Just to Sell, or to Make There as Well? Agency Costs, Multinational Firms and The Mode of Foreign Entry

Benjamin D. Nelson*

November 22, 2009

Abstract

Multinational Corporations enter overseas markets in various ways. We focus on the distinction between Distribution oriented and Production oriented Foreign Direct Investment (FDI). Firms solve a multitask principal-agent problem incentivising imperfectly observable effort in two tasks: production and distribution. Firms optimally assign different tasks to different agents, possibly separating production and distribution across firm boundaries. Firms specialise in their mode of foreign entry according to their productivity and skill intensity, while the host country attracts a mix of FDI modes. Firms engaging in distribution oriented projects are smaller than those involved in production oriented projects. Skill intensity, worse effort monitoring in the South, and smaller foreign markets encourage distribution orientation over production orientation. Trade liberalisation can increase FDI, affecting its composition. These predictions are consistent with recent evidence.

JEL Classifications: F12, F21, F23

Keywords: Multinational firms, Distribution oriented FDI, Firm Heterogeneity

1 Introduction

Multinational Corporations face an array of choices over how they serve foreign markets. Much traction has been gained by studying the choice between producing in the home country for export, and setting up overseas production through Foreign Direct Investment (FDI). Recent empirical work has revealed that multinationals face a richer set of choices than this simple dichotomy would suggest however. In particular, FDI can take different forms. This paper focuses on the choice between exporting, Distribution oriented FDI, and Production oriented FDI. We present a model in which different modes of foreign entry are characterised by differences in the severity of agency costs in two tasks: production and distribution. We argue that our model explains the various observed regularities reported in recent empirical work which remain under-explored by theory.

*Bank of England, Threadneedle Street, London. Email: benjamin.nelson@bankofengland.co.uk. I am extremely grateful to Peter Neary for valuable comments, discussion, and encouragement. Thanks are also due to Volker Nocke and Stephen Redding for an extremely useful set of comments. All errors remain my own. *Note*: the views expressed herein are those solely of the author, and do not necessarily reflect those of the Bank of England or MPC.

We describe a set of empirical findings, drawn from various papers, in more detail in Section 2. We focus on the following. First, many Multinational Corporations (MNCs) have foreign affiliates classified as 'wholesale affiliates', regardless of the parent firm's industry classification. We call this *Distribution oriented FDI* ("D-FDI") to distinguish it from foreign investments intended to establish production facilities, which we call *Production oriented FDI* ("P-FDI"). Data on US, Japanese and French firms reveal that these Distribution oriented affiliates typically account for 15-40% of total affiliate sales. Sales through this mode therefore represent a substantial fraction of total firm activity, and are typically the second largest source of sales after those through affiliates which share the parent's industry. Second, MNCs appear to serve a given foreign market either through D-FDI or through P-FDI, with a minority serving a given foreign market through both. Firms that have both production and distribution oriented foreign affiliates in the same country are engaged in *Integrated FDI* ("I-FDI"). Thus Integrated FDI accounts for a minority of overseas investments. Third, host countries appear to receive both Distribution and Production oriented FDI. So while firms specialise, undertaking different modes in different countries, the host countries themselves do not. Fourth, in Hanson, Mataloni and Slaughter's (2001) sample of US MNCs, the share of sales through Distribution oriented affiliates (i) increases with distance, (ii) decreases in the presence of a common host-home country language, (iii) decreases with host country demand (GDP) and (iv) increases with host country GDP per capita. As noted by the authors, (i) is inconsistent with standard theories of Horizontal FDI to the extent that distance proxies for trade costs.[1] Finally, in Javorcik's (2004) data, MNCs are more likely to engage in Distribution relative to Production oriented FDI where they are more R&D intensive, smaller, and invest in countries with weak Intellectual Property Rights (IPR) regimes.[2]

This set of five empirical findings motivates our model. It is important to note that existing theories of MNCs' foreign entry modes do not address these regularities, in particular the distinction between Distribution and Production oriented FDI. Some theories are inconsistent with them. Further, as noted by Neary (2009), an observed positive association between trade liberalisation and FDI is potentially inconsistent with the standard 'export versus horizontal-FDI' model. Our model addresses this too.

We posit that MNCs must solve a multitask Principal-Agent problem in serving foreign markets. In particular, firms must incentivise risk-averse workers whose actions can be monitored only imperfectly to exert effort on two tasks: production and distribution. As in Holmstrom & Milgrom (1991), we show that firms optimally assign different tasks to different agents, since by doing so they can reduce the costs of providing workers with income insurance in the face of moral hazard. This provides us with a theoretical justification for firms to separate tasks both across countries and across firm boundaries. In particular, we allow Northern MNCs to locate production either at home or in the South, and to serve the Southern market either by setting up an in-house wholesale affiliate there or by licensing distribution to an outside agent. When the firm produces in the North and licenses

[1] Broadly, these predict a *proximity–concentration trade-off*. Higher trade costs increase the marginal benefit of locating close to a foreign market, but this foregoes the economies of scale associated with producing in a single location, which arise in the presence of fixed costs of production. In this way, higher trade costs are supposed to encourage horizontal FDI. See e.g. Brainard (1993).

[2] Previous work has shown that MNCs are less likely to engage in overseas production relative to exports if they are R&D intensive, less productive, and when trade costs are lower. See Brainard (1997), Helpman, Melitz & Yeaple (2004), Norbäck (2001).

distribution, it engages in exporting. If it moves distribution in-house but retains production in the North, it engages in Distribution oriented FDI. If production is moved to the South, but an outside agent is licensed to distribute, the firm undertakes Production oriented FDI. Finally, if both production is located in the South and distribution is undertaken in-house, the firm engages in Integrated FDI.

To match other empirical observations, we allow for two sources of firm heterogeneity. First, firms in different sectors vary according to the skill intensity of their production techniques. This allows firms' vulnerability to weak IPR enforcement to vary. Only the North is endowed with skilled labour, which is used by MNCs to produce product components for final goods from blueprints. We posit that IPR protection is weaker in the South, requiring firms to expend more skilled labour resources per unit of output when production is located offshore in order to preserve their product differentiation. Second, firms have heterogeneous productivities. We embed our multitask Principal-Agent microfoundations in a model of monopolistic competition with firm heterogeneity in order to explore the correlation between firm characteristics, such as size, and modes of foreign entry revealed in the data.

Consistent with the empirical evidence, our model predicts that firms tend to choose either Production or Distribution oriented FDI, but rarely both through Integrated FDI. Increases in firm skill intensity or a weakening of IPR protection in the South bias firms' entry choices towards Distribution FDI and away from modes under which production takes place in the South. Our model predicts that firms undertaking Distribution oriented FDI will be smaller and less productive than those undertaking production in the South. Finally, we show that trade liberalisations can give rise to increases in FDI, in contrast to the standard model of horizontal FDI, which is accounted for by a rise in D-FDI relative to entry modes under which production takes place in the South. Trade costs therefore affect the composition of FDI, not necessarily reducing its level.

1.1 Related literature

Multinational Corporations, and in particular foreign investment, have long been of theoretical and empirical interest, as have the issues surrounding the internalisation decisions of such firms.[3] Dunning (1981) structures a firm's foreign investment decision around the 'OLI' framework, which is used to assess the 'Ownership' (O), 'Location' (L) and 'Internalisation' (I) advantages of different modes of foreign entry. More formal models have been constructed around this framework, emphasising the importance of 'knowledge capital' (Markusen 2002), or knowledge-based assets that create firm level scale economies.[4] Recent work has also focussed on different mechanisms through which Foreign Direct Investment can occur, distinguishing 'greenfield' FDI from cross-border mergers and acquisitions.[5] Rather than focussing on this distinction, our model emphasises that foreign investments are oriented towards particular types of activity, and that there are systematic reasons behind firms' choices over modes of foreign entry.

[3] See, for example, Caves (1971), on FDI.

[4] See, *inter alia*, Markusen (1984), Helpman (1984), Helpman (1985), Horstmann & Markusen (1987*a*), Horstmann & Markusen (1987*b*), Horstmann & Markusen (1996), Brainard (1993), Ethier & Markusen (1996), Markusen (1997), Markusen & Venables (1998) and Markusen & Venables (2000).

[5] See in particular Neary (2007), Nocke & Yeaple (2007), and Nocke & Yeaple (2008).

Various papers have explored the 'black box' of firms' internalisation and location decisions by taking an approach microfounded in the informational and contractual frictions inherent in market-based transactions. In our context, Ethier (1986) calls this "the international economics of information" (p. 808). We take an agency theoretic approach, as do Horstmann & Markusen (1987*a*), Ethier & Horn (1990), Ethier & Markusen (1996) and Horstmann & Markusen (1996). In these papers, the relevant agency frictions arise from a concern for brand reputation, the problem of managerial control in large hierarchical organisations, the problem of firm-specific knowledge dissipation, and the problem of imperfect knowledge of foreign market conditions respectively. Absent precise empirical work on agency theoretic models in an international context, there is plenty of case study/anecdotal evidence in support of these mechanisms as being of concern to MNCs.[6] Our model can be interpreted to capture the first, third and forth of these sources, as they are closely related, but importantly adds firm heterogeneity to the analysis. This feature links our model's predictions directly to recent firm-level data on FDI modes.

This paper proceeds as follows. The empirical motivation is outlined in section 2. Section 3 outlines the model, section 4 presents equilibrium modes of foreign entry and section 5 completes the description of the free entry equilibrium. Section 6 considers some comparative statics, section 7 general equilibrium, and section 8 concludes.

2 Empirical motivation

We seek to explain the key empirical findings of the small literature on Distribution oriented FDI. First, Hanson, Mataloni & Slaughter (2001) explore the expansion strategies of US MNCs. They use firm-level data to examine the patterns of these firms' foreign operations, and highlight some striking results concerning Distribution oriented FDI that have gone relatively unexamined to date.[7] Second, Ando & Kimura (2003) employ Japanese firm-level data to explore the formation of production and distribution networks in East Asia and the rest of the world. Third, Defever (2006) uses firm data from 23 countries to study non-European firms' 'functional fragmentation' across Europe. Finally, Javorcik (2004) uses firm-level data to examine the modes of FDI by multinational firms in a set of transition economies. Synthesising these papers, we highlight the following findings.

The first concerns the significance of wholesale trade (or Distribution oriented) affiliates in MNCs' operations. The Standard Industrial Classification (SIC) division F defines wholesale trade as "establishments or places of business primarily engaged in selling merchandise to retailers; to industrial, commercial, institutional, farm, construction contractors, or professional business users; or to other wholesalers; or acting as agents or brokers in buying merchandise for or selling merchandise to such persons or companies. The chief functions of establishments included in Wholesale Trade are selling goods to trading establishments ... and bringing buyer and seller

[6]Some of the most interesting examples come from economic history. Navaretti & Venables (2004) cite Fitzgerald (1995)'s study of Rowntree, a well-known British confectioner (now part of Nestle). Following expansion into the South African market in 1900, by 1925 Rowntree had set up a joint venture with an overseas producer. Divergent views emerged when the licensed distributor followed a low cost, low effort marketing strategy, in contrast to the wishes of the Rowntree headquarters. The foreign operation was subsequently internalised when Rowntree acquired majority control in 1950. See also Nicholas (1983)' study of agency problems in pre-1939 British manufacturers' FDI, and Carlos & Nicholas (1990) case study of the Hudson's Bay Company's operations in the 17th and 18th centuries. They highlight a concern of early multinationals with, *inter alia*, low 'effort' by foreign-based agents, including, according to Willan (1956) "horedom [sic], incontinency, drunkenness and idleness".

[7]The authors examine foreign affiliates in wholesale trade, termed Distribution oriented FDI.

together."[8] Data from Hanson et al, reproduced in Table 1 (in appendix B), show that for US MNCs in 1998, between around 10-40% of total affiliate sales were accounted for by wholesale trade affiliates. These proportions were highest for Industrial Machinery (38%) and Electrical Equipment (28%). Wholesale affiliates were typically the largest source of affiliate sales after those that shared the same industry classification of the parent company. Table 2 (in appendix B) is taken from Ando and Kimura (2003) and illustrates the channels through which Japanese firms served foreign markets in 2000. The table shows that around 40% of Japanese MNCs' value added was accounted for by production in Japan combined with distribution in overseas markets through a foreign affiliate. As the table shows, this mode of foreign entry therefore explains the majority of the 50% of value added left after horizontal FDI is accounted for. Finally, Table 3 (in appendix B) taken from Defever (2006) shows the structure of new investments by European and non-European firms in the enlarged European Union for 1997-2002. 30% of new investments in the EU15 were in sales and marketing operations, the largest share after production operations which accounted for around 40% of new investments. Taken together, we suggest that theories of FDI that focus exclusively on production oriented operations exclude a significant segment of MNC activity.

Empirical Finding 1: *Wholesale trade affiliates account for a significant proportion of the total sales of Multinational Corporations, the parent firms of which do not operate in the wholesale industry themselves*.

Next, Hanson et al explore the mix of distribution oriented and production oriented foreign operations in their sample. They write that "[t]here are few U.S. multinationals that operate both manufacturing and wholesale-trade affiliates in the same country. For example, in 1998 533 U.S. parents had manufacturing affiliates in Canada, 259 had wholesale-trade affiliates in Canada, 60 had both a manufacturing and a wholesale trade affiliate in Canada, and 732 had either a manufacturing or a wholesale trade affiliate in Canada–but not one of each in Canada. The Canadian case is typical: in most countries only about 10% of U.S. parents choose this strategy, with an even smaller share in the developing countries". We include this apparent specialisation by firms as our next empirical finding.

Empirical Finding 2: *The majority of MNCs appear to serve foreign markets either through Distribution oriented FDI or through Production oriented FDI. A minority perform both (Integrated FDI). MNCs therefore appear to specialise in their modes of foreign entry*.

While firms appear to specialise in the their modes of foreign entry, Hanson et al note that "[c]ountries host affiliates in both manufacturing and wholesale trade, not just all one or all the other." That is, countries do not appear to specialise in the modes of foreign entry they attract to the same extent to which firms themselves choose different entry modes in different countries. They go on to note that "[t]o the extent that U.S. firms face roughly the same host-country market conditions, that finding doesn't support a simple horizontal-FDI explanation". We therefore state

[8] We quote a full SIC definition in the appendix to this chapter.

Empirical Finding 3: *Host countries receive both D-FDI and P-FDI. Thus, while firms appear to specialise, countries do not.*

We turn next to more detailed econometric evidence on MNCs' foreign activities contained in these papers. First, Hanson et al regress the share of total affiliate sales accounted for by wholesale affiliates on host country distance, a common language dummy, host country GDP and host country GDP per capita. In addition, Javorcik includes the two GDP variables as covariates in a probit model of choice of project mode, comparing the impact of these and other variables on the relative probability of distribution and production oriented projects being chosen. Together, the results suggest that

Empirical Finding 4: *(Hanson et al 2001) Sales through Distribution oriented FDI as a share of total affiliate sales appear to*

(a) *increase with distance;*

(b) *decrease with common language;*

(c) *decrease with host country demand (GDP);*

(d) *increase with host country per capita GDP.*

(Javorcik 2004) The probability of choosing Distribution oriented FDI relative to Production oriented FDI appears to

(e) *decrease with host country demand (GDP);*

(f) *increase with host country GDP per capita.*

The increase in sales through D-FDI with distance is inconsistent with existing theories of horizontal FDI to the extent that distance proxies for trade costs. Of course, distance could be capturing something other than trade costs, on which we have more to say below. The negative correlation with common language has no obvious theoretical foundation in existing theories. We suggest below however that common language could capture a dimension of monitoring costs, such that an agency theoretic treatment of FDI grants this finding a new interpretation.

Finally, analysis of the correlations between firm-level attributes and project modes undertaken in Javorcik (2004) provide some finer detail on the firm characteristics that correlate with choices of foreign entry mode. We summarise these findings as

Empirical Finding 5: *(Javorcik 2004) MNCs are less likely to engage in P-FDI relative to D-FDI when they*

(a) *are more R&D and advertising intensive;*

(b) *are smaller;*

(c) *invest in countries with weak IPR protection regimes.*

These systematic patterns in the choices of FDI modes suggest firm-level heterogeneity may be an important predictor of foreign entry mode choice. Next we develop a model that aims to capture the empirical findings described above, before examining some of the model's further predictions.

3 The model

There are two countries, labelled North and South. Each is endowed with a fixed amount of unskilled labour L_k, where k is a country index $k = N, S$. Only the North is endowed with skilled labour K, and only Northern firms produce differentiated varieties.

3.1 Preferences

All agents have the following CARA utility function which they maximise subject to a budget constraint. We specify utility over the outside good q_0 and CES aggregates of differentiated products $Q(z)$ in sector $z \in [0, 1]$ as

$$U = -e^{-\kappa u(.)}, \qquad (1)$$
$$u(.) = q_0 + \mu \int_0^1 \xi(z) \ln Q(z) dz, \qquad (2)$$

where $\int_0^1 \xi(z)dz = 1$, $\kappa > 0$ is the coefficient of absolute risk aversion, $\mu > 0$ determines the relative preference for differentiated varieties, and $Q(z) = \left[\int_{\gamma \in \Gamma_z} q(\gamma)^\rho d\gamma\right]^{1/\rho}$. Γ_z is the set of varieties produced in sector z. We choose the price of the outside good p_0 as the numeraire ($p_0 = 1$). For an agent with nominal income y, the maximisation problem yields demands for the outside good and for a firm in sector z producing differentiated variety γ as $q_0 = y - \mu$ and $q(z, \gamma) = \theta(z)p(z, \gamma)^{-\sigma}$ respectively. $p(.)$ is the price charged by the firm producing variety (z, γ), and $\theta(z) \equiv \frac{\mu\xi(z)}{P(z)^{1-\sigma}}$ captures demand conditions. The term $P(z)$ is the CES price index for sector z, given by $P(z)^{1-\sigma} \equiv \int_{\gamma \in \Gamma_z} p(z, \gamma)^{1-\sigma} d\gamma$.

We substitute the demand functions into the utility function in order to give the indirect utility function in terms of real income $m = y - P$, given by

$$U = -e^{-\kappa m}, \qquad (3)$$

where $P \equiv \mu \int_0^1 \xi(z) \ln P(z)dz + \widetilde{\mu}$ is a composite price index in which $\widetilde{\mu}$ is a constant.[9] If real income m is normally distributed, then its certainty equivalent is $CE[m] = E[m] - \frac{1}{2}\kappa\text{var}[m]$. This certainty equivalent representation will be used to construct workers' participation constraints in the agency problem that follows. Straightforwardly, an increase in the variance of real income m reduces the certainty equivalent value of that income. In other words, as m becomes more risky, it becomes less attractive to the agent, all else being equal.

[9] It is defined as $\widetilde{\mu} \equiv \mu \int_0^1 \xi(z) \ln [\mu\xi(z)]\, dz - \mu$.

3.2 Production

3.2.1 Outside sector

Production in the outside good sector is perfectly competitive, and employs only unskilled labour. The outside good is produced by both North and South and is freely traded. As such, wage rates w_k, $k = N, S$ for unskilled labour in the North and South are fixed according to unskilled labour productivity in that sector. Northern labour is more productive than Southern labour in the outside sector, so that $w_N > w_S$.

3.2.2 Entry, technology and skilled labour

There are $N(z)$ Northern firms producing for the domestic market in each sector, a subset of which will choose to enter the Southern market in equilibrium. They produce differentiated varieties by employing skilled and unskilled labour together with their firm-specific technology. This technology is characterised by two parameters. Productivity φ is drawn independently from the identical distribution $G(\varphi)$ such that $\varphi \in [0, \infty)$. Further, the technology has a skill intensity $z \in [0, 1]$. z is used as a sector index, where sectors with higher z's require more skilled labour per unit of output. The $\{\varphi, z\}$ pair is the firm's *blueprint*. Firms are 'born' knowing their sector index z, and pay an entry cost f_E to learn their productivity draws φ.

Skilled labour uses the blueprint $\{\varphi, z\}$ to produce a component for the firm's final good, which is assembled by unskilled labour. The former is the source of product differentiation: without the use of skilled labour, a firm's output cannot be differentiated. When this is the case, firms produce the outside good and make zero profits in equilibrium. They therefore strictly prefer to employ skilled labour to differentiate their output wherever it gives rise to positive profits.

