DISCUSSION PAPER SERIES

No. 7523

INTEGRATION OF THE NORTH AMERICAN ECONOMY AND NEW-PARADIGM GLOBALISATION

Richard Baldwin

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP7523.asp**

ISSN 0265-8003

INTEGRATION OF THE NORTH AMERICAN ECONOMY AND NEW-PARADIGM GLOBALISATION

Richard Baldwin, The Graduate Institute, Geneva and CEPR

Discussion Paper No. 7523
October 2009

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 7523

October 2009

ABSTRACT

Integration of the North American economy and new-paradigm globalisation

This paper presents the trade-in-tasks conceptual framework and extends it to consider a setting where offshoring occurs between high wage nations and where agglomeration forces are important. It also considers the policy implications ranging from rules of origin and trade facilitation to external trade policy and R&D subsidies. The focus is on policy initiatives that could support the development of North American production platforms.

JEL Classification: F14, F21 and L5
Keywords: globalisation, industrial policy, supply chains, trade in tasks and unbundling

Richard Baldwin
The Graduate Institute
11a Ave de la Paix
1202 Geneva
SWITZERLAND

Email: Baldwin@graduateinstitute.ch

Submitted 21 October 2009

1 INTRODUCTION

Since the dawn of human civilisation, the cost of moving goods, people and ideas has forced the geographical bundling of economy activity. Before the days of easy shipping, communities were obliged to consume what they could make. The gradual reduction of shipping costs, with acceleration from roughly 1850 onwards, meant that factories did not have to be near consumers, and competitive pressures pushed production towards the most efficient locations. The world's production and consumption of goods was unbundled spatially.

This 'first unbundling' brought about many wonders of the modern world. Nations (and regions within nations) started to specialise in the production of certain goods while simultaneously expanding the range of goods consumed. Large cities arose and the concentration of talent and knowhow fostered further innovation and scale economies; the industrial revolution was born along with the rise of mass intra-national and international trade. Up to the mid-1980s, unbundling operated at the level of factories or even whole industries since it was economical to keep all manufacturing stages in close proximity.

Since about the mid 1980s, rapidly falling communication and coordination costs have fostered a second unbundling – this time of the factories themselves. Cheaper, higher quality and more reliable communications reduced the need to perform most manufacturing stages near each other. As with the first unbundling, changing technology opened the door to spatial separation and competitive pressures pushed industry across the threshold. Even more recently, the second unbundling has spread from factories to offices with the result being the outsourcing and offshoring of service-sector jobs.

It is useful to view the first and second unbundling as being described by two paradigms. The old paradigm – essentially traditional trade theory – was useful for understanding the impact of the first unbundling. Understanding the second unbundling requires a new paradigm – what Gene Grossman and Estaban Rossi-Hansberg called "trade in tasks" in their famous Jackson Hole paper (Grossman and Rossi-Hansberg 2006). Even thought the old and new paradigms happily coexist (factories and consumers continue to be separated even as the factories themselves are unbundled), they have quite different implications for how governments should react to globalisation.

As we shall see, the key difference is the level of analysis. In the old paradigm, greater openness tended to affect sectors as a whole and, importantly, the fortunes of sectors tended to be shared with the productive factors used most intensively in the sectors. The standard level of analysis was thus sectors and labour skill-groups. Globalisation occurs with a much finer resolution in the new paradigm, forcing a rethink of the policy prescriptions flowing from the old paradigm.

This paper examine presents the "trade in tasks" conceptual framework and extends it to allow for factors that are critical to the analysis of the development of North American industry (e.g. recognising the fact that Canada and the US are both high-income nations while Mexico is not). It also considers the policy implications for the Government of Canada, identifying the policy levers and policy initiatives that should be examined to support the development of North American economic platforms.

To accomplish these goals, it is necessary to start with the old paradigm, recasting it in a fashion that facilitates comparison with the new paradigm. This is the job of Section 2. The subsequent two sections respectively introduce the new paradigm (trade in tasks), and then extend it to allow for factors critical to the study of North American integration. The next section, Section 5, discusses the policy implications of the extended trade-in-task framework, including the impact of trade facilitation, labour and industrial policies, tariff policies, rules of origin and product standards. The final section

presents critical remarks on Alan Blinder's grim view of the second unbundling in services and concludes with some suggestions for future research and data-gathering efforts.

2 THE OLD PARADIGM

Traditional thinking about globalisation – namely standard trade theory – is based on a comparison of nations' competitiveness sector by sector. The goal is to work out a nation's comparative advantage. To think about this, it is useful to start with a fairly abstract view of the competitiveness of a nation's various sectors. Figure 1 facilitates the analysis.

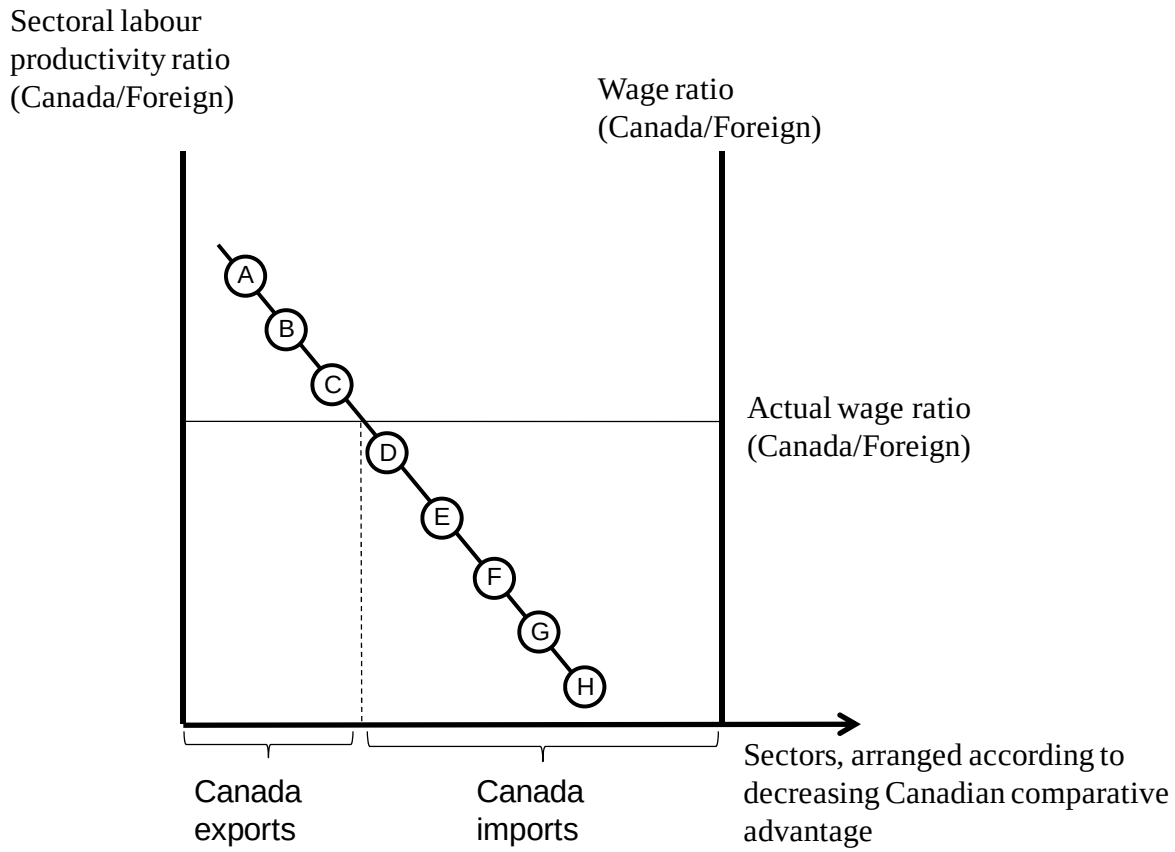


Figure 1: Old paradigm analysis of competitiveness

The diagram lists the sectors along the horizontal axis, ordering them according to their competitiveness. Canada's most competitive sectors are on the left. For instance, the ratio of Canadian to foreign labour productivity is highest for sector A. The least competitive sectors are on the right, e.g. sector H. This measure of competitiveness, however, is incomplete since it does not account for the wage differential. The actual wage gap, i.e. the ratio of Canadian wages to foreign wages is marked with the flat line.

As drawn, Canada's productivity gap more than outweighs the wage gap for sectors A, B and C. That is, given the actual wage ratio (wage gap) and the productivity ratio (productivity gap), Canada can produce sector A, B and C goods more cheaply and thus it exports these good. The other goods are where foreign has a comparative advantage. Canada imports these.

The Figure 1 analysis ignores transportation and other trade costs. Since changes in such cost are a central character in globalisation's drama, we have to modify the diagram to get them into the picture. This is simple, requiring nothing more than the realisation that the competitiveness of a Canadian good

is different in Canadian than it is in the foreign market and vice versa. Specifically, we have to adjust the productivity gap.

The cost of Canadian products inside foreign markets will be higher due to trade costs, so Canada's productivity edge will be dampened by trade costs, and the opposite holds for the competitiveness of foreign products inside Canada. We show this in Figure 2 by having two lines representing the labour productivity ratio: one for the ratio inside Canada (where foreign firms face the disadvantage of having to pay transport costs) and one for the ratio inside the foreign market (where it is the Canadian firms that are disadvantaged by the transport costs).

The implications of this are intuitively obvious – some goods will be made in both nations since local producers are more competitive in both markets given trade costs. In other words, there will be non-traded goods. In the diagram we see that product C is above the wage line for sales inside Canada; as usual, this indicates that Canadian firms will be the low cost producers for the Canadian market. However, product C is below the line in the foreign market, so it foreign firms will be the competitive one in product C in their own market. The same holds for goods D and E, so C, D and E will be non-traded. Using the bundling terminology, transport costs mean that the production and consumption are still bundled nation-by-nation for these sectors; nations consume only what they make.

