

**Wake Up and Smell the Ginseng:
International Trade and the Rise of Incremental Innovation
in Low-Wage Countries**

Diego Puga^{*‡}

IMDEA, Universidad Carlos III and CEPR

Daniel Trefler^{*§}

University of Toronto

First version, August 2005

This version, January 2009

ABSTRACT: Increasingly, a small number of low-wage countries such as China and India are involved in incremental innovation. That is, they are responsible for resolving production-line bugs and suggesting product improvements. We provide evidence of this new phenomenon and develop a model in which there is a transition from old-style product-cycle trade to trade involving incremental innovation in low-wage countries. The model explains why levels of involvement in incremental innovation vary across low-wage countries and across firms within each low-wage country. We draw out implications for sectoral earnings, living standards, the capital account and, foremost, international trade in goods.

Key words: international trade, low-wage country innovation

JEL classification: F1

^{*}We are thankful to Joel Blit for research assistance and to Pol Antràs and Elhanan Helpman for helpful comments on an earlier draft. Both authors thank the Canadian Institute for Advanced Research (CIFAR) for its tremendous support. Funding from the Social Sciences and Humanities Research Council of Canada (SSHRC), Spain's Ministerio de Educación y Ciencia (SEJ2006-09993), the Comunidad de Madrid (PROCIUDAD-CM), and the Centre de Recerca en Economia Internacional is gratefully acknowledged.

[‡]Madrid Institute for Advanced Studies (IMDEA) Social Sciences, Antiguo pabellón central del Hospital de Cantoblanco, Carretera de Colmenar Viejo km. 14, 28049 Madrid, Spain (e-mail: diego.puga@imdea.org; website: <http://diegopuga.org>).

[§]Rotman School of Management and Department of Economics, University of Toronto, 105 Saint George Street, Toronto, Ontario M5S 3E6, Canada (e-mail: dtrefler@rotman.utoronto.ca; website: <http://www.rotman.utoronto.ca/~dtrefler/>). Also affiliated with the National Bureau of Economic Research (NBER).

1. Introduction

When the auto parts giant Delphi Technologies first set up shop in Chihuahua, no one predicted that the Mexican shop-floor engineers at this low-wage affair would soon be introducing minor product changes that would slash product defect rates. While the cumulative effects of these product changes are large, no single innovation is pathbreaking. Even the most sophisticated innovations — those that actually generate patents — are just better mouse traps that incrementally improve on existing auto parts technology. For example, most Delphi-Chihuahua patents improve on the control systems of minor moving parts. These patents are examples of incremental innovation, Rosenberg's (1982*a*) unsung hero of modern economic growth. This paper is about the rise of incremental innovation in a select group of low-wage countries, how it is allowing these countries to export increasingly sophisticated new goods, and why it is leading to a shift in world trade patterns.

Incremental innovation in low-wage countries is not part of the lexicon of international trade theory. Instead, we are glued to Vernon's (1966) product-cycle theory in which products and processes are developed and standardized in rich countries before being moved offshore to low-wage countries. (See also Krugman, 1979, and the more sophisticated approaches of Grossman and Helpman, 1991*b,c* and Antràs, 2005.) Increasingly, though, the first location of production for moderately sophisticated goods is taking place in a handful of low-wage countries such as China. This implies that the many bugs that plague new goods and their production processes are partly and sometimes wholly resolved by local managers and engineers. That is, a handful of low-wage countries are becoming engaged in incremental innovation. If incremental innovation is indeed the unsung hero of modern economic growth, then the long-term implications of this development are significant.

From the perspective of a rich-country firm, there are costs and benefits of involving low-wage countries in incremental innovation. One key benefit is that it frees up valuable innovation resources at home so that they can be focused on 'big ideas' innovation. In addition, low-wage country involvement in incremental innovation allows the firm to locate production in a low-wage country even before products and processes are fully developed and standardized: local engineers are used to help complete the process of standardization. By locating a good's first production line in a low-wage country, the process of lowering labour costs is telescoped forward. Another benefit is that in a world of complex foreign supply chains, a firm that involves local suppliers in incremental innovation can insist that each supplier deliver continual product and process upgrades. Being directly engaged in the production process, suppliers can come up with improvements on the shop floor that would have been more costly for the firm to identify from head office. These additional improvements allow firms such as Delphi to stay a hair's breadth ahead of the competition.

Against these benefits are some significant drawbacks of involving local suppliers in incremental innovation. First and foremost, parts suppliers from low-wage countries typically produce components for complex, interdependent systems in which an incremental improvement in one component is not effective unless other components are also modified. This interdependence

means that parts suppliers do not internalize all of a firm's innovation costs. In the simplest case, when a firm asks a parts supplier to improve a component, the solution will usually entail *residual incompatibilities* with other components of the system, thus forcing the firm to incur the additional expense of bringing other components into line.

We also model a second and more familiar cost of involving a local supplier in incremental innovation. Once involved, the supplier acquires information and expertise that can be used outside the relationship. This makes it necessary for the firm to offer expensive incentives aimed at encouraging the supplier to stay within the relationship.

The modelling core of this paper is about the firm's choice of whether or not to involve a supplier in innovation. Nothing in the formal set-up of the model is specific to incremental innovation as opposed to more substantial innovation. Thus, the reader should feel free to interpret our model in terms of innovation more broadly defined. Our reading of the available evidence, however, persuades us that innovation in low-wage countries is to date dominated by incremental product and process innovation. We embed the firm's choice of whether or not to involve a supplier in innovation into a general equilibrium model. Each firm in the model chooses where to locate production among a set of low-wage countries and whether or not to involve locals in innovation. The general equilibrium choices of firms aggregate to yield predictions about the extent to which a low-wage country will attract rich-country firms and engage in incremental innovation. We then build on these results to generate a rich set of predictions about international trade in goods, earnings and living standards and, to some extent, the capital account.

Empirical Evidence

Evidence on incremental innovation in rich countries is abundant e.g., Rosenberg (1982*b*) and von Hippel (1988). Systematic evidence for developing countries is only just emerging e.g., Sutton (2001, 2004).¹ Fortunately, there are two observable activities that are tied into incremental innovation. First, incremental innovation is most intense during the early stages of the production of new goods, when product bugs are fixed and production processes optimized. This is why learning curves are steepest by far in the earliest stages of production (e.g., Lucas, 1993, and Nahmias, 2005). Thus, one way to track incremental innovation is to track where new goods are produced. Second, while incremental innovation is about minor cumulative improvements to products and processes, occasionally it leads to patents. Thus, a second way to track incremental innovation is to track patents developed by residents in low-wage countries for rich-country corporations.

Some evidence on the first location of production of new goods can be culled from the business strategies of contract manufacturers. The industry is dominated by six giants. Typical of these is

¹ Anecdotal evidence on more substantial innovations in developing countries frequently appears in the press. For example, 3com's 8800 high-end network switch and Nokia's 6108 handset were developed in China. The Organisation for Economic Co-operation and Development (2006) predicts that by the end of 2006 China will be second only to the us in R&D expenditures. However, the incremental innovations on which our paper focuses are rarely the stuff of news reports or formal R&D expenditures.

Sanmina-SCI.² Sanmina-SCI reports that 54 of its 159 plants are exclusively engaged in producing new goods and developing new products. The top locations for these plants are (number of plants in parentheses): US (15); Canada (5); China, Hungary and Sweden (4 each); Finland, Germany and Israel (3 each); and the UK and Mexico (2 each). What stands out from this list is the appearance of three lower-wage countries: China, Hungary and Mexico.

For more systematic evidence showing that this Sanmina-SCI example generalizes we turn to trade data. Following Feenstra and Rose (2000), we can identify where new goods are first produced using US import statistics. The United States classifies imports using 10-digit Harmonized System (HS) codes, which currently cover over 17,000 products and are regularly updated to incorporate new goods. A 10-digit HS good is 'new in year t ' if it was first imported into the US in year t . For example, the United States first imported 20-page per minute laser printers (HS 8471923600) in 1994, primarily from Japan. For another example, the US first imported machinery for making DVDs (HS 8477104000) in 1995, with two-thirds of imports by value coming from Germany, the Netherlands and China.³

Table 1 shows that production of new goods is beginning to happen in low-wage countries. The table reports statistics for innovative new goods. Innovative goods are defined as goods in sectors identified by the US National Science Board (2006) as being medium-high-technology and high-technology sectors i.e., chemicals (including pharmaceuticals), machinery (including electronics), transportation equipment, and instruments. In the 1990-92 period, the US imported 694 innovative new goods. Of these, 472 or 68% were exported by Japan to the US (see column 1). A decade later, of the innovative new goods that first appeared in the 2000-02 period, only 55% of these were exported by Japan to the US (see column 2). The decline of 13% is statistically significant ($t = -15.51$). In contrast, only 17% of innovative new goods arrived from China in 1990-92, whereas 48% arrived from China in 2000-02. This is an increase of 30 percentage points ($t = 35.80$). The other big gainers among poor and middle-income countries were Malaysia, Thailand, India, Mexico and the Philippines.⁴

Changes in the number of new goods could be an artifact of changes in US administrative procedures used to create HS codes. For example, if these procedures became easier over time then all countries would experience an upward trend in the number of new goods. One way of controlling for a common trend associated with changed administrative procedures is to double-difference the data. For example, one could compare China's change of 30% with the Japanese change of

²The six giants are (employment in parentheses): Hon Hai (167,000), Flextronics (92,000), Jabil Circuit (61,000), Solectron (57,000), Sanmina-SCI (53,000), and Celestica (42,000). We focus on Sanmina-SCI because it alone provides information about what each of its plants does in each country. Sanmina-SCI appears typical in that its sales per worker is \$201,000, which is in the middle of the \$166,000-\$244,000 range for the six giants.

³Both 20-page per minute laser printers and machinery for making DVDs were incremental innovations over earlier products (i.e., slower laser printers and other plastic-moulding machinery). Following those incremental innovations, an HS code was created specifically for them. In the case of laser printers, the key is that fast printers (i.e., 20 pages per minute) only appeared in 1994 with the introduction of the Xerox 4520ps and DataProducts LZR2080 printers. Likewise, the HS code for DVD machinery was introduced the same year that an international agreement on standards for DVDs was reached. The first DVD players were brought to market the following year.

⁴Data on US imports at the 10-digit HS level are subject to mis-classification by customs officials and jobbers, which result in minuscule amounts of exports of most goods by many countries. To eliminate such errors, a country is only counted as having exported a new good to the US if it accounted for at least 1% of US imports of that particular new good. Had we not done this, the statistics in table 1 would have been even more supportive of our claims.

	% of New Goods Imported by the United States from Each Country						
	By Number				By Value		
	1990–92	2000–02	Change	<i>t</i>	1990–92	2000–02	Change
(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Rich							
Japan	68%	55%	-13%	(-15.51)	34%	9%	-24%
Canada	46%	46%	0%	(0.31)	25%	7%	-18%
Germany	62%	62%	-1%	(0.86)	10%	7%	-4%
Poor							
China	17%	48%	30%	(35.80)	0%	5%	5%
Malaysia	9%	16%	8%	(9.25)	2%	7%	5%
Thailand	5%	11%	6%	(6.91)	1%	1%	0%
India	6%	10%	5%	(5.42)	0%	0%	0%
Mexico	19%	23%	4%	(4.57)	6%	4%	-2%
Philippines	6%	9%	4%	(4.51)	1%	4%	3%
Indonesia	1%	3%	2%	(2.27)	0%	0%	0%

Source: Authors' calculations based on data from us Bureau of the Census.