The cost of employing skilled labour with intensity z is denoted by the function $r_k(z)$. This contains a unit input requirement z and a skilled wage rate r. The function $r_k(z)$ depends on the country k in which the component produced by skilled labour is employed however. In particular, while $r_N(z) = zr$, $r_S(z) = \beta(z)zr$ where $\beta(z) \geq 1$. That $\beta(z) \geq 1$ embodies the view that MNCs find it costly to employ their technology in foreign countries in which IPR protection is weaker. Firms must then take steps to protect the knowledge embodied in their components, for example through encoding proprietary information, instructions and procedures for assembly, and so on, by using extra skilled labour which is costly. This feature is motivated by e.g. Dietz, Lin & Yang (2005), who describe steps taken by 10 MNCs studied to protect their IP in China. In particular, they write that "[t]he most successful companies ... take strategic and operational action to protect their IP ... One large equipment manufacturer designs and develops hardware in China but produces the related software (in this case, the most valuable IP) abroad. The software, with its source code hidden, is delivered to Chinese engineers ready to plug into the system. By separating functions and keeping technological details secret in this way, the manufacturer significantly reduces the possibility of an IP leak. [But] [d]eveloping software in a country with better IP protection and then transporting it to China adds time, costs, and complexity to the process." That $r_S(z) > r_N(z)$ reflects this view. The ratio $\frac{r_S(z)}{r_N(z)} = \beta(z)$ varies exogenously on $[1, \infty)$, in order to vary the relative cost of using components in production abroad relative to home.[10]

Final production and sales require unskilled labour. The employment of unskilled labour is subject to an agency problem however, the severity of which varies across different foreign *entry modes*. This implies that variable costs vary by entry mode. The agency problem is specified next.

3.2.3 The agency problem

The production and distribution of differentiated varieties requires that the effort of unskilled labour be exerted in the two tasks, denoted p and d. Effort e^t in each task $t = p, d$ can be monitored only imperfectly, and monitoring varies across entry modes. Higher effort in each task reduces the firm's unit labour costs, but is the source of a private disutility to workers. This disutility of effort is convex in the total effort exerted over all tasks according to $c(e) = \frac{1}{2}\left(\sum_{t=p,d} e^t\right)^2$ and is denoted in units of the numeraire good.

Effort in production can be thought to correspond to the diligence with which the component is transformed into the firm's final good. In this sense, higher effort in production reduces the firm's unit costs by reducing errors and wastage or increasing the productive time spent on the job. Similarly, effort in distribution can be thought to contribute positively to the efficiency with which final goods reach the end consumer, for example through improvements in marketing, the organisation of sales and after sales service. Hence distribution is treated as a task that must be performed along with production in order that the firm can serve the end consumer. For simplicity, we model distribution as a component of the firm's unit costs, alongside costs arising from production.[11] Hence higher effort in each task reduces unit costs and so contributes positively to profits.

Consider an unskilled labour 'production-distribution unit' consisting of two workers, indexed by $J = i, j$. The firm observes independent, verifiable noisy signals of the total effort exerted in each task $t = p, d$.[12] Task t's signal x^t contains noisy information on effort given by

$$x^t = e^t + \varepsilon^t, \qquad \varepsilon^t \sim N(0, v^t), \qquad t = p, d, \tag{4}$$

where $e^t = \sum_{J=i,j} e_J^t$ is the total effort exerted in task t. Profits depend on the realisation of these signals net of any wages paid to workers. We specify per unit labour costs $C = r_k(z)\left[\frac{\widetilde{\omega}^p(x^p)}{\varphi} + \frac{\widetilde{\omega}^d(x^d)}{\varphi}\right]$, where the unit wage costs $\widetilde{\omega}^p(x^p)$ and $\widetilde{\omega}^d\left(x^d\right)$ are functions of the effort signals x^p and x^d that the firm observes, and are scaled by

[10] Any trade costs associated with the transportation of components used in production and produced by skilled labour can be captured by $\beta(z)$. Where the components are interpreted as pure knowledge for example, the trade cost component of $\beta(z)$ is small.

[11] There are a number of possible alternatives to this approach. One would be to model distribution as affecting the share of the firm's residual demand it is able to access, therefore affecting variable profits. Though we do not explore this modelling possibility in detail, we conjecture that the results of such an approach would be analogous to the case in which distribution costs affect the firm's unit costs, which we explore here. The latter approach has the advantage of tractability in the context of the agency model we explore in what follows, though the former is likely to be worthy of attention in future work.

[12] This amounts to the view that there is a component of costs that is not perfectly observed and that x^t is the only verifiable information upon which the firm can contract. For example, production could be subject to unforseen interruptions, lowering productivity and raising costs, but this is not directly observable by the firm: low productivity could also be due to low effort. Similarly, distribution could be subject to random congestion effects, reducing the efficiency with which final goods reach the end consumer. The firm does not observe this directly however, and inefficient distribution could equally be due to low effort in sales and distribution.

firm productivity φ. This reflects the fact that (a) higher worker effort generates higher effort signals, increasing labour productivity and so reducing unit costs, and (b) that the firm may contract upon the signals x^t in order to induce higher worker effort (which is not directly observable). For tractability, we specify these unit wage costs as linear functions of the wage contract W_J offered to workers, and the effort signal x^t observed on each task, such that

$$\widetilde{\omega}^p\left(x^p\right)+\widetilde{\omega}^d\left(x^d\right)=\sum_{J=i,j} W_J-\sum_{t=p,d} x^t. \tag{5}$$

Clearly, this assumption implies that unit costs increase linearly with the sum of the wage contracts offered to the two workers in each production-distribution unit, but that unit costs decrease linearly with the sum of the observed effort signals. Firms then choose the parameters of the wage contracts W_J so as to minimise expected unit costs, $C=\frac{r_k(z)}{\varphi}\left[\sum_{J=i,j} W_J-\sum_{t=p,d} x^t\right]$.

Firms offer linear wage contracts of the form $W_J=a_J+\sum_{t=p,d} b_J^t x^t$.[13] We refer to the a_Js as the 'insurance' parameter and the b_J^ts as the 'piece rate' parameters, all to be chosen by the firm. A higher piece rate for worker J on task t links that worker's pay more strongly to the observed signal x^t of total effort on that task. A higher a_J provides the worker with more income 'for sure'.

Contract constraints A firm's choice of contract parameters is subject to two constraints. The first of these is that the wage offered to a given worker must be Individually Rational for the worker to accept. This requires that the certainty equivalent wage offered under employment by the firm must equal that available under employment in the outside good sector. Using the certainty equivalent formulation above, this requires

$$a_J+\sum_{t=p,d} b_J^t e^t-\frac{1}{2}\left(\sum_{t=p,d} e_J^t\right)^2-\frac{1}{2}\kappa\sum_{t=p,d}\left(b_J^t\right)^2 v^t=w_k, \tag{6}$$

where we have used that var$\left[a_J+\sum_{t=p,d} b_J^t x^t\right]=\sum_{t=p,d}\left(b_J^t\right)^2 v^t$. This defines the Individual Rationality (IR) constraint, which requires that a_J satisfies

$$a_J^*=w_k-\sum_{t=p,d} b_J^t e^t+\frac{1}{2}\left(\sum_{t=p,d} e_J^t\right)^2+\frac{1}{2}\kappa\sum_{t=p,d}\left(b_J^t\right)^2 v^t. \tag{IR}$$

This says that the insurance parameter of the wage contract offered workers must compensate them for forgone income in the outside good sector (w_k), the private cost of exerted effort $\frac{1}{2}\left(\sum_{t=p,d} e_J^t\right)^2$, and a risk premium $\frac{1}{2}\kappa\sum_{t=p,d}\left(b_J^t\right)^2 v^t$ which varies with the riskiness of the wage contract offered.

The second constraint the firm faces is that workers respond optimally to the piece rate offered when making their private effort choices, congnizant of the fact that their effort choices will affect the signals observed by the firm. Workers maximise their certainty equivalent wage under employment by the firm through their choice of

[13] We do not derive the optimal shape of wage contracts from first principles. See, however, Edmans & Gabaix (2009) for a recent exploration of the optimality of the linear contracts used here.

effort, giving Incentive Compatibility (IC) constraints of the form

$$b_J^t = e_J^p + e_J^d, \qquad t = p, d, \qquad J = i, j. \tag{IC}$$

These constraints relate the optimal effort choices of workers to the piece rates offered by the firm.

Task separation We show first that under the above conditions:

Lemma 1 *The firm optimally assigns a single task to each agent, setting, without loss of generality,* $b_i^p = 0$, $b_i^d > 0$*, and* $b_j^p > 0$, $b_j^d = 0$.

Proof. See Appendix C.1. ■

This result is obtainable in a wider class of models, as shown by Holmstrom and Milgrom (1991), but applies given the set up of the model used here as well. We provide proof in the appendix, but sketch the intuition of the result here. Consider the firm facing the problem of how best to assign tasks to the two agents, for a given total amount of effort to be exerted. If each agent is given positive incentives (positive piece rates, b_J^t) on both tasks, the variance of the wage payment to each is

$$\text{var}[W_J] = \left(b_J^d\right)^2 v^d + \left(b_J^p\right)^2 v^p. \tag{7}$$

This variance is increasing in the piece rate attached to each task b_J^t. In other words, as the number of tasks on which the agent receives positive piece rates falls, so does the variance of her wage. From IR, all else being equal, this increases the certainty equivalent of the wage contract for the worker when the worker is sufficiently risk averse. When this is the case, in order that IR continues to bind, the firm can then offer a lower 'insurance' parameter a_J, since the firm has to offer the worker less income 'for sure' in order to leave her indifferent between employment by the firm and in the outside good sector.[14] In other words, the firm saves on a 'fixed cost' associated with providing income insurance to the worker if she is assigned fewer tasks. Of course, in order to maintain total effort at a constant level, the firm must still offer positive incentives on at least one task for each agent. In the simple two task case considered here, it is therefore optimal to assign a single task to each agent, by setting, without loss of generality, $b_i^p = 0$, $b_i^d > 0$, and $b_j^p > 0$, $b_j^d = 0$.

This result is particularly interesting in our case, as it implies, *inter alia*, that tasks are not only optimally separated between workers, but might also be optimally separated across firm- and national boundaries. The multitask principal-agent microfoundation therefore provides a strong theoretical rationale for firms to fragment their production processes internationally in a way that optimally mitigates the agency frictions they face in organising production.

[14] In particular, $\frac{\partial a_I^*}{\partial b_I^t} > 0$ if $\kappa > \frac{e^t}{b_I^t v^t}$.

Optimal contract parameters Given that firms optimally assign single tasks to each worker, how should incentives be chosen? The expected variable costs of producing q units of output are

$$\frac{r_k(z)}{\varphi}\left[\sum_{J=i,j} a_J + \sum_{t=p,d} b^t e^t - \sum_{t=p,d} e^t\right] q.$$

The firm's problem is then to solve

$$\min_{b^p,b^d} \frac{r_k(z)}{\varphi}\left[\sum_{J=i,j} a_J + \sum_{t=p,d} b^t e^t - \sum_{t=p,d} e^t\right] q, \tag{8}$$

$$\text{s.t. } a_J = a_J^*, \qquad J = i, j, \tag{9}$$

$$\text{s.t. } e^t = b^t, \qquad t = p, d. \tag{10}$$

Substituting in the IR and the IC constraints and solving the resulting first order conditions for b^t gives the optimal piece rate as

$$b^{t*} = \frac{1}{1+\kappa v^t}, \qquad t = p, d. \tag{11}$$

The strength of incentives is therefore decreasing in worker risk aversion and the variance of measured effort on task t. When either of these variables rise, a given piece rate imposes more risk on the worker's income, reducing certainty equivalent income. In order that their certainty equivalent income be restored to satisfy the IR constraint, the firm must offer the worker an income less exposed to risk. The firm does this by reducing the piece rate on the relevant task, so as to reduce the sensitivity of the worker's pay to the effort signal x^t.

Having chosen the cost minimising contract parameters, the firm chooses a price so as to maximise profits. If variable costs are C, with price equal to $\frac{C}{\rho\varphi}$, variable profits are $\theta(z)\overline{\rho}C^{1-\sigma}\varphi^{\sigma-1}$, where $\overline{\rho} \equiv (1-\rho)\rho^{\sigma-1}$. Using the optimal piece rate and insurance parameters, variable profits are therefore

$$\theta(z)\overline{\rho}r_k(z)^{1-\sigma}\left[\omega_k^p + \omega_k^d\right]^{1-\sigma}\varphi^{\sigma-1}, \tag{12}$$

where $\omega_k^t \equiv w_k - \frac{1}{2}b^{t*}$ is the agency cost on task t net of worker effort, or task t's *net agency cost*. Ex post variable profits are clearly decreasing in net agency costs by $\sigma > 1$. Using (11), we have

Lemma 2 *The optimal piece rate for each task is decreasing in the noise with which effort on that task is monitored. Net agency cost on task t is*

1. *increasing in the variance of measured effort on task t:* $\frac{\partial \omega_k^t}{\partial v^t} > 0$;
2. *increasing in the worker's outside option wage rate:* $\frac{\partial \omega_k^t}{\partial w_k} > 0$;
3. *increasing in worker risk aversion:* $\frac{\partial \omega_k^t}{\partial \kappa} > 0$.

Proof. By $\frac{\partial b^{t*}}{\partial v^t} < 0$, $\frac{\partial b^{t*}}{\partial \kappa} < 0$ and $\frac{\partial \omega^t}{\partial b^{t*}} < 0$, $\frac{\partial \omega_k^t}{\partial w_k} > 0$ (See Appendix C.2). ■

3.2.4 Entry modes

We now introduce two ways in which each task can be performed. In particular, production can be located in the North (N) or the South (S), and the distribution of final goods in the South can be *Licensed* (L) to an outside agent or *Integrated* (I) within the firm. The ability of the firm to monitor effort under each of these arrangements differs, which affects the optimal wage contract parameters and hence variable costs. Effort monitoring under production in the South is harder than under production in the North due to the added difficulties associated with lack of common language, knowledge of local contingencies, business cultures and so on. There is therefore a negative 'border effect' due to international boundaries on effort monitoring. Second, monitoring the effort of an outside agent under licensed distribution is more difficult than monitoring that of an agent undertaking distribution within the boundaries of the firm. There is therefore a negative 'firm boundary effect' associated with moving distribution to an agent outside the firm. Formally, these amount to the assumptions that

$$v_S^p > v_N^p, \qquad v_L^d > v_I^d, \tag{13}$$

where the subscript on each v^t relates to the mode under which that activity is undertaken. As above, worse monitoring implies a given piece rate imposes more income risk on the agent, such that for individual rationality constraints to be satisfied, the strength of incentives must be reduced. This lowers effort and hence raises costs.[15]

Since distribution must always take place in the South in serving Southern consumers, agents involved in distribution operations have the same outside option wage w_S. But for production workers, the outside option wage varies according to the country in which production takes place. Straightforwardly, Northern workers have outside option w_N, and Southern worker have outside option w_S.

If production is located in the North and distribution is licensed to an outside agent, the Northern firm is said to *Export* (X) to the South. Exporting incurs a fixed export cost denoted in units of Northern labour $f_X w_N$, and a variable iceberg trade cost $\tau > 1$, giving export profits of[16]

$$\pi^X(\varphi, z) = \theta(z)\bar{\rho}\tau^{1-\sigma} r_N(z)^{1-\sigma} \left[\omega_N^p + \omega_L^d\right]^{1-\sigma} \varphi^{\sigma-1} - f_X w_N. \tag{14}$$

If production is retained in the North but distribution is moved in-house under Integrated distribution, the Northern firm is said to perform *Distribution oriented FDI* (D). In this case, foreign investment is concentrated solely in establishing a fully owned sales affiliate located in the South, which incurs fixed cost $f_D w_S$ denoted in units

[15] If instead the firm offered a fixed payment to the licencee under Licenced distribution, the results that follow would not change as long as this raised variable costs relative to the case in which distribution is performed in-house. This seems reasonable. In particular, a fixed payment offers no effort incentives to workers under unobservable action - so such a contract would induce lower effort, and hence higher costs than under integrated distribution under which effort incentives are offered.

[16] Note that this implies that variable trade costs τ operate on all labour inputs, including those engaged in distribution activities in the South. In this way, trade costs uniformly dampen firm productivity across all tasks (since, in this case, variable profits are proportional to $\left(\frac{\varphi}{\tau}\right)^{\sigma-1}$), such that trade costs are incurred on all components of value added. An alternative is to model a trade tax factor $\tilde{\tau} > 1$ as impinging on only unskilled labour production costs in the North, such that variable profits under exporting become $\theta(z)\bar{\rho}r_N(z)^{1-\sigma}\left[\tilde{\tau}\omega_N^p + \omega_L^d\right]^{1-\sigma}\varphi^{\sigma-1}$. We detail this case in a supplementary appendix, and show that the results we present in the body of the remainder of this paper do not change under this alternative assumption.

of Southern labour. We assume that f_D is sufficiently large relative to f_X such that $f_D w_S > f_X w_N$. Since production is still located in the North, final goods must be shipped to the South incurring variable trade cost $\tau > 1$ in order to be distributed.[17] Distribution oriented FDI therefore yields profits of

$$\pi^D(\varphi, z) = \theta(z)\overline{\rho}\tau^{1-\sigma} r_N(z)^{1-\sigma} \left[\omega_N^p + \omega_I^d\right]^{1-\sigma} \varphi^{\sigma-1} - f_D w_S. \tag{15}$$

If instead production is shifted to the South but distribution is still performed by a licensed agent, the Northern firm performs *Production oriented FDI* (P). Since distribution is performed by an outside agent, foreign investment is concentrated solely in establishing a fully owned production affiliate in the South, which incurs fixed costs $f_P w_S$. It seems reasonable that the fixed costs associated with setting up a plant dedicated to production in the South should exceed those associated with setting up a distribution only operation in the South, so we posit that $f_P > f_D$. Producing in the South saves on variable trade cost τ, but implies that the firm must take extra steps to preserve its blueprint, raising skilled labour costs to $r_S(z) > r_N(z)$. Profits under Production oriented FDI are therefore

$$\pi^P(\varphi, z) = \theta(z)\overline{\rho} r_S(z)^{1-\sigma} \left[\omega_S^p + \omega_L^d\right]^{1-\sigma} \varphi^{\sigma-1} - f_P w_S. \tag{16}$$

Finally, if production is located in the South, saving on trade costs, and distribution is performed by setting up a fully owned distribution affiliate, the Northern firm performs *Integrated FDI* (I). Here, the firm incurs both fixed costs f_P and f_D, denominated in units of Southern labour, but saves on trade costs in establishing a plant in the South to perform both tasks. Profits under Integrated FDI are

$$\pi^I(\varphi, z) = \theta(z)\overline{\rho} r_S(z)^{1-\sigma} \left[\omega_S^p + \omega_I^d\right]^{1-\sigma} \varphi^{\sigma-1} - (f_P + f_D)\, w_S. \tag{17}$$

We summarise the various modes of foreign entry in Table 4, and the fixed and variable costs associated with each in Table 5. Next, we compute a Northern firm's optimal choice of foreign entry mode, conditional on its sector z and productivity φ.

TABLE 4: Foreign entry modes

		Distribution Mode	
		Integrated	Licensed
Production location	North	*D*	*X*
	South	*I*	*P*

[17] As in the case of exporting, trade costs operate on all components of value added under this assumption. As above, see the supplementary appendix for a discussion of the case in which trade costs impinge only on unskilled labour production costs in the North. We show that the substantive results presented in what follows are not sensitive to this alternative formulation.