By contrast, products A and B are above in the foreign market, indicating that Canada would be the low-cost producer, so Canada exports these; F and G are below inside Canada, so these are the sectors where Canada would be the import.

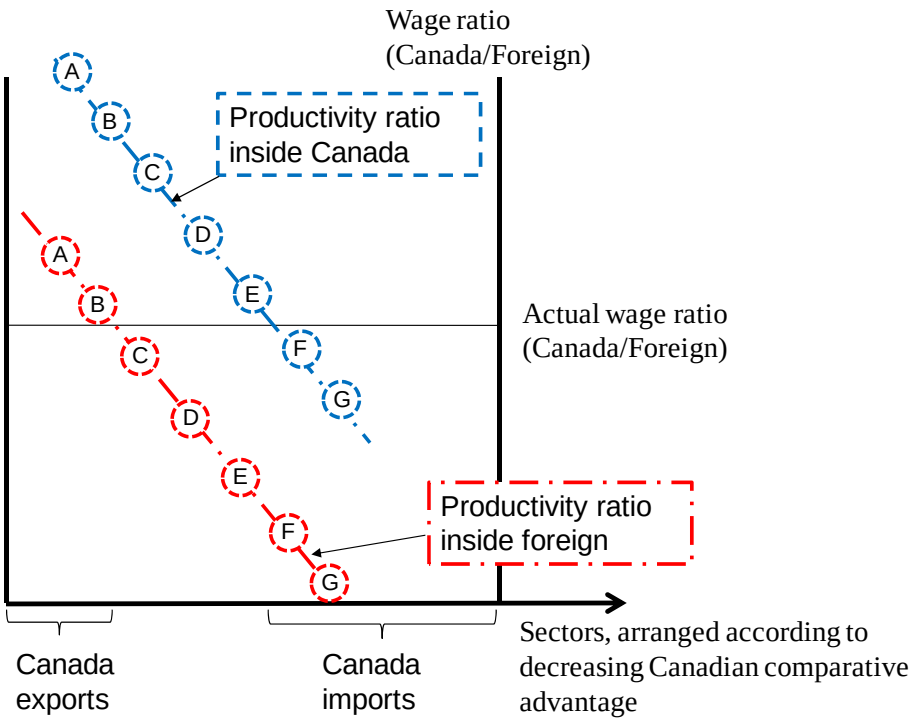


Figure 2: Old paradigm analysis of competitiveness with trade costs

2.1 The impact of falling trade costs: the first unbundling

The last thing to do with this old-paradigm construction is the most crucial. We use this diagram to consider the impact of globalisation, i.e. lowering trade costs. This is done in Figure 3. As trade costs fall, the two lines get closer since the trade cost is less of a factor in determining competitiveness. Naturally the result is an expansion of trade; consider the pattern of this expansion. Canada now becomes competitive in sector C (the trade-cost adjusted productivity ratio in foreign market is now above line for C) and so it starts to export this sector. By the same token, the trade-costs adjusted

productivity ratio is now below the line inside Canada, so the foreigner become competition and Canada starts to import sector D.

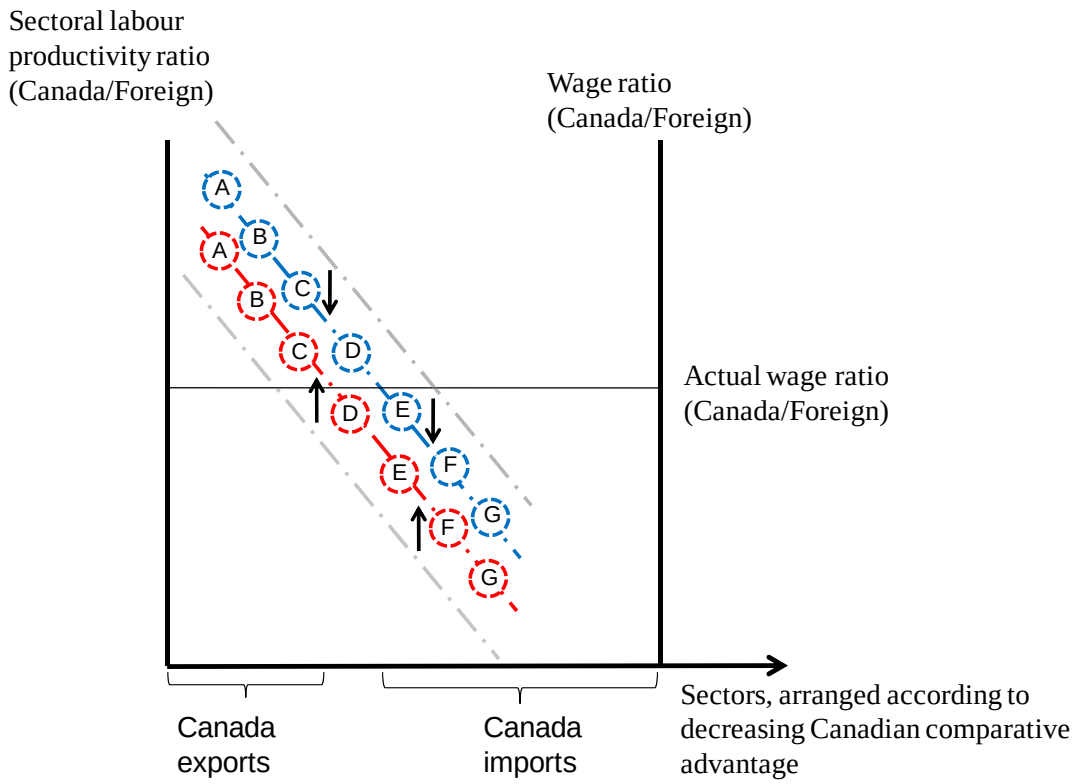


Figure 3: Unbundling in the old paradigm; impact of lower trade costs

2.1.1 Key lessons for old-paradigm policy thinking

While few policy makers would have this diagram in mind, something like it was very evident in shaping their thinking about globalisation, the effects on the economy and what they as policymakers should do about it.

The key point is that globalisation made some of Canada's sectors more competitive and others less. But which one? The winners and losers were not randomly assigned. The new 'winners' from globalisation are sectors that are similar to the ones that were already exported. The 'losers', like sector E, are the sectors that are similar to the sectors where Canada was already uncompetitive. Or to put it differently, the losers will be the Canada's least competitive sectors and the citizens who work in them, while the winners will be the most competitive sectors and those who work in them.

2.2 The appropriate level of analysis: sectors and skill groups

A critical implication of this line of reasoning – a line that most policymakers still work with today – is that globalisation's impact is rather predictable. Policymakers could and did identify 'sunrise' and 'sunset' sectors in advance. They felt they had a rough idea of the identities of the winning and losing sectors. After all, the first unbundling essentially exaggerates the existing pattern of comparative advantage.

For example, as the world opened up, Canadian clothing manufacturers lost out to import competition, and as globalisation proceeded, this trend deepened. The lower trade costs, however, meant the Canadian natural-resource-based and high-tech products gained markets, with the range of such winning sectors expanding as globalised rolled on.

There are a couple of critical assumptions lurking behind this thinking. First, as drawn in the diagram, it assumes that further globalisation lowers trade costs more or less evenly for all sectors. That is, one would not expect a radically different change in the trade costs facing sector D and sector E. Second, the comparative advantage of the sector is roughly related to their factor intensity. For example, it was useful to think of Canada's sunset sectors as marked by unskilled-labour intensity, while the sunrise sectors where marked by skill intensity.

2.3 Policy thinking based on the old paradigm

In the old-paradigm thinking, sectors, or at most firms, are the finest level at which globalisation's impact was felt. More open trade spurred the fortunes of some firms while spiking the fortunes of others but the sector was the finest level of disaggregation worth looking at. Since most firms in a sector stood or fell together, the type of labour used most intensively in the sector typically shared the sector's fortunes. This led governments to organise their globalisation policies around sectors and labour market skill groups.

More specifically, the correlation between current competitiveness and the impact of deeper globalisation demonstrated in Figure 3, led governments to believe they could predict globalisations future impact on the domestic economy. The sectors that 'won' from globalisation were the sectors that were already the most competitive sectors. The 'losing' sectors were the least competitive sectors. Going further, one could roughly associate the most competitive sectors with high-tech, human-capital-intensive sectors, and the least competitive sectors with unskilled-labour-intensive sectors. In turn, one could roughly associate the winners from globalisation as the Canada's high-skilled, high-education workers (and those working in natural-resource based sectors); the losers were, typically, low-skilled, low-education workers.

This had a profound effect on policy thinking in Canada and other industrialised nations around the world. Guided by this old-paradigm worldview, the job of a good policy maker was crystal clear – at least in the abstract. The job is to help the country move resources from the sectors that are likely to lose as the first unbundling continued and shift them into sectors that are likely to win. In the Figure 3 example, the government should be helping to retrain workers who lost their jobs in sector E to become sector C workers. Again roughly speaking, this meant raising skill levels and shifting workers from sunset sectors to sunrise sectors. Skill upgrading, research and development, support for high-tech industries and the information were but some of the natural policy initiatives that flowed from this thinking.

As we shall see below, the new paradigm introduces a line of thinking that should make governments much more cautious about predictions concerning globalisation's winners and losers, thus more cautious about their optimal policy response.