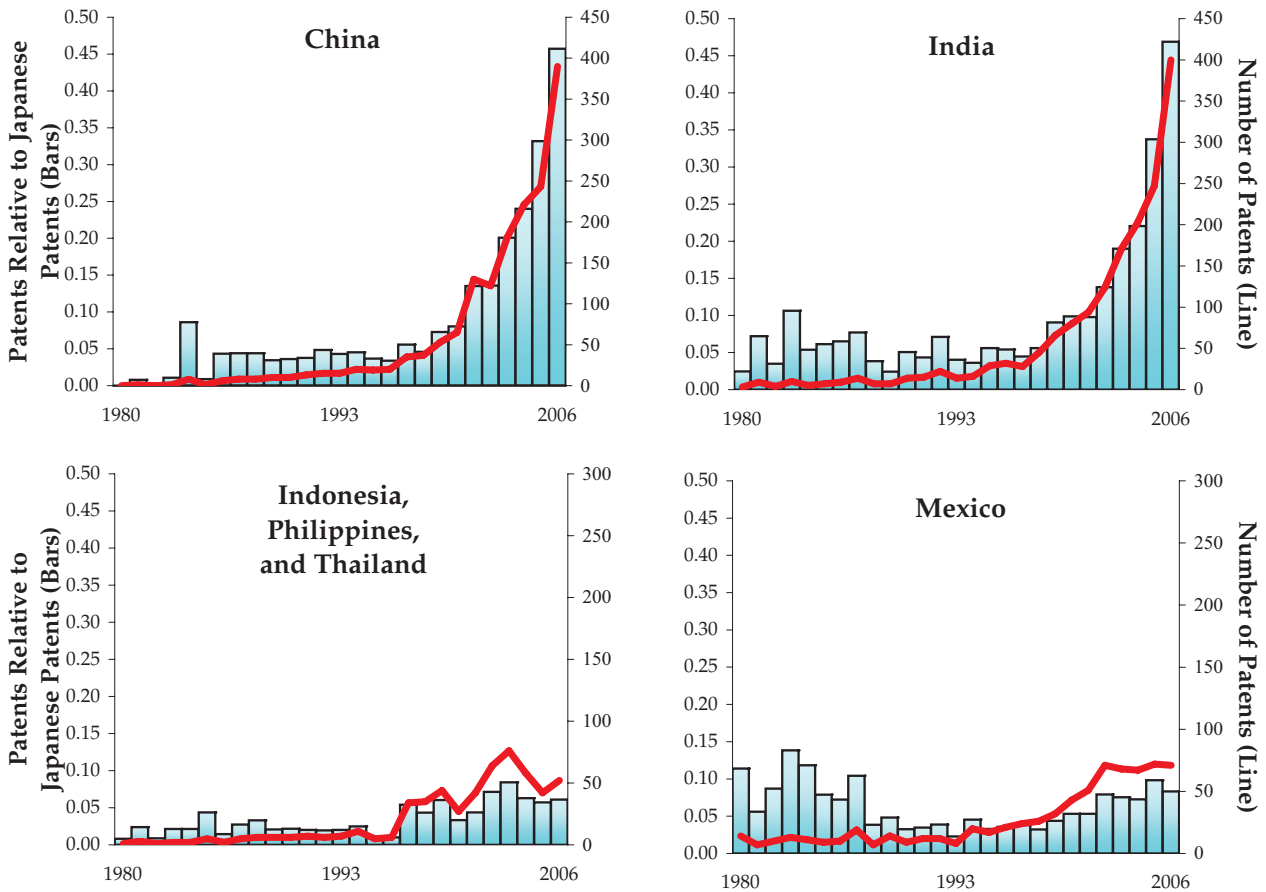
Notes: This table reports statistics on goods that were first imported by the United States either in the period 1990–92 or the period 2000–02. Only new goods in medium-high-technology and high-technology sectors are included (us National Science Board, 2006): chemicals (including pharmaceuticals), machinery (including electronics), transportation equipment, and instruments. Column 1 states that of the innovative new goods first imported by the United States in 1990-92, Japan was an exporter 68% of the time. Column 1 sums to more than 100% because most new goods are imported into the United States from several countries. Column 5 states that of the innovative new goods first imported by the United States in 1990-92, 34% by value was imported from Japan. Column 5 would sum to 100% if all us trading partners were listed.

Table 1: Origin of us Imports of Innovative New Goods

-13%. The difference of 43% has a *t*-statistic of 14.34. Comparing Japan with each of the other poor countries in table 1, the *t*-statistic for each double difference is never less than 4.97. Thus, our conclusion that some poor countries are increasingly the location of first production of innovative goods does not appear to be an artifact of administrative procedures for introducing new HS codes.

In addition to showing where early-stage production takes place, table 1 displays two other features that will be outcomes of our model. First, the list of countries is short. All other middle-income and poor countries never exported substantial numbers of innovative new goods. We have included Indonesia in table 1 in order to make this point. Even though a considerable amount of manufacturing takes place in Indonesia, it is not the location of first production of innovative goods. First location of production is confined to only a handful of countries. Second, new goods are often exported to the us by several countries at once. This would not happen in a quality-ladder model where only a single producer survives e.g., Grossman and Helpman (1991c). In our model, any firm that has developed a new good can bring it to market provided that it successfully engages in incremental innovation. Our focus is not on developing new goods, but on incremental innovation.

Columns 5–7 of table 1 provide an alternative view of the same phenomenon. They report the share of us imports of innovative new goods that originated in the indicated country by value instead of by number. For example, Japan accounted for 34% of us imports of innovative new



Source: Authors' calculations from the United States Patent and Trademark Office website (<http://www.uspto.gov/>).

Notes: Lines (right-hand scale) indicate the *total number* of us-owned patents with at least one inventor who is a resident of the indicated country. Bars (left-hand scale) scale the lines by the number of us-owned patents with at least one inventor who is a resident of Japan.

Figure 1: Share of us-Owned Patents with a Foreign Inventor

goods in 1990–92, but only 9% in 2000–02. In contrast, China's share rose from 0% to 5% in just a decade. Thus, the conclusion that low-wage countries are beginning to be the first location of production comes both from frequency data and trade volume data. Table 1 leaves little doubt that some low-wage countries are now the location of early stages of production and hence of the incremental innovation that is tied to early-stage production.

A second source of evidence on incremental innovation is suggested by our introductory example of Delphi patents generated by Mexican engineers at Delphi's Guadalajara facility. This source is the us patent database (<http://www.uspto.gov/>). For each patent, the database lists whether the patent is owned by a us entity (such as Michigan-based Delphi) and whether any of the inventors reside in a low-wage country. We identified for analysis each us patent that (a) is owned by a us entity i.e., has a us assignee and (b) includes a *resident* of a foreign country in its list of inventors. Perusal of detailed patent applications for this group of patents suggests that the Delphi example is common for patents with an inventor from a low-wage country. That is, many

of these patents do indeed capture incremental innovation.

The top left panel of figure 1 tracks all such patents with at least one inventor that resides in China. (These Chinese residents invariably have Chinese names, which suggests they are not us engineers working in China.) The solid line gives the absolute number of such patents. To facilitate interpretation of the line we also scale it by the corresponding number for Japan i.e., the number of us patents with a us assignee and an inventor residing in Japan. This is shown in the bars and the left-hand axis. Scaling does two things. First, it controls for the explosion of patenting activity in the us since the mid-1990s. Second, it gives a clearer sense of the level of Chinese innovative activity: by 2006 it was almost half that of Japan, a country that is among the most innovative on the planet.

Three features stand out. (1) In recent years there has been an explosion of us-owned patents involving Chinese and Indian inventors. While in 1981 the Chinese bar was zero, in 2006 it was 0.46 which indicates that this measure of Chinese innovative activity is at 46% of Japanese levels.⁵ Similarly, 2006 Indian activity is at 47% of Japanese levels.⁶ (2) In contrast to China and India, many other low-wage countries have barely been touched by this tide of incremental innovation. Even the combined numbers for Indonesia, the Philippines and Thailand are small. As with us imports of innovative new goods, there are only a handful of low-wage countries appearing frequently in the us patent database. (3) Since 2002, patenting has plateaued in Mexico (and some other popular FDI destinations in Eastern Europe and Latin America) as companies such as Delphi have shifted production to China. This is an important general equilibrium effect that we will be modelling.

The patent database does not distinguish between incremental patents and 'fundamental' (non-incremental) patents. A casual examination of the patents underlying figure 1 suggests that the Chinese patents are typically incremental whereas the Japanese patents are a mix of incremental and fundamental. Thus, had we been able to look only at incremental patents the figure 1 surge would be even more pronounced.⁷

Summarizing, us import and patent data both provide consistent and systematic evidence that a small number of low-wage countries are increasingly engaged in incremental product and process innovation. We will show that this has important implications for the rapidly evolving pattern of world trade.

⁵Another way of thinking about scaling by Japan is that it controls for size as measured by GDP. Japanese GDP is the same as that of India and about half that of China. Hence, the Chinese and Indian numbers are large even relative to GDP. Further, the Chinese and Indian numbers have been growing rapidly relative to GDP. Over the 2000–2006 period these patents have grown three times faster than GDP in both countries. Finally, the trends in figure 1 are almost identical when scaling is by total us patents rather than Japanese-inventor patents.

⁶The fact that patent numbers for India are virtually identical to those for China is surprising given India's low values in table 1. The explanation is that Indian patents are mostly related to services.

⁷We have tried to identify incremental patents by equating them with process innovation. This does not work. For one, it is surprisingly hard to identify process innovations. For example, in the Canadian innovation survey used by Lileeva and Trefler (2007), firms are explicitly asked to identify which of their innovations are process innovations. Yet half of the firms report that they cannot: their product innovations are inseparable from their process innovations. For another, many incremental innovations are in fact product innovations. In our introductory Delphi example, electric door locks are made more reliable by improvements to solenoids. This is an incremental product innovation (where the product is solenoids or door locks or autos). In short, it is difficult to equate incremental innovation with process innovation.

Related literature

This paper has many touchstones with the existing international trade literature. Vernon's (1966) product-cycle model and its dynamic Ricardian variants (e.g., Krugman, 1979, Grossman and Helpman, 1991*b,c*, and Antràs, 2005) either assume or predict that innovation occurs exclusively in rich countries. They rule out innovation in low-wage countries. In the absence of local innovation, technologies diffuse to low-wage countries via such channels as imported machinery, FDI, scientific journals, technology licensing and theft (Grossman and Helpman, 1991*a*, Markusen, 2002). A few papers (e.g., Grossman and Helpman, 1991*b,c*) allow agents in low-wage countries to actively invest in acquiring knowledge. However, this knowledge acquisition is just reverse engineering of products originally developed and produced in rich countries.

