provides an explanation for the bigger jump in future profits. It shows the distribution of knowledge of the median worker after 5 periods.[36] Following integration, the distribution shifts to the right. On average, the agent becomes a more knowledgeable manager compared to autarky and hence, his expected profits are higher.

Evidence regarding the impact of multinational production on wages has been mixed. Aitken et al. (1996) report that in Mexico and Venezuela, the wage spillover to domestic firms is negative and significant. On the other hand, Lipsey and Sjöholm (2004) find significant positive wage spillovers to domestic firms in Indonesia. From Figure 1 it seems that following integration, wages could down over at least some knowledge intervals. This result, however, is not inconsistent with the finding of positive wage spillovers. In the above mentioned studies, the reported wage is the average of wages paid by all domestic firms. The average wage depends not only on the level of the wage schedule but also on the distribution of workers. With integration, as workers get matched with better managers and learn more, the mass of the distribution

[36]Given the parameter values, the probability of the agent living for more than 5 periods is extremely small.

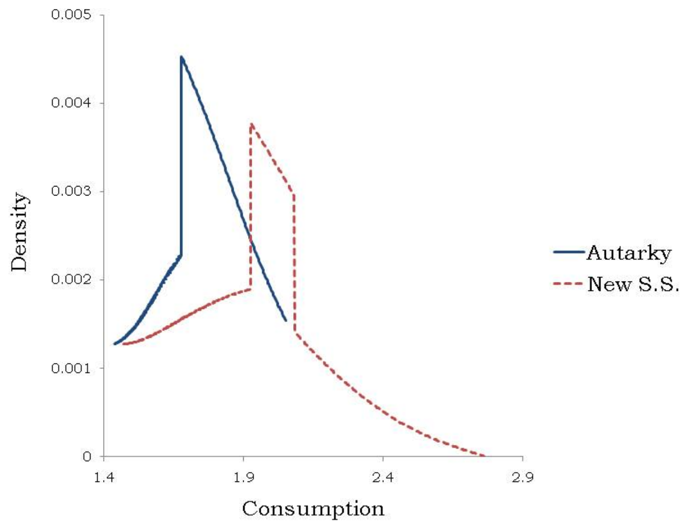

Figure 3: *Consumption distributions*

shifts to the right. Therefore, a lowering of the wage schedule and a higher average wage can go hand in hand.

5.2 Inequality

By generating a non-degenerate consumption distribution, the model also allows me to examine the effect of integration on inequality. Figure 3 plots the home consumption distributions under autarky and integration. δ is set to 0.5 as before. With integration, the distribution stretches out, as mass shifts to the upper tail (The maximum consumption under autarky is 2.05, while that under integration is 2.77). Figure 4 plots the percentage change in the gini coefficient due to integration, as a function of δ. For $\delta = 0.5$, inequality rises by about 40%. But this rise in inequality is not a general phenomenon. For higher values of δ, integration actually leads to a reduction in inequality. Moreover, there is a monotonic relation between inequality and the rate of learning.

When agents learn fast enough, integration gives an advantage to those who are born as the most knowledgeable workers. They work for the most knowledgeable foreign managers, learn a lot, and in turn, become knowledgeable managers in the future. Agents who are born with very little knowledge continue to be matched with the less knowledgeable incumbent home managers and accordingly, learn less - learning amplifies the initial inequality in the economy.[37]

[37]The relation between globalization and inequality has been the subject of a large literature. In a seminal paper,

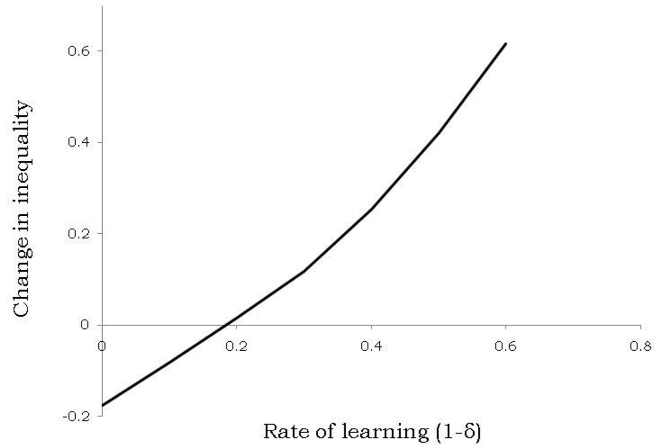

Figure 4: *Change in inequality*

The above predictions could partly explain the experience of India since the early 1990s. The annual FDI inflows to India increased from US $ 654 million to US $ 3083 million between the periods 1993-94 and 1999-00 (Ministry of Commerce and Industry, Government of India). During this time, there has been an increase in inequality, as measured by per-capita expenditure (Deaton and Dreze, 2002). This increase in inequality has occurred not only across regions, but within urban areas as well. If better managers hire more knowledgeable workers, and there is wide discrepancy in the knowledge levels of individuals,[38] then part of this rising inequality could be due to differences in opportunities faced by the Indian workers.

5.3 Pattern of MNE activity

As discussed in Section 4, integration leads to the creation of a new class of home managers, who are as productive as their counterparts in the foreign country. This can be seen in Figure 5 which plots the supply and demand for managers at home in the integrated steady-state equilibrium. The supply of managers is simply the part of the knowledge distribution that lies above the threshold. Recall that the upper bound of the home newborn distribution is 1.5. Hence, there is a discrete drop in the density of newborn agents to the right of 1.5, which explains the

Helpman, Itshoki and Redding (2010) show how wage inequality could go up as a country gradually opens up. In their model, the increase in wage inequality is a consequence of increased dispersion in revenues following opening up, combined with the result that larger firms pay higher wages.

[38]In spite of churning out almost 400,000 engineers every year, only one in four of India's engineers are employable in the software industry (*NYTimes, 17th October, 2006*).