TABLE 5: Cost structure

Organisational Form	*Fixed cost*	*Variable cost*
X	$f_X w_N$	$\frac{\tau r_N(z)(\omega_N^p+\omega_L^d)}{\varphi}$
D	$f_D w_S$	$\frac{\tau r_N(z)(\omega_N^p+\omega_I^d)}{\varphi}$
P	$f_P w_S$	$\frac{r_S(z)(\omega_S^p+\omega_L^d)}{\varphi}$
I	$(f_P+f_D)\, w_S$	$\frac{r_S(z)(\omega_S^p+\omega_I^d)}{\varphi}$

4 Equilibrium foreign entry mode

Consider how firms will sort into organisational forms as productivity changes, for a sector of given skill intensity z. In so doing, assume

$$\tau > \left(\frac{\omega_S^p+\omega_L^d}{\omega_N^p+\omega_I^d}\right)\beta(z). \tag{18}$$

Then

Proposition 3 *There exist productivity cut-offs $\varphi^X(z)$, $\varphi^D(z)$, $\varphi^P(z)$ and $\varphi^I(z)$ such that*

(a) *The least productive Northern firms, for which $\varphi < \varphi^X(z)$, do not enter the Southern market;*

(b) *Of the firms that do enter the Southern market, the least productive enter through exporting. For these firms $\varphi^X(z) < \varphi < \varphi^D(z)$;*

(c) *Firms entering the Southern market through Distribution oriented FDI are more productive than those that enter through exporting. For these firms $\varphi^D(z) < \varphi < \varphi^P(z)$. This follows from $f_X w_N < f_D w_S$, and from $v_I^d < v_L^d$, which implies $\omega_I^d < \omega_L^d$;*

(d) *Firms entering the Southern market through Production oriented FDI are more productive than those entering through Distribution oriented FDI. For these firms $\varphi^P(z) < \varphi < \varphi^I(z)$. This follows from $f_P > f_D$ and for sufficiently high variable trade costs τ (the satisfaction of (18));*

(e) *Firms entering the Southern market through Integrated FDI are more productive than those entering through Production oriented FDI. For these firms $\varphi^I(z) < \varphi$. This follows from $f_D > 0$ and from $v_I^d < v_L^d$, which implies $\omega_I^d < \omega_L^d$.*

Proof. (a) When $\varphi < \varphi^X(z)$, profits on Southern operations are negative; The ordering proposed in (b)—(e) follows when

i. $\frac{\tau r_N(z)\left(\omega_N^p+\omega_L^d\right)}{\varphi} > \frac{\tau r_N(z)\left(\omega_N^p+\omega_I^d\right)}{\varphi}$ and $\frac{r_S(z)\left(\omega_S^p+\omega_L^d\right)}{\varphi} > \frac{r_S(z)\left(\omega_S^p+\omega_I^d\right)}{\varphi}$, which both hold if $\omega_L^d > \omega_I^d$. Since $v_L^d > v_I^d$, this is the case;

ii. (18) is satisfied, such that $\frac{\tau r_N(z)\left(\omega_N^p+\omega_I^d\right)}{\varphi} > \frac{r_S(z)\left(\omega_S^p+\omega_L^d\right)}{\varphi}$;

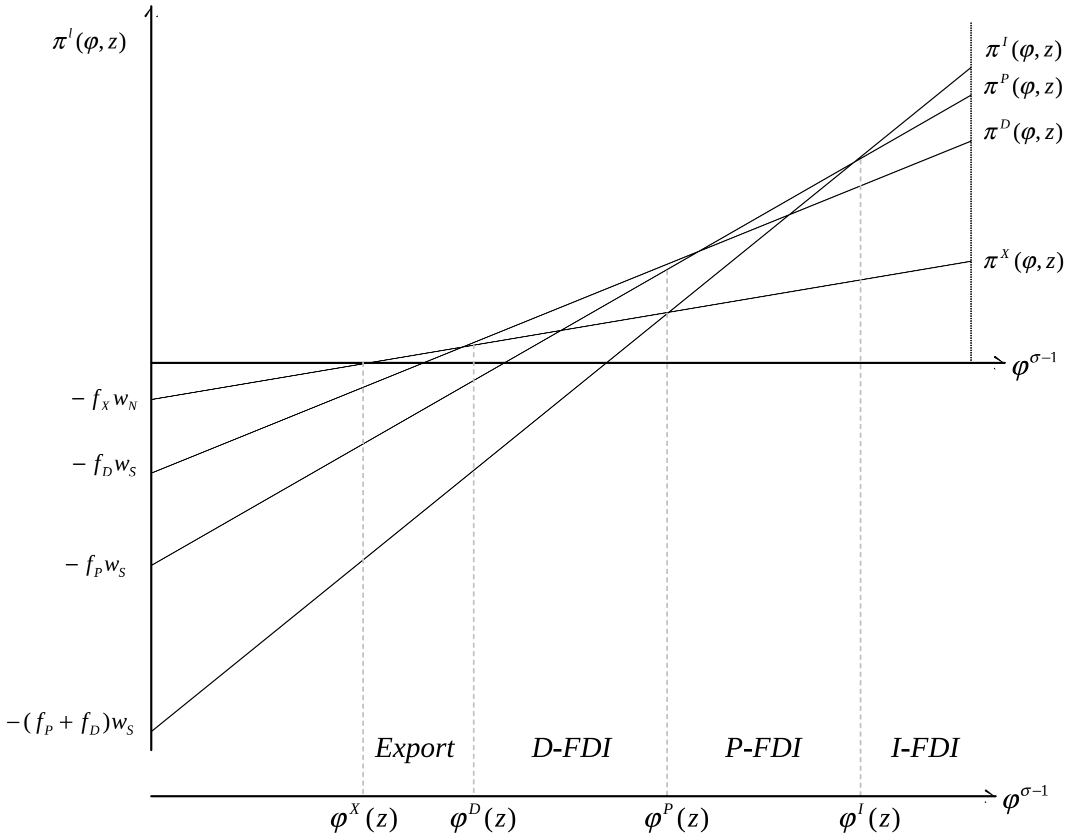

Figure 1: Profits and Southern entry modes as a function of productivity

iii. $f_X w_N < f_D w_S$ and $0 < f_D < f_P$, such that fixed costs are ranked according to $f_X w_N < f_D w_S < f_P w_S < (f_P + f_D) w_S$.

■

Proposition 3, links immediately to our empirical findings above. In particular:

- Firms specialise in different modes of FDI, as in Empirical Finding 2, which vary systematically with productivity, illustrated in Figure 1;
- The host country receives a mixture of FDI types, as in Empirical Finding 3;
- Firms engaging in Distribution oriented FDI are smaller (produce less output and generate lower revenues) than those engaging in modes of FDI involving Southern production, P and I, as in Empirical Finding 5 (b).

As in the standard horizontal FDI framework, there exists a proximity-concentration trade-off of sorts in our model. In particular, lower variable costs in the South are driven by the avoidance of trade costs, which when (18) holds offsets other location and monitoring disadvantages. On top of this trade-off however is one involving the organisation of distribution. In-house distribution features lower variable costs due to better monitoring than outsourced distribution to a licensed agent. This comes at higher fixed costs however. These higher fixed costs restrict modes of in-house distribution to the more productive firms for given production location decisions.

We derive the productivity thresholds for different organisational forms formally in the appendix (See Appendix C.3). We state them here, for a given sector z, as

$$\begin{aligned}
\varphi^X(z) &= \left(\frac{f_x w_N}{\theta(z)\overline{\rho}}\right)^{\frac{1}{\sigma-1}} \tau r_N(z)\left(\omega_N^p+\omega_L^d\right), \\
\varphi^D(z) &= \left(\frac{f_d w_S - f_x w_N}{\theta(z)\overline{\rho}}\right)^{\frac{1}{\sigma-1}} \frac{\tau r_N(z)}{\left[\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}-\left(\omega_N^p+\omega_L^d\right)^{1-\sigma}\right]^{\frac{1}{\sigma-1}}}, \\
\varphi^P(z) &= \left(\frac{f_p w_S - f_d w_S}{\theta(z)\overline{\rho}}\right)^{\frac{1}{\sigma-1}} \frac{\beta(z) r_N(z)}{\left[\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}-\left(\frac{\tau}{\beta(z)}\right)^{1-\sigma}\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\right]^{\frac{1}{\sigma-1}}}, \\
\varphi^I(z) &= \left(\frac{f_d w_S}{\theta(z)\overline{\rho}}\right)^{\frac{1}{\sigma-1}} \frac{\beta(z) r_N(z)}{\left[\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\right]^{\frac{1}{\sigma-1}}}.
\end{aligned} \tag{19}$$

By Proposition 3, we have that $\varphi^I(z) > \varphi^P(z) > \varphi^D(z) > \varphi^X(z)$. For a given skilled wage rate, we show that the sector price index can be written as $P(z)^{1-\sigma} = n(z)p\left[\widetilde{\varphi}(z)\right]^{1-\sigma}$, where $\widetilde{\varphi}(z)$ is the average productivity of MNCs in sector z (see Appendix C.4). This in turn is a function of the export cut-off $\varphi^X(z)$, which we determine by free entry, next.

5 Free entry equilibrium

We allow for the free entry of Northern firms into the Southern market. $N(z)$ Northern firms serve the domestic market, $n(z)$ of which pay the entry cost to enter the Southern market. Entry is rational when

$$\int_{\varphi^X(z)}^{\infty} \pi^l(\varphi, z)dG(\varphi) \geq f_E w_N, \tag{20}$$

in which $l = X, D, P, I$ over the relevant productivity ranges, and which must hold with equality in equilibrium. Both $\varphi^X(z)$ and $\pi^l(\varphi, z)$ are functions of sector level demand $\theta(z)$. We show that the left hand side of (20) is monotonically increasing in $\theta(z)$ (see Appendix C.5). Equilibrium sector level demand $\theta^*(z)$ is therefore uniquely determined by

$$\int_{\varphi^X[\theta^*(z)]}^{\infty} \pi^l(\varphi, \theta^*(z), z)dG(\varphi) = f_E w_N, \tag{21}$$

in which $l = X, P, D, I$ as above. With $\theta^*(z)$ determined, the price index is given by $\left[P^*(z)\right]^{\sigma-1} = \frac{\mu\xi(z)}{\theta^*(z)}$ for a given skilled wage rate, and we obtain the equilibrium foreign entry cut-off $\varphi^{X*}(z)$ by using these in (19). This pins down the number of successful entrants into the Southern market as $n^*(z) = N(z)\left\{1 - G\left[\varphi^{X*}(z)\right]\right\}$.[18]

We proceed further with our analysis under the assumption of productivities distributed according to a Pareto distribution with shape parameter δ.[19] When this is the case, we show that

Lemma 4 *With Pareto distributed productivity, the export cut-off satisfies* $\frac{\partial \varphi^{X*}(z)}{\partial \xi(z)} < 0$*, together with* $\frac{\partial \varphi^{X*}(z)}{\partial \beta(z)} < 0$ *and* $\frac{\partial \varphi^{X*}(z)}{\partial \tau} > 0$ *when (18) holds.*

Proof. See Appendix C.6. ∎

6 Cross-sector comparisons

We first consider some sector-level comparative statics, holding the skilled wage rate constant. In particular, we consider changes in a sector's skill sensitivity, monitoring technology, trade costs and IPR enforcement. To facilitate comparisons, we show in the appendix the entry mode cut-offs can be related to the export cut-off according to

$$\varphi^{l*}(z) = \left[\Lambda^l(z)\right]^{\frac{1}{\sigma-1}} \times \varphi^{X*}(z), \qquad l = D, P, I, \tag{22}$$

where the 'wedges' are

$$\Lambda^D(z) \equiv \frac{f_d w_S - f_x w_N}{f_x w_N} \frac{1}{\left(\frac{\omega_N^p + \omega_I^d}{\omega_N^p + \omega_L^d}\right)^{1-\sigma} - 1}, \tag{23}$$

$$\Lambda^P(z) \equiv \frac{f_p w_S - f_d w_S}{f_x w_N} \frac{1}{\left(\frac{\beta(z)\left(\omega_S^p + \omega_L^d\right)}{\tau\left(\omega_N^p + \omega_L^d\right)}\right)^{1-\sigma} - 1}, \tag{24}$$

$$\Lambda^I(z) \equiv \frac{f_d w_S}{f_x w_N} \frac{\left(\frac{\tau}{\beta(z)}\right)^{1-\sigma}\left(\omega_N^p + \omega_L^d\right)^{1-\sigma}}{\left(\omega_S^p + \omega_I^d\right)^{1-\sigma} - \left(\omega_S^p + \omega_L^d\right)^{1-\sigma}}. \tag{25}$$

6.1 Changes in sector skill sensitivity

The parameter $\beta(z)$ measures the extent to which the costs of employing the blueprints provided by Northern skilled labour increase when production is located in the South. We consider an increase in $\beta(z)$ to represent an

[18] Writing the price index in terms of average productivity $\widetilde{\varphi}(z)$ yields $P(z)^{1-\sigma} = n(z)p\left[\widetilde{\varphi}(z)\right]^{1-\sigma}$. Average productivity is given by

$$\widetilde{\varphi}(z)^{\sigma-1} \equiv [\tau r_N(z)]^{1-\sigma}\left[\int_{\varphi^X(z)}^{\varphi^D(z)}\left(\omega_N^p + \omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}\chi(\varphi,z)d\varphi + \int_{\varphi^D(z)}^{\varphi^P(z)}\left(\omega_N^p + \omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}\chi(\varphi,z)d\varphi\right]$$
$$+ [\beta(z)r_N(z)]^{1-\sigma}\left[\int_{\varphi^P(z)}^{\varphi^I(z)}\left(\omega_S^p + \omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}\chi(\varphi,z)d\varphi + \int_{\varphi^I(z)}^{\infty}\left(\omega_S^p + \omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}\chi(\varphi,z)d\varphi\right],$$

where $\chi(\varphi,z) \equiv g(\varphi)/\left\{1 - G[\varphi^{X^*}(z)]\right\}$ is the *ex post* productivity distribution of Northern firms entering the Southern market. This implies that $\widetilde{\varphi}(z) = \widetilde{\varphi}\left[\varphi^X(z)\right]$. Equilibrium average productivity is then given by $\widetilde{\varphi}^*(z) = \widetilde{\varphi}\left[\varphi^{X*}(z)\right]$.

[19] With $\varphi \sim$ *Pareto*(δ) on $[1,\infty)$, we have

$$g(\varphi) = \frac{\delta}{\varphi^{\delta+1}},$$

where $\delta > \sigma - 1$ is required for the size distribution of firms to have finite mean in equilibrium.

increase in a sector's *skill sensitivity*. It is straightforward that

$$\frac{\partial \Lambda^D(z)}{\partial \beta(z)} = 0, \quad \frac{\partial \Lambda^P(z)}{\partial \beta(z)} > 0, \quad \frac{\partial \Lambda^I(z)}{\partial \beta(z)} > 0, \tag{26}$$

such that increases in skill sensitivity increase the Production oriented FDI cut-off relative to the Distribution oriented FDI cut-off, expanding the range of productivities over which D is chosen relative to modes under which production is located in the South. In other words, the probability of entering through Distribution oriented FDI relative to modes under which Southern based production takes place increases as a sector's skill sensitivity increases.[20,21]

An increase in a sector's skill sensitivity also has an endogenous impact on the level of the export cut off. In particular, by Lemma 4, increased skill sensitivity reduces the export cut-off, allowing firms of lower productivity to serve the Southern market. By $n^*(z) = N(z)\left\{1 - G\left[\varphi^{X*}(z)\right]\right\}$ this has the further implication that more firms enter the Southern market in sectors of high skill sensitivity. The intuitive reason is that higher skilled labour costs in such sectors lead to higher prices, which blunt competition and allow less productive firms to profitably enter the Southern market.

These changes also translate into analogous changes in the relative shares of affiliate sales. Total affiliate sales under Distribution oriented FDI can be written

$$R^D(z) \equiv \int_{\varphi^D(z)}^{\varphi^P(z)} p^D(\varphi, z) q(z) n(z) \chi(\varphi, z) d\varphi,$$

where $\chi(\varphi, z) \equiv g(\varphi)/\left\{1 - G\left[\varphi^{X*}(z)\right]\right\}$ is the *ex post* productivity distribution of firms entering the Southern market. Using analogous expressions for affiliate sales through Production- and Integrated-FDI, we can write

$$\frac{R^D(z)}{R^P(z) + R^I(z)} = \left(\frac{\beta(z)}{\tau}\right)^{\sigma-1} \frac{V^D(z)}{V^P(z) + V^I(z)}, \tag{28}$$

[20] This can be seen straightforwardly as follows. With the Pareto distribution function $G(\varphi) = 1 - \left(\frac{1}{\varphi}\right)^{\delta}$, the probability of undertaking D-FDI is $\Pr(D) = G\left[\varphi^P(z)\right] - G\left[\varphi^D(z)\right]$ while production occurs in the South with probability $\Pr(S) = 1 - G\left[\varphi^P(z)\right]$. Then

$$\begin{aligned} \frac{\Pr(D)}{\Pr(S)} &= \frac{G\left[\varphi^P(z)\right] - G\left[\varphi^D(z)\right]}{1 - G\left[\varphi^P(z)\right]} \\ &= \frac{\left(\frac{1}{\Lambda^D(z)}\right)^{\frac{\delta}{\sigma-1}} - \left(\frac{1}{\Lambda^P(z)}\right)^{\frac{\delta}{\sigma-1}}}{\left(\frac{1}{\Lambda^P(z)}\right)^{\frac{\delta}{\sigma-1}}}. \end{aligned}$$

Then since $\frac{\partial \Lambda^D(z)}{\partial \beta(z)} = 0$, $\frac{\partial \Lambda^P(z)}{\partial \beta(z)} > 0$, we have that $\frac{\partial}{\partial \beta(z)}\left(\frac{\Pr(D)}{\Pr(S)}\right) > 0$.

[21] Within the set of firms that locate production in the South there is also a compositional effect of an increase in skill sensitivity. In particular

$$\frac{\partial\left(\frac{\Lambda^I(z)}{\Lambda^P(z)}\right)}{\partial \beta(z)} < 0 \tag{27}$$

indicating that the cut-off for Integrated FDI falls relative to that for Production oriented FDI. Thus conditional on producing in the South, increases in skill sensitivity increase the probability of undertaking Integrated FDI relative to Production oriented FDI.

where

$$V^l(z) \equiv \left(\omega_k^p + \omega_h^d\right)^{1-\sigma} \int_{\varphi^l(z)}^{\varphi^{l'}(z)} \varphi^{\sigma-1} g(\varphi) d\varphi, \qquad l = D, P, I, \quad k = N, S, \quad h = I, L \tag{29}$$

(where l' indexes the entry mode that appears next in the ranking, and $\varphi^{l'}(z) \longrightarrow \infty$ when $l = I$.) In the appendix we show that

$$\frac{\partial}{\partial \beta(z)} \left(\frac{R^D(z)}{R^D(z) + R^P(z) + R^I(z)} \right) > 0, \tag{30}$$

such that:

Proposition 5 *More skill sensitive sectors feature*

1. *an unambiguously higher probability of entering the Southern market through Distribution oriented FDI relative to modes under which production is located in the South;*
2. *a higher share of affiliate sales taking place through Distribution oriented FDI.*

Proof. See Appendix C.7. ■

Part 1 of Proposition 5 is consistent with Empirical Finding 5 (a). In particular, in Javorcik's data, an increase in R&D expenditure reduced the probability of an MNC engaging in P-FDI relative to D-FDI. Part 2 of the Proposition provides a further testable implication of the model. Underpinning both is the idea that greater skill sensitivity raises an MNC's variable costs, and by proportionately more when production is located in the South owing to IPR concerns. This reduces the firm's ability to cover the larger fixed costs of achieving proximity to the Southern market through P-FDI relative to D-FDI. Hence an increase in skill sensitivity shifts the extensive margin of firms away from the former towards the latter.

6.2 Changes in Southern monitoring

6.2.1 Production monitoring

Suppose first that there is an increase in the variance of measured effort in production in the South, such that effort monitoring becomes more difficult there. Above we showed that $\frac{\partial \omega_l^t}{\partial v_l^t} > 0$, such that net agency costs rise under worse monitoring. The reason is that lower piece rates are required in order to reduce the riskiness of worker income, which by incentive compatibility also reduces the return to workers of exerting effort. It is straightforward therefore that worse Southern monitoring reduces profits associated with entry modes under which production is located in the South, which penalises in particular the least productive firms producing in the South, causing them to shift production back home. In particular

$$\frac{\partial \Lambda^D(z)}{\partial v_S^p} = 0, \qquad \frac{\partial \Lambda^P(z)}{\partial v_S^p} > 0, \qquad \frac{\partial \Lambda^I(z)}{\partial v_S^p} > 0. \tag{31}$$

The result is an unambiguous increase in the probability of entering through Distribution oriented FDI relative to FDI under which Southern production is chosen when effort under the latter becomes harder to monitor. This shift is reflected in the share of affiliate sales accounted for by D-FDI, which we show in the appendix changes according to

$$\frac{\partial}{\partial v_S^p}\left(\frac{R^D(z)}{R^D(z)+R^P(z)+R^I(z)}\right)>0,$$

indicating that sales through D-FDI affiliates increase as a share of total affiliate sales when Southern production monitoring becomes more difficult.

6.2.2 Distribution monitoring

Suppose next that effort monitoring under licensed distribution relationships becomes more difficult relative to effort monitoring under integrated distribution. By $\frac{\partial \omega_l^t}{\partial v_l^t} > 0$ net agency costs under licensed distribution rise relative to those under integrated distribution. This implies

$$\frac{\partial \Lambda^D(z)}{\partial v_L^d}<0, \quad \frac{\partial \Lambda^P(z)}{\partial v_L^d}>0, \quad \frac{\partial \Lambda^I(z)}{\partial v_L^d}<0, \tag{32}$$

such that the probability of setting up a Distribution oriented Southern operation unambiguously increases relative to exporting and modes under which production is located in the South. Exporting and Production oriented FDI are both subject to the negative effects of worse monitoring under licensed distribution, shifting the cut-offs in favour of integrated distribution, which favours D-FDI and I-FDI at the expense of P-FDI. As for worse production monitoring, we show in the appendix that worse distribution monitoring under licensing implies

$$\frac{\partial}{\partial v_L^d}\left(\frac{R^D(z)}{R^D(z)+R^P(z)+R^I(z)}\right)>0,$$

when f_p is not too big. Intuitively, sales through D-FDI and I-FDI both increase when licensed monitoring becomes more difficult, since both of these entry modes enjoy comparatively lower distribution costs. Hence the rise in $R^I(z)$ dampens the relative increase in sales taking place through D-FDI. This effects is smaller however when f_p is smaller: when this is the case, sales through I-FDI are already relatively large, and so the proportional increase in $R^I(z)$ is smaller. This ensures that $\frac{R^D(z)}{R^D(z)+R^P(z)+R^I(z)}$ rises overall when v_L^d rises, or effort under licensed distribution becomes harder to monitor.