Our paper also fits into the literature on incomplete contracts and trade. However, in order to focus on our novel contribution we sidestep the two most important questions addressed by this literature. The first is about the choice between vertical integration and outsourcing in general equilibrium (e.g., McLaren, 2000, Grossman and Helpman, 2002, 2003, Antràs, 2003, Nunn and Trefler, 2008). The second is about the choice between sourcing inputs from the North or from the South (e.g., Grossman and Helpman, 2005).⁸ We do not tackle these two questions, choosing instead to focus on the conditions that promote incremental innovation in low-wage countries. Our starting point is that a Northern firm has already decided to locate production in the South. We model the firm's choices about (i) which of several low-wage countries to enter and (ii) whether or not to involve a local agent in incremental innovation.⁹

Notice that we are silent on whether we are dealing with FDI or offshore outsourcing. What matters to us is only whether incremental innovation is being done. However, we will be assuming that if residual incompatibilities within the match are large, the principal takes control of incremental innovation. This is appropriate in a long-term relationship i.e., in an FDI relationship. In a short-term offshore outsourcing relationship it would be more appropriate to assume that the principal walks away from relationships with high residual incompatibilities. In the real world there is frequently a medium-term relationship, namely, that the principal takes a minority position in the agent's firm. Our assumption is appropriate for such medium-term relationships as well.

Our model has a second cost of doing incremental innovation in low-wage countries, namely, the possibility that a local parts supplier could take the knowledge gained from incremental innovation and use it outside of the relationship. This means that our paper is also related to work on contract enforcement and trade. See the seminal work of Ethier and Markusen (1996) as well as Markusen (2002) and Nunn (2007). Weak contract enforcement is one of several possible reasons why a local parts supplier can use proprietary knowledge outside of the relationship.

⁸Several papers combine both questions in order to deal with the choice between integrated home production, domestic outsourcing, FDI, and offshore outsourcing, or some subset of these (e.g., Antràs and Helpman, 2004, Grossman and Helpman, 2004, and also Antràs, 2005, who integrates this choice in a product-cycle model). Spencer (2005), Trefler (2006) and Helpman (2006) review this literature.

⁹In an earlier draft of this paper (Puga and Trefler, 2005) there was a third choice, namely, whether to delegate control of incremental innovation decisions to the local agent. This aspect of the paper built on our earlier closed-economy work on organizational choice with incremental innovation (Puga and Trefler, 2002) which in turn built on Aghion and Tirole (1997). This choice and the common elements with Puga and Trefler (2002) have been eliminated in the current version of the paper in order to simplify the analysis.

Ethier and Markusen (1996) are interested in the choice between exporting and producing abroad for goods whose production has already been standardized. Producing abroad allows the firm to benefit from lower wages, but at a cost: it also allows the local agent to *passively* absorb the firm's technology and steal it. This is an important insight. In contrast, we are interested in *active* participation of local agents in the incremental innovation of new goods.

There are also some interesting comparisons between our work and that of Casella and Rauch (2002) and Rauch and Trindade (2003). Their models elegantly focus on search and matching frictions associated with imperfect information in a single foreign market. In contrast, our model focuses on agency problems associated with incremental innovation in multiple foreign markets. They are interested in how imperfect information reduces trade between the domestic country and the foreign country whereas we are interested in the implications of incremental innovation for the choice between foreign locations. There is however a similarity. Both business networks and agent innovation increase the value of matching in a foreign country. This makes foreign countries more attractive places in which to produce.

The paper is organized as follows. Section 2 sets up the model. Section 3 analyzes the decision by a single rich-country firm about whether to involve her low-wage country partner in incremental innovation. Sections 5 and 6 embed this choice in a general equilibrium international trade framework with multiple rich-country firms and alternative low-wage country locations. This yields predictions about the extent to which different low wage countries will engage in incremental innovation and attract rich-country firms, as well as predictions about international trade in goods, the capital account, earnings and living standards. Section 7 concludes.

2. Set-up

We have in mind a situation in which a US firm with an existing product has decided to produce it in China. Changes in market conditions and available technologies force the US firm to incrementally improve the product or its production process so that it remains profitable. We refer to the American firm as the principal, denoted by 'she' or a subscript p . The firm has a Chinese partner, who we refer to as the agent, denoted by 'he' or a subscript a . We are not concerned here with how the US principal came to have the asset that allows her to produce (although we will discuss free entry in section 6) or why she has decided to produce abroad. Nor are we concerned here with the mode of entry into China. (As discussed above, it is easiest to think of the Chinese partner as the senior manager or engineer of a US-owned subsidiary.) Our analysis focuses on the extent to which American firms involve the managers and engineers of Chinese plants in the innovation process. This will provide the building blocks for our general equilibrium analysis of incremental innovation and international trade.

Incremental innovation

The American principal can develop the incremental innovation in the United States and ask the Chinese agent merely to implement it. We call this 'Principal Innovation' and denote it by superscript PI . Alternatively, the American principal can assign the role of developing the incremental

innovation to her agent in China. We call this ‘Agent Innovation’ and denote it by superscript AI . Developing the incremental innovation requires ‘creative’ effort either from the principal under principal innovation or from the agent under agent innovation. Let e_i be the innovation effort of i ($i = p, a$) where $e_i \in [0, 1]$. Innovation effort level e_i leads to a successful incremental innovation with probability e_i and to no innovation with probability $1 - e_i$.

Incremental innovations designed by the Chinese agent are different from those designed by the American principal. In particular, interactions between interdependent components of a complex system imply that an incremental innovation designed by the Chinese agent creates some residual incompatibilities for the American principal. The principal must then exert some ‘debugging’ effort to sort out these residual incompatibilities. The extent of residual incompatibilities is parameterized by $r \in [0, 1]$. We assume that debugging effort by the principal is the only characteristic distinguishing one incremental innovation from another. Every successfully developed and debugged incremental innovation allows production of one unit of output. Each unit of output yields the same total profit π (where π is endogenously determined in general equilibrium). For simplicity, no additional inputs are required other than those already provided by the principal and agent.

Both the creative effort e_i required to develop an incremental innovation and the debugging effort required to sort out residual incompatibilities r are costly because they eat into an individual’s unit endowment of leisure. Leisure for the principal and the agent are:

$$l_p = \begin{cases} 1 - e_p & \text{under principal innovation,} \\ 1 - r & \text{under agent innovation,} \end{cases} \quad (1)$$

$$l_a = \begin{cases} 1 & \text{under principal innovation,} \\ 1 - e_a & \text{under agent innovation.} \end{cases}$$

Preferences are Cobb-Douglas over consumption and leisure and homothetic over all goods. Thus, indirect utility for economic actor i ($i = p, a$) is

$$U_i = \left(\frac{y_i}{P}\right)^\alpha l_i^\beta \quad (2)$$

where y_i is i ’s income l_i is i ’s leisure and P is the relevant price index. In the appendix we show that none of our propositions depend on the values of α and β , so in the main text we set $\alpha = \beta = 1$, which yields particularly simple expressions.¹⁰

Residual incompatibilities in a developing-country context

When a firm looks for a supplier in a low-wage country, a notable feature is the scarcity of information about whether the supplier can meet the expectations of the firm. Thus, the firm and the supplier do not know r until after they have been matched for some time and committed resources

¹⁰Setting $\alpha = \beta = 1$ has the advantage of making our core expressions linear in r . Otherwise they would be linear in an increasing transformation of r , $1 - (1 - r)^\beta$. Thus, all of our propositions also hold with $\beta \neq 1$. As will become obvious below, the only modification needed when $\beta \neq 1$ is that $\bar{r} \equiv 1 - \frac{1}{2} \frac{1-w/\pi}{1-\lambda}$ of proposition 1 must be redefined as $\bar{r} \equiv 1 - \frac{\beta}{1+\beta} \left(\frac{1-w/\pi}{1-\lambda}\right)^{\alpha/\beta}$. See the appendix for details.

to the relationship. Evidence on this abounds. At the most general level, Rajan and Zingales (2003) argue that the inaccessibility of information about enterprises in low-wage countries is the primary reason for underdevelopment — it makes investment in a relationship with developing-country enterprises too risky. At a more micro level, Sutton (2004) studies Chinese and Indian first-tier suppliers to car makers such as GM. In an industry where the main purpose of incremental innovation is to reduce product defect rates (as in our Delphi example), 45% of Chinese suppliers and 57% of Indian suppliers fail to deliver the defect rate improvements needed to achieve international best-practice targets. It is common for a firm to take 3 years to figure out whether its supplier can achieve these targets i.e., three years before r is known with certainty. Indeed, Sutton (2004) describes an interview with a firm that took 5 years to arrive at this conclusion. At the end of this process of discovery the firm must make a decision about whether to leave incremental innovation in the hands of the supplier (agent innovation in our model) or whether to take control of incremental innovation (principal innovation in our model).

This discussion suggests three important features that we must model. (1) There is wide heterogeneity in the residual incompatibilities r created by different Chinese suppliers. (2) In any match there is uncertainty about the extent of the residual incompatibilities, uncertainty which can take years to resolve. Restated, r may be unknown for a long time. (3) As a consequence of this delay, it is often better for a US firm in a high- r match to take control of incremental innovation rather than to search for and switch to another supplier.

We formalize these three features in a stylized way. Each supplier is characterized by a residual incompatibility r . There is supplier heterogeneity that is captured by a cumulative distribution function $F(r)$ over residual incompatibilities r . The firm is matched with a supplier that is randomly drawn according to F . We completely telescope the process of discovering r : r is learned immediately after matching. What is important for this paper is that if a firm finds itself in a bad match (i.e., a high- r match) it is more profitable to take control of innovation than to incur the costs of searching for a new supplier and bringing him up to speed on the project. We make this extreme by assuming the former always happens. Thus, after r is learned the firm must stay with the supplier and choose one of two organizational forms (agent innovation or principal innovation).¹¹

We now turn to aspects of contractual incompleteness related to residual incompatibilities. Contracts are incomplete in the sense that they cannot be contingent on innovation effort. The agent (Chinese supplier) has no liquid assets so the principal (US firm) cannot ask for lump-sum transfers from the agent. The contract only specifies what the payment to the agent will be conditional on whether or not there is production and whether or not the agent will be involved in incremental innovation. After the contract is signed, r is revealed and the principal decides whether or not to offer the agent the choice of getting involved in innovation.

One can imagine that if the agent were involved in innovation his payment would be conditional on residual incompatibilities r . However, there are two features of incremental innovation in developing countries that lead us to rule out r -contingent contracts. First, residual incompat-

¹¹In contrast, when there is a bad match in the models of Casella and Rauch (2002) and Rauch and Trindade (2003) then the principal abandons the match and returns to the domestic market. This leads to important insights about how the level of international trade is affected by business networks and information flows. It also leads to novel insights about trade and inequality.

ibilities will be difficult to verify by an outside party. Second, to be able to contribute to the incremental innovation, the agent needs confidential technical and/or marketing specifications from the American firm. Much of this information will typically be non-codifiable information passed on by us managers and engineers to their Chinese counterparts. This information will almost certainly be valuable outside the relationship. Thus, once the Chinese agent has this information, the American principal will have to pay an earnings premium in order to retain the Chinese agent. This premium will not necessarily reflect the size of the residual incompatibilities created by the agent's innovation, but rather the value of the information that the agent has acquired by being involved in innovation.