2.3.1 Diagrammatic analysis of winners and losers

The difference between the old and new paradigms can be made clearer by introducing a simple diagram that helps connects the fortunes of sectors and skill groups. Figure 4 is the diagram.

We start with the left panel of the diagram. Here the wage of unskilled workers, w, is on the vertical axis and that of skilled workers, v, is on the horizontal. For simplicity's sake, there are only two sectors, the Y sector whose pricing is especially sensitive to the price of skilled labour (since it is skill-intensive) and the X-sector whose price is especially sensitive to unskilled wages. This sensitive is easy to see. The Y-sector pricing equation shows the combinations of w and v that allow Y-sector firms to match the market price. Plainly any increase in either w or v must be matched by a reduction in the other if price competitiveness is to be maintained. But note that a small increase in the skilled wage, v, requires a larger decrease in w – that's because Y is skill intensive. Similarly, X is unskilled

labour intensive, so a 1% increase in w would require a more than 1% drop in v to allow X-sector firms to remain competitive with foreign producers.

The combination of skilled and unskilled wages where both sectors are competitive is marked by the point E; the equilibrium wages are marked as w^o and v^o.

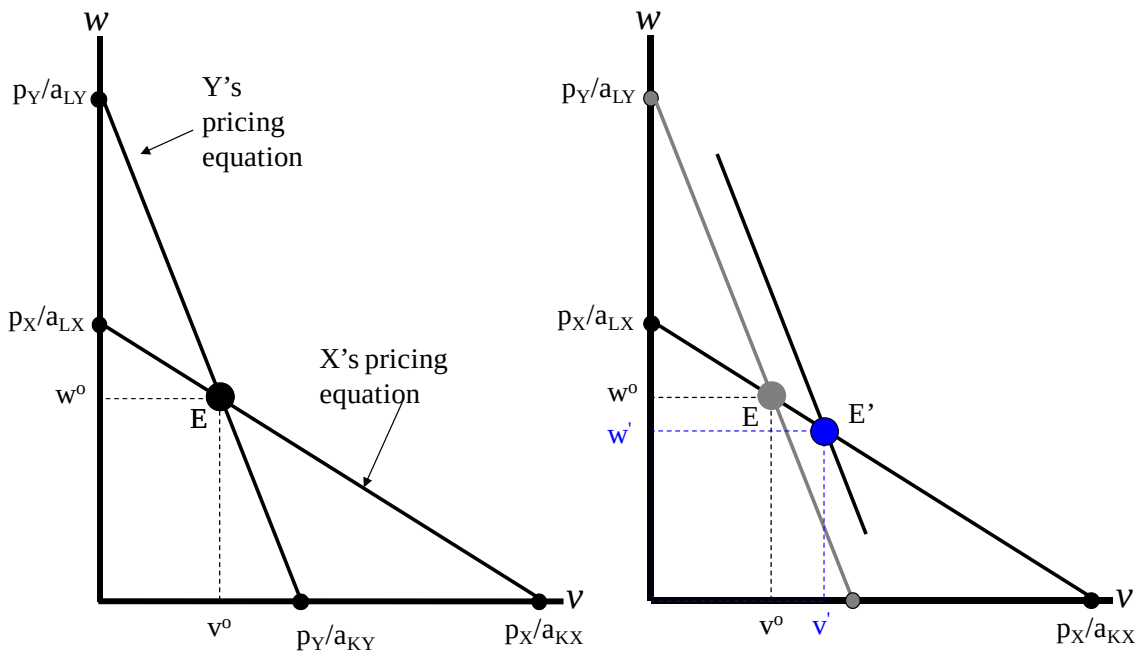


Figure 4: Sectors and the fate of skill groups: first unbundling

The purpose of the diagram is to allow us to connect the fate of skill groups to the sectors in which they are intensively employed. The left panel does this. In this case, we assume that Y is the export sector, so lower trade barriers, natural and manmade, favour Y. Specifically, as Y-sector firms get better access to foreign markets, the sector adjusts along two dimensions, first, the sector produces and sells more, and second, it sees a higher price net of trade costs.

In the diagram, this favourable export-sector development shows up as a shift out in the Y-sector price line. That is, the sector can now maintain competitiveness even after paying some combination of higher v and/or w. The situation in the import competing sector, the numeraire X sector, doesn't change. This tells us that the w and v must move in opposite directions if both sectors are to remain competitive after the further market opening.

The new intersection, point E', shows the new combination of w and v that allows both sectors to be competitive. The result – a result we foretold with verbal reasoning above – is that the factor used intensively in the export sector gains from globalisation while the factor used intensively in the import sectors loses.

This, in diagrams, is the correlation between sectoral fates and skill group fates; a correlation that is at the heart of most nation's thinking on the effects of globalisation.

3 THE NEW PARADIGM: SECOND UNBUNDLING AND TRADE IN TASKS

As manufactures account for 70% of global trade, the nature of trade and the nature of manufacturing are inexorably linked. Both the first and second unbundlings fostered and were fostered by radical changes how things are made.

3.1 Nature of manufacturing, nature of trade, and the first unbundling

Before the industrial revolution, manufactured goods were basically handicrafts. One of the most sophisticated 18th century machines – rifles – were constructed one at a time by highly skilled craftsmen using hand tools. The workshops making them were geographically dispersed across nations, roughly in line with the location of consumers; trade flows were modest.

Whitney, Ford and gigantic factories

In 1801, Eli Whitney came up with the notion of standardising parts to the extent that they were interchangeable. Rifles could be made faster, cheaper and with less skilled workers. The resulting gains in competitiveness gave rise to large manufacturing corporations that put many smaller arms makers out of business. The resulting geographical concentration of rifle making separated factories and consumers, spurring long distance trade (both intra- and inter- national) of the first-unbundling type.

A century later, the Ford Motor Company greatly refined assembly-line mass production. The Ford method was much faster and used less manpower than 19th century manufacturing techniques, but worked best at massive scales of production. This further stimulated first-unbundling trade as the competitiveness of Ford's products forced smaller automotive factories around the world to close – thus increasing the distance between automakers and most auto buyers.

The Ford method faced important organisational challenges. To keep things moving smoothly and reliably, producing a car every 3 minutes, Ford spatially concentrated the production of almost everything. What he couldn't concentrate, he bought so as to better control. He owned rubber plantations, coal mines and forests as well as the ships and railroad cars that transport them to his plant. The famous River Rouge plant employed about 100,000 workers in the early 20th century.

This hyper concentration came at a cost. It meant that almost every stage of producing a model-T had to be done with labour and capital located in Michigan. There would have been a financial gain from unbundling production stages and locating where factor costs were better suited to each stage's demands, but this was impossible. Coordination costs were the bottleneck. Coordinating complex activities over long distances was impossible at the time. Transportation was slower and less reliable; telecommunications were only for emergencies. To ensure that parts and components were ready when needed, North American labour, capital and technology were spatially bundled in one place.

3.2 Unbundling and the coordination revolution

Geographically separating various production stages became more attractive as it became less costly to coordinate complex tasks across distance. Falling trade costs – the combination of lower tariffs and lower freight costs – played some role, but not a dominate one (Hummels 2007). As Figure 5 shows, trade costs (the combination of freight rates and tariffs) did fall in this period, but for most sectors the reduction was less than 5% from 1982 to 1992. Regular surface shipping did not get much cheaper but the growing density of shipping lines made surface shipping easier and more reliable. The price of air cargo fell, but again not spectacularly (WTO 2008).

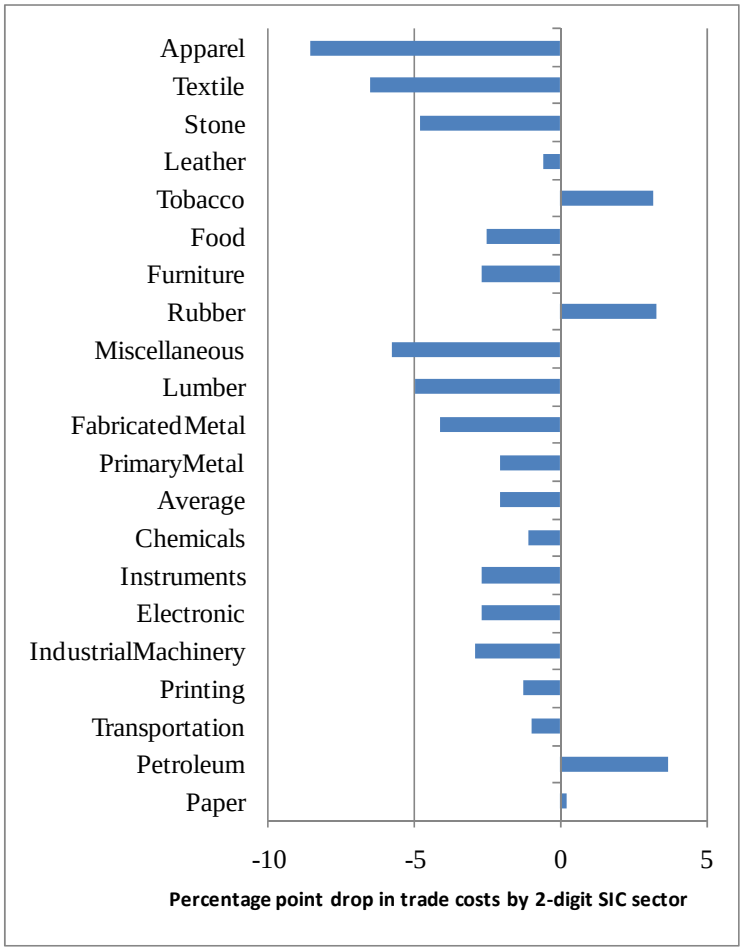


Figure 5: Drop in trade costs 1982 -1992 by SIC sector.