6.2.3 Combined changes in Southern monitoring

Combining the above two, under which monitoring production effort in the South and distribution effort under licensing both become more difficult, such that $\partial v_S^p = \partial v_L^d = \partial v_S$ yields

$$\frac{\partial \Lambda^D(z)}{\partial v_S}<0, \quad \frac{\partial \Lambda^P(z)}{\partial v_S}>0, \quad \frac{\partial \Lambda^I(z)}{\partial v_S}\lesseqgtr 0.$$

The effects of the simultaneous change in v_S^p and v_L^d go in the same direction for the D-FDI and P-FDI wedges. I-FDI is subject to competing effects: on the one hand, production in the South is more costly, but on the other, distribution undertaken in-house is more attractive. Combining the results for affiliate sales clearly implies that

$$\frac{\partial}{\partial v_S}\left(\frac{R^D(z)}{R^D(z)+R^P(z)+R^I(z)}\right) > 0,$$

when f_p is not too large, the intuition for which is analogous to the above. Using these results we state

Proposition 6 *Worse Southern monitoring in production and in licensed distribution relative to in-house distribution*

1. *increases the productivity range of firms undertaking Distribution oriented FDI relative to (i) exporting and (ii) modes under which production is based in the South;*
2. *increases the share of affiliate sales that take place through Distribution oriented FDI when f_p is not too big.*

Proof. See Appendix C.8. ■

One question is how to relate this result to the empirical findings. One possible proxy for monitoring costs is the presence of a common language. When the host country features a common language with the source country, one might reasonably expect monitoring costs to be somewhat lower. Indeed, the regression results of Hanson et al (2001) (Empirical Finding 4 (b) above) indicate that the presence of a common language is associated with a decrease in the share of affiliate sales accounted for by Distribution oriented FDI. This is consistent with Proposition 6, the converse of which states that better Southern monitoring e.g. in the form of a common language, should decrease Distribution oriented FDI relative to FDI modes under which production takes place in the South. Another possible proxy for monitoring costs is distance. Proposition 6 is therefore consistent with Empirical Finding 4 (a), which argued that affiliate sales through D-FDI as a share of total affiliate sales increase with distance. To the extent that these variables are accepted as empirical proxies for monitoring costs, our agency theoretic approach grants a new interpretation of, for example, the effects of common language dummies on the choice of foreign entry mode.

6.3 Changes in Southern demand

What is the impact of a change in the level of demand in the South on the relative prevalence of Northern MNCs' entry modes? Consider a preference shock $\partial\xi(z)$ that raises the demand for a given sector's output. This lowers the export cut-off $\frac{\partial\varphi^{X*}(z)}{\partial\xi(z)} < 0$. We show further that

$$\frac{\partial\varphi^{I*}(z)}{\partial\xi(z)} < \frac{\partial\varphi^{P*}(z)}{\partial\xi(z)} < \frac{\partial\varphi^{D*}(z)}{\partial\xi(z)} < \frac{\partial\varphi^{X*}(z)}{\partial\xi(z)} < 0, \tag{33}$$

such that all cut-offs fall, but in a way so as to change the probability of observing particular organisational forms. In particular, the range of firms choosing exporting is compressed, while the range choosing Integrated FDI expands. This implies

Proposition 7 *Increases in host country demand reduce the range of productivities over which Distribution oriented FDI is chosen relative to modes of entry under which production takes place in the South.*

Proof. See Appendix C.9. ■

This accords well with empirical finding 4 (e) above – that the probability of choosing D-FDI relative to modes under which production is based in the South falls as host country demand rises.[22] Here, higher final demand raises firm revenues and so helps MNCs to overcome the larger fixed costs associated with production in the South, relative to distribution-only projects in the South.

6.4 Trade liberalisation

Consider an isolated change in the variable trade cost τ in sector z. From the wedges above, we have

$$\frac{\partial \Lambda^D(z)}{\partial \tau} = 0, \quad \frac{\partial \Lambda^P(z)}{\partial \tau} < 0, \quad \frac{\partial \Lambda^I(z)}{\partial \tau} < 0, \tag{34}$$

such that, intuitively, increases in trade costs increase the range of firms producing in the South relative to the export cut-off.

Trade liberalisations have implications for sector equilibrium through free entry. In particular, by Lemma 4, $\frac{\partial \varphi^{X*}(z)}{\partial \tau} > 0$, indicating that the least productive exporters no longer do so as trade costs rise. Conversely, as trade costs fall, so does the export cut-off, allowing firms of lower productivity to enter the Southern market. By $\frac{\partial \Lambda^D}{\partial \tau} = 0$ and $\varphi^{D*}(z) = \left(\Lambda^D\right)^{\frac{1}{\sigma-1}} \varphi^{X*}(z) > \varphi^{X*}(z)$, we have that

$$\frac{\partial \varphi^{D*}(z)}{\partial \tau} > \frac{\partial \varphi^{X*}(z)}{\partial \tau}, \tag{35}$$

[22] This is true when the slope of the distribution function at the two cut-offs $\varphi^P(z)$ and $\varphi^D(z)$ does not differ by too much, which in turn requires that δ is not too large (i.e. that there exists sufficient mass in the right hand tail of the productivity distribution). To see this, note that the probability of observing D-FDI relative to observing production in the South is

$$\frac{\Pr(D)}{\Pr(S)} = \frac{G\left[\varphi^P(z)\right] - G\left[\varphi^D(z)\right]}{1 - G\left[\varphi^P(z)\right]}.$$

Then

$$\frac{\partial}{\partial \xi(z)}\left(\frac{\Pr(D)}{\Pr(S)}\right) = \frac{\frac{\partial G}{\partial \varphi^P(z)}\frac{\partial \varphi^P(z)}{\partial \xi(z)} - \frac{\partial G}{\partial \varphi^D(z)}\frac{\partial \varphi^D(z)}{\partial \xi(z)}}{1 - G\left[\varphi^P(z)\right]} + \frac{\frac{\Pr(D)}{\Pr(S)}}{1 - G\left[\varphi^P(z)\right]}\frac{\partial G}{\partial \varphi^P(z)}\frac{\partial \varphi^P(z)}{\partial \xi(z)},$$

which is negative if

$$\begin{aligned} \frac{\partial G}{\partial \varphi^P(z)}\frac{\partial \varphi^P(z)}{\partial \xi(z)} &< \frac{\partial G}{\partial \varphi^D(z)}\frac{\partial \varphi^D(z)}{\partial \xi(z)} \\ \Leftrightarrow \frac{1}{\varphi^P(z)^{\delta+1}}\frac{\partial \varphi^P(z)}{\partial \xi(z)} &< \frac{1}{\varphi^D(z)^{\delta+1}}\frac{\partial \varphi^D(z)}{\partial \xi(z)}, \end{aligned}$$

which, by $\frac{\partial \varphi^P(z)}{\partial \xi(z)} < \frac{\partial \varphi^D(z)}{\partial \xi(z)}$ but $\varphi^D(z) < \varphi^P(z)$ is true if δ is not too large.

such that increases in trade costs raise the cut-off for D by more than that for exporting. By $\frac{\partial \Lambda^D}{\partial \tau} = 0$ and $\frac{\partial \Lambda^P}{\partial \tau} < 0$, we know that increases in trade costs cause $\frac{\varphi^{P*}(z)}{\varphi^{D*}(z)}$ to fall, such that the range of firms undertaking Distribution oriented FDI is squeezed when trade costs rise. Of course, the converse is that trade liberalisations increase the range of firms undertaking Distribution oriented FDI. Reductions in trade barriers do not therefore imply a reduction in FDI per se, but rather alter the composition of FDI towards Distribution oriented projects and away from those under which production is based in the South. The simple reason is that as trade costs fall, Northern firms bring production back home in order to enjoy the cost savings associated with the cheaper employment of skilled labour blueprints in the North, saving on IPR related costs. But in contrast to standard horizontal FDI models, Northern firms retain a presence in the South by owning a distribution oriented sales affiliate. Trade liberalisation thereby changes the composition of FDI received by the South, without reducing its level.

For affiliate sales, we show in the appendix that

$$\frac{\partial}{\partial \tau}\left(\frac{R^D(z)}{R^D(z)+R^P(z)+R^I(z)}\right) < 0, \tag{36}$$

such that trade liberalisations lead to a rise in affiliate sales taking place through Distribution oriented FDI relative to Production oriented FDI. The intuition above, namely that trade liberalisations reduce the ‘proximity’ benefit to producing in the South, applies. But, again, in contrast to standard horizontal FDI models, many Northern firms retain a Southern presence, increasing the share of affiliated sales that take place through D-FDI. In sum, we have:

Proposition 8 *Trade liberalisation*

1. *raises the probability of Northern firms entering the Southern market through Distribution oriented FDI relative to modes under which production is based in the South;*

2. *raises the share of affiliate sales under Distribution oriented FDI relative to Production oriented and Integrated FDI when f_p is not too big.*

Proof. See Appendix C.10. ■

This result is illustrated in Figure 2, consistent with observed increases in FDI contemporaneous with decreases in trade costs highlighted by Neary (2009). This model emphasises that trade liberalisation affects the composition of FDI, not necessarily reducing its level.

6.5 Improvements in the South’s IPR regime

The final sector-level change we consider is an improvement in the South’s IPR regime. As for changes in skill sensitivity, changes in the security of IPR in the South change the relative cost of employing blueprints supplied by skilled Northern labour in the South relative to the North, $\beta(z)$. In particular, improvements in the South’s IPR

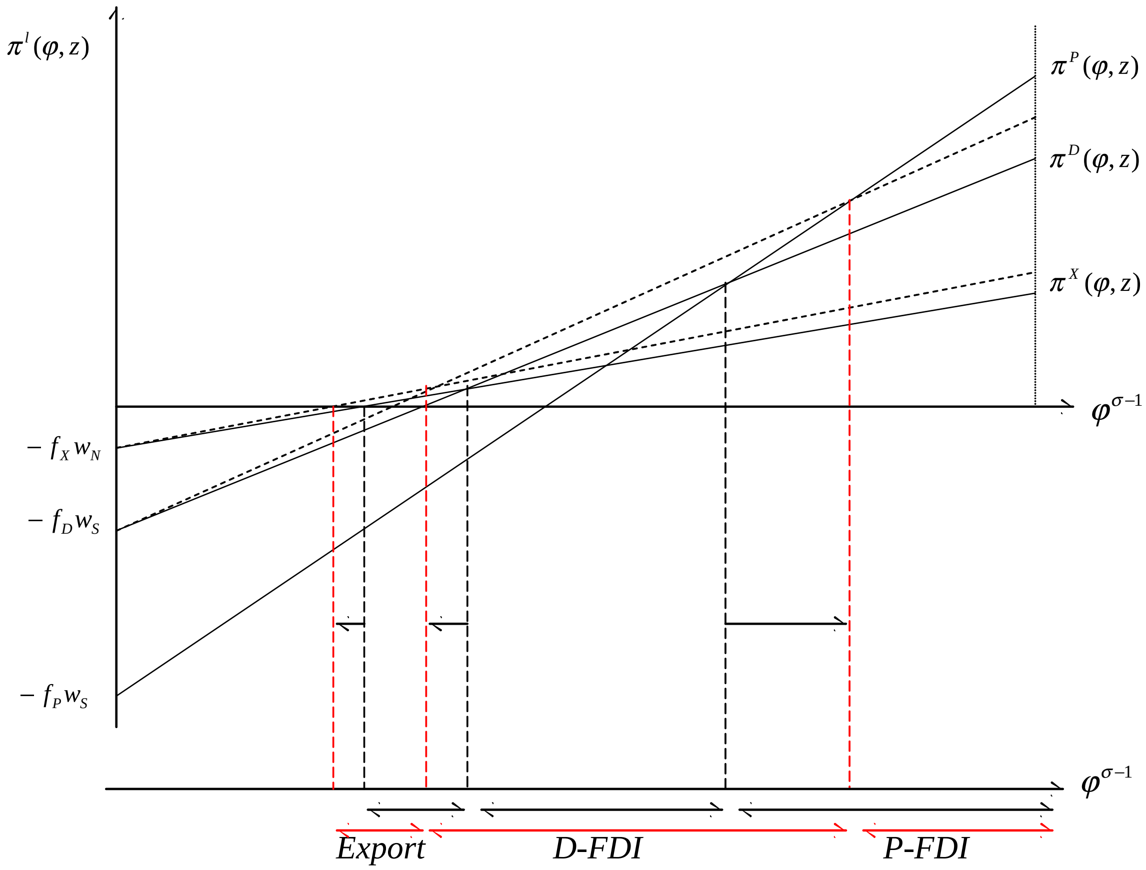

Figure 2: Trade liberalisation (Note: I-FDI omitted from diagram)

reduce $\beta(z)$. We show in the appendix however that

$$\frac{\partial^2 \Lambda^P(z)}{\partial \beta(z)^2} > 0, \quad \frac{\partial^2 \Lambda^I(z)}{\partial \beta(z)^2} > 0, \tag{37}$$

where the first inequality holds when $\beta(z)$ is sufficiently large and the second holds if $\sigma > 2$. Recall that $\frac{\partial \Lambda^D(.)}{\partial \beta(z)} = 0$. Then

Proposition 9 *Improvements in the South's IPR regime increase the probability of producing in the South relative to D-FDI, and by more the higher is the sector's skill sensitivity.*

Proof. See Appendix C.11. ■

Intuitively, improvements in the South's IPR regime have greater marginal value for firms in sectors that are highly skill sensitive. In such sectors, a larger range of firms find it profitable to shift production to the South when its IPR regime improves.

7 General equilibrium

So far we have taken the Northern skilled wage as given. In general equilibrium the skilled wage is determined by the demand for skilled labour by Northern firms serving the domestic market together with demand from Northern firms that enter the Southern market. Above we specified skilled labour costs as $r_N(z) = zr$ when production is based in the North, where r is the skilled wage and z indexes the unit input requirement, and $r_S(z) = \beta(z)zr$ when production is based in the South, where $\beta(z) \geq 1$ reflects the IPR related costs of employing skilled labour blueprints under Southern production. We denote skilled labour demand by firms serving the domestic market by $d(z,r)$, and by those firms that also serve the Southern market by $d^*(z,r)$. In the appendix we show that total demand in sector z for skilled labour is

$$D(z,r) = d(z,r) + d^*(z,r), \tag{38}$$

which satisfies $\frac{\partial D(z,r)}{\partial r} < 0$. Total demand across all sectors is then

$$\overline{D}(r) = \int_0^1 D(z,r)dz, \tag{39}$$

which satisfies $\frac{\partial \overline{D}(r)}{\partial r} < 0$. Factor market clearing for skilled labour, with total supply K, gives the equilibrium skilled wage straightforwardly as

$$\overline{D}(r^*) = K, \tag{40}$$

which closes the model.

By $\frac{\partial \overline{D}(r)}{\partial r} < 0$, an increase in the North's endowment of skilled labour K implies a reduction in the equilibrium skilled wage, $\frac{\partial r^*}{\partial K} < 0$. But this has no implications for industry equilibrium; in particular, the export cut-off $\varphi^X(z)$ is invariant to r in equilibrium.[23]

8 Conclusion

Sales through distribution oriented wholesale affiliates constitute a significant proportion of total sales of MNCs, yet have gone relatively unexamined in theory. We have presented a model in which MNCs overcome a multitask agency problem in determining their optimal mode of foreign entry. Firms optimally separate tasks between agents, with distinct groups exerting effort in production and distribution. Effort monitoring varies depending on whether production is located in the North or the South, and whether distribution is performed by an in-house agent or is licensed to an outside agent. Where monitoring is better, effort net of agency costs is higher, reducing variable costs and raising profits.

Firms choose different entry modes according to their productivities, with, *inter alia*, firms choosing distribution oriented FDI being less productive than those choosing FDI modes under which production is located in the South, which accords with recent evidence. The equilibrium choices of MNCs involve a trade-off between variable costs, including the agency costs associated with imperfect effort monitoring, and fixed costs, which are larger when production is located in the South than just distribution alone.

The model is consistent with various stylised empirical findings. While firms specialise in their mode of foreign entry, the host country attracts a mix of different modes of FDI. MNCs that are more skill sensitive are more likely to enter through distribution oriented FDI, rather than modes under which production takes place in the South. Worse effort monitoring in the South biases firms' choices towards distribution oriented projects, away from production oriented projects. To the extent that variables such as common language capture an element of monitoring costs, this is consistent with the empirical finding of Hanson et al (2001) that the presence of a common language encourages production oriented FDI relative to distribution oriented FDI. Larger host country demand encourages production based projects over distribution based projects. And trade liberalisation can increase FDI, contrary to the standard horizontal-FDI view, biasing it towards a distribution orientation away from a production orientation.

A Wholesale Trade: SIC Definition

The following definition is drawn from the definition of SIC division F 'wholesale trade', quoted on the US Department of Labor website.

SIC division F: Wholesale Trade.

[23]For example, if r rises, all firms' costs rise. This raises the price index by exactly the same amount, such that the two effects exactly cancel in their impact on firm profits, leaving the cut-offs unchanged. Of course, changes in K, which change r, have implications for welfare and inequality, but these follow trivially, in particular given the fixed wage of unskilled labour in the outside sector.

This division includes establishments or places of business primarily engaged in selling merchandise to retailers; to industrial, commercial, institutional, farm, construction contractors, or professional business users; or to other wholesalers; or acting as agents or brokers in buying merchandise for or selling merchandise to such persons or companies.

The chief functions of establishments included in Wholesale Trade are selling goods to trading establishments, or to industrial, commercial, institutional, farm, construction contractors, or professional business users; and bringing buyer and seller together. In addition to selling, functions frequently performed by wholesale establishments include maintaining inventories of goods; extending credit; physically assembling, sorting, and grading goods in large lots; breaking bulk and redistribution in smaller lots; delivery; refrigeration; and various types of promotion such as advertising and label designing.

The principal types of establishments included are: (1) merchant wholesalers-wholesalers who take title to the goods they sell, such as wholesale merchants or jobbers, industrial distributors, voluntary group wholesalers, exporters, importers, cash-and-carry wholesalers, drop shippers, truck distributors, retailer cooperative warehouses, terminal elevators, cooperative buying associations, and assemblers, buyers or cooperatives engaged in the marketing of farm products; (2) sales branches and sales offices (but not retail stores) maintained by manufacturing, refining or mining enterprises apart from their plants or mines for the purpose of marketing their products; and (3) agents, merchandise or commodity brokers, and commission merchants.

Establishments primarily engaged in selling merchandise to construction contractors, institutions, industrial users, or businesses are included in Wholesale Trade with a few exceptions. These exceptions are made necessary because of sales to both the general public for personal or household consumption and to businesses, industrial users, or construction contractors. These exceptions are lumber yards; paint, glass, and wallpaper stores, typewriter stores; stationery stores; and gasoline service stations which are classified in Retail Trade, Division G.

However, establishments that sell similar products only to institutions, industrial users, and establishments that sell merchandise for use exclusively by business establishments or to other wholesalers are classified in Wholesale Trade. Establishments primarily engaged in selling such merchandise as plumbing equipment; electrical supplies; used automobile parts; and office furniture are classified in Wholesale Trade, even if a higher proportion of their sales is made to individuals for household use. Establishments primarily engaged in the wholesale distribution of used products are classified on the basis of the products sold.

Guidelines for the classification of establishments primarily engaged in the wholesale distribution and construction or installation of equipment manufactured by other establishments are outlined in the Introduction to Division C, Construction.