We capture this by making the payment to an agent who is involved in innovation independent of r . We express this payment as a fraction λ of profits π . Since, conditional on production, π is exogenous to the relationship, the issue is what determines λ . We think about this as follows. Once the agent has come up with an incremental innovation the agent could walk away from the relationship and produce on his own. However, there is a surplus if the principal and agent maintain their relationship. The outcome of bargaining over this surplus can be described as an ex-post sharing rule whereby the agent receives a share λ of profits and the principal a share $1 - \lambda$. For instance, suppose the principal and agent successfully develop an incremental innovation. If they produce together, joint profits are π . If they go their separate ways, the principal obtains a fraction γ_p of the total profits of joint production, net of the additional costs she must incur to be able to produce without the agent. Likewise, the agent obtains a fraction γ_a of the total profits of joint production, net of the additional costs he must incur to be able to produce without the principal. The surplus in the relationship is $\pi(1 - \gamma_p - \gamma_a)$, which we assume is positive.¹² If, for example, we use the Nash bargaining solution with equal weights to allocate the surplus, when the agent is involved in incremental innovation he gets $\frac{1}{2}(1 + \gamma_a - \gamma_p)\pi$ while the principal keeps $\frac{1}{2}(1 - \gamma_a + \gamma_p)\pi$. In this case, under agent innovation ex-post bargaining leaves the agent with a share of profits $\lambda = \frac{1}{2}(1 + \gamma_a - \gamma_p)$. This determines λ . The key point is that λ is independent of r .¹³

If the agent is not involved in innovation he does not create any residual incompatibilities so that his payment is also independent of r . We express this payment as a wage w .

It follows that income for the agent and principal, conditional on production, is

$$y_a = \begin{cases} w & \text{under principal innovation,} \\ \lambda\pi & \text{under agent innovation,} \end{cases} \quad (3)$$

$$y_p = \pi - y_a .$$

3. The choice over the agent's involvement in innovation

The level of agent involvement in innovation will depend on the earnings premium $\lambda\pi/w$ and on the importance of residual incompatibilities r . This section describes exactly how.

¹²Note that at this point the effort levels affecting leisure have already been incurred and therefore the surplus depends only on profits within and outside the relationship.

¹³We actually do not even need that λ is independent of r . See footnote 15 below.

Principal innovation

Under principal innovation, with probability e_p an incremental innovation is developed by the principal which yields incomes $y_a = w$ and $y_p = \pi - w$. With probability $(1 - e_p)$ no incremental innovation is developed and $y_a = y_p = 0$.¹⁴ Substituting this and (1) into (2), yields the following expected utility levels for the principal and the agent:

$$\mathbf{EU}_p^{PI} = e_p \frac{\pi - w}{P} (1 - e_p), \quad (4)$$

$$\mathbf{EU}_a^{PI} = e_p \frac{w}{P}. \quad (5)$$

The problem for the principal is to choose e_p in order to maximize \mathbf{EU}_p^{PI} , which yields $e_p = 1/2$. Substituting this equilibrium effort level into (4) and (5) yields equilibrium expected utility levels:

$$\mathbf{EU}_p^{PI} = \frac{1}{4} \frac{\pi - w}{P}, \quad (6)$$

$$\mathbf{EU}_a^{PI} = \frac{1}{2} \frac{w}{P}. \quad (7)$$

Agent innovation

Under agent innovation, substituting (1) and (3) into (2), yields the following expected utility levels for the principal and the agent:

$$\mathbf{EU}_p^{AI} = e_a \frac{(1 - \lambda)\pi}{P} (1 - r), \quad (8)$$

$$\mathbf{EU}_a^{AI} = e_a \frac{\lambda\pi}{P} (1 - e_a). \quad (9)$$

The problem for the agent is to choose e_a in order to maximize \mathbf{EU}_a^{AI} , which yields $e_a = 1/2$. Substituting this equilibrium effort level into (8) and (9) yields equilibrium expected utility levels:

$$\mathbf{EU}_p^{AI} = \frac{1}{2} \frac{(1 - \lambda)\pi}{P} (1 - r), \quad (10)$$

$$\mathbf{EU}_a^{AI} = \frac{1}{4} \frac{\lambda\pi}{P}. \quad (11)$$

The principal's choice

The principal chooses her preferred organizational form — principal innovation or agent innovation.

Proposition 1 (Extent of the agent's involvement in innovation) *The principal prefers agent innovation to principal innovation if and only if*

$$r < \bar{r} \equiv 1 - \frac{1}{2} \frac{1 - w/\pi}{1 - \lambda}. \quad (12)$$

¹⁴The principal is protected by limited liability, so if there is no incremental innovation and thus no production the principal does not need to pay the agent's wage.

Proof The principal prefers agent innovation to principal innovation if and only if $EU_p^{AI} > EU_p^{PI}$, where EU_p^{AI} is given by equation (10) and EU_p^{PI} by equation (6). \square

Recall that agent innovation requires costly effort on the part of the agent. The agent is willing to incur this effort because he receives an earnings premium or ‘bonus’ that makes it worthwhile. That is, the agent’s earnings with innovation ($\lambda\pi$) are sufficiently higher than earnings without innovation (w). When we introduce the full general equilibrium model, we establish conditions on primitives which ensure that

$$\lambda\pi \geq 2w . \quad (13)$$

From equations (7) and (11) this ensures that the agent prefers agent innovation over principal innovation when this is offered.¹⁵

4. General equilibrium

We now embed our previous discussion into a general equilibrium model. For concreteness, let each principal be a US auto designer or manufacturer who has decided to produce auto parts (or autos for short) in a low-wage country. Since we are not interested in the choice between producing in high-wage versus low-wage countries we assume that all auto production takes place in low-wage countries. For clarity we consider only two low-wage countries.¹⁶ Recall from table 1 and figure 1 that while China is increasingly involved in incremental innovation, Indonesia is not. With this in mind we call the two countries China and Indonesia.

There are two additional sectors: apparel and fine-chemical drug manufacture. Apparel is only produced in China and Indonesia and serves as an alternative occupation for agents. This will be needed to endogenize the number of agents in each country. Drugs are only produced in the United States and serve as an alternative occupation for principals. This is needed to endogenize the total number of principals.

Apparel is a homogenous good produced under perfect competition with raw labour i.e., without innovation effort e_i . Apparel production is subject to diminishing returns to labour — think of this as capturing a fixed factor such as land. To avoid scale effects associated with the

¹⁵Is there any additional scope for renegotiation between the principal and the agent? Consider first the case where $r > \bar{r}$. The agent prefers to be involved in innovation (because this will allow him to steal the critical technology) and so might be willing to make a lump-sum payment to the principal. We have ruled this out by assuming that the agent is liquidity constrained. (Even if he had the cash, it would be an odd model if the agent were allowed to pay the principal for the opportunity to steal the principal’s technology.) Now consider the case where $r < \bar{r}$. Both parties strictly prefer agent innovation to principal innovation. We have already assumed that the surplus generated by agent innovation is split according to a Nash bargain with equal weights, which leaves the agent with a share $\lambda = (1 + \gamma_a - \gamma_p)/2$ of profits. This assumes that what the agent and principal get in the event of bargaining breakdown ($\gamma_a\pi$ and $\gamma_p\pi$, respectively) is independent of r . This independence is not necessary for our results. In particular, if one thinks that agents in low- r relationships also tend to have better outside options (higher γ_a) then one could allow γ_a to be decreasing in r . None of our propositions would be altered by this provided that γ_a does not decrease too rapidly in r . (If it decreases too rapidly then the principal perversely likes high- r agents — their bargaining power is so weak that it dominates any innovation-related issues.) To see this more formally, consider figure 2. The key restriction we require on how γ_a varies with r is that neither EU_a^{AI} nor EU_p^{AI} are increasing in r . To recap, one could extend the model to allow for more sophisticated renegotiation while still retaining all of our results.

¹⁶As will become clear, it requires virtually no change in our analysis to allow for any finite number of low-wage agent countries and high-wage principal countries.

fixed factor we assume that China and Indonesia have the same technology and are the same size, each having a workforce of size L . We denote the (endogenous) number Chinese nationals who choose to become agents in the auto sector by m , so that $L - m$ is employment in the Chinese apparel sector. We choose apparel as the numéraire so that its price is unity. Let $w_A(L - m)$ be a function denoting the marginal product and wage of labour in the apparel sector when $L - m$ workers are employed in the apparel sector. By diminishing returns to labour $w'_A < 0$. Asterisks will denote Indonesian variables. In particular m^* is the endogenous number of Indonesian agents and $w_A(L - m^*)$ is the wage in the apparel sector. All our results refer to diversified equilibria, that is, equilibria in which both China and Indonesia produce both autos and apparel. (See the appendix for a discussion of what is required for this to be the case.)

The endogenous number of agents $m + m^*$ must equal the endogenous number of us principals M . The principals' alternative occupation is drug manufacture which is set up analogously to apparel. Since the size of the United States plays absolutely no interesting role we also set the us population to L . Let $w_D(L - M)$ be the marginal product of labour in the drug sector when $L - M$ workers are employed in it. By diminishing returns to labour $w'_D < 0$.

We next turn to returns in the auto sector. Recall that in the partial equilibrium model there is a cut-off \bar{r} such that when r is below \bar{r} the agent is engaged in incremental innovation and when r is above \bar{r} the principal takes charge of incremental innovation.¹⁷ We now assume that there is heterogeneity in the value of residual incompatibilities r across principal-agent matches. This generates heterogeneity in organizational forms and agents' involvement in innovation.¹⁸ To model heterogeneity we assume that r is randomly drawn from a distribution with support $[0,1]$. Let F and F^* be the cumulative distribution functions of r in China and Indonesia, respectively. Both F and F^* are assumed differentiable. We focus our analysis on the case in which agent innovation and principal innovation co-exist in one and possibly both low-wage countries. This requires that the cut-off \bar{r} defined by equation (12) lies strictly between 0 and 1.¹⁹ Principals and agents must decide whether to enter the auto sector and, in the case of principals, whether to enter China or Indonesia. Only after this decision is made are residual incompatibilities within matches discovered. Taking expectations over r , expected utility from entering the Chinese auto sector is given by

$$EU_i = \underbrace{\int_0^{\bar{r}} EU_i^{AI} dF(r)}_{\text{Agent Innovation}} + \underbrace{\int_{\bar{r}}^1 EU_i^{PI} dF(r)}_{\text{Principal Innovation}}, \quad i = p, a. \quad (14)$$

The Indonesian equivalent of equation (14) replaces F with F^* .

We can now state the conditions for factor-market clearing. Consider equation (14) from the perspective of a Chinese agent. Substituting in the values of EU_a^{AI} and EU_a^{PI} given respectively by

¹⁷ \bar{r} depends on the earnings of agents under different organizational forms, something which we now endogenize.

¹⁸This heterogeneity has the flavour of Melitz (2003) where exporters and non-exporters coexist because of cross-plant differences in productivity. Likewise, Antràs and Helpman (2004) use Melitz's productivity heterogeneity to generate the coexistence of integrated home production, domestic outsourcing, FDI and offshore outsourcing.