Source: Bernard, Jenson and Schott (2003, Table 1).

More important are advances in information and communications technology (ICT) in explaining the dramatic drop in the cost of organising complex activities over distances. This showed up in many ways. The price of an old-fashioned telephone call plummeted, along with deregulation, falling computing costs, and the swift drop in the cost of fibre optic transmission rates. New forms of communication appeared and rapidly transformed the workplace. Faxes became standard equipment. Cellular phone usage exploded. The telecommunications network also became denser and more reliable as it became cheaper. Above all, the internet – first email and then web-based technology – revolutionised the sharing of information over distance. In 1984, there were 1024 internet hosts in the world; by 1995, the number was 6.6 million rising to 106.8 million in 2000.

Interacting with cheaper communications costs was the spectacular fall in the price of computing power. Things that required a Cray super computer in 1984 could be performed on a high powered PC by the mid 1980s. This encouraged the development and widespread use of information-management software (ranging from excel spreadsheets to sophisticated database programmes). Cheap and reliable telecommunications combined with information manage software and desktop computers to run them completely transformed the difficulty of organising group-work across space. Stages of production that had to be performed in close proximity – within walking distance to facilitate face-to-face coordinate of innumerable small glitches – could now be dispersed without an enormous drop in efficiency or timeliness. Working methods and product designs were also shifted in reaction to the spatial separation, typically in ways that made production more modular.

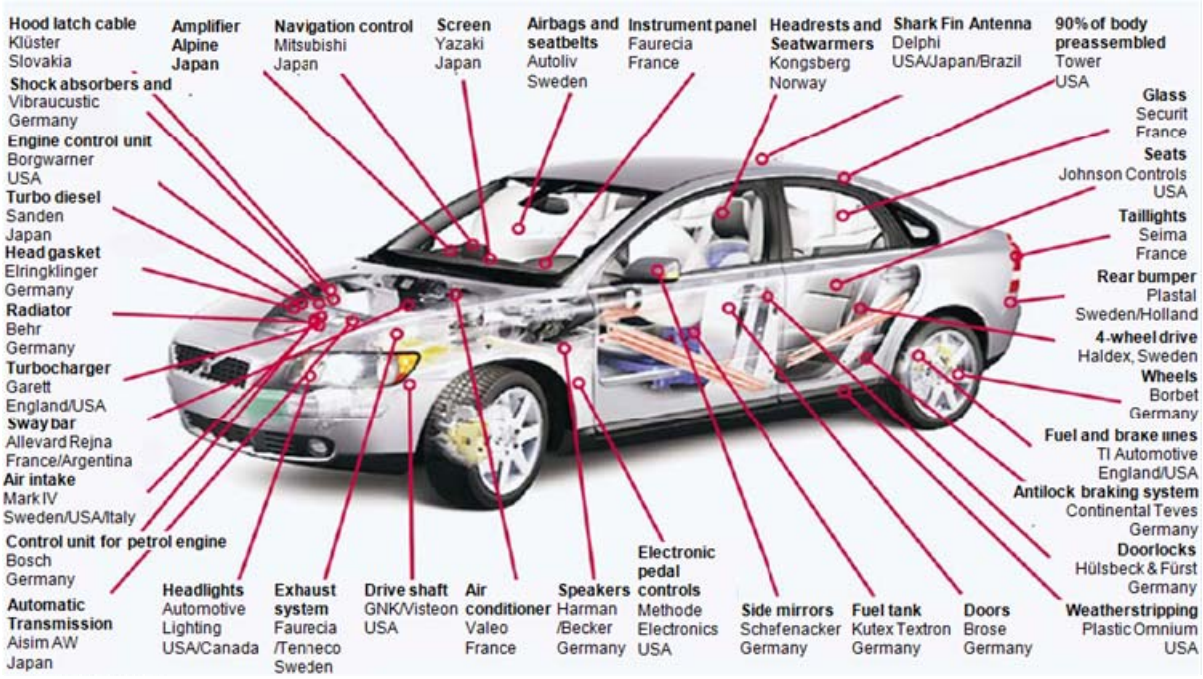


Figure 6: Where are the components of the Volvo S40 made?

Note: thanks to Shon Ferguson for translation from Swedish.

The second unbundling is a result of this lower communication costs. Things that had to be done in various bays in the same factory in order to reduce delays due to miscommunications could now be done in separate factories located far from each other. In essence, the production bays became their own factories and were dispersed to locations that had factor prices and other characteristics better suited to the particular needs of the production stage.

An example of the second unbundling can be seen in Figure 6. This shows where the parts of the "Swedish" Volvo S40, are made. The navigation control and screen is made in Japan, the side mirror and fuel tank in Germany, the air conditioner in France, the headlights in the US and Canada, the fuel and brake lines in England, the hood latch cable in Germany. Some parts are even made in Sweden (airbag and seat beats). These 'parts' are themselves made up of many parts and components, whose production location is likely to be equally dispersed. For example, the air conditioner will have to have a compressor, motor and a control centre, each of which may be made by a different company in a different nation.

The diagram makes clear that Henry Ford's spatial concentration of production is finished. Manufacturing stages that used to be done by the same company in the same factory are now dispersed around the world. Sometimes these are owned or controlled by the original manufacturer, but often they are owned by independent suppliers.

Much is regional not global

It is important to note that many of these international supply chains are regional, not global. The cost and unpredictable delays involved in intercontinental shipping still matters. Moreover, coordination in the same time zone is easier and more reliable. An addition factor that has fostered regionalisation over globalisation is that the fact that the cost of moving key managers and technicians has not fall radically. Even if plane fares have come down, the opportunity cost of the managers' time has actually risen. If a Canadian firm puts a factory in Mexico, the manager may have to spend a whole day to hold a 1 hour face-to-face meeting. If the factory is in China, the time-cost will be more like one whole workweek.

The first large-scale production unbundling started in the mid 1980s and took place over very short distances. The Maquiladora programme created ‘twin plants’, one on the US side of the border and one on the Mexican side. Although the programme existed since 1965, it only boomed in the 1980s with employment growing at 20% annually from 1982-89 (Dallas Fed 2002, Feenstra and Hanson 1996). Another second unbundling started in East Asia at about the same time (and for the same reasons). In this region distances are short compared to the vast wage differences (Tokyo and Beijing are about 90 minutes apart by plane, yet in the 1980s the average Japanese income was 40 times the Chinese average). In Europe, the second unbundling was stimulated first by the EU accession of Spain and Portugal in 1986, and then by the emergence of Central and Eastern European nations.

3.3 The trade in tasks conceptual framework

To organise our thinking about the second unbundling, it is useful to explain the basic determinants of whether a particular task is performed at home or abroad. This is not difficult as it boils down to cost savings.

Consider a task that requires some skilled and some unskilled labour. If the firm organises production such that the task is performed domestically, then the cost of the task will be:

$$\begin{pmatrix}\text{Domestic}\\ \text{task}\\ \text{cost}\end{pmatrix} = \begin{pmatrix}\text{Domestic}\\ \text{unskilled}\\ \text{wage}\end{pmatrix}\begin{pmatrix}\text{Domestc}\\ \text{unskilled}\\ \text{requirement}\end{pmatrix} + \begin{pmatrix}\text{Domestic}\\ \text{skilled}\\ \text{wage}\end{pmatrix}\begin{pmatrix}\text{Domestic}\\ \text{skilled}\\ \text{requirement}\end{pmatrix}$$

The cost of the task if the firm buys it from abroad would be quite similar but note that now foreign wages and foreign input requirements would be used. There is also additional costs that would arise from coordinating the production with one of the tasks taking place far away:

$$\begin{pmatrix}\text{Foreign}\\ \text{task}\\ \text{cost}\end{pmatrix} = \begin{pmatrix}\text{Foreign}\\ \text{unskilled}\\ \text{wage}\end{pmatrix}\begin{pmatrix}\text{Foreign}\\ \text{unskilled}\\ \text{requirement}\end{pmatrix} + \begin{pmatrix}\text{Foreign}\\ \text{skilled}\\ \text{wage}\end{pmatrix}\begin{pmatrix}\text{Foreign}\\ \text{skilled}\\ \text{requirement}\end{pmatrix} + \begin{pmatrix}\text{Offshoring}\\ \text{costs}\end{pmatrix}$$

The last terms encompasses all manner of coordination and trade costs.

In the trade-in-task framework introduced by Gene Grossman and Estaban Rossi-Hansberg at the Jackson Hole conference in 2006, the key determinant of unbundling is the cost of performing each task at home or abroad. In one version of their theory, they allow firms to use home-country technology when employing foreign workers abroad. In this case the ‘Foreign task cost’ involves foreign wages, but Home labour requirements – a factor that have interesting implications for R&D policy (Sector 5).

Determinants of offshoring costs: Unpredictability

The cost offshoring a task will depend upon a large number of factors. Easily observable things such as telecommunication, transport costs, and the cost of delays or uncertain deliveries are just one type. Perhaps the dominant cost has to do with the increased cost of managing a more complex production structure.

It is not a random outcome that the production of goods and services is undertaken in factories and offices throughout the world. Spatially clustering production stages, i.e. packaging tasks in offices and factories, is done to make easier and cheaper to produce what the firm sells. The problem is that economists really do not understand the ‘glue’ that binds production stages and tasks together. The standard approach, production functions, is a black box; one assumes that certain amounts of productive factors are combined to produce a certain amount of output. Given this lack of modelling –

to say nothing of a lack of empirical work in the area – economists cannot really pretend to understand how expensive it would be to offshore various bits of a production process. Worse yet, the problem cannot really be considered task-by-task since the offshoring of some tasks will typically change the cost of offshoring other tasks.