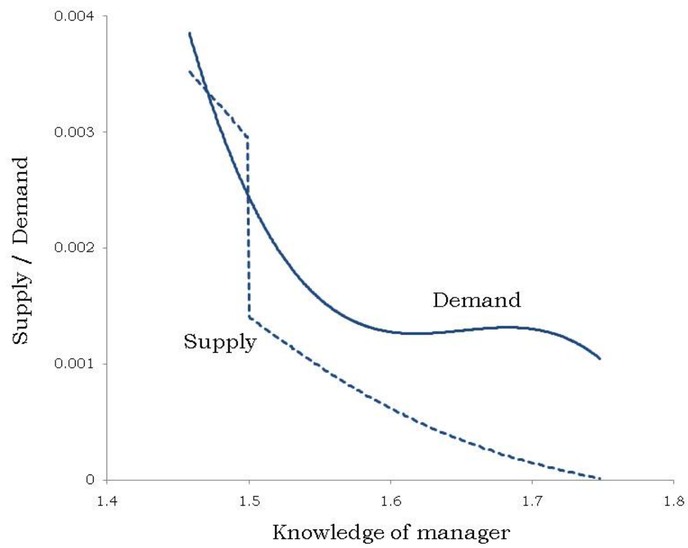

Figure 5: *Supply and demand for managers*

discontinuity at 1.5. The demand for managers is obtained by looking at the number of workers of each type and the demand for manager per worker.[39]

Figure 5 suggests that, in the integrated steady-state equilibrium, there are home managers who are as knowledgeable as their foreign counterparts. At the same time, the supply of home managers is not sufficient to meet the demand. In the new equilibrium, some of the foreign managers hire home workers and hence, home firms and foreign MNEs operate together.[40] Figure 5 also throws light on the pattern of multinational activity. The supply of home managers falls short of demand, and there are *almost* no home MNEs in this equilibrium. Here I follow Antràs et al. (2006) in assuming that a manager hires workers in the other country only if he strictly prefers doing so. Moreover, most of the MNEs operating at home are the most productive foreign firms. Thus, the MNEs, on average, are bigger and more productive than the home firms. PAM implies that a worker in a MNE, on average, is more knowledgeable than a worker in a home firm. Therefore, the former employees of MNEs are also, on average, more productive managers, consistent with the findings of Görg and Strobl (2005).

[39]The demand for managers per worker is simply the reciprocal of the span of control. This is a special feature of the span of control depending only on the knowledge of the worker.

[40]See Markusen and Venables (1999) for the case where FDI leads to the development of local industry, but is driven out as the industry develops enough.

6 Conclusion

Despite both formal and anecdotal evidence on knowledge spillover from foreign multinationals, we do not have a good understanding of, for example, how the distribution of knowledge changes and occupational choice evolves as a country gradually integrates with the rest of the world. This paper takes a step in that direction. I show that allowing domestic workers to learn from foreign managers not only generates novel predictions about lifetime earnings profiles and consumption inequality among other things, ignoring the dynamic nature of knowledge diffusion could lead us to draw incorrect conclusions about welfare. Much remains to be done, however.

As pointed out in the Introduction, in order to focus on matching and occupational choice, I assume that the transfer of knowledge is a costless process. In reality, firms spend a substantial part of their resources on training their workforce. If there is some degree of mobility of workers, firms' incentive to train their workers might be affected, which in turn would affect the diffusion of knowledge. How would the entry of multinationals affect domestic workers and firms in this setting?

This paper highlights a new channel through which domestic managers can gain indirectly from foreign multinationals, even in the absence of any flow of knowledge between them. But mobility of workers from more productive multinationals to less productive domestic firms could be another channel through which domestic managers gain. Workers with experience in multinationals could move to domestic firms for various reasons. A domestic firm could lure a multinational trained worker with higher wages. Or, in a frictional labor market, a multinational trained worker could accept a job in a small domestic firm, rather than wait to be matched with a more productive firm. Whatever be the case, domestic firms benefit from the superior knowledge that these workers bring along with them. I leave the examination of such interesting scenarios for future work.

References

Abowd, J. M. and Kramarz, F. (1999), The analysis of labor markets using matched employer-employee data, *in* O. Ashenfelter and D. Card, eds, 'Handbook of Labor Economics', 1 edn, Vol. 3, Part B, Elsevier, chapter 40, pp. 2629–2710.

Aitken, B., Harrison, A. and Lipsey, R. E. (1996), 'Wages and foreign ownership: A comparative study of Mexico, Venezuela, and the United States', *Journal of International Economics* **40**(3-4), 345 – 371.

Aitken, B. J. and Harrison, A. E. (1999), 'Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela', *American Economic Review* **89**(3), 605–618.

Antràs, P. (2003), 'Firms, Contracts and Trade Structure', *Quarterly Journal of Economics* **118**(4), 1375–1418.

Antràs, P., Garicano, L. and Rossi-Hansberg, E. (2006), 'Offshoring in a Knowledge Economy', *Quarterly Journal of Economics* **121**(1).

Antràs, P. and Helpman, E. (2004), 'Global Sourcing', *Journal of Political Economy* **112**(3), 552–580.

Aw, B. Y., Chung, S. and Roberts, M. J. (2000), 'Productivity and turnover in the export market: Micro-level evidence from the Republic of Korea and Taiwan (China)', *World Bank Economic Review* **14**(1), 65–90.

Balsvik, R. (2011), 'Is Labor Mobility a Channel for Spillovers from Multinationals? Evidence from Norwegian Manufacturing', *The Review of Economics and Statistics* **93**(1), 285–297.

Barba Navaretti, G. and Venables, A. (2004), *Multinational Firms in the World Economy*, Princeton University Press.

Beaudry, P. and Francois, P. (2010), 'Managerial Skills Acquisition and the Theory of Economic Development', *Review of Economic Studies* **77**(1), 90–126.

Becker, G. S. (1962), 'Investment in Human Capital: A Theoretical Analysis', *Journal of Political Economy* **70**(s5), 9.

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

Biesebroeck, J. V. (2005), 'Exporting raises productivity in sub-Saharan African manufacturing firms', *Journal of International Economics* **67**(2), 373–391.

Bloom, N. and Van Reenen, J. (2010), 'Why do management practices differ across firms and countries?', *Journal of Economic Perspectives* **24**(1), 203–24.

Boyd, J. H. and Prescott, E. C. (1987), 'Dynamic coalitions: Engines of growth', *American Economic Review* **77**(2), 63–67.

Burstein, A. T. and Monge-Naranjo, A. (2009), 'Foreign Know-How, Firm Control, and the Income of Developing Countries', *Quarterly Journal of Economics* **124**(1), 149–195.