¹⁹The appendix discusses parameter restrictions required to simultaneously ensure both this and the agent participation constraint in equation (13).

(11) and (7) yields expected *ex ante* utility for a Chinese national from being in the auto sector:

$$\mathbf{EU}_a(w, F) = \underbrace{\frac{1}{4} \frac{\lambda \pi}{P} F(\bar{r}(w))}_{\text{Agent Innovation}} + \underbrace{\frac{1}{2} \frac{w}{P} [1 - F(\bar{r}(w))]}_{\text{Principal Innovation}} \quad (15)$$

where, from proposition 1, $\bar{r}(w) = 1 - \frac{1}{2} \frac{1-w/\pi}{1-\lambda}$. The Indonesian equivalent of equation (15) replaces F with F^* and w with w^* to yield $\mathbf{EU}_a(w^*, F^*)$. If China produces both apparel and autos then its nationals must obtain the same *ex-ante* expected utility whether they work in the apparel sector or are agents in the auto sector. Utility from working in the apparel sector is w_A/P . A Chinese agent is *ex-ante* indifferent between working in the apparel and auto sectors when

$$\mathbf{EU}_a(w, F) = \frac{w_A(L - m)}{P}. \quad (16)$$

The corresponding indifference condition for Indonesian nationals is

$$\mathbf{EU}_a(w^*, F^*) = \frac{w_A(L - m^*)}{P}. \quad (17)$$

These two equations are central to our analysis.²⁰

Turning to the principals' occupation and location decisions, each American principal must employ either a Chinese or Indonesian agent to produce one unit of autos. In choosing between locating in China and Indonesia, each principal compares *ex ante* expected returns of entering each country. Given that all successfully developed and debugged innovations yield the same total profit, from the point of view of a principal there are only two differences between China and Indonesia: (a) the distributions F and F^* and (b) auto sector wages, w and w^* . Substituting the values of \mathbf{EU}_p^{AI} and \mathbf{EU}_p^{PI} in equations (10) and (6), respectively, into (14) yields expected *ex ante* utility for an American principal from entering the Chinese auto sector:

$$\mathbf{EU}_p(w, F) = \underbrace{\int_0^{\bar{r}(w)} \frac{1}{2} \frac{(1-\lambda)\pi}{P} (1-r) dF(r)}_{\text{Agent Innovation}} + \underbrace{\frac{1}{4} \frac{\pi - w}{P} [1 - F(\bar{r}(w))]}_{\text{Principal Innovation}}. \quad (18)$$

Note that r enters equation (18) in two ways. From proposition 1, r determines whether each agent is involved in innovation or not. Second, under agent innovation \mathbf{EU}_p^{AI} also depends on the costs of debugging residual incompatibilities r . The principal's corresponding return from entering Indonesia is $\mathbf{EU}_p(w^*, F^*)$. If principals operate in both China and Indonesia, expected *ex-ante* returns must be equalized across the two countries:

$$\mathbf{EU}_p(w, F) = \mathbf{EU}_p(w^*, F^*). \quad (19)$$

This equation is also central to our analysis.

²⁰Agents who choose the auto sector and get a low draw of r are *ex-post* better off than workers in the apparel sector, whereas agents who get a high draw of r are *ex-post* worse off than workers in the apparel sector. As in models à la Harris and Todaro (1970), we assume that after r is realized it is not possible for workers with a bad draw to relocate to the apparel sector. Alternatively, we could introduce the need to obtain a degree prior to working in the auto sector and finding out the realization of r . The earnings of agents would then have to be higher in order to compensate for the cost of an education. For a high enough cost of education, even an agent in a high- r match will be *ex post* better off in autos than in apparel.

In addition to becoming a principal in the auto sector, us nationals also have the choice of working in the drug sector. A us national is *ex-ante* indifferent between working in the drug and auto sectors when

$$\mathbf{E}U_p(w, F) = P_D \frac{w_D(L - M)}{P}, \quad (20)$$

where P_D is the price of drugs.

This completes the discussion of factor markets. Agents are indifferent between autos and apparel. Principals are indifferent between autos in China, autos in Indonesia, and drugs in the United States.

Finally, we need conditions for product market clearing. We make four additional assumptions. (a) Demand for apparel and drugs slopes downwards. (b) Autos are differentiated goods with a symmetric imperfectly competitive market structure. (c) Each auto firm that ends up producing earns common profits π that are decreasing in the number of producers. (d) Preferences are internationally identical and homothetic. Assumption (a) together with diminishing returns to labour in apparel and drugs ensure that entry into a low-wage country's auto sector increases returns in the local apparel sector and in the us drug sector, thus increasing the attractiveness of the alternative occupations of both local agents and us principals. Assumptions (b) and (c) ensure that entry into auto production does not make further entry even more attractive. Together with (d), assumptions (b) and (c) are required only to determine equilibrium trade flows in corollary 4.2 and some secondary results related to entry in proposition 5.

While this structure may seem complicated, our key results about the location choices of American principals, local involvement in incremental innovation, wages and well-being are derived from only three equations: (16), (17), and (19). This is because all we need to know is how differences between F and F^* lead to differences in wages ($w > w^*$) and differences in the location of American principals ($m > m^*$). All of our core results flow from these two inequalities.

5. International trade and incremental innovation

The difference between China and Indonesia

Our choice of country labels reflects the fact that both China and Indonesia are low-wage recipients of substantial foreign investments. However, unlike Indonesia, China has become a prominent location for first production of new innovative goods (table 1) and for inventors listed in patents owned by us multinationals (figure 1).

To bring the results into stark relief we allow for only one difference between China and Indonesia. When a principal goes to China, she is less likely to face large residual incompatibilities. Mathematically, we assume that F^* first-order stochastically dominates F . This means that F^* is right-shifted relative to F : $F^* \leq F$ for all r . To avoid the uninteresting possibility that $F^* = F$ for all $r < \bar{r}$ we assume that $F^* < F$ near $r = 0$.²¹ As is well known, first-order stochastic dominance implies $\int_0^1 u(r) dF^*(r) \leq \int_0^1 u(r) dF(r)$ for any non-increasing function $u(r)$.

²¹More formally, $dF(r)/dr > dF^*(r)/dr$ when evaluated at $r = 0$.

Although we believe the assumption that F^* first-order stochastic dominates F to be the obvious one, it is worth reviewing the evidence for what the Chinese distribution looks like relative to countries such as Indonesia, Thailand and the Philippines. First, relative to Indonesian engineers, Chinese engineers receive training that allows them to work more effectively with us engineers. Part of this is the high quality of Chinese engineering schools. The *Times Higher Education Supplement* places 2 Chinese engineering schools in the top 15 worldwide and another 6 in the top 100. In contrast, no Indonesian, Thai or Philippine school makes this list.²² Also, there is a large number of Chinese nationals who graduated from us engineering schools and moved back to China. Among foreign-born scientist and engineering students who are enrolled in us schools but have no firm plans to stay in the United States, 25% are from China whereas only 1% are from Indonesia, Thailand and the Philippines combined.²³ The large number of Chinese with us engineering degrees makes it easier to initiate contacts (credentialism) and communicate engineering solutions.

Second, Chinese engineers likely have better specific industrial training than their Southeast Asian counterparts. They have been nurtured by the Chinese diaspora in Hong Kong, Taiwan, Singapore and the United States. This diaspora has invested heavily in bringing Chinese manufacturing plants up to snuff. Kerr (2008) shows that the large ethnic Chinese and ethnic Indian research communities based in the United States have greatly facilitated knowledge diffusion and increased output in innovative sectors back in China and India. (Kerr uses data on ethnic inventor names appearing in us patents, data on foreign citations to these patents, as well as data on migration and production patterns.) There is also significant evidence that Chinese plants are adopting Western management techniques which emphasize quality control and information flow. This can be seen, for example, in the prevalence of ISO 9001 certificates, a standard reference for quality management practices in business-to-business dealings. As of December 2003, China had a stock of almost 100,000 ISO 9001 certificates compared to only 3,449 for Indonesia, Thailand and the Philippines combined. Even controlling for differences in country size, this is a huge difference in certifications.²⁴

Third, China is a major FDI destination for us firms not just because of its low-wages, but also because of the size of its internal market. This means that firms in China are producing for domestic consumption, a fact that puts Chinese engineers in closer proximity to customers. For instance, Nokia designed its 6108 handset in Beijing to optimize its Chinese text messaging capabilities. It has since had several other handsets designed in Beijing for different markets. The size of China's market and the nascent sophistication of its consumers flips Vernon's (1966) argument on its head. Proximity to discerning consumers is a key driver of Vernon's argument for why innovation occurs in rich countries. Now it is an argument for why incremental innovation occurs in China.

For these three reasons it is appropriate to assume that F^* first order stochastic dominates F . Note that all three reasons apply almost as much to India as they do to China.

²²Rankings are published in *Times Higher Education Supplement*, 5 November 2004, available at <http://www.thes.co.uk/worldrankings/>.

²³Authors' calculations based on data in Johnson (1998) and <http://sestat.nsf.gov>.

²⁴Data are from ISO Central Secretariat (2003).

Wages as an equilibrating mechanism

The wage w paid under principal innovation is the key equilibrating mechanism in our model. Our first result discusses equilibrium wage differences in the auto sector between China and Indonesia.

Proposition 2 (Cross-country wage differences) *Suppose $F^* \succ_{FOSD} F$. Then in any diversified equilibrium, auto wages are higher in China than in Indonesia, $w > w^*$.*

Proof Suppose, contrary to proposition 2, that $w \leq w^*$. Differentiating equation (18) yields $\partial \mathbf{EU}_p(w, F) / \partial w = -[1 - F(\bar{r})] / 4P$.²⁵ By assumption, F places positive mass in the neighbourhood of $r = 1$ so that $F(\bar{r}) < 1$ and $\partial \mathbf{EU}_p(w, F) / \partial w < 0$. Hence $w \leq w^*$ implies $\mathbf{EU}_p(w, F) \geq \mathbf{EU}_p(w^*, F)$. Let

$$u_p(r, w) = \begin{cases} \mathbf{EU}_p^{AI}(r) = \frac{1}{2} \frac{(1-\lambda)\pi}{P} (1-r) & \text{for } r < \bar{r}, \\ \mathbf{EU}_p^{PI}(w) = \frac{1}{4} \frac{\pi-w}{P} & \text{for } r \geq \bar{r}. \end{cases} \quad (21)$$

(See equations 8 and 10.) Note that $u_p(r, w)$ is a non-increasing function of r and is strictly decreasing on $[0, \bar{r})$. Since $F^* \succ_{FOSD} F$, the definition of first-order stochastic dominance implies $\mathbf{EU}_p(w^*, F) = \int_0^1 u_p(r, w^*) dF(s) > \int_0^1 u_p(r, w^*) dF^*(s) = \mathbf{EU}_p(w^*, F^*)$. Strict inequality comes from the fact that, by assumption, $dF(0)/dr > dF^*(0)/dr$, which means that there is a subinterval on which both $F^* < F$ and u_p is strictly decreasing. Combining $\mathbf{EU}_p(w, F) \geq \mathbf{EU}_p(w^*, F)$ with $\mathbf{EU}_p(w^*, F) > \mathbf{EU}_p(w^*, F^*)$ implies $\mathbf{EU}_p(w, F) > \mathbf{EU}_p(w^*, F^*)$, a violation of equation (19). Hence $w > w^*$. \square

The basic insight is straightforward. Consider the top panel of figure 2. It plots the expected utility for principals facing different values of residual incompatibilities r in China (solid line) and Indonesia (dashed line). That is, the solid line plots $u_p(r, w)$, which equation (21) defines as $\mathbf{EU}_p^{AI}(r)$ to the left of \bar{r} and $\mathbf{EU}_p^{PI}(w)$ to the right of \bar{r} . The dashed line is the corresponding curve for Indonesia, $u_p(r, w^*)$. Note that $\mathbf{EU}_p(w, F)$ and $\mathbf{EU}_p(w^*, F^*)$ are integrals over r of these curves. A core feature of the principal's problem is that her expected utility is decreasing in r under agent innovation. This reflects the fact that the principal prefers working with an agent whose involvement in innovation creates fewer residual incompatibilities. Since, as illustrated in the bottom panel of figure 2, Indonesia's distribution of residual incompatibilities F^* first order stochastically dominates China's F , if $w = w^*$ then each principal would strictly prefer China over Indonesia. To ensure that each principal is *ex-ante* indifferent between locating in these two countries, lower wages are needed in Indonesia to offset the higher expected residual incompatibilities.