B Tables

TABLE 1: Distribution of Total Affiliate Sales of US Firms, 1998, by Industry

		Affiliate Primary Industry											
		Petroleum	Food	Chem.	Metals	Ind. Mach.	Elec. Equip.	Transp. Equip.	Other mfg.	Wholesale	FIRE*	Services	Other inds.
	Petroleum	84.4	0.0	4.0	0.1	0.2	0.0	0.0	0.2	1.3	0.6	0.7	8.5
	Food	0.0	70.4	0.2	0.1	0.0	0.7	0.0	9.1	17.0	0.5	1.5	0.6
	Chem.	3.7	1.6	60.7	0.2	0.3	0.1	0.0	8.6	22.6	1.3	0.3	0.5
	Metals	0.3	0.7	8.3	59.9	4.3	2.0	1.3	5.9	11.0	0.8	1.6	4.0
Parent	Ind. Mach.	1.3	0.0	0.1	1.3	41.7	1.0	0.4	1.0	37.7	0.6	9.9	4.9
Primary	Elec. Equip.	0.0	0.0	0.1	3.0	2.1	59.1	2.7	1.8	28.1	0.9	0.8	1.5
Ind.	Transp. Equip.	0.0	0.0	1.4	0.6	4.2	4.5	68.5	1.8	9.7	7.1	1.6	0.5
	Other mfg.	0.1	0.1	3.0	0.7	4.9	1.2	2.4	64.5	19.7	0.6	1.6	1.1
	Wholesale	0.2	14.4	1.1	0.4	4.1	3.1	1.4	4.8	64.8	0.8	1.8	3.1
	FIRE*	23.1	0.4	0.7	0.6	0.1	0.0	0.2	0.5	1.7	64.9	1.9	5.8
	Services	0.0	0.0	0.2	0.8	0.8	0.4	0.0	0.8	3.3	1.3	86.1	6.2
	Other inds.	3.3	0.2	0.5	0.3	0.0	0.2	0.0	0.5	2.2	0.6	4.3	88.0

Source: Reproduced from Hanson et al (2001), Table 9. *FIRE=Finance, Insurance, Real Estate and Leasing industry group.

TABLE 2: Modes for Japanese firms' sales abroad, 2000

Mode	Value Added (¥m)	%
Produce in Japan and export	3,132,287	8.2
Produce in Japan and distribute through Foreign Affiliate in ROW*	15,203,158	39.6
Produce in ROW and sell locally	19,723,339	51.4
Other	335,898	0.9
Total	38,394,682	

Source: Ando and Kimura (2003) (data from MITI database).

* We combine Ando and Kimura's "ROW" and "Asia" in our definition of Rest of the World (ROW).

TABLE 3: Extensions of investments and new investments by European and non-European firms in the Enlarged EU, 1997-2001

	Number of new investments by function in EU15 and CEE8*					
	EU15	*% of EU15*	*Accession Countries*	*% Accession Co's*	*Total*	*% Total*
Headquarters	840	9.0	19	1.0	859	7.7
R&D	946	10.1	56	3.1	1,002	9.0
Production	3,912	41.8	1,304	71.6	5,216	46.6
Logistics	816	8.7	142	7.8	958	8.6
Sales and Marketing	2,849	30.4	299	16.4	3,148	28.1
Total	9,363		1,820		11,183	

Source: Defever (2006)

*Central and Eastern European 8 (CEE8): Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, Slovenia

C Proofs

C.1 Proof of Lemma 1: Task separation

Proof. Write the firm's expected variable costs as

$$\frac{r_k(z)}{\varphi} \times \left[\begin{array}{c} a_i^* + a_j^* + b_i^d \left(e_i^d + e_j^d\right) + b_j^d \left(e_i^d + e_j^d\right) + b_i^p \left(e_i^p + e_j^p\right) + b_j^p \left(e_i^p + e_j^p\right) \\ -\left(e_i^p + e_j^p\right) - \left(e_i^d + e_j^d\right) \end{array} \right] q.$$

IR gives a_J^* according to

$$a_J^* = w_k - \sum_{t=p,d} b_J^t e^t + \frac{1}{2} \left(\sum_{t=p,d} e_J^t \right)^2 + \frac{1}{2} \kappa \sum_{t=p,d} \left(b_J^t\right)^2 v^t,$$

which we use to write variable costs as

$$\frac{r_k(z)}{\varphi} \left[\begin{array}{c} 2w_k + \frac{1}{2}\left(e_i\right)^2 + \frac{1}{2}\kappa \sum_{t=p,d} \left(b_i^t\right)^2 v^t + \frac{1}{2}\left(e_j\right)^2 + \frac{1}{2}\kappa \sum_{t=p,d} \left(b_j^t\right)^2 v^t \\ -e_i - e_j \end{array} \right] q.$$

where we have used that $e_J = e_J^p + e_J^d$, $J = i, j$. Hold each agent's total effort constant at $\overline{e}_J$. The cost of implementing this effort level with positive piece rates on each task for each agent is then

$$\frac{r_k(z)}{\varphi} \left[\begin{array}{c} 2w_k + \frac{1}{2}\left(\overline{e}_i\right)^2 + \frac{1}{2}\kappa \sum_{t=p,d} \left(b_i^t\right)^2 v^t + \frac{1}{2}\left(\overline{e}_j\right)^2 + \frac{1}{2}\kappa \sum_{t=p,d} \left(b_j^t\right)^2 v^t \\ -\overline{e}_i - \overline{e}_j \end{array} \right] q \equiv \zeta(\overline{e}_i, \overline{e}_j).$$

This exceeds the cost of implementing the same effort levels $(\overline{e}_i, \overline{e}_j)$ under which a single task is assigned to each agent, under which $b_i^p = 0, b_i^d > 0$, and $b_j^p > 0$, $b_j^d = 0$, given by

$$\frac{r_k(z)}{\varphi} \left[\begin{array}{c} 2w_k + \frac{1}{2}\left(\overline{e}_i\right)^2 + \frac{1}{2}\kappa \left(b_i^d\right)^2 v^d + \frac{1}{2}\left(\overline{e}_j\right)^2 + \frac{1}{2}\kappa \left(b_j^p\right)^2 v^p \\ -\overline{e}_i - \overline{e}_j \end{array} \right] q \equiv \zeta'(\overline{e}_i, \overline{e}_j).$$

This follows from

$$\zeta'(\overline{e}_i, \overline{e}_j) < \zeta(\overline{e}_i, \overline{e}_j),$$
$$\Leftrightarrow \left(b_i^d\right)^2 v^d + \left(b_j^p\right)^2 v^p < \sum_{t=p,d} \left(b_i^t\right)^2 v^t + \sum_{t=p,d} \left(b_j^t\right)^2 v^t,$$

indicating that the variance of workers' pay is lower under scheme $\zeta'(\overline{e}_i, \overline{e}_j)$ than scheme $\zeta(\overline{e}_i, \overline{e}_j)$, in which the former features $b_i^p = 0, b_i^d > 0$, and $b_j^p > 0$, $b_j^d = 0$. This lowers the total cost to the firm of incentivising fixed effort levels $(\bar{e}_i, \bar{e}_j)$. ■

C.2 Proof of Lemma 2: Optimal contract parameters

Proof. Given that the firm assigns one task to each agent, it chooses the optimal incentive parameters b_j^p and b_j^d. Since each agent performs only one task, we drop the agent sub-indices. The firm's problem is

$$\min_{b^p, b^d} \frac{r_k(z)}{\varphi}\left[a_i + a_j + b^d e^d + b^d e^d - e^p - e^d\right] q,$$
$$\text{s.t. } a_J = a_J^*, \quad J = i, j,$$
$$\text{s.t. } e^t = b^t, \quad t = p, d.$$

We use

$$\begin{aligned} a_i^* &= w_k - b^p e^p + \frac{1}{2}\left(e^p\right)^2 + \frac{1}{2}\kappa\left(b^p\right)^2 v^p, \\ a_j^* &= w_k - b^d e^d + \frac{1}{2}\left(e^d\right)^2 + \frac{1}{2}\kappa\left(b^d\right)^2 v^d, \end{aligned}$$

to write the firm's problem as

$$\min_{b^p, b^d} \frac{r_k(z)}{\varphi}\left[w_k + w_k + \sum_{t=p,d}\frac{1}{2}\left(b^t\right)^2 + \sum_{t=p,d}\frac{1}{2}\kappa\left(b^t\right)^2 v^t - \sum_{t=p,d} b^t\right] q,$$

which has first order conditions

$$-1 + b^t + \kappa b^t v^t = 0, \quad t = p, d,$$

such that

$$b^{t*} = \frac{1}{1 + \kappa v^t}.$$

Minimised costs are then

$$\begin{aligned} & \frac{r_k(z)}{\varphi}\left[w_k + w_k + \sum_{t=p,d}\frac{1}{2}\left(b^{t*}\right)^2\left(1 + \kappa v^t\right) - \sum_{t=p,d} b^{t*}\right] q \\ = \quad & \frac{r_k(z)}{\varphi}\left[\sum_{t=p,d}\omega_k^t\right] q, \end{aligned}$$

where $\omega_k^t \equiv w_k - \frac{1}{2}b^{t*} = w_k - \frac{1}{2}\frac{1}{1+\kappa v^t}$. We have the properties that

$$\frac{\partial \omega_k^t}{\partial v^t} > 0, \quad \frac{\partial \omega_k^t}{\partial w_k} > 0, \quad \frac{\partial \omega_k^t}{\partial \kappa} > 0,$$

establishing the lemma. ■

C.3 Cut-offs

For the export cut-off, we have

$$\pi_L^N(\varphi,z)=\theta(z)\overline{\rho}\tau^{1-\sigma}r_N(z)^{1-\sigma}\left(\omega_N^p+\omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}-f_x w_N=0,$$

such that

$$\left[\varphi^X(z)\right]^{\sigma-1}=\frac{f_x w_N}{\theta(z)\overline{\rho}}\tau^{\sigma-1}r_N(z)^{\sigma-1}\left(\omega_N^p+\omega_L^d\right)^{\sigma-1}.$$

The export curve crosses the curve for D when

$$\theta(z)\overline{\rho}\tau^{1-\sigma}r_N(z)^{1-\sigma}\left(\omega_N^p+\omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}-f_x w_N$$
$$=\theta(z)\overline{\rho}\tau^{1-\sigma}r_N(z)^{1-\sigma}\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}-f_d w_S,$$

giving

$$\left[\varphi^D(z)\right]^{\sigma-1}=\frac{f_d w_S-f_x w_N}{\theta(z)\overline{\rho}}\frac{\tau^{\sigma-1}r_N(z)^{\sigma-1}}{\left[\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}-\left(\omega_N^p+\omega_L^d\right)^{1-\sigma}\right]}.$$

The D curve crosses the P curve when

$$\theta(z)\overline{\rho}\tau^{1-\sigma}r_N(z)^{1-\sigma}\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}-f_d w_S$$
$$=\theta(z)\overline{\rho}r_S(z)^{1-\sigma}\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}-f_p w_S,$$

such that

$$\left[\varphi^P(z)\right]^{\sigma-1}=\frac{f_p w_S-f_d w_S}{\theta(z)\overline{\rho}}\frac{r_N(z)^{\sigma-1}}{\left[\beta(z)^{1-\sigma}\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}-\tau^{1-\sigma}\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\right]}.$$

Finally, the P curve crosses the I curve when

$$\theta(z)\overline{\rho}r_S(z)^{1-\sigma}\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}-f_p w_S$$
$$=\theta(z)\overline{\rho}r_S(z)^{1-\sigma}\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}-\left(f_p+f_d\right)w_S,$$

such that

$$\left[\varphi^I(z)\right]^{\sigma-1}=\frac{f_d w_S}{\theta(z)\overline{\rho}}\frac{\beta(z)^{\sigma-1}r_N(z)^{\sigma-1}}{\left[\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\right]}.$$

But note that the cut-offs contain an endogenous variable $\theta(z)$. So we relate them using $\left[\varphi^X(z)\right]^{\sigma-1} = \frac{f_x w_N}{\theta(z)\rho}\tau^{\sigma-1} r_N(z)^{\sigma-1}\left(\omega_N^p +\right.$ to the export cut-off according to

$$\begin{aligned}
\frac{\left[\varphi^D(z)\right]^{\sigma-1}}{\left[\varphi^X(z)\right]^{\sigma-1}} &= \Lambda^D(z), \\
\Lambda^D(z) &\equiv \frac{f_d w_S - f_x w_N}{f_x w_N} \frac{1}{\left(\frac{\omega_N^p+\omega_I^d}{\omega_N^p+\omega_L^d}\right)^{1-\sigma} - 1}, \qquad (41) \\
\frac{\left[\varphi^P(z)\right]^{\sigma-1}}{\left[\varphi^X(z)\right]^{\sigma-1}} &= \Lambda^P(z), \\
\Lambda^P(z) &\equiv \frac{f_p w_S - f_d w_S}{f_x w_N} \frac{1}{\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma} - 1}, \qquad (42) \\
\frac{\left[\varphi^I(z)\right]^{\sigma-1}}{\left[\varphi^X(z)\right]^{\sigma-1}} &= \Lambda^I(z), \\
\Lambda^I(z) &\equiv \frac{f_d w_S}{f_x w_N} \frac{\left(\frac{\tau}{\beta(z)}\right)^{1-\sigma}\left(\omega_N^p+\omega_L^d\right)^{1-\sigma}}{\left(\omega_S^p+\omega_I^d\right)^{1-\sigma} - \left(\omega_S^p+\omega_L^d\right)^{1-\sigma}}, \qquad (43)
\end{aligned}$$

Then $\Lambda^I(z) > \Lambda^P(z) > \Lambda^D(z)$ for all z if

$$\frac{f_d w_S}{f_p w_S - f_d w_S} > \frac{\left(\omega_S^p+\omega_I^d\right)^{1-\sigma} - \left(\omega_S^p+\omega_L^d\right)^{1-\sigma}}{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma} - \left(\frac{\tau}{\beta(z)}\left(\omega_N^p+\omega_L^d\right)\right)^{1-\sigma}}$$

and

$$\frac{f_p w_S - f_d w_S}{f_d w_S - f_x w_N} > \frac{\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma} - 1}{\left(\frac{\omega_N^p+\omega_I^d}{\omega_N^p+\omega_L^d}\right)^{1-\sigma} - 1}$$

which require that fixed cost differences be large enough relative to variable cost differences.

C.4 Equilibrium sector price index

In sector equilibrium we take the Northern skilled wage rate as given. The prices of sector z goods in the South are aggregated according to $P(z)^{1-\sigma} = \int_{\varphi^X(z)}^{\infty} p_l^k(\varphi,z)^{1-\sigma} \times n(z)\chi(\varphi,z)d\varphi$ where $\chi(\varphi,z) \equiv \frac{g(\varphi)}{1-G[\varphi^X(z)]}$. Equilibrium prices are given by $p^X(\varphi,z) = \frac{\tau r_N(z)\left(\omega_N^p+\omega_L^d\right)}{\rho\varphi}$, $p^D(\varphi,z) = \frac{\tau r_N(z)\left(\omega_N^p+\omega_I^d\right)}{\rho\varphi}$, $p^P(\varphi,z) = \frac{r_S(z)\left(\omega_S^p+\omega_L^d\right)}{\rho\varphi}$,

$p^I(\varphi,z) = \frac{r_S(z)\left(\omega_S^p+\omega_I^d\right)}{\rho\varphi}$. This gives

$$\begin{aligned}
& P(z)^{1-\sigma} \\
= \;& n(z)\left[\frac{\tau r_N(z)}{\rho}\right]^{1-\sigma} \\
& \times\left[\int_{\varphi^X(z)}^{\varphi^D(z)}\left(\omega_N^p+\omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}\chi(\varphi,z)d\varphi+\int_{\varphi^D(z)}^{\varphi^P}\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}\chi(\varphi,z)d\varphi\right] \\
& +n(z)\left[\frac{\beta(z)r_N(z)}{\rho}\right]^{1-\sigma} \\
& \times\left[\int_{\varphi^P(z)}^{\varphi^I(z)}\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}\chi(\varphi,z)d\varphi+\int_{\varphi^I(z)}^{\infty}\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}\chi(\varphi,z)d\varphi\right] \\
= \;& n(z)p\left[\widetilde{\varphi}(z)\right]^{1-\sigma},
\end{aligned}$$

where

$$\begin{aligned}
& \widetilde{\varphi}(z)^{\sigma-1} \\
\equiv \;& \left[\tau r_N(z)\right]^{1-\sigma} \\
& \times\left[\int_{\varphi^X(z)}^{\varphi^D(z)}\left(\omega_N^p+\omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}\chi(\varphi,z)d\varphi+\int_{\varphi^D(z)}^{\varphi^P}\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}\chi(\varphi,z)d\varphi\right] \\
& +\left[\beta(z)r_N(z)\right]^{1-\sigma} \\
& \times\left[\int_{\varphi^P(z)}^{\varphi^I(z)}\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}\chi(\varphi,z)d\varphi+\int_{\varphi^I(z)}^{\infty}\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}\chi(\varphi,z)d\varphi\right],
\end{aligned}$$

is average productivity adjusted for agency, trade, and IPR costs. Next, use

$$\begin{aligned}
n(z) &= N(z)\left\{1-G\left[\varphi^X(z)\right]\right\}, \\
\chi(\varphi,z) &= g(\varphi)\left\{1-G\left[\varphi^X(z)\right]\right\}^{-1},
\end{aligned}$$

and $r_N(z)=zr$, $r_S(z)=\beta(z)zr$ to write

$$\begin{aligned}
& P(z)^{1-\sigma} \\
= \;& N(z)\left[\frac{\tau zr}{\rho}\right]^{1-\sigma} \\
& \times\left[\left(\omega_N^p+\omega_L^d\right)^{1-\sigma}\int_{\varphi^X(z)}^{\varphi^D(z)}\varphi^{\sigma-1}g(\varphi)d\varphi+\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\int_{\varphi^D(z)}^{\varphi^P(z)}\varphi^{\sigma-1}g(\varphi)d\varphi\right] \\
& +N(z)\left[\frac{\beta(z)zr}{\rho}\right]^{1-\sigma} \\
& \times\left[\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\int_{\varphi^P(z)}^{\varphi^I(z)}\varphi^{\sigma-1}g(\varphi)d\varphi+\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\int_{\varphi^I(z)}^{\infty}\varphi^{\sigma-1}g(\varphi)d\varphi\right].
\end{aligned}$$

Next, use

$$
\begin{aligned}
\varphi^D(z) &= \left[\Lambda^D(z)\right]^{\frac{1}{\sigma-1}} \varphi^X(z), \\
\varphi^P(z) &= \left[\Lambda^P(z)\right]^{\frac{1}{\sigma-1}} \varphi^X(z), \\
\varphi^I(z) &= \left[\Lambda^I(z)\right]^{\frac{1}{\sigma-1}} \varphi^X(z),
\end{aligned}
$$

to write

$$
\begin{aligned}
& P(z)^{1-\sigma} \\
= \; & N(z)\left[\frac{\tau z r}{\rho}\right]^{1-\sigma} \times \left[\begin{array}{c} \left(\omega_N^p+\omega_L^d\right)^{1-\sigma} \int_{\varphi^X(z)}^{\Lambda^D(z)^{\frac{1}{\sigma-1}}\varphi^X(z)} \varphi^{\sigma-1} g(\varphi) d\varphi \\ +\left(\omega_N^p+\omega_I^d\right)^{1-\sigma} \int_{\Lambda^D(z)^{\frac{1}{\sigma-1}}\varphi^X(z)}^{\Lambda^P(z)^{\frac{1}{\sigma-1}}\varphi^X(z)} \varphi^{\sigma-1} g(\varphi) d\varphi \end{array}\right] \\
& + N(z)\left[\frac{\beta(z) z r}{\rho}\right]^{1-\sigma} \times \left[\begin{array}{c} \left(\omega_S^p+\omega_L^d\right)^{1-\sigma} \int_{\Lambda^P(z)^{\frac{1}{\sigma-1}}\varphi^X(z)}^{\Lambda^I(z)^{\frac{1}{\sigma-1}}\varphi^X(z)} \varphi^{\sigma-1} g(\varphi) d\varphi \\ +\left(\omega_S^p+\omega_I^d\right)^{1-\sigma} \int_{\Lambda^I(z)^{\frac{1}{\sigma-1}}\varphi^X(z)}^{\infty} \varphi^{\sigma-1} g(\varphi) d\varphi \end{array}\right].
\end{aligned}
$$

And then make the following assumption:

Assumption: $\varphi \sim$ *Pareto*(δ) *on* $[1, \infty]$*, such that*

$$
g(\varphi) = \frac{\delta}{\varphi^{\delta+1}},
$$

where $\delta > \sigma - 1$.