Carmichael, L. (1983), 'Firm-specific Human Capital and Promotion Ladders', *The Bell Journal of Economics* **14**(1), pp. 251–258.

Chari, V. V. and Hopenhayn, H. (1991), 'Vintage Human Capital, Growth, and the Diffusion of New Technology', *Journal of Political Economy* **99**(6), 1142–65.

Clerides, S. K., Lach, S. and Tybout, J. R. (1998), 'Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco', *The Quarterly Journal of Economics* **113**(3), pp. 903–947.

Coe, D. and Helpman, E. (1995), 'International R&D spillovers', *European Economic Review* **39**(5), 859–887.

Coe, D. T., Helpman, E. and Hoffmaister, A. W. (1997), 'North-South R&D spillovers', *Economic Journal* **107**(440), 134–49.

Deaton, A. and Dreze, J. (2002), 'Poverty and Inequality in India: A Re-examination', *Economic and Political Weekly* **37**(36), pp. 3729–3748.

Dix-Carneiro, R. (2011), Trade liberalization and labor market dynamics, Unpublished manuscript, Princeton University.

Easterly, W. (2001), *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics*, The MIT Press.

Filson, D. and Franco, A. (2006), 'Spin-outs: Knowledge diffusion through employee mobility', *Rand Journal of Economics* **37**(4), 841–60.

Fosfuri, A., Motta, M. and Rønde, T. (2001), 'Foreign direct investment and spillovers through workers' mobility', *Journal of International Economics* **53**(1), 205 – 222.

Garetto, S. (2008), Input Sourcing and Multinational Production, mimeo, University of Chicago.

Garicano, L. (2000), 'Hierarchies and the Organization of Knowledge in Production', *Journal of Political Economy* **108**(5), 874–904.

Giarratana, M., Pagano, A. and Torrisi, S. (2004), The Role of Multinational Firms in the Evolution of the Software Industry in India, Ireland and Israel, *in* A. Arora and A. Gambardella, eds, 'From Underdogs to Tigers: Bridging the Gap', Oxford University Press, New York, pp. 207–235.

Glass, A. J. and Saggi, K. (2002), 'Multinational Firms and Technology Transfer', *The Scandinavian Journal of Economics* **104**(4), 495–513.

Goldberg, P. K. and Pavcnik, N. (March 2007), 'Distributional Effects of Globalization in Developing Countries', *Journal of Economic Literature* **45**, 39–82(44).

Görg, H. and Strobl, E. (2005), 'Spillovers from foreign firms through worker mobility: An empirical investigation', *Scandinavian Journal of Economics* **107**(4).

Griffith, R., Redding, S. J. and Simpson, H. (2003), Productivity convergence and foreign ownership at the establishment level, Cep discussion papers, Centre for Economic Performance, LSE.

Grossman, G. M. and Helpman, E. (2003), 'Outsourcing versus FDI in Industry Equilibrium', *Journal of the European Economic Association* **1**(2/3), 317–327.

Hashimoto, M. (1981), 'Firm-specific Human Capital as a Shared Investment', *The American Economic Review* **71**(3), pp. 475–482.

Haskel, J. E., Pereira, S. C. and Slaughter, M. J. (2007), 'Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms?', *The Review of Economics and Statistics* **89**(3), 482–496.

Helpman, E. (1984), 'A Simple Theory of International Trade with Multinational Corporations', *Journal of Political Economy* **92**(3), 451–471.

Helpman, E., Itshoki, O. and Redding, S. (2010), 'Inequality and Unemployment in a Global Economy', *Econometrica* **78**(4), 1239–1283.

Holmes, T. J. and Schmitz, James A., J. (1990), 'A Theory of Entrepreneurship and its Application to the Study of Business Transfers', *Journal of Political Economy* **98**(2), 265–294.

IMF (2007), *World Economic Outlook, October 2007: Globalization and Inequality*, International Monetary Fund.

Javorcik, B. S. (2004), 'Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages', *The American Economic Review* **94**(3), pp. 605–627.

Jovanovic, B. and Rob, R. (1989), 'The Growth and Diffusion of Knowledge', *Review of Economic Studies* **56**(4), 569–582.

Keller, W. (1998), 'Are international R&D spillovers trade-related?: Analyzing spillovers among randomly matched trade partners', *European Economic Review* **42**(8), 1469 – 1481.

Keller, W. (2002), 'Geographic localization of international technology diffusion', *American Economic Review* **92**(1), 120–142.

Keller, W. (2004), 'International technology diffusion', *Journal of Economic Literature* **42**(3), pp. 752–782.

Keller, W. and Hovhannisyan, N. (2010), International business travel: An engine of innovation?, CEPR Discussion Papers 7829, C.E.P.R. Discussion Papers.

Keller, W. and Yeaple, S. R. (2009*a*), Gravity in the weightless economy, CEPR Discussion Papers 7553, C.E.P.R. Discussion Papers.

Keller, W. and Yeaple, S. R. (2009*b*), 'Multinational Enterprises, International Trade, and Productivity Growth: Firm-Level Evidence from the United States', *The Review of Economics and Statistics* **91**(4), 821–831.

Kerr, W. R. (2008), 'Ethnic scientific communities and international technology diffusion', *The Review of Economics and Statistics* **90**(3), 518–537.

Klepper, S. (2002), 'The capabilities of new firms and the evolution of the US automobile industry', *Industrial and Corporate Change* **11**(4), 645–666.

Klepper, S. and Sleeper, S. (2005), ‘Entry by Spinoffs’, *Management Science* **51**(8), 1291–1306.

Kremer, M. (1993), ‘The O-ring Theory of Economic Development’, *The Quarterly Journal of Economics* **108**(3), 551–575.

Kremer, M. and Maskin, E. (2006), Globalization and Inequality, mimeo, Harvard University.

Kremer, M. and Maskin, E. S. (1996), Wage inequality and segregation by skill, NBER Working Papers 5718, National Bureau of Economic Research, Inc.

Lipsey, R. E. and Sjöholm, F. (2004), ‘FDI and Wage Spillovers in Indonesian Manufacturing’, *Review of World Economics* **140**(2), 287–310.