Involvement of Chinese and Indonesian agents in innovation

We saw empirically in figure 1 that Chinese agents are more involved in incremental innovation than their Indonesian counterparts. We now show that the smaller expected residual incompatibilities created by Chinese agents imply that a higher fraction of Chinese agents is involved in

²⁵We are using the fact that $\bar{r}(w)$ is defined to satisfy $\frac{1}{2}(1-\lambda)\pi(1-\bar{r}) = \frac{1}{4}(\pi-w)$. Thus, the derivative of $\mathbf{EU}_p(w, F)$ with respect to \bar{r} is zero.

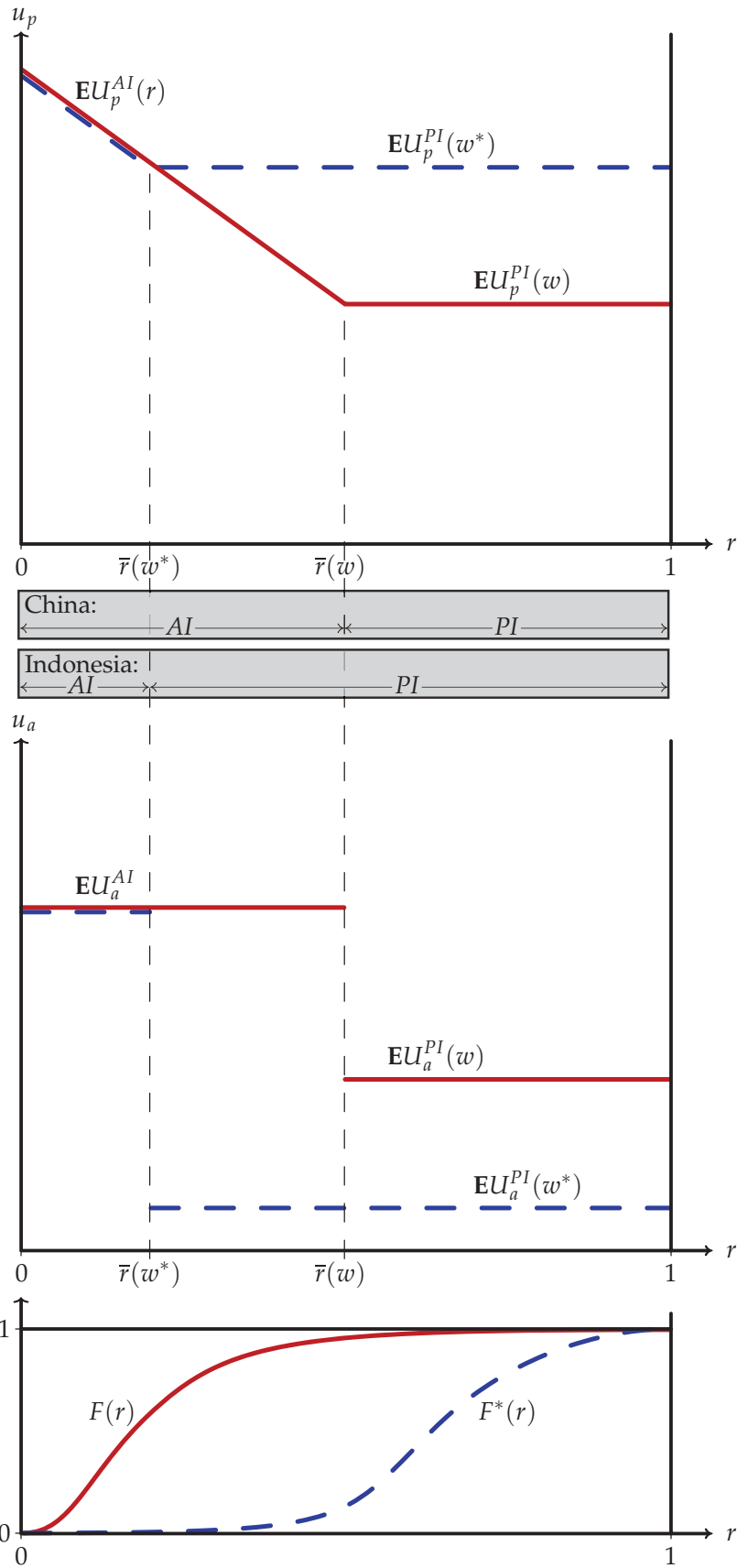


Figure 2: Principals' utility, agents' utility and distribution of residual incompatibilities

innovation. More remarkably, even among agents who create identical residual incompatibilities, Chinese agents are more involved in incremental innovation than their *identical* Indonesian counterparts.

Proposition 3 (General equilibrium involvement in innovation) *Suppose $F^* \succ_{FOSD} F$. Then in any diversified equilibrium $\bar{r}(w) > \bar{r}(w^*)$. This implies the following. (1) Chinese agents have a higher probability of being involved in incremental innovation than Indonesian agents. (2) Consider a Chinese match and an Indonesian match that have identical residual incompatibilities r with $\bar{r}(w^*) < r < \bar{r}(w)$. Then only the Chinese agent will be involved in incremental innovation. The Indonesian agent will not be. Further, the Chinese agent will be paid more and have higher utility than the Indonesian agent.*

Proof By equation (12), $\partial \bar{r}(w)/\partial w > 0$. By proposition 2, $w > w^*$, so that $\bar{r}(w) > \bar{r}(w^*)$. Consider part (1). Since $\bar{r}(w) > \bar{r}(w^*)$, first order stochastic dominance implies that F places more mass on the interval $[0, \bar{r}(w))$ than F^* places on the interval $[0, \bar{r}(w^*))$. That is, Chinese agents have a higher probability of being involved in innovation. Consider part (2). The result that only the Chinese agent will be involved in incremental innovation follows immediately from proposition 1 and $\bar{r}(w^*) < r < \bar{r}(w)$. Regarding utility and income, agents' voluntary participation in innovation implies that $EU_a^{AI}(w) \geq EU_a^{PI}(w)$, i.e., that equation (13) holds. Since $w > w^*$, $EU_a^{PI}(w) > EU_a^{PI}(w^*)$. Hence, $EU_a^{AI}(w) = \frac{1}{4} \frac{\lambda \pi}{P} > \frac{1}{2} \frac{w^*}{P} = EU_a^{PI}(w^*)$, i.e., the Chinese agent has higher utility. Consequently, $\lambda \pi > w^*$, i.e., the Chinese agent is paid more. \square

Part (2) of proposition 3 and to some extent part (1) operate through the endogenous general equilibrium wage differences between China and Indonesia. Agents performing basic auto tasks are paid more in China than in Indonesia so that the additional monetary cost of including workers in incremental innovation is lower in China ($\lambda \pi - w < \lambda \pi - w^*$). Thus, the maximum residual incompatibilities under which agents are involved in incremental innovation is also higher in China: $\bar{r}(w) > \bar{r}(w^*)$. Therefore, even agents who create *identical* residual incompatibilities are more involved in incremental innovation in China than in Indonesia. This is a general equilibrium effect and it highlights that the cost of protecting intellectual property rights depends not only on formal institutions but also on income levels. Here formal institutions are part of what determines λ , the share of profits that agents can capture by threatening to run away with the knowledge acquired while involved in incremental innovation. We have assumed that λ is the same in China and Indonesia. Nevertheless, the lower wage for workers performing basic auto tasks in Indonesia means that Indonesian workers have to be paid a higher earnings premium not to try to appropriate the profits from a successful incremental innovation. As a result, workers with a value of r that would have seen them involved in incremental innovation in China are kept out of the innovation process in Indonesia.

We could easily introduce institutional differences across countries by having λ differ between China and Indonesia. In that case, a higher λ would reduce r (see equation (12)) and, through general equilibrium interactions, also wages (by a similar argument to proposition 2) and the share of auto firms locating in the country (by a similar argument to proposition 4 below). Thus countries with weaker protection of intellectual property (higher λ) are less involved in incremental innovation, have lower income, and attract less foreign direct investment.

The location of American principals

The better distribution of Chinese agents makes China a more attractive location than Indonesia for the same wages. However, wages are lower in Indonesia and this offsets China's advantage from the point of view of principals. We now wish to see what the trade-off between a better distribution of agents in China and lower wages in Indonesia means for the relative number of principals in each country. The next proposition states that China must end up with more than half of all principals.