For example, consider a 'team' of tasks that is spatial clustered in a single office. To be concrete, say there are 'n' tasks – each performed by one worker – that must be performed to produce the intermediate input (say a marketing report) which is itself fed into a larger production process. Coordinating the n tasks requires each worker to talk, say, once a day with every other worker. Turning to offshoring possibilities, assume that offshoring entails a fixed cost per task offshored, and that each of the tasks could be performed more cheaply in India.

But what about coordination costs? Talking face-to-face is more efficient in terms of time than e-communicating. Keeping all the tasks in the same office reduces coordination costs – but this is true whether the office is in Canada or India. In particular, coordination costs are maximised when half the tasks are done in India and half in Canada. Now what this means is that wage savings plus extra coordination cost may not make offshoring one task worthwhile. However, if the coordination cost among a group of tasks falls, the offshoring decision can face a tipping point. Offshoring of tasks happens in a lumpy fashion. In this simple example, no tasks are offshored for all coordination costs up to a certain level, but beyond that point all tasks are offshored.

Another key source of unpredictability could come from cluster economies. In both services and manufacturing, tasks are subject to backward and forward linkages. That is, there is a tendency to cluster certain tasks together spatially to improve efficiency and gain better access to customers. In this sort of world, the international allocation of tasks can be subject to multiple equilibrium with the possibility that small changes can shift the economy between these equilibrium. For example, it could be that few tasks are offshored since the local production of these tasks create agglomeration economies that make local production competitive. However, if enough task get offshored to erode the agglomeration economies, all the rest of the task may also then be offshored.

The range of possibilities is quite large as policy analyses in the New Economic Geography show (see Baldwin et al 2003). When agglomeration economies are important, marginal changes can lead to very large shifts.

3.3.1 Is trade in tasks good or bad?

In 2004, Greg Mankiw, who was then Chairman of US Council of Economic Advisors, announced to the US business media that offshoring was just like trade in goods: "More things are tradable than were tradable in the past, and that's a good thing."[1] Mankiw was in good company since trade theorists have long modelled the second unbundling, i.e. fragmentation, as if it were just like trade in new goods.[2]

A central insight in the Mankiw-offshoring literature is that one can think of offshoring as technical progress in final goods. The intuition is dead easy. Unbundling production processes – i.e. allowing trade in intermediate goods and services – opens new opportunities for arranging final-good production more efficiently. The extra opportunities must mean that the same quantity of primary resources can produce a higher value of final goods. That, of course, is just the definition of technological progress in final goods, and this is why offshoring tends to act like technological

[1] Quoted in Blinder (2006).

[2] For example Dixit and Grossman (1982), Ron Jones and co-authors: (Jones and Findlay (2000, 2001), Jones and Kierzkowski (1990, 1998, 2000), Jones and Marjit (1992); Deardorff 1989, Venables (1999), Markusen (2005). These papers present a bouquet of special cases in which many expected and unexpected things can happen. For an even older tradition see Batra and Casas (1973).

progress in final goods. While the productivity improvement is guaranteed at the global level, national gains are subject to the usual provisos concerning terms of trade, factor intensive reversals, etc. This ancient insight is very helpful in placing offshoring models in the broader context of trade theory.[3] It is also a useful way to explain the potential gains from offshoring to non-specialists.

A second central insight in the Mankiw-offshoring literature concerns the impact of offshoring on wages. In general the literature concludes that there is nothing that can be said in general. The impact depends upon the factor intensity of the offshored task and the factor-intensity of the sector doing the offshoring. The point of these results was to dispel the common perception that offshoring the production of labour intensive goods to low-wage nations definitely harms low-skilled workers in the offshoring nation.

The fundamental economic logics of these two key insights are considered in turn.

Offshoring as technical progress

The core economic logic of the offshoring-as-technical-progress insight can be most directly illustrated in a very simple framework where there are no gains from trade in final goods. That is there are two nations, but only one final good and only one factor of production, labour. The production f the final good involves two 'stages' or 'tasks.'

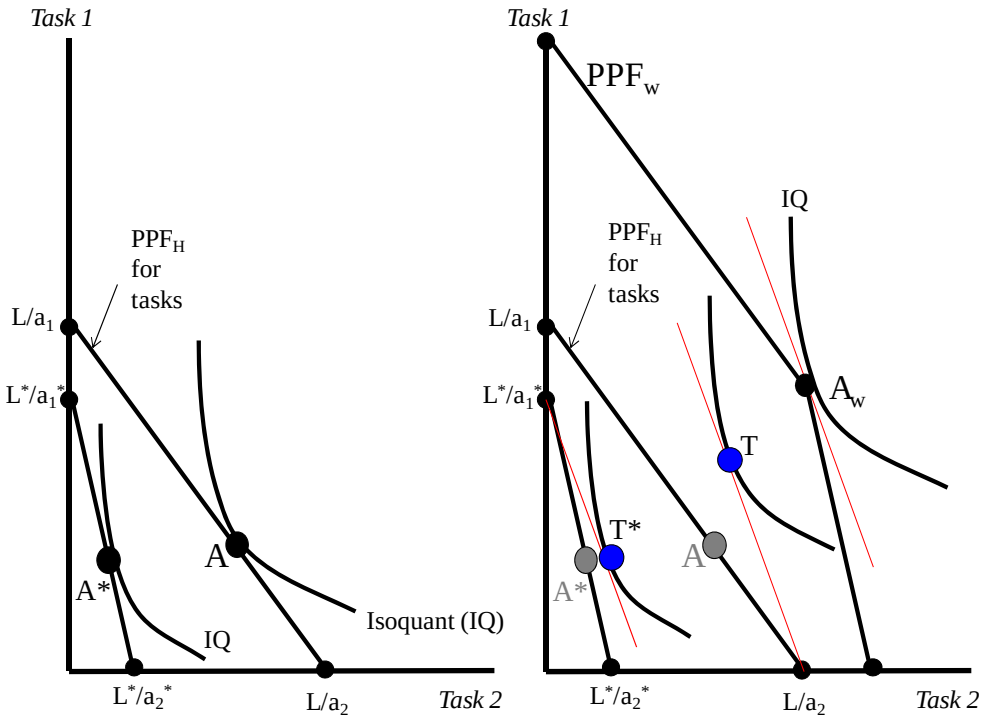


Figure 7: Trade in tasks as technological progress

To study the welfare effects of Mankiw-offshoring, it is useful to introduce the standard Ricardian diagram where there are two types of tasks, task 1 and task 2, one final good and two nations in Figure 7. As usual the amount of total amount of the tasks that can be produced by each nation shown with the production possibility frontier (PPF) for Home and Foreign. The tasks, however, cannot be directly

[3] Jones and Kierzkowski (1990) point out that it can be gleaned from Adam Smith's work; they also quote the 1928 American Economics Association Presidential address by Allyn Young: "… over a large part of the field of industry an increasingly intricate nexus of specialised undertakings has inserted itself between the producer of raw materials and the consumer of the final product." The insight is quite explicit in Jones and Kierzkowski (2000 p.13) and implicit in the diagrammatic analysis in Jones and Kierzkowski (1998).

consumed; they are combined into the single final good; graphically, this is shown as an 'isoquant', i.e. the combination of task 1 and 2 that can make a given amount of the final good.

To see how much Home makes without trade in tasks, we search for the highest isoquant that respects Home task-production constraint, namely the PPF. The answer is at point A in the left panel. Note that:

- A similar exercise reveals that Foreign would be at point A* without trade in tasks.
- The implicit prices of task 1 and 2 in Home and Foreign are set in their local markets and equal to the slopes of their respective PPFs.
- There would be no trade between these nations since wages would adjust to make each nation equally competitive in producing the final good.

When trade in tasks becomes possible, nations can trade the two intermediate tasks 1 and 2 as well as final good X. This situation is described by the right panel where the world PPF, marked PPF_w, becomes the relevant constrain on the production of final good X. (For simplicity, we assume away trade costs for tasks and goods in the diagram, so this is a switch from prohibitive task-trade costs to zero task-trade costs.)

At the world level, the optimal combination of task 1 and 2 is shown by the point A_w and the relative price of task 1 and 2 are now established on the world market by the slope of the isoquant at A_w. The world relative price lies between the two no-trade prices (as it must if all labour is to be employed). This change in prices makes Home task-1 production uncompetitive, so all Home task-1 production is offshored and all Home labour shifts from to task-2 production. The change in relative prices makes Foreign task 2 production uncompetitive, so all foreign task 2 production is offshored.

The right panel shows how trade in tasks shifts the <u>final-good</u> production point from points A and A* to O and O* (production of the final good is like consumption in the class 2-good Ricardian model). Note that the isoquant tangent to O and O* are higher than the isoquants tangent to A and A*.

The result is just like technological progress in both nations. Trade in tasks allows Home and Foreign to produce more of the final good with the same amount of primary factors. In both nations' labour forces becomes more productive when the productivity measured as final-good output per hour.

3.3.2 Wage effects of offshoring

Once we realise that offshoring is like technological progress, we can explore the general equilibrium wage effects of offshoring using a diagram like Figure 4. The result is shown in Figure 8. Since offshoring can occur in both sectors and in tasks that are both skilled and unskilled labour intensive, the new price lines will, in general, be shifted out. The new intersection, however, is can imply offshoring raises skilled wages while lowering unskilled wages (as at point E2), raising both (point E1), or raising unskilled wages while lowering skilled wages (point E3).