Lipsey, R. E. and Sjöholm, F. (2005), The impact of inward FDI on host countries: Why such different answers?, *in* T. H. Moran, E. M. Graham and M. Blomström, eds, ‘Does foreign direct investment promote development?’, Peterson Institute, pp. 23–44.

Loecker, J. K. D. (2007), ‘Do exports generate higher productivity? Evidence from Slovenia’, *Journal of International Economics* **73**(1), 69–98.

Malchow-Møller, N., Markusen, J. R. and Schjerning, B. (2007), Foreign Firms, Domestic Wages, Working paper, NBER.

Markusen, J. R. and Trofimenko, N. (2009), ‘Teaching locals new tricks: Foreign experts as a channel of knowledge transfers’, *Journal of Development Economics* **88**(1), 120 – 131.

Markusen, J. R. and Venables, A. J. (1999), ‘Foreign direct investment as a catalyst for industrial development’, *European Economic Review* **43**(2), 335 – 356.

Monge-Naranjo, A. (2008), Foreign Firms, Domestic Entrepreneurial Skills and Development, mimeo, Penn State University.

Ohnsorge, F. and Trefler, D. (2007), ‘Sorting It Out: International Trade with Heterogeneous Workers’, *Journal of Political Economy* **115**(5), 868–892.

Oman, C. (2000), *Policy competition for foreign direct investment: A study of competition among governments to attract FDI*, OECD.

Poole, J. (2011), Knowledge transfers from multinational to domestic firms: Evidence from worker mobility, Forthcoming, *Review of Economics and Statistics*.

Ramondo, N. (2008), Size, Geography and Multinational Production, mimeo, University of Texas at Austin.

Rodríguez-Clare, A. (1996), 'Multinationals, Linkages, and Economic Development', *American Economic Review* **86**(4), 852–73.

Rodríguez-Clare, A. (2007), Trade, Diffusion and the Gains from Openness, Working paper, NBER.

Rosen, S. (1972), 'Learning and Experience in the Labor Market', *Journal of Human Resources* **7**(3), 326–342.

Rosen, S. (1982), 'Authority, Control, and the Distribution of Earnings', *Bell Journal of Economics* **13**(2), 311–323.

Sattinger, M. (1993), 'Assignment Models of the Distribution of Earnings', *Journal of Economic Literature* **31**(2), 831–880.

Appendix

Proof of Proposition 2: Let the number of people being born every period be normalized to 1. Cohort t at time t are all newborns. All agents in $[\underline{k}, k_A^*]$ are workers. The measure of these agents is $\int_{\underline{k}}^{k_A^*} \phi_H(k)dk$. A worker with knowledge k demands $\frac{1}{n(k)}$ managers. Therefore, the total demand for managers by cohort t workers is $\int_{\underline{k}}^{k_A^*} \frac{\phi_H(k)}{n(k;\beta)}dk$. The supply of cohort t managers is simply the measure of agents in $[k_A^*, \overline{k}_H]$. This is given by $\int_{k_A^*}^{\overline{k}_H} \phi_H(k)dk$. Let us consider the distribution of cohort $t-1$ agents at time t. A fraction $1-\delta$ of every type of agent in $[\underline{k}, k_A^*]$ survive in period t. Out of the ones that survive, a fraction π of every type of agent do not learn and remain where they are. Hence, the total demand for managers by cohort $t-1$ workers is $\int_{\underline{k}}^{k_A^*} \frac{\pi(1-\delta)\phi_H(k)}{n(k;\beta)}dk$. Similarly, a fraction $1-\delta$ of the cohort $t-1$ managers in $[k_A^*, \overline{k}]$ survive in period t. These agents do not learn. Moreover, $(1-\theta)(1-\delta)\int_{\underline{k}}^{k_A^*} \phi_H(k)dk$ agents move into this interval from $[\underline{k}, k_A^*]$. They are the cohort $t-1$ agents who were workers in period $t-1$ but become managers in period t. Therefore, the supply of cohort $t-1$ managers is $(1-\delta)[\int_{k_A^*}^{\overline{k}_H} \phi_H(k)dk + (1-\pi)\int_{\underline{k}}^{k_A^*} \phi_H(k)dk]$. The supply and demand for managers in other cohorts can be obtained in a similar fashion. Adding up the demand for managers and the supply of managers in each cohort, we get

$$\text{Demand for managers } = \frac{1}{1-\theta(1-\delta)} \int_{\underline{k}}^{k_A^*} \frac{\phi_H(k)}{n(k;\beta)}dk$$

$$\text{Supply of managers} = \frac{1}{\delta(1-\theta(1-\delta))}[(1-\theta(1-\delta))\int_{k_A^*}^{\overline{k}_H} \phi_H(k)dk + (1-\theta)(1-\delta)\int_{\underline{k}}^{k_A^*} \phi_H(k)dk]$$

In equilibrium, supply must equal demand. Equating the above two expressions and after a bit of algebra, we obtain the following :

$$\int_{\underline{k}}^{k_A^*} \frac{\phi_H(k)}{n(k;\beta)}dk = \int_{k_A^*}^{\overline{k}_H} \phi_H(k)dk + (1-\theta)(\frac{1-\delta}{\delta})$$

In order to derive the properties of k_A^*, we use the Implicit Function Theorem. Differentiating the above equation w.r.t. π,

$$\frac{\partial k_A^*}{\partial \pi}\frac{\phi_H(k_A^*)}{n(k_A^*;\beta)} = -\frac{\partial k_A^*}{\partial \pi}\phi_H(k_A^*) - (\frac{1-\delta}{\delta})$$

Therefore,

$$\frac{\partial k_A^*}{\partial \pi}[\frac{\phi_H(k_A^*)}{n(k_A^*;\beta)} + \phi_H(k_A^*)] = -(\frac{1-\delta}{\delta})$$

Since the LHS is positive while the RHS is negative, $\frac{\partial k_A^*}{\partial \pi} < 0$. In a similar fashion it can be shown that

$\frac{\partial k_A^*}{\partial \delta} < 0$. Differentiating the labour market clearing condition w.r.t. β,

$$\frac{\partial k_A^*}{\partial \beta}\frac{\phi_H(k_A^*)}{n(k_A^*;\beta)} + \int\limits_{\underline{k}}^{k_A^*} \frac{-\phi_H(k)}{n(k;\beta)^2}\frac{\partial n(k;\beta)}{\partial \beta}dk = -\frac{\partial k_A^*}{\partial \beta}\phi_H(k_A^*)$$