Proposition 4 (Location of American principals) *Suppose $F^* \succ_{FOSD} F$. Then in any diversified equilibrium, more American principals locate in China than in Indonesia, $m > m^*$.*

Proof We know from proposition 2 that auto wages are higher in China than in Indonesia ($w > w^*$). We first show this implies that auto-sector expected utility (before r is known) is also higher in China. We use this to show that apparel wages are also higher in China: $w_A(L - m) > w_A(L - m^*)$. By diminishing returns this implies $L - m < L - m^*$ or $m > m^*$. Differentiating equation (15) yields

$$\frac{\partial \mathbf{EU}_a(w, F)}{\partial w} = \frac{1}{2P}[1 - F(\bar{r})] + \left[\frac{1}{4} \frac{\lambda \pi}{P} - \frac{1}{2} \frac{w}{P} \right] \frac{dF(\bar{r})}{d\bar{r}} \frac{\partial \bar{r}}{\partial w}. \quad (22)$$

By assumption, F places positive mass in the neighbourhood of $r = 1$ so that $F(\bar{r}) < 1$. The term in brackets, $[\lambda \pi / 4P - w / 2P]$, is $\mathbf{EU}_a^{AI}(w) - \mathbf{EU}_a^{PI}(w)$ which is non-negative (otherwise the agent would turn down the principal's request to assist in innovation). Since F is a cumulative density function, $dF(\bar{r})/d\bar{r} \geq 0$. By equation (12), $\partial \bar{r} / \partial w > 0$. Hence, $\partial \mathbf{EU}_a(w, F) / \partial w > 0$. By proposition 2, $w > w^*$ so that $\mathbf{EU}_a(w, F) > \mathbf{EU}_a(w^*, F)$. Let

$$u_a(r, w) = \begin{cases} \mathbf{EU}_a^{AI} = \frac{1}{4} \frac{\lambda \pi}{P} & \text{for } r < \bar{r}, \\ \mathbf{EU}_a^{PI}(w) = \frac{1}{2} \frac{w}{P} & \text{for } r \geq \bar{r}. \end{cases} \quad (23)$$

Note that, by equation (13), $u_a(r, w)$ is a non-increasing function of r . Since $F^* \succ_{FOSD} F$, the definition of first-order stochastic dominance implies $\mathbf{EU}_a(w^*, F) = \int_0^1 u_a(r, w^*) dF(s) \geq \int_0^1 u_a(r, w^*) dF^*(s) = \mathbf{EU}_a(w^*, F^*)$. Combining this inequality with $\mathbf{EU}_a(w, F) > \mathbf{EU}_a(w^*, F)$ implies $\mathbf{EU}_a(w, F) > \mathbf{EU}_a(w^*, F^*)$. From equations (16) and (17),

$$\frac{w_A(L - m)}{P} = \mathbf{EU}_a(w, F) > \mathbf{EU}_a(w^*, F^*) = \frac{w_A(L - m^*)}{P}. \quad (24)$$

or

$$w_A(L - m) > w_A(L - m^*). \quad (25)$$

By diminishing returns to labour $w'_A < 0$. Hence $m > m^*$. \square

Corollary 4.1 (Apparel wages and utility) *Suppose $F^* \succ_{FOSD} F$. Then in any diversified equilibrium, wages in the apparel sector are higher in China than in Indonesia, $w_A(L - m) > w_A(L - m^*)$, and expected utility in both the apparel and auto sectors is higher in China than in Indonesia.*

The basic insight works off the agent's indifference between the apparel and auto sectors. Consider the middle panel of figure 2, which is the agent's counterpart to the top panel. It plots

the expected utility for agents facing different values of residual incompatibilities r in China (solid line) and Indonesia (dashed line). That is, the solid line plots $u_a(r, w)$, which equation (23) defines as EU_a^{AI} to the left of \bar{r} and $EU_a^{PI}(w)$ to the right of \bar{r} . The dashed line is the corresponding curve for Indonesia, $u_a(r, w^*)$. Note that $EU_a(w, F)$ and $EU_a(w^*, F^*)$ are integrals over r of these curves.

Since Chinese agents are paid more than Indonesian agents under principal innovation ($w > w^*$), the Chinese profile lies above the Indonesian profile for $r > \bar{r}(w)$. Moving left of $r = \bar{r}(w)$ the profile jumps up because of the agent's participation constraint. This raises the Chinese profile even higher above the Indonesian profile in the interval $r \in [\bar{r}(w^*), \bar{r}(w)]$. Thus, Chinese agents are better off than Indonesian agents both because they have a higher profile and because the Chinese distribution puts more weight on the higher outcomes to the left. Since an agent's returns must be equalized across the apparel and auto sectors, Chinese apparel wages $w_A(L - m)$ must also be higher than Indonesian apparel wages $w_A(L - m^*)$. In general equilibrium, for wages to be higher in China, there must be more agents and hence more American principals in China than in Indonesia i.e., $m > m^*$. As a corollary, higher Chinese apparel wages also imply that utility in the apparel sector is higher in China than in Indonesia i.e., $w_A(L - m)/P > w_A(L - m^*)/P$. Finally, since agents are indifferent between sectors, expected utility in the auto sector must also be higher in China than in Indonesia.

Notice that a larger share of principals go to where agents are most expensive. This result never occurs in standard product-cycle models where us principals always locate in the lowest-wage country. In our model us firms take into account wages *and* the ability of local agents to participate in incremental innovation. In general equilibrium, there are sufficient additional principals in China relative to Indonesia to make wages differ just enough to offset China's advantages in incremental innovation. These advantages stem from the lower average residual incompatibilities created by Chinese agents and also from the lower general equilibrium earnings premium Chinese agents must be paid to prevent them from walking away. Higher wages and greater chances of becoming involved in incremental innovation attract more Chinese nationals into the auto sector, where they are better off than their Indonesian counterparts. Note also that the higher income and welfare of Chinese agents are driven by more than just China's lower average residual incompatibilities. Open-economy general equilibrium wage adjustments are central. In fact, to use standard international trade terminology, there is no conditional factor price equalization: agents who create identical residual incompatibilities can end up involved in incremental innovation and earning a premium in China but not in Indonesia.

Production patterns and trade flows

We finally establish the pattern of world production and trade flows. There will be trade in goods as well as a capital account transaction, namely, repatriated profits from us activities in China. Since China has a larger number of American investments, China has more auto production and less apparel production than Indonesia. These production patterns together with identical homothetic preferences and the assumptions we made about market structure imply the following result about equilibrium trade flows.

Corollary 4.2 (Trade flows) *Suppose $F^* \succ_{FOSD} F$. Then in any diversified equilibrium trade flows are characterized as follows:*

- *Autos: The United States imports more autos from China than from Indonesia. China and Indonesia export autos to each other, but China is a net auto exporter to Indonesia.*
- *Apparel: Indonesia exports apparel to the United States. If China also exports apparel to the United States, it exports less than does Indonesia. If China does not export apparel to the United States, it imports apparel from Indonesia.*
- *Drugs: The United States exports drugs to both China and Indonesia.*
- *Capital Account: The United States runs a trade deficit that is financed with a capital account surplus against both China and Indonesia. This capital account surplus comes about because principals repatriate to the us their share of profits from auto sales in China and India.*

6. From the product cycle to the rise of innovation in low-wage countries

In January 1992, China's Deng Xiaoping visited the nascent special economic zone of Shenzhen as part of his now famous *Nanxun* or Southern Tour. His purpose was revolutionary — to praise the efficiency of capitalist firms operating in this and similar zones. He announced the expansion of the export-processing zone program and the liberalization of the foreign investment regime to allow more foreign companies to operate in China. The rest is history. The new investment regime has led to massive entry of foreigner entrepreneurs into Chinese manufacturing. In addition, as shown in table 1 and figure 1, Chinese manufacturing facilities are increasingly moving beyond the production of mature goods and getting more and more involved in incremental innovation. In this section, we examine the consequences of China's opening up to foreign investment.

Consider a world in which initially all American principals locate in Indonesia because they are not allowed into China. Suppose now that Deng Xiaoping makes his Southern Tour and American principals are allowed to enter China. Once China opens up, what is the equilibrium distribution of American principals between China and Indonesia, to what extent are agents in each country involved in innovation, and what are the consequences of this for income levels and trade patterns?

Proposition 5 (The consequences of China's opening up) *Suppose $F^* \succ_{FOSD} F$. Suppose that, starting from an initial situation where all American principals locate in Indonesia because they are not allowed into China, China opens up to American firms. Once American principals are allowed to enter China, in any diversified equilibrium Indonesia ends up with fewer American principals and lower wages in both sectors than before China's opening up to American firms. Relative to Indonesia, China ends up with more principals and higher wages in both sectors. Chinese agents are involved in incremental innovation more often than their Indonesian counterparts. Even agents creating residual incompatibilities $r \in (\bar{r}(w^*), \bar{r}(w))$ are involved in innovation in China but not in Indonesia.*

Proof Initially, Indonesia hosts all principals. After China's opening to American principals, by proposition 4, Indonesia hosts less than one half of all principals. If the total number of principals

has not increased, Indonesia must have fewer principals than before. We next show that the same result holds even if the total number of principals increases following China's opening. Entry of principals increases their earnings in the alternative occupation. To reestablish indifference of us nationals between becoming principals and engaging in an alternative occupation, $EU_p(w^*, F^*)$ must be higher than before China's opening. By assumption, entry of principals lowers auto sector profits π . By equation (18), this tends to reduce $EU_p(w^*, F^*)$. Thus, the only way for $EU_p(w^*, F^*)$ to end up being higher than before China's opening is by having auto wages in Indonesia w^* fall (recall that $EU_p(w^*, F^*)$ is decreasing in w^*). Recall that $EU_a(w^*, F^*)$ is increasing in w^* , so the reduction of auto wages in Indonesia makes Indonesian agents worse off. By equation (15) so does the fall in profits. Thus, to keep Indonesian nationals indifferent between being agents in the auto sector and working in their alternative occupation in the apparel sector, Indonesian apparel wages $w_A(L - m^*)$ must also fall. Since $w'_A < 0$, this implies that the number of principals operating in Indonesia m^* is lower after than before China's opening. The higher Chinese wages and the greater Chinese involvement in innovation follow from propositions 2 and 3. \square

Prior to China's entry, the world looks close to a product cycle model: most innovation is done in the United States and the production of the standardized good is done in Indonesia. After China's opening up to American investments, most American firms locating in Indonesia and some of those locating in China continue this product-cycle pattern of developed-country innovation followed by low-wage standardized production. However, many American firms begin involving their Chinese agents in incremental innovation. These were precisely the patterns implied by our numbers on first locations of innovative new products and on patents in the introduction.

Recall from figure 1 that the rise of Chinese and Indian patenting has coincided with a plateauing of or even a decline in patenting by Indonesia, the Philippines, Thailand and Mexico. Proposition 5 explains this. It states that the appearance of China on the world scene has a negative impact on investment in Indonesia. Notice that the proposition does *not* state that China's entry into world markets reduces Indonesian welfare. We have not tracked any of the traditional gains from trade so it is still possible that Indonesia benefits from China's entry. Our main point is simply that these traditional gains from trade for Indonesia will be offset, at least in part, by the departure of American principals from Indonesia. This problem has been commented on by many in the press who point out that many firms are moving operations to China from other low-wage countries such as Indonesia and Mexico.

7. Conclusions

To our mind, a central feature behind the recent success of China and India in international markets has been the ability of these countries to deliver shop-floor incremental innovation to foreign buyers operating complex supply chains. Firms in rich countries need their suppliers to produce high-quality goods — goods that are reliable, have low failure rates and incorporate the latest demands of an ever-changing marketplace.

In the old product-cycle view, all innovation, including incremental innovation, is done in the North. The Northern-designed factory is shipped to the South without any ensuing technical problems. However, the claim that all innovation is done in the North is no longer tenable. We provided the first systematic evidence on incremental innovation in low-wage countries using data on the location of first production for innovative new goods and data on patents developed by low-wage country inventors for us corporations.

To model this phenomenon we introduced the notion of residual incompatibilities. We showed that if residual incompatibilities were above a threshold \bar{r} then the foreign agent did not engage in incremental innovation. Below the threshold there was incremental innovation. This resulted in a model in which the heterogeneity of residual incompatibilities within matches induced a heterogeneity of incremental innovation across countries and even within countries.