Then define $\sigma - \delta - 1 \equiv \psi_1 < 0$ and $\frac{\sigma-\delta-1}{\sigma-1} \equiv \psi_2 < 0$, evaluating the integrals according to

$$
\delta \int_{\Lambda^D \varphi^X(z)}^{\Lambda^P \varphi^X(z)} \varphi^{\sigma-\delta-2} d\varphi = \frac{\delta}{\psi_1}\left[\left\{\left(\Lambda^P\right)^{\psi_2} - \left(\Lambda^D\right)^{\psi_2}\right\} \varphi^X(z)^{\psi_1}\right],
$$

and so on. This gives

$$
\begin{aligned}
& P(z)^{1-\sigma} \\
= \; & N(z)\left[\frac{\tau z r}{\rho}\right]^{1-\sigma} \frac{\delta}{\psi_1} \times \left[\begin{array}{c} \left(\omega_N^p+\omega_L^d\right)^{1-\sigma}\left[\left\{\left(\Lambda^D\right)^{\psi_2} - 1\right\} \varphi^X(z)^{\psi_1}\right] \\ +\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\left[\left\{\left(\Lambda^P\right)^{\psi_2} - \left(\Lambda^D\right)^{\psi_2}\right\} \varphi^X(z)^{\psi_1}\right] \end{array}\right] \\
& + N(z)\left[\frac{\beta(z) z r}{\rho}\right]^{1-\sigma} \frac{\delta}{\psi_1} \times \left[\begin{array}{c} \left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left[\left\{\left(\Lambda^I\right)^{\psi_2} - \left(\Lambda^P\right)^{\psi_2}\right\} \varphi^X(z)^{\psi_1}\right] \\ -\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\left(\Lambda^I\right)^{\psi_2} \varphi^X(z)^{\psi_1} \end{array}\right],
\end{aligned}
$$

such that

$$
P(z)^{\sigma-1} = \frac{\left[\frac{zr}{\rho}\right]^{\sigma-1}}{N(z)\varphi^X(z)^{\psi_1}\left[\tau^{1-\sigma}\Delta_N + \beta(z)^{1-\sigma}\Delta_S\right]},
$$

where

$$\Delta_S \equiv \frac{\delta}{\psi_1}\left[\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left\{\left(\Lambda^I\right)^{\psi_2}-\left(\Lambda^P\right)^{\psi_2}\right\}-\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\left(\Lambda^I\right)^{\psi_2}\right],$$

$$\Delta_N \equiv \frac{\delta}{\psi_1}\left[\left(\omega_N^p+\omega_L^d\right)^{1-\sigma}\left\{\left(\Lambda^D\right)^{\psi_2}-1\right\}+\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\left\{\left(\Lambda^P\right)^{\psi_2}-\left(\Lambda^D\right)^{\psi_2}\right\}\right].$$

These change according to

$$\begin{aligned}
\frac{\partial\Delta_S}{\partial\beta(z)} &= \frac{\delta}{\psi_1}\times\left[\begin{array}{c}\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left\{\psi_2\left(\Lambda^I\right)^{\psi_2-1}\frac{\partial\Lambda^I}{\partial\beta(z)}-\psi_2\left(\Lambda^P\right)^{\psi_2-1}\frac{\partial\Lambda^P}{\partial\beta(z)}\right\}\\ -\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\psi_2\left(\Lambda^I\right)^{\psi_2-1}\frac{\partial\Lambda^I}{\partial\beta(z)}\end{array}\right]\\
&= \underbrace{\frac{\delta\psi_2}{\psi_1}}_{+}\times\left[\begin{array}{c}\left[\underbrace{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}}_{-}\right]\left\{\left(\Lambda^I\right)^{\psi_2-1}\underbrace{\frac{\partial\Lambda^I}{\partial\beta(z)}}_{+}\right\}\\ -\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left(\Lambda^P\right)^{\psi_2-1}\underbrace{\frac{\partial\Lambda^P}{\partial\beta(z)}}_{+}\end{array}\right]<0,\\
\frac{\partial\Delta_S}{\partial\tau} &= \underbrace{\frac{\delta\psi_2}{\psi_1}}_{+}\times\left[\begin{array}{c}\left[\underbrace{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}}_{-}\right]\left(\Lambda^I\right)^{\psi_2-1}\underbrace{\frac{\partial\Lambda^I}{\partial\tau}}_{-}\\ -\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left(\Lambda^P\right)^{\psi_2-1}\underbrace{\frac{\partial\Lambda^P}{\partial\tau}}_{-}\end{array}\right]>0,\\
\frac{\partial\Delta_N}{\partial\beta(z)} &= \underbrace{\frac{\delta\psi_2}{\psi_1}}_{+}\left[\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\left(\Lambda^P\right)^{\psi_2-1}\underbrace{\frac{\partial\Lambda^P}{\partial\beta(z)}}_{+}\right]>0,\\
\frac{\partial\Delta_N}{\partial\tau} &= \underbrace{\frac{\delta\psi_2}{\psi_1}}_{+}\left[\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\left(\Lambda^P\right)^{\psi_2-1}\underbrace{\frac{\partial\Lambda^P}{\partial\tau}}_{-}\right]<0,
\end{aligned}$$

since $\frac{\partial\Lambda^D}{\partial\beta(z)},\frac{\partial\Lambda^D}{\partial\tau}=0$, which we use later.

C.5 Free entry

Entry is rational when

$$\int_{\varphi^X(z)}^{\infty}\pi^l(\varphi,z)dG(\varphi)\geq f_E w_N.$$

Use $\varphi^X(z)=\left(\frac{f_x w_N}{\theta(z)\bar{\rho}}\right)^{\frac{1}{\sigma-1}}\tau r_N(z)\left(\omega_N^p+\omega_L^d\right)$ to write $\varphi^X[\theta(z)]$ where $\frac{\partial\varphi^X[\theta(z)]}{\partial\theta(z)}<0$, and write $\pi^l(\varphi,z)=\pi^l(\varphi,\theta(z),z)$ where $\frac{\partial\pi^l(\varphi,\theta(z),z)}{\partial\theta(z)}>0$. Then $\frac{\partial}{\partial\theta(z)}\int_{\varphi^X(z)}^{\infty}\pi^l(\varphi,z)dG(\varphi)>0$. The LHS of (20) is therefore monotonically increasing in sector level demand $\theta(z)$. The equilibrium value $\theta^*(z)$ is then determined when (20) holds with equality, or

$$\int_{\varphi^X[\theta^*(z)]}^{\infty} \pi^l(\varphi, \theta^*(z), z) dG(\varphi) = f_E w_N.$$

Then with $\theta^*(z)$ determined, the price index is given by $[P^*(z)]^{\sigma-1} = \frac{\mu\xi(z)}{\theta^*(z)}$. The equilibrium export cut-off $\varphi^{X*}(z)$ determines equilibrium average productivity $\widetilde{\varphi}^*(z)$, which sets the average price in sector z, $p\left[\widetilde{\varphi}^*(z)\right]$. The number of entrants into the Southern market is then determined by $n^*(z) = \left(\frac{P^*(z)}{p[\widetilde{\varphi}^*(z)]}\right)^{1-\sigma}$.

C.6 Proof of Lemma 3: Equilibrium export cut-off

Proof. Using $\left[\varphi^X(z)\right]^{\sigma-1} = \frac{f_x w_N}{\theta(z)\rho}\tau^{\sigma-1} r_N(z)^{\sigma-1}\left(\omega_N^p + \omega_L^d\right)^{\sigma-1}$, $\theta(z) = \mu\xi(z)P(z)^{\sigma-1}$, the expression for $P(z)^{\sigma-1}$ above, and $\sigma - \delta - 1 \equiv \psi_1$ to give

$$\left[\varphi^X(z)\right]^{\delta} = \frac{f_x w_N N(z)}{\mu\xi(z)(1-\rho)}\left(\omega_N^p + \omega_L^d\right)^{\sigma-1}\left[\Delta_N + \left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\Delta_S\right], \tag{44}$$

the equilibrium value of $\varphi^X(z)$.

A change in $\beta(z)$ affects the RHS of (44). In particular

$$\frac{\partial\varphi^X(z)}{\partial\beta(z)} < 0,$$

if

$$\frac{\partial\Delta_N}{\partial\beta(z)} + \left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\frac{\partial\Delta_S}{\partial\beta(z)} + \frac{(1-\sigma)}{\beta(z)}\left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\Delta_S < 0$$

which is negative if

$$\frac{\partial\Delta_N}{\partial\beta(z)} + \left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\frac{\partial\Delta_S}{\partial\beta(z)} < 0$$

or

$$\underbrace{\frac{\delta\psi_2}{\psi_1}}_{+}\left[\begin{array}{c}\left[\left(\omega_N^p + \omega_I^d\right)^{1-\sigma} - \left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\left(\omega_S^p + \omega_L^d\right)^{1-\sigma}\right]\left(\Lambda^P\right)^{\psi_2-1}\underbrace{\frac{\partial\Lambda^P}{\partial\beta(z)}}_{+} \\ + \left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\left[\underbrace{\left(\omega_S^p + \omega_L^d\right)^{1-\sigma} - \left(\omega_S^p + \omega_I^d\right)^{1-\sigma}}_{-}\right]\left\{\left(\Lambda^I\right)^{\psi_2-1}\underbrace{\frac{\partial\Lambda^I}{\partial\beta(z)}}_{+}\right\}\end{array}\right] < 0,$$

for which a sufficient condition is $\left(\omega_N^p + \omega_I^d\right)^{1-\sigma} - \left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\left(\omega_S^p + \omega_L^d\right)^{1-\sigma} < 0$, or

$$\tau > \left(\frac{\omega_S^p + \omega_L^d}{\omega_N^p + \omega_I^d}\right)\beta(z),$$

which is none other than (18). Thus when (18) holds, $\frac{\partial\varphi^X(z)}{\partial\beta(z)} < 0$.

For changes in trade costs, the RHS of (44) changes according to

$$\begin{aligned} & \frac{\partial}{\partial \tau}\left[\Delta_N + \tau^{\sigma-1}\beta(z)^{1-\sigma}\Delta_S\right] \\ = \; & \frac{\partial \Delta_N}{\partial \tau} + (\sigma-1)\frac{\tau^{\sigma-1}}{\tau}\beta(z)^{1-\sigma}\Delta_S + \tau^{\sigma-1}\beta(z)^{1-\sigma}\frac{\partial \Delta_S}{\partial \tau}, \end{aligned}$$

sufficient for this to be positive is that

$$\begin{aligned} & \frac{\partial \Delta_N}{\partial \tau} + \left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\frac{\partial \Delta_S}{\partial \tau} > 0 \\ \Leftrightarrow & \left[\left(\omega_N^p + \omega_I^d\right)^{1-\sigma} - \left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\left(\omega_S^p + \omega_L^d\right)^{1-\sigma}\right]\left(\Lambda^P\right)^{\psi_2-1}\underbrace{\frac{\partial \Lambda^P}{\partial \tau}}_{-} \\ & + \left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\left[\underbrace{\left(\omega_S^p + \omega_L^d\right)^{1-\sigma} - \left(\omega_S^p + \omega_I^d\right)^{1-\sigma}}_{-}\right]\left(\Lambda^I\right)^{\psi_2-1}\underbrace{\frac{\partial \Lambda^I}{\partial \tau}}_{-} > 0, \end{aligned}$$

which, in turn, is positive if $\left(\omega_N^p + \omega_I^d\right)^{1-\sigma} - \left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\left(\omega_S^p + \omega_L^d\right)^{1-\sigma} < 0$ or

$$\tau > \beta(z)\frac{\left(\omega_S^p + \omega_L^d\right)}{\left(\omega_N^p + \omega_I^d\right)},$$

which is none other than (18). So when (18) holds, the RHS of (44) increases when τ rises, such that $\frac{\partial \varphi^X(z)}{\partial \tau} > 0$.

For changes in demand, from (44), we have $\frac{\partial \varphi^X(z)}{\partial \xi(z)} < 0$. ■

C.7 Proof of Proposition 5: Changes in skill sensitivity

Proof. Changes in skill sensitivity occur when $\beta(z)$ changes for a given sector z. This affects the cutoffs for P and I. From (41), (42) and (43),

$$\frac{\partial \Lambda^D(z)}{\partial \beta(z)} = 0, \quad \frac{\partial \Lambda^P(z)}{\partial \beta(z)} > 0, \quad \frac{\partial \Lambda^I(z)}{\partial \beta(z)} > 0.$$

From $\varphi^D(z) = \left[\Lambda^D(z)\right]^{\frac{1}{\sigma-1}}\varphi^X(z)$, where $\Lambda^D(z)$ is independent of $\beta(z)$, it follows that

$$\begin{aligned} \frac{\partial \varphi^D(z)}{\partial \beta(z)} &= \Lambda^D(z)^{\frac{1}{\sigma-1}}\frac{\partial \varphi^X(z)}{\partial \beta(z)} < 0, \\ \left|\frac{\partial \varphi^D(z)}{\partial \beta(z)}\right| &> \left|\frac{\partial \varphi^X(z)}{\partial \beta(z)}\right|, \end{aligned}$$

by $\Lambda^D(z) > 1$ such that $\varphi^D(z)$ falls and by more than $\varphi^X(z)$ when $\beta(z)$ rises.

For $\varphi^P(z) = \left[\Lambda^P(z)\right]^{\frac{1}{\sigma-1}}\varphi^X(z)$ and $\varphi^I(z) = \left[\Lambda^I(z)\right]^{\frac{1}{\sigma-1}}\varphi^X(z)$ there are two competing effects. On the one hand $\frac{\partial \varphi^X(z)}{\partial \beta(z)} < 0$, but against this goes $\frac{\partial \Lambda^j(z)}{\partial \beta(z)} > 0$. But since $\Lambda^D(z)$ does not change and $\frac{\partial \Lambda^j(z)}{\partial \beta(z)} > 0$ for

$j = P, I$, it must be that $\varphi^P(z)$ and $\varphi^I(z)$ rise relative to $\varphi^D(z)$. Formally

$$\begin{aligned}
\left(\frac{\varphi^P(z)}{\varphi^D(z)}\right)^{\sigma-1} &= \frac{\Lambda^P(z)}{\Lambda^D(z)}, \\
\frac{\partial\left(\frac{\varphi^P(z)}{\varphi^D(z)}\right)^{\sigma-1}}{\partial\beta(z)} &= \frac{1}{\Lambda^D(z)}\frac{\partial\Lambda^P(z)}{\partial\beta(z)} > 0,
\end{aligned}$$

and

$$\begin{aligned}
\left(\frac{\varphi^I(z)}{\varphi^D(z)}\right)^{\sigma-1} &= \frac{\Lambda^I(z)}{\Lambda^D(z)}, \\
\frac{\partial\left(\frac{\varphi^I(z)}{\varphi^D(z)}\right)^{\sigma-1}}{\partial\beta(z)} &= \frac{1}{\Lambda^D(z)}\frac{\partial\Lambda^I(z)}{\partial\beta(z)} > 0,
\end{aligned}$$

such that increases in skill sensitivity raise the cut-off for production in the South relative to the cut-off for Distribution oriented FDI.

Further

$$\begin{aligned}
\frac{\Lambda^I(z)}{\Lambda^P(z)} &= \frac{f_d w_S}{f_p w_S - f_d w_S}\frac{\left(\omega_S^p + \omega_L^d\right)^{1-\sigma} - \left(\frac{\tau}{\beta(z)}\left(\omega_N^p + \omega_L^d\right)\right)^{1-\sigma}}{\left(\omega_S^p + \omega_I^d\right)^{1-\sigma} - \left(\omega_S^p + \omega_L^d\right)^{1-\sigma}}, \\
\frac{\partial\left(\frac{\varphi^I(z)}{\varphi^P(z)}\right)^{\sigma-1}}{\partial\beta(z)} &= \frac{\partial\left(\frac{\Lambda^I(z)}{\Lambda^P(z)}\right)}{\partial\beta(z)} < 0,
\end{aligned}$$

such that $\frac{\varphi^I(z)}{\varphi^D(z)}$ falls as skill sensitivity $\beta(z)$ rises.

Affiliate Sales. The value of total sales R in sector z, $R(z)$, is

$$\begin{aligned}
R(z) &= \int_{\varphi^X(z)}^{\infty} p^l(\varphi, z)q(z)n(z)\chi(\varphi, z)d\varphi \\
&= \frac{n(z)\theta(z)}{1 - G[\varphi^X(z)]}\int_{\varphi^X(z)}^{\infty} p^l(\varphi, z)^{1-\sigma}g(\varphi)d\varphi.
\end{aligned}$$

Sales under Distribution oriented FDI are

$$R^D(z) = \frac{n(z)\theta(z)}{1 - G[\varphi^X(z)]}\left[\frac{\tau r_N(z)}{\rho}\right]^{1-\sigma}\int_{\varphi^D(z)}^{\varphi^P(z)}\left(\omega_N^p + \omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}g(\varphi)d\varphi.$$

Total sales under production in the South (P+I) are

$$\begin{aligned}
R^{P+I}(z) = {} & \frac{n(z)\theta(z)}{1 - G[\varphi^X(z)]}\left[\frac{r_S(z)}{\rho}\right]^{1-\sigma} \\
& \times\left[\int_{\varphi^P(z)}^{\varphi^I(z)}\left(\omega_S^p + \omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}g(\varphi)d\varphi + \int_{\varphi^I(z)}^{\infty}\left(\omega_S^p + \omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}g(\varphi)d\varphi\right],
\end{aligned}$$

such that relative sales are

$$\begin{aligned}
&\frac{R^D(z)}{R^{P+I}(z)} \\
= \;& \frac{[\tau r_N(z)]^{1-\sigma}\int_{\varphi^D(z)}^{\varphi^P(z)}\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}g(\varphi)d\varphi}{[r_S(z)]^{1-\sigma}\left[\int_{\varphi^P(z)}^{\varphi^I(z)}\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}g(\varphi)d\varphi+\int_{\varphi^I(z)}^{\infty}\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}g(\varphi)d\varphi\right]} \\
= \;& \left[\frac{\tau}{\beta(z)}\right]^{1-\sigma}\frac{V^D(z)}{V^P(z)+V^I(z)},
\end{aligned}$$

where

$$\begin{aligned}
V^D(z) &\equiv \int_{\varphi^D(z)}^{\varphi^P(z)}\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}g(\varphi)d\varphi, \\
V^P(z) &\equiv \int_{\varphi^P(z)}^{\varphi^I(z)}\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\varphi^{\sigma-1}g(\varphi)d\varphi, \\
V^I(z) &\equiv \int_{\varphi^I(z)}^{\infty}\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\varphi^{\sigma-1}g(\varphi)d\varphi.
\end{aligned}$$

Using the Pareto distribution

$$\begin{aligned}
V^D(z) &= \frac{\delta}{\psi_1}\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\left\{\varphi^P(z)^{\psi_1}-\varphi^D(z)^{\psi_1}\right\}, \\
V^P(z) &= \frac{\delta}{\psi_1}\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left\{\varphi^I(z)^{\psi_1}-\varphi^P(z)^{\psi_1}\right\}, \\
V^I(z) &= \frac{\delta}{\psi_1}\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\left[-\varphi^I(z)^{\psi_1}\right].
\end{aligned}$$

Then we have that

$$\begin{aligned}
\frac{V^D(z)}{V^P(z)+V^I(z)} &= \frac{\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\left\{\varphi^P(z)^{\psi_1}-\varphi^D(z)^{\psi_1}\right\}}{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left\{\varphi^I(z)^{\psi_1}-\varphi^P(z)^{\psi_1}\right\}+\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\left[-\varphi^I(z)^{\psi_1}\right]} \\
&= \frac{\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\left\{\left(\Lambda^P\right)^{\psi_2}-\left(\Lambda^D\right)^{\psi_2}\right\}}{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left\{\left(\Lambda^I\right)^{\psi_2}-\left(\Lambda^P\right)^{\psi_2}\right\}+\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\left[-\left(\Lambda^I\right)^{\psi_2}\right]} \\
&= \frac{\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\left\{\left(\Lambda^P\right)^{\psi_2}-\left(\Lambda^D\right)^{\psi_2}\right\}}{\left[\underbrace{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}}_{-}\right]\left(\Lambda^I\right)^{\psi_2}-\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left(\Lambda^P\right)^{\psi_2}} \\
&= \frac{\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\left\{\left(\Lambda^D\right)^{\psi_2}-\left(\Lambda^P\right)^{\psi_2}\right\}}{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left(\Lambda^P\right)^{\psi_2}-\left[\underbrace{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}}_{-}\right]\left(\Lambda^I\right)^{\psi_2}},
\end{aligned}$$

Then by $\frac{\partial\Lambda^D}{\partial\beta(z)}=0$, $\frac{\partial\Lambda^P}{\partial\beta(z)}>0$, $\frac{\partial\Lambda^I}{\partial\beta(z)}>0$, and $\psi_2<0$, the numerator rises when $\beta(z)$ rises (since $\left(\Lambda^P\right)^{\psi_2}$ falls), while the denominator falls when $\beta(z)$ rises by

$$\begin{aligned}
&\frac{\partial}{\partial\beta(z)}\left\{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left(\Lambda^P\right)^{\psi_2}-\left[\underbrace{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}}_{-}\right]\left(\Lambda^I\right)^{\psi_2}\right\} \\
&= \psi_2\left\{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left(\Lambda^P\right)^{\psi_2-1}\frac{\partial\Lambda^P}{\partial\beta(z)}-\underbrace{[.]}_{-}\left(\Lambda^I\right)^{\psi_2-1}\frac{\partial\Lambda^I}{\partial\beta(z)}\right\}<0,
\end{aligned}$$

by $\psi_2<0$. Then $\frac{\partial}{\partial\beta(z)}\left(\frac{V^D(z)}{V^P(z)+V^I(z)}\right)>0$ such that

$$\frac{\partial}{\partial\beta(z)}\left(\frac{R^D(z)}{R^{P+I}(z)}\right)>0.$$

Note also that this implies

$$\frac{\partial}{\partial\beta(z)}\left(\frac{R^D(z)}{R^D(z)+R^{P+I}(z)}\right)>0,$$

since

$$\begin{aligned}
\frac{\partial}{\partial\beta(z)}\left(\frac{R^D(z)}{R^D(z)+R^{P+I}(z)}\right) &= \frac{\partial}{\partial\beta(z)}\left(\frac{\frac{R^D(z)}{R^{P+I}(z)}}{\frac{R^D(z)}{R^{P+I}(z)}+1}\right) \\
&= \frac{\frac{\partial}{\partial\beta(z)}\left(\frac{R^D(z)}{R^{P+I}(z)}\right)}{\frac{R^D(z)}{R^{P+I}(z)}+1}-\frac{\frac{R^D(z)}{R^{P+I}(z)}}{\left(\frac{R^D(z)}{R^{P+I}(z)}+1\right)^2}\frac{\partial}{\partial\beta(z)}\left(\frac{R^D(z)}{R^{P+I}(z)}\right) \\
&= \frac{\frac{\partial}{\partial\beta(z)}\left(\frac{R^D(z)}{R^{P+I}(z)}\right)}{\frac{R^D(z)}{R^{P+I}(z)}+1}\left[\frac{1}{\frac{R^D(z)}{R^{P+I}(z)}+1}\right]>0.
\end{aligned}$$

■

C.8 Proof of Proposition 6: Changes in Southern monitoring

Proof. Production monitoring. From (42) we have that $\frac{\partial\Lambda^P(z)}{\partial\omega_S^p}>0$. Use $\frac{\partial\omega_S^p}{\partial v_S^p}>0$ then to argue that $\frac{\partial\Lambda^P(z)}{\partial v_S^p}>0$, or that worse monitoring in the South in production reduces the incentive for firms to locate production there, such that fewer firms do so. From (43) $\frac{\partial\Lambda^I(z)}{\partial\omega_S^p}>0$ so $\frac{\partial\Lambda^I(z)}{\partial v_S^p}>0$. From (41) it is clear that $\frac{\partial\Lambda^D(z)}{\partial v_S^p}=0$.