Re-arranging terms, we have

$$\frac{\partial k_A^*}{\partial \beta}[\frac{\phi_H(k_A^*)}{n(k_A^*;\beta)} + \phi_H(k_A^*)] = \int\limits_{\underline{k}}^{k_A^*} \frac{\phi_H(k)}{n(k;\beta)^2}\frac{\partial n(k;\beta)}{\partial \beta}dk$$

Since $\frac{\partial n(k;\beta)}{\partial \beta} > 0$, both the LHS and the RHS are positive. Therefore, $\frac{\partial k_A^*}{\partial \beta} > 0$

Lemma 1. *If a knowledge distribution G first-order stochastic dominates another distribution H, then $k_G^* > k_H^*$, where k_G^* and k_H^* are the thresholds under G and H respectively.*

Proof. Let G f.o.s.d. H. Let g and h be the corresponding densities. Also, let $\xi(k)$ be the demand for manager per worker, where the worker has knowledge k. Since the span of control is only a function of the worker's knowledge, a worker with knowledge k works in a firm of size $n(k)$. Hence $\xi(k)$ is simply the reciprocal of $n(k)$. Therefore $\xi'(k) < 0$ (this follows from $n'(k) > 0$). Also, let k^* be the threshold under H.

We shall prove the lemma by contradiction. Let k^* also be the threshold for G. We can have two possibilities - (i) $g(k) < h(k)$ for all $k < k^*$. In this case, the demand for managers under $G = \int_{\underline{k}}^{k^*} \xi(k)g(k)dk < \int_{\underline{k}}^{k^*} \xi(k)h(k)dk =$ demand for managers under H. But the supply of managers under $G = 1 - G(k^*) > 1 - H(k^*) =$supply of managers under H. Hence at k^*, there is an excess supply of managers under G. This means that the threshold for G must be greater than k^*. (ii) There are n intervals $A_i \subset [\underline{k}, k^*]$, $i = 1, ...n$ such that

$$\begin{aligned} g(k) &> h(k) \forall k \in A_i, \forall i \\ g(k) &< h(k) \text{ otherwise} \end{aligned}$$

Rank the $A_i s$ such that $A_i < A_j \Rightarrow maxA_i < minA_j$. We proceed as follows - We know that $\underline{k} < minA_1 = a_1(say)$ (otherwise H would f.o.s.d. G). Let $B = [\underline{k}, a_1]$. Then it must be the case that $g(k) < h(k)$ for all $k \in B$. G f.o.s.d. H implies that

$$\int_B g(k)dk + \int_{A_1} g(k)dk < \int_{A_1} h(k)dk + \int_B h(k)dk$$

Re-arranging the above equation,

$$\int_{A_1} [g(k) - h(k)]dk < \int_B [h(k) - g(k)]dk$$

Multiplying both sides by $\xi(a_1)$,

$$\int_{A_1} \xi(a_1)[g(k) - h(k)]dk < \int_B \xi(a_1)[h(k) - g(k)]dk$$

Now, $\xi'(k) < 0$ implies that $\xi(a_1) < \xi(k)\forall k \in B$ and $\xi(a_1) > \xi(k)\forall k \in A_1$. Replacing $\xi(a_1)$ in the above equation,

$$\int_{A_1} \xi(k)[g(k) - h(k)]dk < \int_B \xi(k)[h(k) - g(k)]dk$$

Here we are using the fact that $h(k) - g(k) > 0\forall k \in B$ and $g(k) - h(k) > 0\forall k \in A_1$. We re-arrange again to obtain

$$\int_B \xi(k)g(k)dk + \int_{A_1} \xi(k)g(k)dk < \int_{A_1} \xi(k)h(k)dk + \int_B \xi(k)h(k)dk$$

The LHS and the RHS are the demand for managers by workers in $B \cup A_1$ under G_1 and G_2 respectively. Define $maxA_1 = a_1'$ and $minA_2 = a_2$. Let $C = [a_1', a_2]$. G f.o.s.d. H implies that

$$\int_B g(k)dk + \int_{A_1} g(k)dk + \int_C g(k)dk + \int_{A_2} g(k)dk < \int_B h(k)dk + \int_{A_1} h(k)dk + \int_C h(k)dk + \int_{A_2} h(k)dk$$

Re-arranging, we have

$$\int_{A_2} [g(k) - h(k)]dk < (\int_B [h(k) - g(k)]dk - \int_{A_1} [g(k) - h(k)]dk) + \int_C [h(k) - g(k)]dk$$

Multiplying both sides by $\xi(a_2)$,

$$\int_{A_2} \xi(a_2)[g(k) - h(k)]dk < \xi(a_2)(\int_B [h(k) - g(k)]dk - \int_{A_1} [g(k) - h(k)]dk) + \int_C \xi(a_2)[h(k) - g(k)]dk$$

Since $\xi'(k) < 0$, we have

$$\begin{aligned} \int_{A_2} \xi(k)[g(k) - h(k)]dk \quad &< \quad \xi(a_2)(\int_B [h(k) - g(k)]dk - \int_{A_1} [g(k) - h(k)]dk) \\ &\quad + \int_C \xi(k)[h(k) - g(k)]dk \end{aligned}$$

Again, using $\xi'(k) < 0$ in the above inequality

$$\begin{aligned} LHS \quad &< \quad \xi(a_1)(\int_B [h(k) - g(k)]dk - \int_{A_1} [g(k) - h(k)]dk) \\ &\quad + \int_C \xi(k)[h(k) - g(k)]dk \\ &\quad \int_B \xi(k)[h(k) - g(k)]dk - \int_{A_1} \xi(k)[g(k) - h(k)]dk \\ &\quad + \int_C \xi(k)[h(k) - g(k)]dk \end{aligned}$$

Re-arranging gives us that the demand for managers by workers in $B \cup A_1 \cup C \cup A_2$ under G is less than that under H. We can repeat this argument by expanding the set till we reach k^*. But then we have shown that the demand for managers under G is less than that under H. However the supply of managers under G is greater than that under H. Therefore, at k^*, there is an excess supply of managers under G. Hence the threshold under G has to be greater than k^*. $\square$