Differences across countries are due to differences in the distribution of residual incompatibilities i.e., due to differences between F and F^* . Improvements in information and communication technologies will allow principals to better communicate with agents and reduce residual incompatibilities everywhere. Thus, we expect to see a trend where product-cycle trade is replaced to some extent by trade in which agents are involved in incremental innovation and new products are first produced in low-wage countries. However, some countries such as China and India have developed world-class engineering schools, send more students for training to the United States, have a more committed diaspora, emphasize standards and quality control, or get a head-start in working with foreign multinationals by developing products catered to a particularly large local market. All of these developments shift the distribution of matches in these countries towards lower values of residual incompatibilities much faster, and serve to simultaneously attract more foreign firms and to increase the proportion of those firms that involve locals in innovation relative to other low-wage countries. As a result, the liberalizing of the international trade regime in China and India has led to a vast inflow of foreign investments into these countries. This has led to the growth of increasingly sophisticated, high-quality Chinese and Indian exports. It has also led to problems for countries such as Indonesia and Mexico that were once the major recipients of Western FDI.

The rise of incremental innovation in some low-wage countries is a fact of rapidly growing importance. This paper is the first to explain its implications for international trade. As we have shown, the implications are significant — it's time to wake up and smell the ginseng.

Appendix

Parameter restrictions

Our analysis has focused on the richest possible case in which (1) agent innovation and principal innovation can both arise for some values of r , (2) agents willingly participate in innovation when asked, and (3) production remains diversified in both China and Indonesia. This appendix discusses what is required to ensure this. By proposition 1, a principal chooses agent innovation over principal innovation if and only if $r < \bar{r}$. Since $r \in [0,1]$, with $\bar{r} < 0$ a principal would prefer principal innovation for all r , while with $\bar{r} > 1$ a principal would prefer agent innovation

for all r . Thus, agent innovation and principal innovation are both possible preferred choices for a principal for some values of r if and only if $0 < \bar{r} < 1$, where \bar{r} is given by proposition 1. This requires $0 < 1 - \frac{1}{2} \frac{1-w/\pi}{1-\lambda} < 1$. Agents always accept this choice if and only equation (13) holds, so that $\lambda\pi \geq 2w$. These two conditions put together reduce to $2\lambda - 1 < \frac{w}{\pi} \leq \frac{\lambda}{2}$. For this to be satisfied, we must have $2\lambda - 1 < \frac{\lambda}{2}$, which simply requires $\lambda < \frac{2}{3}$. In addition, we need $\frac{w}{\pi}$ to fall somewhere in the interval $(2\lambda - 1, \frac{\lambda}{2}]$. From equations (16) and (17), we know that the agent's wage w is directly related to the wage in the alternative apparel sector w_A . Thus, by changing the endowment of (apparel-sector) land we can always shift w_A so that $\frac{w}{\pi}$ lies in the required interval. Finally, we have also focused on situations in which both China and Indonesia keep some production in the auto and the apparel sectors. This is akin to the usual restriction in trade models of being inside the cone of diversification. To ensure apparel production in China and Indonesia, we assume $w_A(0) = \infty$. To have auto production in both countries, we also require differences between F and F^* not to be so extreme that principals would *ex-ante* prefer locating in China even when Indonesian wages are zero.

Generalized utility coefficients

In this appendix we relax the assumption of $\alpha = \beta = 1$ in the utility function of equation (2). Then the principal's expected utility with principal innovation that was described by equation (4) must be changed to

$$\mathbf{EU}_p^{PI} = e_p \left(\frac{\pi - w}{P} \right)^\alpha (1 - e_p)^\beta.$$

Optimizing with respect to effort level e_p yields $e_p = 1/(1 + \beta)$. Plugging this back into the previous equation and simplifying yields

$$\mathbf{EU}_p^{PI} = \frac{\beta^\beta}{(1 + \beta)^{1+\beta}} \left(\frac{\pi - w}{P} \right)^\alpha. \quad (\text{A1})$$

The agent's expected utility with agent innovation that was described by equation (9) must be changed to

$$\mathbf{EU}_a^{AI} = e_a \left(\frac{\lambda\pi}{P} \right)^\alpha (1 - e_a)^\beta.$$

Optimizing with respect to effort level e_a yields $e_a = 1/(1 + \beta)$.

The principal's expected utility with agent innovation that was described by equation (8) must be changed to

$$\begin{aligned} \mathbf{EU}_p^{AI} &= e_a \left(\frac{(1 - \lambda)\pi}{P} \right)^\alpha (1 - r)^\beta, \\ &= \frac{1}{1 + \beta} \left(\frac{(1 - \lambda)\pi}{P} \right)^\alpha (1 - r)^\beta. \end{aligned} \quad (\text{A2})$$

The principal chooses agent innovation whenever $\mathbf{EU}_p^{AI} > \mathbf{EU}_p^{PI}$. From equations (A1)-(A2), this inequality reduces to

$$r < \bar{r} \equiv 1 - \frac{\beta}{1 + \beta} \left(\frac{1 - w/\pi}{1 - \lambda} \right)^{\alpha/\beta}.$$

All the results of the paper go through with \bar{r} of equation (12) redefined in this way.

References

- Aghion, Philippe and Jean Tirole. 1997. Formal and real authority in organizations. *Journal of Political Economy* 105(1):1–29.
- Antràs, Pol. 2003. Firms, contracts, and trade structure. *Quarterly Journal of Economics* 118(4):1374–1418.
- Antràs, Pol. 2005. Incomplete contracts and the product cycle. *American Economic Review* 95(4):1054–1073.
- Antràs, Pol and Elhanan Helpman. 2004. Global sourcing. *Journal of Political Economy* 112(3):552–580.
- Casella, Alessandra and James Rauch. 2002. Anonymous market and group ties in international trade. *Journal of International Economics* 58(1):19–47.
- Ethier, Wilfred J. and James R. Markusen. 1996. Multinational firms, technology diffusion and trade. *Journal of International Economics* 41(1-2):1–28.
- Feenstra, Robert C. and Andrew K. Rose. 2000. Putting things in order: Patterns of trade dynamics and growth. *Review of Economics and Statistics* 82(3):369–82.
- Grossman, Gene and Elhanan Helpman. 1991a. *Innovation and Growth in the World Economy*. Cambridge, MA: MIT Press.
- Grossman, Gene M. and Elhanan Helpman. 1991b. Endogenous product cycles. *Economic Journal* 101(408):1214–1229.
- Grossman, Gene M. and Elhanan Helpman. 1991c. Quality ladders and product cycles. *Quarterly Journal of Economics* 106(2):557–586.
- Grossman, Gene M. and Elhanan Helpman. 2002. Integration versus outsourcing in industry equilibrium. *Quarterly Journal of Economics* 117(1):85–120.
- Grossman, Gene M. and Elhanan Helpman. 2003. Outsourcing versus FDI in industry equilibrium. *Journal of the European Economic Association* 1(2-3):317–327.
- Grossman, Gene M. and Elhanan Helpman. 2004. Managerial incentives and the international organization of production. *Journal of International Economics* 63(2):237–262.
- Grossman, Gene M. and Elhanan Helpman. 2005. Outsourcing in a global economy. *Review of Economic Studies* 72(1):135–159.
- Harris, John R. and Michael P. Todaro. 1970. Migration, unemployment and development — 2-sector analysis. *American Economic Review* 60(1):126–142.
- Helpman, Elhanan. 2006. Trade, FDI, and the organization of firms. *Journal of Economic Literature* 44(3):589–630.
- ISO Central Secretariat. 2003. *The ISO Survey of ISO 9001:2000 and ISO 14001 Certificates – 2003*. Geneva, Switzerland: ISO Central Secretariat.

- Johnson, Jean M. 1998. *Statistical Profiles of Foreign Doctoral Recipients in Science and Engineering: Plans to Stay in the United States*. Arlington, VA: National Science Foundation, Division of Science Resources Studies. NSF 99-304.
- Kerr, William R. 2008. Ethnic scientific communities and international technology diffusion. *Review of Economics and Statistics* 90(3):518–537.
- Krugman, Paul R. 1979. A model of innovation, technology transfer, and the world distribution of income. *Journal of Political Economy* 87(2):253–266.
- Lileeva, Alla and Daniel Trefler. 2007. Improved market access to foreign markets raises plant-level productivity... for some plants. Working Paper 13297, National Bureau of Economic Research.
- Lucas, Robert E., Jr. 1993. Making a miracle. *Econometrica* 61(2):251–272.
- Markusen, James R. 2002. *Multinational Firms and the Theory of International Trade*. Cambridge: MIT Press.
- McLaren, John. 2000. 'Globalization' and vertical structure. *American Economic Review* 90(5):1239–1254.
- Melitz, Marc J. 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71(6):1695–1725.
- Nahmias, Steven. 2005. *Production and Operations Analysis*. Fifth edition. Toronto: McGraw-Hill Irwin.
- Nunn, Nathan. 2007. Relationship-specificity, incomplete contracts, and the pattern of trade. *Quarterly Journal of Economics* 122(2):569–600.
- Nunn, Nathan and Daniel Trefler. 2008. The boundaries of the multinational firm: An empirical analysis. In Elhanan Helpman, Dalia Marin, and Thierry Verdier (eds.) *The Organization of Firms in a Global Economy*. Cambridge, MA: Harvard University Press.
- Organisation for Economic Co-operation and Development. 2006. *Science, Technology and Industry Outlook, 2006*. Paris: Organisation for Economic Co-operation and Development Press.
- Puga, Diego and Daniel Trefler. 2002. Knowledge creation and control in organizations. Working Paper 9121, National Bureau of Economic Research.
- Puga, Diego and Daniel Trefler. 2005. Wake up and smell the ginseng: The rise of incremental innovation in low-wage countries. Working Paper 11571, National Bureau of Economic Research.
- Rajan, Raghuram G. and Luigi Zingales. 2003. *Saving Capitalism from the Capitalists: Unleashing the Power of Financial Markets to Create Wealth and Spread Opportunity*. Princeton: Princeton University Press.
- Rauch, James and Vitor Trindade. 2003. Information, international substitutability, and globalization. *American Economic Review* 93(3):775–791.
- Rosenberg, Nathan. 1982a. Technological interdependence in the American economy. In *Inside the Black Box: Technology and Economics*, chapter 3. Cambridge: Cambridge University Press, 55–80.
- Rosenberg, Nathan. 1982b. Learning by using. In *Inside the Black Box: Technology and Economics*, chapter 6. Cambridge: Cambridge University Press, 120–140.

- Spencer, Barbara J. 2005. International outsourcing and incomplete contracts. *Canadian Journal of Economics* 38(4):1107–1135.
- Sutton, John. 2001. Rich trades, scarce capabilities: Industrial development revisited. *Proceedings of the British Academy* 111:245–273.
- Sutton, John. 2004. The auto-component supply chain in China and India: A benchmarking study. Processed, London School of Economics.
- Trefler, Daniel. 2006. Service offshoring: Threats and opportunities. In Susan M. Collins and Lael Brainard (eds.) *Brookings Trade Forum 2005: Offshoring White-Collar Work*. Washington DC: Brookings Institution Press, 35–60.
- US National Science Board. 2006. *Science and Engineering Indicators 2006*. Arlington, VA: United States National Science Foundation.
- Vernon, Raymond. 1966. International investment and international trade in product cycle. *Quarterly Journal of Economics* 80(2):190–207.
- von Hippel, Eric. 1988. *The Sources of Innovation*. New York: Oxford University Press.