This is one of the fundamental differences between the new and the old paradigms. As offshoring can affect both sectors, it is not clear which groups will gain or loss from further globalisation.[4] More precisely, each sector is initially a bundle of tasks and the sector's factor intensity is the average intensity of all its constituent tasks. As unbundling proceeds, tasks are reallocated internationally roughly in line with comparative advantage. However, the process proceeds in both sectors, so the relative change in factor productivity – and thus the wage effects – are not clear cut.

[4] The paper that rekindled academic interest in North America over offshoring, or 'trade in tasks', Grossman and Rossi-Hansberg (2006) argued that offshoring unskilled intensive tasks would ambiguously raise the wage of unskilled workers, but this turned out to be a specially case that arose from there many special assumptions (Baldwin and Robert-Nicoud 2007).

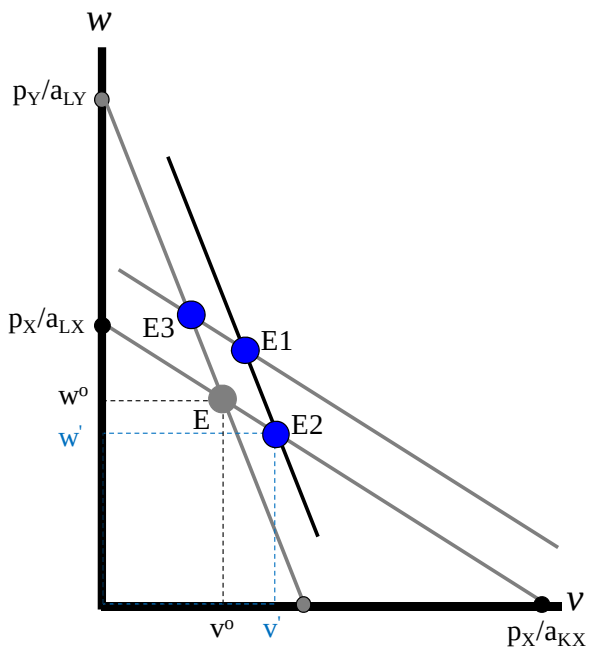


Figure 8: Ambiguous wage effects of offshoring

3.4 What's really new? Globalisation with higher resolution

As far as policy making is concerned, there are three really new things going on here.

1. Unpredictability. The winners and losers from globalisation are much harder to predict.

By their very nature, lower trade costs for goods tend to affect all traded goods in roughly similar ways and this is why one could tell which sectors would win from further reductions in trade costs. Governments felt they could predict which sectors would win and lose from future globalisation. This changes when the main barrier is the cost of coordinating complex processes across distance (trading ideas). Here it is difficult to identify winning and losing tasks, so we do not really understand the 'glue' that binds such tasks together in the first place. Knowing the direct cost of telecommunications is not enough since it interacts in complex and poorly understood ways with the nature of the task and the task's interconnectedness with other tasks.

2. Suddenness.

A job which 3 years ago was considered absolutely safe – say a German computer programmer designing custom software for a Landesbank – may today be offshored to India, or outsourced to a German software firm that offshores the job to India. The deep reason for this suddenness lies in the nature of complex interactions within factories and offices. Telecommunication costs have fallen rapidly but the impact has been quite different for different tasks. This may be due to the organisation of tasks within offices and factories. This organisation has changed more slowly. At some point – what might be called the tipping point – cheap communication costs line up with new management technology and a new task can be offshored to a lower cost location.

3. Individuals not firms, sectors or skill groups.

In the first unbundling, one could view firms as black-box bundles of tasks since firm-against-firm competition was globalisation's finest level of resolution. New paradigm globalisation suggests that the forces of globalisation will achieve a far finer resolution, at the level of tasks. This means that particular workers in particular firms in a given sector could suffer from globalisation while others in the same firm and same educational attainment prosper. New paradigm competition is on a much more

individual basis and this has some implications for policy. Policies design to help sectors may miss globalisation's losers entirely.

In addition to these new features that are important from a policy perspective, it would seem that there are two additional features that change the classic economic analysis of globalisation. These are:

4. Big versus little firm effects.

At present, offshoring of services has been much more aggressively pursued by large firms, probably due to economies of scale or scope involved in offshoring. To the extent that it lowers the costs of big firms, offshoring alters the balance of big-versus-small firm competition in domestic and export markets. This has many implications. For example, suppose one were trying to work out how many jobs had been 'lost' to offshoring. Given the shift on big-small firm competition, it is not enough to simply count the number of, say, data-entry jobs offshored by large companies. The competitive edge gained by large companies will force small firms in the same nation and same industry to downsize and/or go out of business. This suggests the estimates would be too low. On the other hand, the large firm's gain in competitive would typically boost their sales and this would favour job-creation in other tasks. Offshoring data-entry jobs may lead a large truck manufacturer to hire more production workers. This suggests the direct estimates of job loss from offshoring are overestimated. One would need a new-paradigm model to account for such intra-sectoral effects properly. Of course, one could simple assume that offshoring lowered the marginal cost of big firms in a standard heterogeneous firms model, but this would starting the story halfway through. It would not provide an analysis of the connection between the fundamental change (easier trade in tasks) to its affects.

5. Us versus them effects.

Another set of issues concerns international intra-sectoral competition. For example, suppose the home nation forbids outsourcing of data-entry jobs in an attempt to 'save jobs'. If other nations allow their firms to offshore, the home nation firms will find themselves at a competitive disadvantage. The expected result of from this would be a reduction in home firms' production, so in the end the policy could end up indirectly 'destroying' even more data-entry jobs rather offshoring would 'destroy' directly.

3.5 Example of production unbundling: The auto sector

The auto sector provides a clear example of the basic logic behind the trade in task conceptual framework, although a number of industry-specific features come into play.

The first point is that apart from Japanese producers, the first unbundling has not happened in this industry on a global scale – at least nowhere near to the extent it has happened in other sectors. Looking at the top ten auto companies, we see that for 7 of the 10, their regional share in global sales lines up fairly well with their regional share of global production.

Table 1: Limited first unbundling, regional production and sales, 2006.

	GM	Ford	Daimler Chrysler	Renault	PSA	VW	Fiat	Toyota	Nissan	Honda
Regional production share	50	43	58	75	70	66	55	56	41	37
Regional sales share	54	55	58	62	62	56	53	26	22	20

Source: Adapted by author from Sturgeon, Memedovic, Van Biesebroeck and Gereffi (2008), Table 3.

Sturgeon, Memedovic, Van Biesebroeck and Gereffi (2008) argue that this feature is due to the politics of the sector and the high visibility and sensitivity of the "Made in" label. Unlike toasters, TVs are

microwaves, consumers and policymakers care a great deal about where their cars are made. Thus even without explicit trade barriers, auto makers have generally chosen to keep the manufacturing of cars bundled spatially with the consumption, at least at the regional level.

A second unique feature is that while the second unbundling has been important, it has been spatially limited – again at the regional rather than global level. For example, vehicle produced in North America grew 40% from 1991 to 2005, but sales of the largest 150 suppliers in North America almost tripled (Sturgeon, Van Biesebroeck and Gereffi, 2009). In Canada, almost 80% of the supplier companies are located in Ontario (about 60%), or Quebec (about 20%), with exports being even more spatially concentrated. In the US, Detroit and the nearby mid-West are still the dominant locations for assemble and parts. There certainly been a move, especially by non-North American assembly plants to set up in the South. But given the US interstate highway system, many of these plants are a day's drive from the Detroit-Ontario cluster of suppliers, or two at the most.

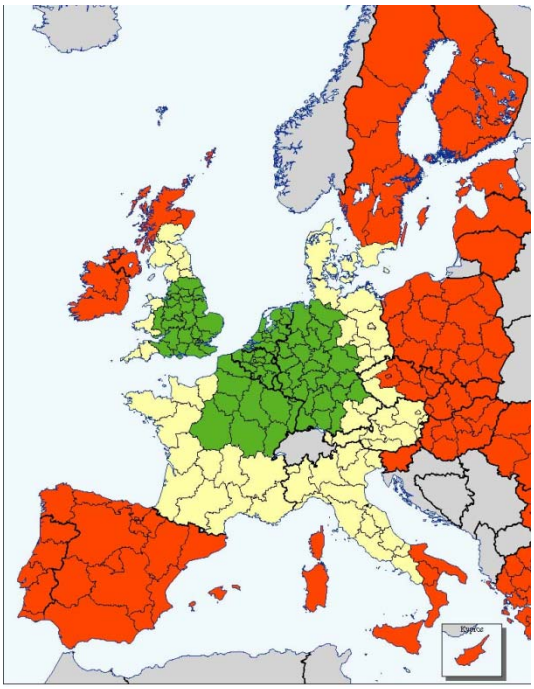
Table 2: Driving time from Detroit to foreign assembly plants opened since 1998.

Company	Location	Driving time from Detroit
GM/Suzuki	Ingersoll, ON	2h34
Toyota	Woodstock, ON	2h42
Honda	East Liberty, OH	2h59
Honda	Greensburg, IN	5h02
Subaru	Lafayette, IN	5h10
Toyota	Princeton, IN	7h36
GM	Spring Hill, TN	8h51
BMW	Spartenburg, SC	10h52
Daimler-Benz	Vance, AL	11h45
Honda	Lincoln, AL	11h52
Kia	Troup County, GA	12h24
Hyundai	Hope Hull, AL	12h37
Nissan	Canton, MS	14h11
Toyota	San Antonio, TX	22h18

Source: Driving time taken from Google Maps.