Proof of Proposition 3: The derivation of the threshold is the same as in the proof of Proposition 2. Equating the supply of managers and the demand for managers, we have

$$\int_{\underline{k}}^{k_I^*} \frac{\phi_H(k) + \phi_F(k)}{2n(k;\beta)} dk = \int_{k_I^*}^{\overline{k}_F} \frac{\phi_H(k) + \phi_F(k)}{2} dk + (1-\theta)(\frac{1-\delta}{\delta})$$

Re-arranging, we have

$$\int_{\underline{k}}^{k_I^*} \frac{\phi_H(k)}{n(k;\beta)} dk - \int_{k_I^*}^{\overline{k}_H} \phi_H(k) dk - (1-\theta)(\frac{1-\delta}{\delta}) = \int_{k_I^*}^{\overline{k}_H} \phi_F(k) dk + (1-\theta)(\frac{1-\delta}{\delta}) - \int_{\underline{k}}^{k_I^*} \frac{\phi_F(k)}{n(k;\beta)} dk$$

Not that the LHS is the excess demand for managers in the Home country if the threshold is k_I^*, while the RHS is the excess supply of managers in the Foreign country if the threshold is k_I^*. If $k_I^* = k_A^*$,where I have dropped the subscript H from $k_{A,H}^*$ for simplicity, the LHS is equal to 0, i.e.

$$\int_{\underline{k}}^{k_A^*} \frac{\phi_H(k)}{n(k;\beta)} dk - \int_{k_A^*}^{\overline{k}_H} \phi_H(k) dk = (1-\theta)(\frac{1-\delta}{\delta})$$

Since $\phi_F(k)$ f.o.s.d. $\phi_H(k)$, from Lemma 1, we know that

$$\int_{\underline{k}}^{k_A^*} \frac{\phi_F(k)}{n(k;\beta)} dk - \int_{k_A^*}^{\overline{k}_H} \phi_F(k) dk < (1-\theta)(\frac{1-\delta}{\delta})$$

Therefore, for $k_I^* = k_A^*$, the RHS is positive. But this means that $k_I^* \neq k_A^*$. In particular, since the LHS is increasing in k_I^* and the RHS is decreasing, it must be the case that $k_I^* > k_A^*$.

Proof of Proposition 5: We know that an allocation A is a Pareto improvement over allocation B if $u(x_i^A) \geq u(x_i^B)$ for all i, and $u(x_j^A) > u(x_j^B)$ for some j. This suggests that in order to show that A *is not* a Pareto improvement over B, it is sufficient to show that $\exists$ individuals 1 and 2 s.t. $u(x_1^A) \geq u(x_1^B) \Rightarrow u(x_2^A) < u(x_2^B)$ and vice versa. From Lemma 1, we have

$$k_{A,NL}^* < k_{I,NL}^*$$

where NL refers to no-learning. If there are incumbent firms in the Home economy, this also means that

$$k_{I,NL}^* < \overline{k}_H$$

The above inequality suggests that under Integration, there are incumbent Home managers who continue to operate ($k \in [k^*_{I,NL}, \overline{k}_H]$). At the same time, under Autarky, $m^{-1}_{A,NL}(k^*_{A,NL}) = \underline{k} \Rightarrow m^{-1}_{A,NL}(k^*_{I,NL}) > \underline{k}$ (follows from PAM). While under Integration, $m^{-1}_{A,NL}(k^*_{I,NL}) = \underline{k}$, i.e., under Integration, the manager with knowledge $k^*_{I,NL}$ has a worse match. The present value of $k^*_{I,NL}$ is just the period profits $\pi_{I,NL}(k^*_{I,NL})$ divided by δ.

$$\begin{aligned}
\pi_{A,NL}(k^*_{I,NL}) &= (f(k^*_{I,NL}) - w_{A,NL}(m^{-1}_{A,NL}(k^*_{I,NL})))n(m^{-1}_{A,NL}(k^*_{I,NL})) \\
&\geq (f(k^*_{I,NL}) - w_{A,NL}(k))n(k) \quad \forall k \\
&\geq (f(k^*_{I,NL}) - w_{A,NL}(\underline{k}))n(\underline{k})
\end{aligned}$$

Note that $\pi_{I,NL}(k^*_{I,NL}) = (f(k^*_{I,NL}) - w_{I,NL}(\underline{k}))n(\underline{k})$. Therefore, the relation between $\pi_{A,NL}(k^*_{I,NL})$ and $\pi_{I,NL}(k^*_{I,NL})$ depends on the relation between $w_{A,NL}(\underline{k})$ and $w_{I,NL}(\underline{k})$. Let us consider the following cases -

(a) $w_{A,NL}(\underline{k}) < w_{I,NL}(\underline{k})$: In this case, $\pi_{A,NL}(k^*_{I,NL}) > \pi_{I,NL}(k^*_{I,NL}) \Rightarrow \underline{k}$ is strictly better-off under Integration but $k^*_{I,NL}$ is strictly worse-off.

(b) $w_{A,NL}(\underline{k}) \gg w_{I,NL}(\underline{k})$: Then it is possible to have, $\pi_{A,NL}(k^*_{I,NL}) < \pi_{I,NL}(k^*_{I,NL}) \Rightarrow k^*_{I,NL}$ is strictly better-off under Integration but $\underline{k}$ is strictly worse-off.