Distribution monitoring. When v_L^d increases, ω_L^d increases, so from (41), (42) and (43), $\frac{\partial\Lambda^D(z)}{\partial v_L^d}<0$,

$$\begin{aligned}
\frac{\partial\Lambda^I(z)}{\partial v_L^d} &= \frac{f_d w_S}{f_x w_N}\left[\begin{array}{c}\frac{(1-\sigma)}{\left(\omega_N^p+\omega_L^d\right)}\frac{\left(\frac{\tau}{\beta(z)}\right)^{1-\sigma}\left(\omega_N^p+\omega_L^d\right)^{1-\sigma}}{\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}} \\ +\frac{\left(\frac{\tau}{\beta(z)}\right)^{1-\sigma}\left(\omega_N^p+\omega_L^d\right)^{1-\sigma}}{\left[\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\right]^2}\frac{(1-\sigma)}{\left(\omega_S^p+\omega_L^d\right)}\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\end{array}\right]\frac{\partial\omega_L^d}{\partial v_L^d} \\
&= \Lambda^I(z)\frac{(1-\sigma)}{\left(\omega_N^p+\omega_L^d\right)}\left[1+\underbrace{\frac{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}}{\left[\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\right]}}_{+}\right]\frac{\partial\omega_L^d}{\partial v_L^d}<0,
\end{aligned}$$

and

$$
\begin{aligned}
\frac{\partial \Lambda^P(z)}{\partial v_L^d} &= \frac{f_p w_S - f_d w_S}{f_x w_N} \frac{1}{\left[\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma} - 1\right]^2} (-1) \\
&\quad \times \left[(1-\sigma)\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{-\sigma}\left(\frac{\beta(z)}{\tau\left(\omega_N^p+\omega_L^d\right)} - \frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\left[\tau\left(\omega_N^p+\omega_L^d\right)\right]^2}\tau\right)\right] \frac{\partial \omega_L^d}{\partial v_L^d} \\
&= \Lambda^P(z) \frac{(\sigma-1)}{\left[\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma} - 1\right]} \left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{-\sigma} \\
&\quad \times \frac{\beta(z)}{\tau\left(\omega_N^p+\omega_L^d\right)}\left(1 - \frac{\left(\omega_S^p+\omega_L^d\right)}{\left(\omega_N^p+\omega_L^d\right)}\right) \frac{\partial \omega_L^d}{\partial v_L^d} \\
&> 0 \text{ if } 1 - \frac{\left(\omega_S^p+\omega_L^d\right)}{\left(\omega_N^p+\omega_L^d\right)} > 0,
\end{aligned}
$$

which is the case iff $\omega_N^p > \omega_S^p$, or if w_N is sufficiently high relative to w_S.

Combined changes in Southern monitoring. If v_S^p and v_L^d rise simultaneously, denoted by a change ∂v_S, combining the above gives $\frac{\partial \Lambda^D(z)}{\partial v_S} < 0$, $\frac{\partial \Lambda^P(z)}{\partial v_S} > 0$ and $\frac{\partial \Lambda^I(z)}{\partial v_S} \lesseqgtr 0$.

Affiliate Sales. Use

$$
\frac{V^D(z)}{V^P(z)+V^I(z)} = \frac{\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\left\{\left(\Lambda^D\right)^{\psi_2} - \left(\Lambda^P\right)^{\psi_2}\right\}}{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left(\Lambda^P\right)^{\psi_2} - \left[\underbrace{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma} - \left(\omega_S^p+\omega_I^d\right)^{1-\sigma}}_{-}\right]\left(\Lambda^I\right)^{\psi_2}}.
$$

Changes in v_S^p. By $\frac{\partial \Lambda^D}{\partial v_S^p} = 0$, $\frac{\partial \Lambda^P}{\partial v_S^p} > 0$ and $\frac{\partial \Lambda^I}{\partial v_S^p} > 0$, together with $\frac{\partial \omega_S^p}{\partial v_S^p} > 0$ and $\frac{\partial}{\partial v_S^p}\left[\left(\omega_S^p+\omega_L^d\right)^{1-\sigma} - \left(\omega_S^p+\omega_I^d\right)^{1-\sigma}\right] > 0$, we have that

$$
\frac{\partial}{\partial v_S^p}\left(\frac{R^D(z)}{R^{P+I}(z)}\right) > 0.
$$

This also implies

$$
\frac{\partial}{\partial v_S^p}\left(\frac{R^D(z)}{R^D(z)+R^{P+I}(z)}\right) > 0.
$$

Changes in v_L^d. Use

$$
\frac{V^D(z)}{V^P(z)+V^I(z)} = \frac{\left(\omega_N^p+\omega_I^d\right)^{1-\sigma}\left\{\left(\Lambda^D\right)^{\psi_2} - \left(\Lambda^P\right)^{\psi_2}\right\}}{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\left(\Lambda^P\right)^{\psi_2} - \left[\underbrace{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma} - \left(\omega_S^p+\omega_I^d\right)^{1-\sigma}}_{-}\right]\left(\Lambda^I\right)^{\psi_2}}.
$$

By $\psi_2 < 0$, $\frac{\partial \Lambda^D}{\partial v_L^d} < 0$ and $\frac{\partial \Lambda^P}{\partial v_L^d} > 0$ imply the numerator increases when v_L^d increases. The denominator falls overall iff

$$\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\psi_2\frac{\left(\Lambda^P\right)^{\psi_2}}{\Lambda^P}\frac{\partial\Lambda^P}{\partial v_L^d}-\left[\underbrace{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}}_{-}\right]\psi_2\frac{\left(\Lambda^I\right)^{\psi_2}}{\Lambda^I}\frac{\partial\Lambda^I}{\partial v_L^d}$$
$$+(1-\sigma)\left(\omega_S^p+\omega_L^d\right)^{-\sigma}\underbrace{\frac{\partial\omega_L^d}{\partial v_L^d}}_{+}\underbrace{\left[\left(\Lambda^P\right)^{\psi_2}-\left(\Lambda^I\right)^{\psi_2}\right]}_{+}<0,$$

for which it is sufficient that

$$\psi_2\left\{\begin{array}{c}\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\frac{\left(\Lambda^P\right)^{\psi_2}}{\Lambda^P}\frac{\partial\Lambda^P}{\partial v_L^d}\\ -\left[\underbrace{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}}_{-}\right]\frac{\left(\Lambda^I\right)^{\psi_2}}{\Lambda^I}\frac{\partial\Lambda^I}{\partial v_L^d}\end{array}\right\}<0,$$

$$\Leftrightarrow\left(\frac{\omega_S^p+\omega_I^d}{\omega_S^p+\omega_L^d}\right)^{\sigma-1}>\frac{-\frac{\left(\Lambda^I\right)^{\psi_2}}{\Lambda^I}\frac{\partial\Lambda^I}{\partial v_L^d}}{\left[\frac{\left(\Lambda^P\right)^{\psi_2}}{\Lambda^P}\frac{\partial\Lambda^P}{\partial v_L^d}-\frac{\left(\Lambda^I\right)^{\psi_2}}{\Lambda^I}\frac{\partial\Lambda^I}{\partial v_L^d}\right]},$$

into which substitute expressions for $\frac{\partial\Lambda^I}{\partial v_L^d}$ and $\frac{\partial\Lambda^P}{\partial v_L^d}$ to give

$$\left(\frac{\omega_S^p+\omega_I^d}{\omega_S^p+\omega_L^d}\right)^{\sigma-1}>\frac{1}{\frac{\left(\Lambda^P\right)^{\psi_2}}{\left(\Lambda^I\right)^{\psi_2}}\frac{\underbrace{\frac{\left(\omega_S^p+\omega_L^d\right)^{-\sigma}}{\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma}-1}}_{+}\left(\frac{\beta(z)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma}\underbrace{\left(\omega_N^p-\omega_S^p\right)}_{+}}{1+\underbrace{\frac{\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}}{\left[\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\right]}}_{+}}+1},$$

which is satisfied for $\frac{\left(\Lambda^P\right)^{\psi_2}}{\left(\Lambda^I\right)^{\psi_2}}$ high enough, or Λ^P small enough relative to Λ^I (by $\psi_2<0$), which is satisfied by, for example, f_p sufficiently small, since

$$\frac{\Lambda^P}{\Lambda^I}=\frac{f_p-f_d}{f_d}\frac{1}{\left[\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma}-1\right]}\frac{1}{\frac{\left(\frac{\tau}{\beta(z)}\right)^{1-\sigma}\left(\omega_N^p+\omega_L^d\right)^{1-\sigma}}{\left(\omega_S^p+\omega_I^d\right)^{1-\sigma}-\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}}}.$$

When this is the case

$$\frac{\partial}{\partial v_L^d}\left(\frac{R^D(z)}{R^{P+I}(z)}\right)>0,$$

which implies that

$$\frac{\partial}{\partial v_L^d}\left(\frac{R^D(z)}{R^D(z)+R^{P+I}(z)}\right)>0.$$

Combined changes in v_S^p and v_L^d. Combining the above, when f_p is not too large, for $\partial v_S=\partial v_S^p=\partial v_L^d$ we have that

$$\frac{\partial}{\partial v_S}\left(\frac{R^D(z)}{R^{P+I}(z)}\right) > 0,$$

which implies that

$$\frac{\partial}{\partial v_S}\left(\frac{R^D(z)}{R^D(z)+R^{P+I}(z)}\right) > 0.$$

■

C.9 Proof of Proposition 7: Changes in host country demand

Proof. Consider a preference shock $\partial\xi(z)$ that raises demand in a sector. From (44), we have

$$\frac{\partial\varphi^X(z)}{\partial\xi(z)} < 0,$$

indicating a lowering of the export cut-off when demand increases. Further,

$$\begin{aligned}
\frac{\partial\varphi^D(z)}{\partial\xi(z)} &= \left(\Lambda^D\right)^{\frac{1}{\sigma-1}}\frac{\partial\varphi^X(z)}{\partial\xi(z)} < 0,\\
\left|\frac{\partial\varphi^D(z)}{\partial\xi(z)}\right| &> \left|\frac{\partial\varphi^X(z)}{\partial\xi(z)}\right| \text{ by } \left(\Lambda^D\right)^{\frac{1}{\sigma-1}} > 1,\\
\frac{\partial\varphi^P(z)}{\partial\xi(z)} &= \left(\Lambda^P\right)^{\frac{1}{\sigma-1}}\frac{\partial\varphi^X(z)}{\partial\xi(z)} < 0,\\
\left|\frac{\partial\varphi^P(z)}{\partial\xi(z)}\right| &> \left|\frac{\partial\varphi^D(z)}{\partial\xi(z)}\right| \text{ by } \left(\Lambda^P\right)^{\frac{1}{\sigma-1}} > \left(\Lambda^D\right)^{\frac{1}{\sigma-1}},\\
\frac{\partial\varphi^I(z)}{\partial\xi(z)} &= \left(\Lambda^I\right)^{\frac{1}{\sigma-1}}\frac{\partial\varphi^X(z)}{\partial\xi(z)} < 0,\\
\left|\frac{\partial\varphi^I(z)}{\partial\xi(z)}\right| &> \left|\frac{\partial\varphi^P(z)}{\partial\xi(z)}\right| \text{ by } \left(\Lambda^I\right)^{\frac{1}{\sigma-1}} > \left(\Lambda^P\right)^{\frac{1}{\sigma-1}},
\end{aligned}$$

so we have that

$$\frac{\partial\varphi^I(z)}{\partial\xi(z)} < \frac{\partial\varphi^P(z)}{\partial\xi(z)} < \frac{\partial\varphi^D(z)}{\partial\xi(z)} < \frac{\partial\varphi^X(z)}{\partial\xi(z)} < 0,$$

implying that the range of firms undertaking D shrinks relative to that undertaking production in the South as host country demand rises. ■

C.10 Proof of Proposition 8: Trade liberalisation

Proof. From (41), (42) and (43)

$$\frac{\partial\Lambda^D(z)}{\partial\tau} = 0, \quad \frac{\partial\Lambda^I(z)}{\partial\tau} < 0, \quad \frac{\partial\Lambda^P(z)}{\partial\tau} < 0,$$

such that increases in trade costs increase the range of firms producing in the South relative to the export cut-off.

Above we showed that $\frac{\partial\varphi^X(z)}{\partial\tau} > 0$. Use $\varphi^D(z) = \left[\Lambda^D(z)\right]^{\frac{1}{\sigma-1}}\varphi^X(z)$ in which $\Lambda^D(z)$ is independent of τ

to argue

$$\begin{aligned} \frac{\partial \varphi^D(z)}{\partial \tau} &= \Lambda^D(z)^{\frac{1}{\sigma-1}} \frac{\partial \varphi^X(z)}{\partial \tau} > \frac{\partial \varphi^X(z)}{\partial \tau} \\ \Rightarrow \frac{\partial \varphi^D(z)}{\partial \tau} &> \frac{\partial \varphi^X(z)}{\partial \tau}, \end{aligned}$$

by $\Lambda^D(z) > 1$, such that increases in trade costs increase the D cut-off by more than the export cut-off. Conversely trade liberalisations decrease the D cut-off by more than the export cut-off.

Next, look at $\left(\frac{\varphi^P(z)}{\varphi^D(z)}\right)^{\sigma-1} = \frac{\Lambda^P(z)}{\Lambda^D(z)}$ to argue

$$\frac{\partial \left(\frac{\varphi^P(z)}{\varphi^D(z)}\right)^{\sigma-1}}{\partial \tau} = \frac{1}{\Lambda^D(z)} \frac{\partial \Lambda^P(z)}{\partial \tau} < 0,$$

indicating that $\varphi^P(z)$ falls relative to $\varphi^D(z)$. So increases in trade costs squeeze the range of firms undertaking D. The converse is that trade liberalisations increase the range of productivities over which firms choose D relative to modes under which production is based in the South.

Finally use $\left(\frac{\varphi^I(z)}{\varphi^D(z)}\right)^{\sigma-1} = \frac{\Lambda^I(z)}{\Lambda^D(z)}$ so

$$\frac{\partial \left(\frac{\varphi^I(z)}{\varphi^D(z)}\right)^{\sigma-1}}{\partial \tau} = \frac{1}{\Lambda^D(z)} \frac{\partial \Lambda^I(z)}{\partial \tau} < 0,$$

such that the I cut-off falls relative to D. Finally

$$\frac{\Lambda^I}{\Lambda^P} = \frac{f_d}{f_p - f_d} \frac{\left(\omega_S^p + \omega_L^d\right)^{1-\sigma} - \left(\frac{\tau}{\beta(z)}\left(\omega_N^p + \omega_L^d\right)\right)^{1-\sigma}}{\left(\omega_S^p + \omega_I^d\right)^{1-\sigma} - \left(\omega_S^p + \omega_L^d\right)^{1-\sigma}},$$

so

$$\frac{\partial \left(\frac{\varphi^I(z)}{\varphi^P(z)}\right)^{\sigma-1}}{\partial \tau} = \frac{\partial \left(\frac{\Lambda^I(z)}{\Lambda^P(z)}\right)}{\partial \tau} > 0,$$

such that I cut off rises relative to P.