This clustering can in part be due to the nature of the good – vehicles and their main parts are bulky, heavy and can be fragile, so shipping costs are more important than in industries such as electronics. This also means that using air cargo is not feasible. The vagaries of sea shipments as have favoured road and rail transport and these tend to work best at a regional level. Another aspect of this "regionalisation not globalisation" is the dominance of a few large firms which has hindered standardisation of major parts and components via the development of global norms. This lack of generic parts and subsystems as hindered the emergence of worldwide scale economies – thus allowing more, smaller, local parts suppliers to remain competitive.

Figure 9: Europe's core and periphery regions



Source: Author's adaption from Schürmann and Talaat (2000).

Moreover, as parts and components are quite model-specific, and because transportation is relatively difficult and expensive, the unbundling of tasks at the factory level has not taken place over vast distances. Widespread adoption of lean production techniques and increasing product variety tends to foster spatial clustering of parts production and final assembly.

3.5.1 Europe

Europe's economic geography is very stark, with a traditional division into 'core' and the 'periphery' regions (even though distance is continuous). The primary driver of this spatial segmentation is accessibility to large markets. Due to variations in road and rail links – teamed with the spatial concentration of population in the core – firms located in the core have much easier access to consumers and suppliers. Put simply, core regions have good access to Europe's markets while peripheral regions do not. Figure 9 establishes boundaries for core, intermediate and periphery regions based on driving time for trucks to major markers. While any such calculation must be somewhat arbitrary, the resulting map lines up well with most experts' notion of Europe's core and periphery regions.

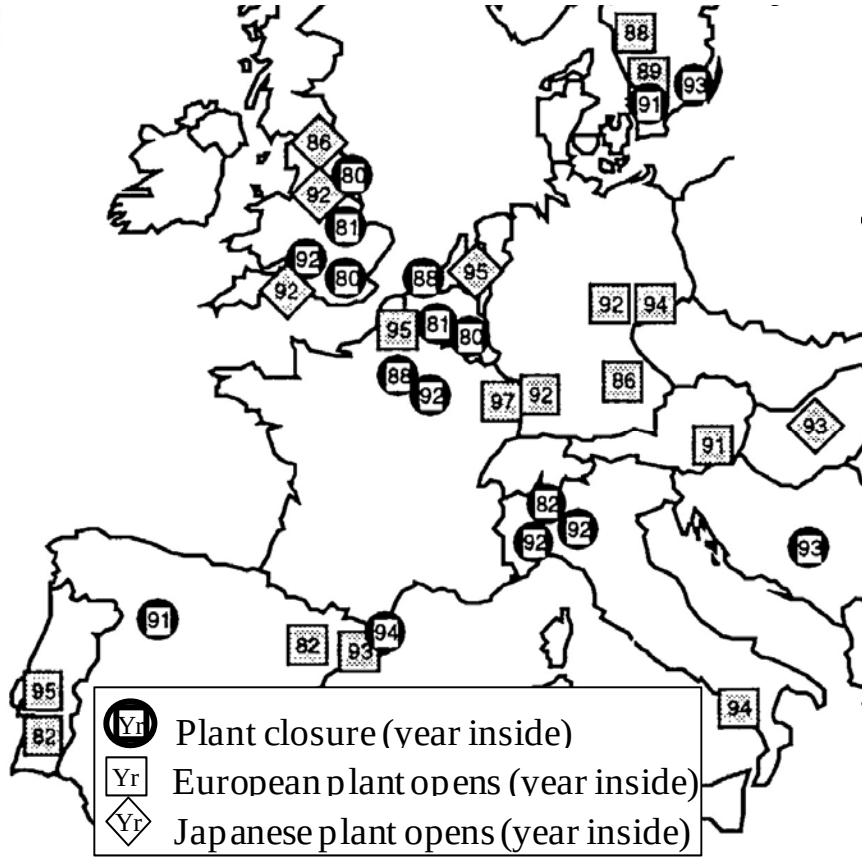


Figure 10: Changing geography of European auto sector.

Source: Author's modification of map in Bordenave and Lung (1996).

As concerns the auto industry, Europe experienced a clear spatial dispersion during the second unbundling, as in North America. Indeed the pattern is remarkably similar to that of North America but the dimension of dispersion away from the core was East and West in Europe instead of North and South. Figure 10 shows that pattern of plant openings and closures in Europe from the early 1980s to 1997. While the 'core' region experienced many plant closures, it also saw some new plants opening. Nevertheless the economic geography of the industry has clearly become more dispersed during the second unbundling (which started in the mid 1980s). Production on the Iberian Peninsula and in Central Europe expanded notably. As in the US, most of the 'transplanted' non-European producers – firms that had no history of producing in the core and thus were not subject to geographical path-dependencies – tended to set up in the periphery regions, first in the UK and than in Slovakia.

Of course many things were happening in Europe during this period other than production unbundling. The Velvet Revolutions in Central and Eastern Europe opened up vast new tracts of land to the possibility of foreign direct investment. European nations were also embarked on a massive liberalisation of trade flows. Although all industrial tariffs had been eliminated since the mid 1970s in Western Europe, the EU's Single Market Programme lowered trade cost s by tackling non tariff barriers and implementing trade facilitation actions such as removing border controls and harmonising customs and taxation procedures.

4 RELEVANCE TO NORTH AMERICAN ECONOMIC INTEGRATION

The trade-in-task theory was developed by Grossman and Rossi-Hansberg (2006) primarily to examine the offshoring driven by low wages, which was the "issue du jour" in the US at the time. This must be modified when thinking about the implications of the framework for Canada's industry.

The focus on large wage differences is misplaced in the US-Canada context, although it is still relevant in the broader NAFTA context. The US and Canada are both rich nations with sophisticated industrial firms in a range of sectors. Although wages are not equalised – and generally speaking Canada's productivity adjusted wages are lower –wages are not massively different. A far more important problem with Grossman and Rossi-Hansberg's 'new paradigm' is that fact that it ignores market size issues. A dominant element affecting the location of industry in North America is the huge market-size advantage possessed by the US. Since this factor is completely assumed away in existing trade-in-task theory, the theory must be extended to allow us to study the interactions between trade costs, agglomeration economies and economic integration.

4.1 The trade-in-tasks framework when market size matters

The mainstream framework for studying the impact of market size on industrial location is the so-called New Economic Geography literature launched by Paul Krugman in the 1990s (e.g. Krugman 1991). We briefly review the logic of this framework before discussing how to integrate it with the trade-in-tasks framework.

4.1.1 A New Economic Geography primer

The focus of the new economic geography (NEG) is on firms' location decisions. These decisions rest on the balance of two sets of forces – dispersion forces and agglomeration forces.

Dispersion forces, as their name suggests, favour the geographic dispersion of economic activity. These forces are generally driven by some sort of congestion broadly defined. Most of these congestion factors (land rent, commuting time, etc.) are rather local and thus not directly of concern in this paper; three dispersion forces are important.

- Labour-market congestion.

Industrialisation tends to push us wages and this tends to discourage further agglomeration. This is an important issue in the US, Germany, Japan, and increasingly China.

- Local-market competition.

This reflects the fact that having many industrial firms located in a particular region tends to increase the degree of competition for customers in the local market; this tends to encourage firms to spread out.

Importantly, local-market competition depends upon trade barriers. For example, in the extreme case where a nation's markets were perfectly open to international competition would see global, not local competition, but short of this, trade barriers of all kinds tend to make local competition a more important consideration. This fact creates a direct link between industry location and all manner of trade barriers – ranging from tariffs, to standards, to border security checks. This linkage will play a key role in the policy discussion in Section 5. Agglomeration forces counteract dispersion forces.

- Standard comparative advantage.

Nations are not all equally good at producing all things, or to phrase it in standard 'old paradigm' terms – nations have different comparative advantages. The sources of these differences can range from resource endowments, to technological differences, and natural geography. These constitute dispersion forces since other things equal they imply that some types of economic activity should be done in all nations. At a sector level, however, the sources of comparative advantage tend to encourage clustering by sector. In the traditional trade framework, countries become more specialised as trade costs fall. For example, as trade barriers come down, an ever larger share of clothing production shifts to China. From the global perspective, however, this might look like the clustering of apparel production, but it is not driven by agglomeration economics.

Agglomeration forces

An agglomeration force is said to exist when the spatial concentration of economic activity creates forces that encourage further spatial concentration. There are many agglomeration forces, but some of them only operate on a very local scale (like the knowledge spillovers that explain why universities departments and government departments are typically clustered in a given building). This level of spatial clustering, however, is not relevant to this paper. The two agglomeration forces we consider are supply-side and demand-side circular causality; they operate at a continent-wide scale and are directly affected by trade costs (and thus affected by policy choices including tariffs, border infrastructure, etc.).