(c) $w_{A,NL}(\underline{k}) = w_{I,NL}(\underline{k})$: In this case, $\pi_{A,NL}(k^*_{I,NL}) \geq \pi_{I,NL}(k^*_{I,NL})$. This is not a negation of Pareto improvement. However let us choose the agent with knowledge $k^*_{I,NL}+\epsilon$ such that $m^{-1}_{I,NL}(k^*_{I,NL}+\epsilon) < m^{-1}_{A,NL}(k^*_{I,NL}+\epsilon)$. Since $m(.)$ is continuous, we can always find such an ϵ. Moreover, since $m(.)$ is a function, its inverse must be strictly monotonic. Hence $m^{-1}_{I,NL}(k^*_{I,NL}+\epsilon) > m^{-1}_{I,NL}(k^*_{I,NL}) = \underline{k}$. Now

$$w'_{A,NL}(\underline{k}) = \frac{(f(k'_{A,NL}) - w_{A,NL}(\underline{k}))n'(\underline{k})}{n(\underline{k})} < \frac{(f(k'_{I,NL}) - w_{I,NL}(\underline{k}))n'(\underline{k})}{n(\underline{k})} = w'_{I,NL}(\underline{k})$$

Combined with $w_{A,NL}(\underline{k}) = w_{I,NL}(\underline{k})$, this means that in the neighborhood of $k = \underline{k}$, $w_{A,NL}(k) < w_{I,NL}(k)$. Hence,

$$\begin{aligned}
\pi_{A,NL}(k^*_{I,NL}+\epsilon) &= (f(k^*_{I,NL}+\epsilon) - w_{A,NL}(m^{-1}_{A,NL}(k^*_{I,NL}+\epsilon)))n(m^{-1}_{A,NL}(k^*_{I,NL}+\epsilon)) \\
&\geq (f(k^*_{I,NL}+\epsilon) - w_{A,NL}(m^{-1}_{I,NL}(k^*_{I,NL}+\epsilon)))n(m^{-1}_{I,NL}(k^*_{I,NL}+\epsilon))
\end{aligned}$$

Using the fact that $w_{A,NL}(k^*_{I,NL}+\epsilon) < w_{I,NL}(k^*_{I,NL}+\epsilon)$, we have

$$\pi_{A,NL}(k^*_{I,NL}+\epsilon) > (f(k^*_{I,NL}+\epsilon) - w_{I,NL}(m^{-1}_{I,NL}(k^*_{I,NL}+\epsilon)))n(m^{-1}_{I,NL}(k^*_{I,NL}+\epsilon)) = \pi_{I,NL}(k^*_{I,NL}+\epsilon)$$

Therefore $k^*_{I,NL}+\epsilon$ is strictly worse-off. Hence, for all the 3 cases (a), (b) and (c), we have shown that at least one individual is worse-off. Since these cases are exhaustive, the result follows.

Proof of Proposition 6: We shall proceed as follows - First, we shall find the condition under which $\overline{k}_H$ is better-off under Integration. Since $\overline{k}_H$ is a manager under both Autarky and Integration, we have to show that $V_{M,I}(\overline{k}_H) > V_{M,A}(\overline{k}_H)$. Since $\overline{k}_H$ is matched with k^*_A under Autarky and $\widetilde{k}$ under Integration, this implies that

$$[f(\overline{k}_H) - w_A(k^*_A)]n(k^*_A) < [f(\overline{k}_H) - w_I(\widetilde{k})]n(\widetilde{k})$$

Re-arranging, we have

$$f(\overline{k}_H)[n(k_A^*) - n(\widetilde{k})] < w_A(k_A^*)n(k_A^*) - w_I(\widetilde{k})n(\widetilde{k})$$

Now,

$$\begin{aligned} w_A(k_A^*)n(k_A^*) - w_I(\widetilde{k})n(\widetilde{k}) \quad = \quad & \frac{\delta + (1-\theta)(1-\delta)}{\delta}[f(k_A^*)n(\underline{k})\frac{n(k_A^*)}{1+n(k_A^*)} - f(k_I^*)n(\underline{k})\frac{n(k_I^*)}{1+n(k_I^*)}] \\ & + \frac{(1-\theta)(1-\delta)}{\delta}[f(\overline{k}_F)n(k_I^*)\frac{n(k_I^*)}{1+n(k_I^*)} - f(\overline{k}_H)n(k_A^*)\frac{n(k_A^*)}{1+n(k_A^*)}] \\ & + \frac{\delta + (1-\theta)(1-\delta)}{\delta}[\frac{n(k_A^*)}{1+n(k_A^*)}\int\limits_{\underline{k}}^{k_A^*} f(m_A(k))n'(k)dk \end{aligned}$$

$$+\frac{n(k_I^*)}{1+n(k_I^*)}\int\limits_{\widetilde{k}}^{k_I^*} f(m_I(k))n'(k)dk] - \int\limits_{\underline{k}}^{\widetilde{k}} f(m_I(k))n'(k)dk$$

Let us consider each term on the RHS.

$$f(k_A^*)n(\underline{k})\frac{n(k_A^*)}{1+n(k_A^*)} - f(k_I^*)n(\underline{k})\frac{n(k_I^*)}{1+n(k_I^*)} > f(\underline{k})n(\underline{k})\frac{n(\underline{k})}{1+n(\underline{k})} - f(\overline{k}_H)n(\underline{k})\frac{n(\overline{k}_H)}{1+n(\overline{k}_H)}$$

$$f(\overline{k}_F)n(k_I^*)\frac{n(k_I^*)}{1+n(k_I^*)} - f(\overline{k}_H)n(k_A^*)\frac{n(k_A^*)}{1+n(k_A^*)} > f(\overline{k}_F)n(\underline{k})\frac{n(\underline{k})}{1+n(\underline{k})} - f(\overline{k}_H)n(\overline{k}_H)\frac{n(\overline{k}_H)}{1+n(\overline{k}_H)}$$

$$-\int\limits_{\underline{k}}^{\widetilde{k}} f(m_I(k))n'(k)dk > -f(\overline{k}_H)[n(\overline{k}_H) - n(\underline{k})]$$

$$\frac{n(k_A^*)}{1+n(k_A^*)}\int\limits_{\underline{k}}^{k_A^*} f(m_A(k))n'(k)dk > 0, \ \frac{n(k_I^*)}{1+n(k_I^*)}\int\limits_{\widetilde{k}}^{k_I^*} f(m_I(k))n'(k)dk > 0$$

Replacing the inequalities in the above equation,

$$\begin{aligned} w_A(k_A^*)n(k_A^*) - w_I(\widetilde{k})n(\widetilde{k}) \quad > \quad & \frac{\delta + (1-\theta)(1-\delta)}{\delta}[f(\underline{k})n(\underline{k})\frac{n(\underline{k})}{1+n(\underline{k})} - f(\overline{k}_H)n(\underline{k})\frac{n(\overline{k}_H)}{1+n(\overline{k}_H)}] \\ & + \frac{(1-\theta)(1-\delta)}{\delta}[f(\overline{k}_F)n(\underline{k})\frac{n(\underline{k})}{1+n(\underline{k})} - f(\overline{k}_H)n(\overline{k}_H)\frac{n(\overline{k}_H)}{1+n(\overline{k}_H)}] \\ & - f(\overline{k}_H)[n(\overline{k}_H) - n(\underline{k})] \\ = \quad & A(say) \end{aligned}$$