Affiliate Sales. Use

$$\frac{V^D(z)}{V^P(z) + V^I(z)} = \frac{\left(\omega_N^p + \omega_I^d\right)^{1-\sigma} \left\{\left(\Lambda^D\right)^{\psi_2} - \left(\Lambda^P\right)^{\psi_2}\right\}}{\left(\omega_S^p + \omega_L^d\right)^{1-\sigma} \left(\Lambda^P\right)^{\psi_2} - \left[\underbrace{\left(\omega_S^p + \omega_L^d\right)^{1-\sigma} - \left(\omega_S^p + \omega_I^d\right)^{1-\sigma}}_{-}\right] \left(\Lambda^I\right)^{\psi_2}},$$

in which the numerator falls when τ rises, by $\frac{\partial \Lambda^P(z)}{\partial \tau} < 0$, and in which the denominator rises when τ rises if

$$\begin{aligned}
&\left(\omega_S^p + \omega_L^d\right)^{1-\sigma} \psi_2 \left(\Lambda^P\right)^{\psi_2 - 1} \frac{\partial \Lambda^P(z)}{\partial \tau} - \left[\begin{array}{c} \left(\omega_S^p + \omega_L^d\right)^{1-\sigma} \\ - \left(\omega_S^p + \omega_I^d\right)^{1-\sigma} \end{array} \right] \times \psi_2 \left(\Lambda^I\right)^{\psi_2 - 1} \frac{\partial \Lambda^I(z)}{\partial \tau} \\
= \;& \psi_2 \left\{ \left(\omega_S^p + \omega_L^d\right)^{1-\sigma} \times \left[\begin{array}{c} \left(\Lambda^P\right)^{\psi_2 - 1} \frac{\partial \Lambda^P(z)}{\partial \tau} \\ - \left(\Lambda^I\right)^{\psi_2 - 1} \frac{\partial \Lambda^I(z)}{\partial \tau} \end{array} \right] + \left(\omega_S^p + \omega_I^d\right)^{1-\sigma} \left(\Lambda^I\right)^{\psi_2 - 1} \frac{\partial \Lambda^I(z)}{\partial \tau} \right\} \\
> \;& 0,
\end{aligned}$$

which requires that

$$\left(\omega_S^p + \omega_L^d\right)^{1-\sigma} \left[\left(\Lambda^P\right)^{\psi_2 - 1} \frac{\partial \Lambda^P(z)}{\partial \tau} - \left(\Lambda^I\right)^{\psi_2 - 1} \frac{\partial \Lambda^I(z)}{\partial \tau}\right] + \left(\omega_S^p + \omega_I^d\right)^{1-\sigma} \left(\Lambda^I\right)^{\psi_2 - 1} \frac{\partial \Lambda^I(z)}{\partial \tau} < 0,$$
$$\Leftrightarrow \left(\Lambda^P\right)^{\psi_2 - 1} \frac{\partial \Lambda^P}{\partial \tau} < \left(\Lambda^I\right)^{\psi_2 - 1} \frac{\partial \Lambda^I(z)}{\partial \tau},$$

in which

$$\begin{aligned}
\frac{\partial \Lambda^P}{\partial \tau} &= \frac{f_p w_S - f_d w_S}{f_x w_N} \frac{(-1)}{\left[\left(\frac{\beta(z)\left(\omega_S^p + \omega_L^d\right)}{\tau\left(\omega_N^p + \omega_L^d\right)}\right)^{1-\sigma} - 1\right]^2} (\sigma - 1) \left(\frac{\beta(z)\left(\omega_S^p + \omega_L^d\right)}{\left(\omega_N^p + \omega_L^d\right)}\right)^{1-\sigma} \frac{\tau^{\sigma-1}}{\tau} \\
&= \Lambda^P(z) \frac{1}{\left(\frac{\beta(z)\left(\omega_S^p + \omega_L^d\right)}{\tau\left(\omega_N^p + \omega_L^d\right)}\right)^{1-\sigma} - 1} (1 - \sigma) \left(\frac{\beta(z)\left(\omega_S^p + \omega_L^d\right)}{\left(\omega_N^p + \omega_L^d\right)}\right)^{1-\sigma} \frac{\tau^{\sigma-1}}{\tau}, \qquad (45) \\
\frac{\partial \Lambda^I}{\partial \tau} &= \frac{f_d w_S}{f_x w_N} \frac{\left(\frac{\tau}{\beta(z)}\right)^{1-\sigma} \left(\omega_N^p + \omega_L^d\right)^{1-\sigma}}{\left(\omega_S^p + \omega_I^d\right)^{1-\sigma} - \left(\omega_S^p + \omega_L^d\right)^{1-\sigma}} (1 - \sigma) \frac{1}{\tau} \\
&= \Lambda^I(z)(1 - \sigma) \frac{1}{\tau}, \qquad (46)
\end{aligned}$$

so $\left(\Lambda^P\right)^{\psi_2 - 1} \frac{\partial \Lambda^P(z)}{\partial \tau} < \left(\Lambda^I\right)^{\psi_2 - 1} \frac{\partial \Lambda^I(z)}{\partial \tau}$ requires that

$$\begin{aligned}
&\frac{1}{\left(\frac{\beta(z)\left(\omega_S^p + \omega_L^d\right)}{\tau\left(\omega_N^p + \omega_L^d\right)}\right)^{1-\sigma} - 1} \left(\frac{\beta(z)\left(\omega_S^p + \omega_L^d\right)}{\tau\left(\omega_N^p + \omega_L^d\right)}\right)^{1-\sigma} \\
&\qquad > \left(\frac{f_d w_S}{f_p w_S - f_d w_S} \frac{\left(\omega_S^p + \omega_L^d\right)^{1-\sigma} - \left(\frac{\tau}{\beta(z)}\left(\omega_N^p + \omega_L^d\right)\right)^{1-\sigma}}{\left(\omega_S^p + \omega_I^d\right)^{1-\sigma} - \left(\omega_S^p + \omega_L^d\right)^{1-\sigma}}\right)^{\psi_2},
\end{aligned}$$

which in turn requires that f_p be sufficiently small (by $\psi_2 < 0$). When this is the case,

$$\frac{\partial}{\partial \tau}\left(\frac{V^D(z)}{V^P(z) + V^I(z)}\right) < 0,$$

such that sales under D fall relative to those under P and I when trade costs rise. This implies that

$$\frac{\partial}{\partial\tau}\left(\frac{V^D(z)}{V^D(z)+V^P(z)+V^I(z)}\right)<0.$$

Equivalently, when trade is liberalised, the share of affiliate sales through D rises. Trade liberalisation therefore shifts the content of FDI away from P oriented projects and towards D oriented projects. ■

C.11 Proof of Proposition 9: Changes in Southern IPR enforcement

Proof. From (41), (42) and (43)

$$\begin{aligned}
\frac{\partial\Lambda^P(z)}{\partial\beta(z)} &= \Lambda^P(z)\frac{(\sigma-1)}{\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma}-1}\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{-\sigma}\left(\omega_S^p+\omega_L^d\right) \\
\frac{\partial^2\Lambda^P(z)}{\partial\beta(z)^2} &= \frac{\partial\Lambda^P}{\partial\beta(z)}\frac{(\sigma-1)}{\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma}-1}\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{-\sigma}\left(\omega_S^p+\omega_L^d\right) \\
&\quad+\Lambda^P\frac{(\sigma-1)}{\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma}-1}\left(\frac{\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{-\sigma}(-\sigma)\beta(z)^{-\sigma-1}\left(\omega_S^p+\omega_L^d\right) \\
&\quad+\Lambda^P\frac{(\sigma-1)}{\left[\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma}-1\right]^2}\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{-\sigma}\left(\omega_S^p+\omega_L^d\right) \\
&\quad\times(-1)(1-\sigma)\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{-\sigma}\left(\omega_S^p+\omega_L^d\right) \\
&= \Lambda^P\frac{(\sigma-1)}{\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma}-1}\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{-\sigma}\left(\omega_S^p+\omega_L^d\right) \\
&\quad\times\left[\frac{2(\sigma-1)}{\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma}-1}\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{-\sigma}\left(\omega_S^p+\omega_L^d\right)-\sigma\beta(z)^{-\sigma-1}\right] \\
&> 0,
\end{aligned}$$

if $\frac{2(\sigma-1)}{\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma}-1}\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{-\sigma}\left(\omega_S^p+\omega_L^d\right)-\sigma\beta(z)^{-\sigma-1}>0$, or

$$\frac{2\left[\tau\left(\omega_N^p+\omega_L^d\right)\right]^{\sigma}\left(\omega_S^p+\omega_L^d\right)^{1-\sigma}\beta(z)}{\left(\frac{\beta(z)\left(\omega_S^p+\omega_L^d\right)}{\tau\left(\omega_N^p+\omega_L^d\right)}\right)^{1-\sigma}-1}>\frac{\sigma}{\sigma-1},$$

which requires $\beta(z)$ high enough.

Next

$$\begin{aligned}\frac{\partial \Lambda^I(z)}{\partial \beta(z)} &= \Lambda^I(z)\frac{(\sigma-1)}{\beta(z)},\\ \frac{\partial^2 \Lambda^I(z)}{\partial \beta(z)^2} &= \frac{\partial \Lambda^I(z)}{\partial \beta(z)}\frac{(\sigma-1)}{\beta(z)} - \Lambda^I(z)\,(\sigma-1)\,\frac{1}{\beta(z)^2}\\ &= \Lambda^I(z)\,(\sigma-1)\,\frac{1}{\beta(z)^2}\,(\sigma-2) > 0,\end{aligned}$$

iff $\sigma > 2$.

Then increases in β have magnified effects in sectors of higher skill sensitivity. Conversely, improvements in the South's IPR regime reduces $\Lambda^P(z)$ and $\Lambda^I(z)$ by relatively more in sectors of greater skill sensitivity. ■

C.12 General Equilibrium

C.12.1 Deriving skilled labour demand

We determine the Northern skilled wage rate. Let $r_N(z) = zr$ and $r_S(z) = \beta zr$, $\beta > 1$. z is then the unit skilled labour requirement. Skilled labour demand comes from two sources: firms producing for the domestic market and those producing for the Southern market. For domestic firms in sector z with integrated distribution, charging price of $p_d(\varphi, z) = \frac{r_N(z)\left(\omega_N^p + \omega_I^d\right)}{\rho\varphi}$, demand is $q(\varphi, z) = \theta_d(z)p_d(\varphi, z)^{-\sigma}$ where $\theta_d(z) \equiv \mu\xi(z)P_d(z)^{\sigma-1}$ denotes domestic demand and $P_d(z)^{\sigma-1}$ is the domestic price index. This is given by $P_d(z)^{1-\sigma} = \int_1^\infty p_d(\varphi, z)^{1-\sigma}N(z)g(\varphi)d\varphi$ such that

$$\begin{aligned}P_d(z)^{1-\sigma} &= N(z)\left[\frac{r_N(z)\left(\omega_N^p + \omega_I^d\right)}{\rho}\right]^{1-\sigma}\int_1^\infty \varphi^{\sigma-1}g(\varphi)d\varphi\\ &= N(z)p_d[\widetilde{\varphi}_d(z)]^{1-\sigma},\\ \widetilde{\varphi}_d(z)^{\sigma-1} &\equiv \int_1^\infty \varphi^{\sigma-1}g(\varphi)d\varphi.\end{aligned}$$

With $r_N(z) = zr$ this can be written $P_d(z)^{1-\sigma} = N(z)\left[\frac{zr\left(\omega_N^p + \omega_I^d\right)}{\rho\widetilde{\varphi}_d(z)}\right]^{1-\sigma}$ such that

$$\begin{aligned}P_d(z) &= N(z)^{\frac{1}{1-\sigma}}\left[\frac{zr\left(\omega_N^p + \omega_I^d\right)}{\rho\widetilde{\varphi}_d(z)}\right],\\ \frac{\partial P_d(z)}{\partial r} &> 0,\end{aligned}$$

implying $\frac{\partial\theta_d(z)}{\partial r} > 0$. Skilled labour demand from firms in sector z producing for the domestic market is then

$$\begin{aligned}
d(z,r) &= \int_1^\infty q(\varphi,z)zN(z)g(\varphi)d\varphi \\
&= \theta_d(z)N(z)\left[\frac{zr\left(\omega_N^p+\omega_I^d\right)}{\rho}\right]^{-\sigma} z\widehat{\varphi}, \\
\widehat{\varphi} &\equiv \int_1^\infty \varphi^\sigma g(\varphi)d\varphi.
\end{aligned}$$

Use that

$$\begin{aligned}
\theta_d(z) &= \mu\xi(z)\left[N(z)^{\frac{1}{1-\sigma}}\left(\frac{zr\left(\omega_N^p+\omega_I^d\right)}{\rho\widetilde{\varphi}_d(z)}\right)\right]^{\sigma-1} \\
&= \mu\xi(z)\frac{1}{N(z)}\left[\frac{zr\left(\omega_N^p+\omega_I^d\right)}{\rho\widetilde{\varphi}_d(z)}\right]^{\sigma-1},
\end{aligned}$$

such that

$$\begin{aligned}
d(z,r) &= \mu\xi(z)\frac{1}{N(z)}\left[\frac{zr\left(\omega_N^p+\omega_I^d\right)}{\rho\widetilde{\varphi}_d(z)}\right]^{\sigma-1} N(z)\left[\frac{zr\left(\omega_N^p+\omega_I^d\right)}{\rho}\right]^{-\sigma} z\widehat{\varphi} \\
&= \mu\xi(z)\frac{\rho}{r\omega_N^p}\frac{\widehat{\varphi}}{\widetilde{\varphi}_d(z)^{\sigma-1}}, \\
\frac{\partial d(z,r)}{\partial r} &< 0.
\end{aligned}$$

For Northern MNCs serving the Southern market,

$$d^*(z,r) = \int_{\varphi^X(z)}^{\varphi^P(z)} q(\varphi,z)zn(z)\chi(\varphi,z)d\varphi + \int_{\varphi^P(z)}^\infty q(\varphi,z)\beta(z)zn(z)\chi(\varphi,z)d\varphi.$$

Use the price index to write

$$\begin{aligned}
& d^*(z,r) \\
=\ & \int_{\varphi^X(z)}^{\varphi^P(z)} \theta(z)p(\varphi,z)^{-\sigma}zn(z)\chi(\varphi,z)d\varphi + \int_{\varphi^P(z)}^\infty \theta(z)p(\varphi,z)^{-\sigma}\beta(z)zn(z)\chi(\varphi,z)d\varphi \\
=\ & \mu\xi(z)\left[\frac{zr}{\rho}\right]^{\sigma-1}\left\{\varphi^X(z)^{\psi_1}\left[\tau^{1-\sigma}\Delta_N+\beta(z)^{1-\sigma}\Delta_S\right]\right\}^{-1} \\
& \times z\left[\int_{\varphi^X(z)}^{\varphi^P(z)} p(\varphi,z)^{-\sigma}g(\varphi)d\varphi + \beta(z)\int_{\varphi^P(z)}^\infty p(\varphi,z)^{-\sigma}g(\varphi)d\varphi\right].
\end{aligned}$$

Firm prices are $p^X(\varphi,z)=\frac{\tau r_N(z)\left(\omega_N^p+\omega_L^d\right)}{\rho\varphi}$, $p^D(\varphi,z)=\frac{\tau r_N(z)\left(\omega_N^p+\omega_I^d\right)}{\rho\varphi}$, $p^P(\varphi,z)=\frac{r_S(z)\left(\omega_S^p+\omega_L^d\right)}{\rho\varphi}$, $p^I(\varphi,z)=\frac{r_S(z)\left(\omega_S^p+\omega_I^d\right)}{\rho\varphi}$. Write

$$\begin{aligned}
& \int_{\varphi^X(z)}^{\varphi^P(z)} p(\varphi,z)^{-\sigma} g(\varphi)d\varphi + \beta(z)\int_{\varphi^P(z)}^{\infty} p(\varphi,z)^{-\sigma} g(\varphi)d\varphi \\
= \; & r^{-\sigma}\left[\int_{\varphi^X(z)}^{\varphi^P(z)} \left[\frac{\tau z\left(\omega_N^p+\omega_l^d\right)}{\rho\varphi}\right]^{-\sigma} g(\varphi)d\varphi + \beta(z)\int_{\varphi^P(z)}^{\infty}\left[\frac{\beta(z) z\left(\omega_S^p+\omega_l^d\right)}{\rho\varphi}\right]^{-\sigma} g(\varphi)d\varphi\right] \\
= \; & r^{-\sigma}\times A,
\end{aligned}$$

where

$$\begin{aligned}
A \;&\equiv\; \int_{\varphi^X(z)}^{\varphi^P(z)} \left[\frac{\tau z\left(\omega_N^p+\omega_l^d\right)}{\rho\varphi}\right]^{-\sigma} g(\varphi)d\varphi + \beta(z)\int_{\varphi^P(z)}^{\infty}\left[\frac{\beta(z) z\left(\omega_S^p+\omega_l^d\right)}{\rho\varphi}\right]^{-\sigma} g(\varphi)d\varphi, \\
l \;&=\; L, I
\end{aligned}$$

so

$$d^*(z,r)=\frac{\mu\xi(z)z^\sigma A}{\rho^{\sigma-1} r\varphi^X(z)^{\psi_1}\left[\tau^{1-\sigma}\Delta_N+\beta(z)^{1-\sigma}\Delta_S\right]}.$$

Recall that in equilibrium

$$\left[\varphi^X(z)\right]^\delta=\frac{f_x w_N N(z)}{\mu\xi(z)(1-\rho)}\left(\omega_N^p+\omega_L^d\right)^{\sigma-1}\left[\Delta_N+\left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\Delta_S\right],$$

which is independent of r. So we can write

$$d^*(z,r)=\frac{\mu\xi(z)z^\sigma A\tau^{\sigma-1}}{\rho^{\sigma-1} r\left\{\frac{f_x w_N N(z)}{\mu\xi(z)(1-\rho)}\left(\omega_N^p+\omega_L^d\right)^{\sigma-1}\right\}^{\frac{\sigma-\delta-1}{\delta}}\left[\Delta_N+\left(\frac{\tau}{\beta(z)}\right)^{\sigma-1}\Delta_S\right]^{\frac{\sigma-1}{\delta}}},$$

giving

$$\frac{\partial d^*(z,r)}{\partial r}<0.$$

Then total skilled labour demand in sector z is

$$\begin{aligned}
D(z,r) \;&=\; d(z,r)+d^*(z,r), \\
\frac{\partial D(z,r)}{\partial r} \;&<\; 0,
\end{aligned}$$

and total skilled labour demand is

$$\overline{D}(r)=\int_0^1 D(z,r)dz.$$

The equilibrium skilled wage is then determined by

$$K = \overline{D}(r^*),$$

such that $\frac{\partial r*}{\partial K} < 0$.

References

Ando, M. & Kimura, F. (2003), ‘The Formation of International Production and Distribution Networks in East Asia’, *NBER Working Paper* .

Brainard, S. (1993), ‘A simple theory of multinational corporations and trade with a trade-off between proximity and concentration’, *NBER working paper 4583* .

Brainard, S. (1997), ‘An empirical assessment of the proximity-concentration trade-off between multinational sales and trade’, *American Economic Review* pp. 520–544.

Carlos, A. & Nicholas, S. (1990), ‘Agency problems in early chartered companies: the case of the Hudson’s Bay Company’, *Journal of Economic History* pp. 853–875.

Caves, R. (1971), ‘International corporations: the industrial economics of foreign investment’, *Economica* pp. 1–27.

Defever, F. (2006), ‘Functional fragmentation and the location of multinational firms in the enlarged Europe’, *Regional Science and Urban Economics* **36**(5), 658–677.

Dietz, M., Lin, S. & Yang, L. (2005), ‘Protecting intellectual property in China’, *The McKinsey Quarterly* **3**.

Dunning, J. (1981), *International production and the multinational enterprise*, London: Geoge Allen and Unwin.

Edmans, A. & Gabaix, X. (2009), ‘Detail-Independent Optimal Contracts, with an Application to CEO Incentives’, *MIT, Mimeo* .

Ethier, W. (1986), ‘The multinational firm’, *Quarterly Journal of Economics* pp. 805–834.

Ethier, W. & Horn, H. (1990), ‘Managerial control of international firms and patterns of direct investment’, *Journal of International Economics* **28**(1-2), 25–45.

Ethier, W. & Markusen, J. (1996), ‘Multinational firms, technology diffusion and trade’, *Journal of International Economics* **41**(1-2), 1–28.

Fitzgerald, R. (1995), *Rowntree and the marketing revolution, 1862-1969*, Cambridge University Press.

Hanson, G., Mataloni, R. & Slaughter, M. (2001), 'Expansion strategies of US multinational firms', *NBER Working Paper* .

Helpman, E. (1984), 'A simple theory of international trade with multinational corporations', *Journal of Political Economy* pp. 451–471.

Helpman, E. (1985), 'Multinational corporations and trade structure', *Review of Economic Studies* pp. 443–457.

Helpman, E., Melitz, M. & Yeaple, S. (2004), 'Export versus FDI with heterogeneous firms', *American Economic Review* **94**(1), 300–316.

Holmstrom, B. & Milgrom, P. (1991), 'Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design.', *Journal of Law, Economics and Organization* **7**, 24–52.

Horstmann, I. & Markusen, J. (1987*a*), 'Licensing versus direct investment: a model of internalization by the multinational enterprise', *Canadian Journal of Economics* pp. 464–481.

Horstmann, I. & Markusen, J. (1987*b*), 'Strategic investments and the development of multinationals', *International Economic Review* pp. 109–121.

Horstmann, I. & Markusen, J. (1996), 'Exploring new markets: direct investment, contractual relations and the multinational enterprise', *International Economic Review* pp. 1–19.

Javorcik, B. (2004), 'The composition of foreign direct investment and protection of intellectual property rights: Evidence from transition economies', *European Economic Review* **48**(1), 39–62.

Markusen, J. (1984), 'Multinationals, multi-plant economies, and the gains from trade', *Journal of International Economics* **16**(3-4), 205–226.

Markusen, J. (1997), 'Trade versus investment liberalization', *NBER working paper 6231* .

Markusen, J. (2002), *Multinational firms and the theory of international trade*, MIT press.

Markusen, J. & Venables, A. (1998), 'Multinational firms and the new trade theory', *Journal of International Economics* **46**(2), 183–203.

Markusen, J. & Venables, A. (2000), 'The theory of endowment, intra-industry and multi-national trade', *Journal of international economics* **52**(2), 209–234.

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

Neary, J. P. (2007), 'Cross-border mergers as instruments of comparative advantage', *Review of Economic Studies* **74**(4), 1229–1257.

Neary, J. P. (2009), ‘Trade costs and foreign direct investment’, *International Review of Economics and Finance* **18**(2), 207–218.

Nicholas, S. (1983), ‘Agency contracts, institutional modes, and the transition to foreign direct investment by British manufacturing multinationals before 1939’, *Journal of Economic History* pp. 675–686.

Nocke, V. & Yeaple, S. (2007), ‘Cross-border mergers and acquisitions vs. greenfield foreign direct investment: The role of firm heterogeneity’, *Journal of International Economics* **72**(2), 336–365.

Nocke, V. & Yeaple, S. (2008), ‘An assignment theory of foreign direct investment’, *Review of Economic Studies* **75**(2), 529–557.

Norbäck, P. (2001), ‘Multinational firms, technology and location’, *Journal of international economics* **54**, 449–469.

Willan, T. (1956), *The early history of the Russia Company, 1553-1603*, Manchester University Press.