Furthermore,

$$\begin{aligned} f(\overline{k}_H)[n(k_A^*) - n(\widetilde{k})] &< f(\overline{k}_H)[n(\overline{k}_H) - n(\underline{k})] \\ &= B(say) \end{aligned}$$

Hence the sufficient condition for $\overline{k}_H$ to be strictly better-off under Integration is that $A > B$. After a bit of algebra, this condition reduces to

$$(1-\theta)(\frac{1-\delta}{\delta}) > \frac{2f(\overline{k}_H)[n(\overline{k}_H) - n(\underline{k})] + n(\underline{k})[\mu(\overline{k}_H)f(\overline{k}_H) - \mu(\underline{k})f(\underline{k})]}{\mu(\underline{k})[f(\overline{k}_F) + f(\underline{k})]n(\underline{k}) - \mu(\overline{k}_H)[n(\underline{k}) + n(\overline{k}_H)]f(\overline{k}_H)}$$

where $\mu(\underline{k}) = \frac{n(\underline{k})}{1+n(\underline{k})}$ and $\mu(\overline{k}_H) = \frac{n(\overline{k}_H)}{1+n(\overline{k}_H)}$. Of course, this only ensures that $\overline{k}_H$ is strictly better off. We need to show that every Home agent can be made better off. Notice that for $k \in [\underline{k}, k_A^*]$, agents are workers under both regimes. For $k \in [k_A^*, k_I^*]$, agents are workers under Integration but managers under Autarky. Finally for $k \in [k_I^*, \overline{k}_H]$, agents are managers under both regimes. In the steady-state, $V_{M,i}(k) = \frac{1}{\delta}\pi_i(k), i \in \{A, I\} \Rightarrow V'_{M,i}(k) = \frac{1}{\delta}\pi'_i(k) = \frac{1}{\delta}f'(k)n(m_i^{-1}(k))$. For $k \in [k_I^*, \overline{k}_H]$,

$$\begin{aligned} m_I^{-1}(k) &< m_A^{-1}(k) \Rightarrow \frac{1}{\delta}f'(k)n(m_I^{-1}(k)) < \frac{1}{\delta}f'(k)n(m_A^{-1}(k)) \\ &\Rightarrow V'_{M,I}(k) < V'_{M,A}(k) \end{aligned}$$

Suppose $V_{M,I}(\overline{k}_H) > V_{M,A}(\overline{k}_H)$. Since $V_{M,A}(.)$ is decreasing at a faster rate than $V_{M,I}(.)$ in the neighborhood $[k_I^*, \overline{k}_H]$, this implies that $V_{M,I}(k) > V_{M,A}(k)$ for $k \in [k_I^*, \overline{k}_H]$. In particular, $V_{M,I}(k_I^*) > V_{M,A}(k_I^*)$. For $k \in [k_A^*, k_I^*]$,

$$V_{W,I}(k) = \frac{1}{\delta + (1-\theta)(1-\delta)}w_I(k) + \frac{(1-\theta)(1-\delta)}{\delta(\delta + (1-\theta)(1-\delta))}f(m_I(k))n(k)$$

$$\Rightarrow V'_{W,I}(k) = \frac{1}{\delta + (1-\theta)(1-\delta)}f(m_I(k))n'(k) + \frac{(1-\theta)(1-\delta)}{\delta(\delta + (1-\theta)(1-\delta))}f'(m_I(k))n(k)m'_I(k)$$

Also, $V'_{M,A}(k) = \frac{1}{\delta}f'(k)n(m_A^{-1}(k))$. When $\delta = 1$, $V'_{W,I}(k) = f(m_I(k))n'(k)$ and $V'_{M,A}(k) = f'(k)n(m_A^{-1}(k))$. Now,

$$\frac{f'(k)}{f(m_I(k))} \geq \frac{f'(\underline{k})}{f(m_I(k))} > \frac{f'(\underline{k})}{f(\overline{k}_H)} > \frac{n'(\underline{k})}{n(\underline{k})} \geq \frac{n'(m_A^{-1}(k))}{n(m_A^{-1}(k))} > \frac{n'(k)}{n(m_A^{-1}(k))}$$

Hence, $V'_{M,A}(k) > V'_{W,I}(k)$. Therefore, $\exists\ \delta_1$ s.t. $\forall \delta > \delta_1$, $V'_{M,A}(k) > V'_{W,I}(k)$ and hence $V_{W,I}(k) > V_{M,A}(k)$ for $k \in [k_A^*, k_I^*]$. In particular, $V_{W,I}(k_A^*) > V_{M,A}(k_A^*)$. For $k \in [\underline{k}, k_A^*]$,

$$V'_{W,A}(k) = \frac{1}{\delta + (1-\theta)(1-\delta)}f(m_A(k))n'(k) + \frac{(1-\theta)(1-\delta)}{\delta(\delta + (1-\theta)(1-\delta))}f'(m_A(k))n(k)m'_A(k)$$

and

$$V'_{W,I}(k) = \frac{1}{\delta + (1-\theta)(1-\delta)} f(m_I(k))n'(k) + \frac{(1-\theta)(1-\delta)}{\delta(\delta + (1-\theta)(1-\delta))} f'(m_I(k))n(k)m'_I(k)$$

When $\delta = 1$, $V'_{W,A}(k) = f(m_A(k))n'(k) > f(m_I(k))n'(k) = V'_{W,I}(k)$. Therefore, $\exists\ \delta_2$ s.t. $\forall \delta > \delta_2$, $V'_{W,A}(k) > V'_{W,I}(k)$ and hence $V_{W,I}(k) > V_{w,A}(k)$ for $k \in [k^*_A, k^*_I]$. Hence, if we choose $\delta^* = \max\{\delta_1, \delta_2\}$, $\forall \delta > \delta^*$, $V_{W,I}(k) > V_{w,A}(k)$ for $k \in [\underline{k}, \overline{k}_H]$.