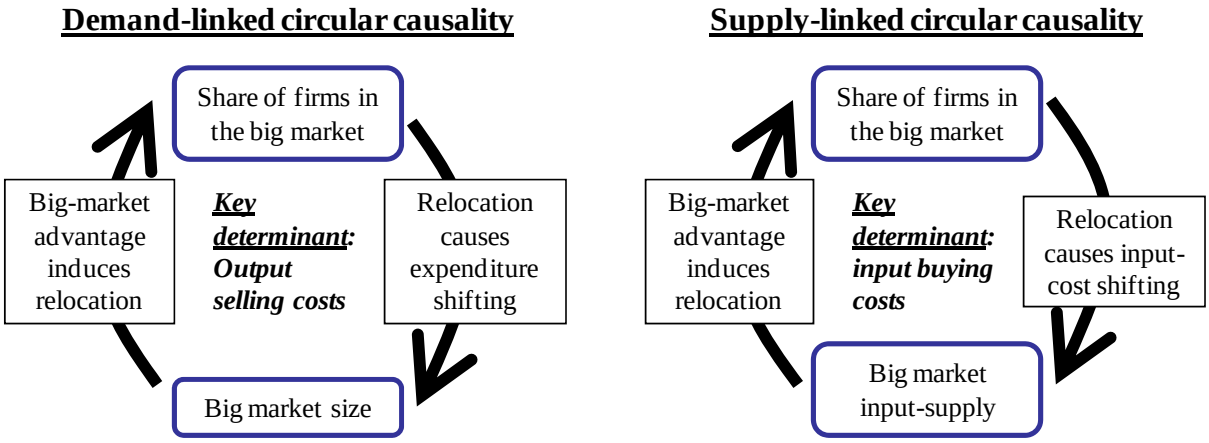


Figure 11: Circular causality & agglomeration forces

Demand-linked circular causality rests on market-size issues. Firms want to locate where they have good access to a large number of customers, like the US, in order to reduce selling costs (where selling costs include everything from shipping charges, border delays and import duties to back-and-forth communication with customers, etc.). Firms buy inputs from other firms, so firm relocation affects market size and thus the causality becomes circular. If are no dispersion forces are in operation to counteract this agglomeration force, all economic activity ends up in the big market. If all factors of product are mobile across borders, this force would tend to completely empty out small regions via factor migration, however in the international setting we usual ignore massive cross-border movements of labour. (This demand-linked circular causality is a key factor in the rapid rural-urban migration observed globally; as internal transport costs fall, firms create jobs near big cities since they want to be near their customers; people move to the cities since that is where the good jobs are, and the cycle begins again.) This is illustrated in the left panel of Figure 11.

The second major type of agglomeration force is the so-called input-cost-linked circular causality, or 'supply linkages'. This is the agglomeration force most relevant to production unbundling in the North American setting since it deals directly with supply chains. Manufacturing firms in modern industrial economies buy many inputs from other firms – machinery, parts and components, specialized services such as marketing, accounting, IT etc. Since it is cheaper to find and buy such input from firms that are

nearby, the presence of many firms in a location tends to reduce manufacturing cost of doing business in that location, other things equal.

Again, this leads to circular causality (see the right panel of the diagram). If many firms are already in the big market, then doing business in the big market will – all else equal – be cheaper and this will attract firms that in turn make the site more attractive from the input-cost perspective. If there were no dispersion forces, this circular causality would empty out the small market entirely. (Inside nations, this goes a long way to explaining the spatial clustering of sectors, e.g. the chemicals sector and the automobile sector.)

Note in a somewhat dated terminology demand and supply linkages are called 'backward and forward linkages'.

4.2 The locational effects of liberalisation

The focus here is on trade and industrial policy and one of the most direct effects of such policies is on trade costs. We therefore turn to studying the connections between trade costs and the location of industry in the NEG framework.

The first thing to observe is that lower trade costs reduce the strength of demand-linked and supply-linked agglomeration forces. Put simply, demand-linked agglomeration is driven by firms' desired to minimise the cost of selling to their customers by locating in the big market (customers here can mean either final consumers or other firms). As selling costs – including freight, border costs, two-way communication with customers, etc. – fall the incentive to locate in the big region diminishes. Likewise, supply-linked agglomeration is motivated by a desire to reduce the cost of buying intermediate inputs. As distance-related buying costs fall, the importance of being geographically close to suppliers shrinks.

The distance-related dispersion forces also get weaker as trade costs fall. The key distance-related dispersion force is the local competition effect. Here again, reduced trade cost reduce the advantage of being located far from your competitors. Indeed if trade were to be become costless, the local competition effect would disappear as the degree of competition would be the same regardless of where firms were located.

Importantly, many dispersion forces do not diminish with distance. For example, the labour market congestion effect – the tendency of industrial wages to rise in nations with relatively high industry GDP shares – is not directly related to distance or trade costs. Other dispersion forces actually get stronger as trade gets freer. Comparative advantage is one; the trend for labour-intensive industry to move to labour abundant nations, for example, gets stronger as trade costs come down.

To illustrate these relationships, Figure 12 plots the forces against the freeness of trade. It shows that both agglomeration and dispersion forces erode with trade freeness but that at totally free trade – i.e. costless trade – the dispersion forces would prevail. Quite simply, a world with costless trade would resemble classic trade theory where each nation's resources were fully employed and "each nation makes what it does best and trades for the rest." If this happened, industry would be far more evenly spread across the globe than it is now, where a handful of nations produce most of the world's manufactured goods.

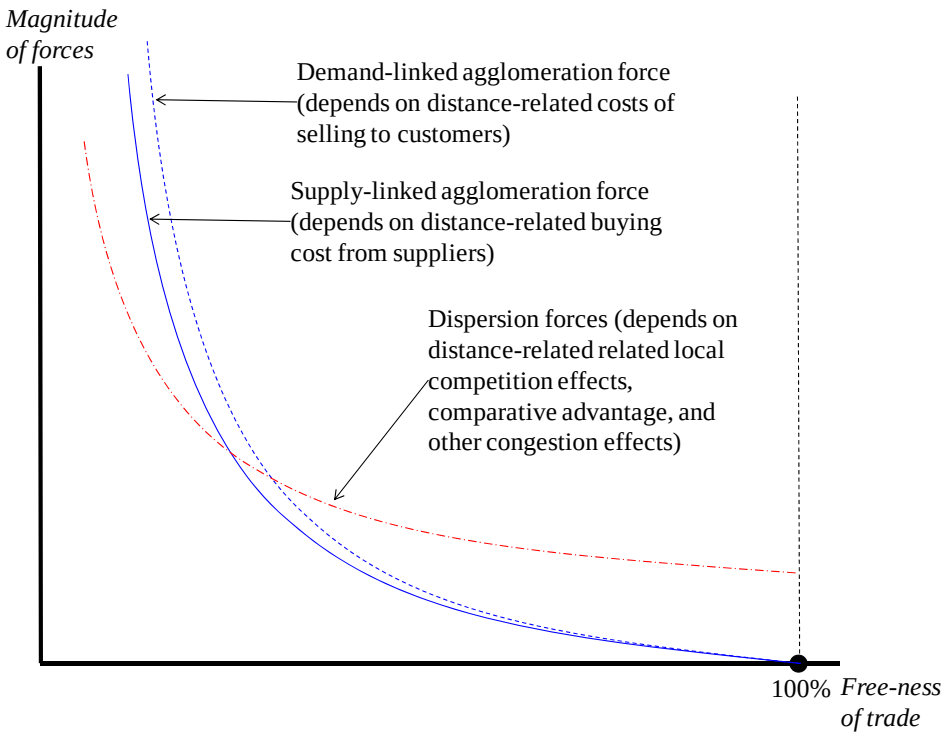


Figure 12: Trade costs and strength of agglomeration and dispersion forces

4.2.1 Determining the spatial equilibrium

Discussion of these forces and the impact of freer trade prepares the ground for the main goal of this section, the study of the spatial equilibrium. As we shall see, the share of industry in the big region adjusts to balance agglomeration and dispersion forces much like a price adjusts to balance supply and demand.

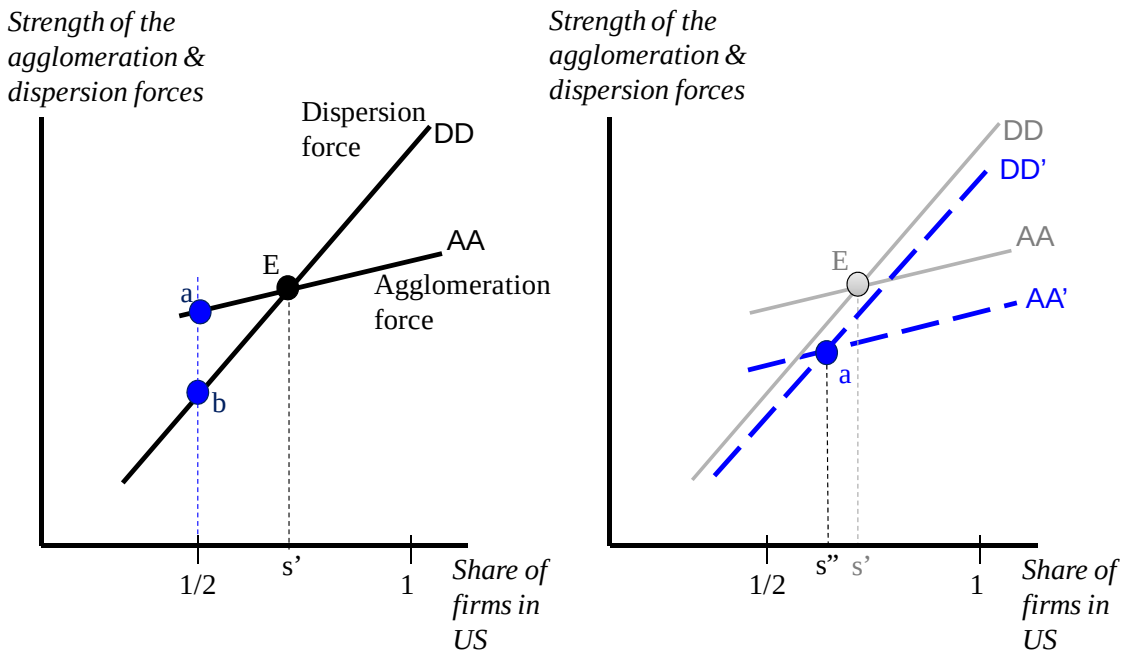


Figure 13: The locational equilibrium diagram

To see this, it helps to have a simple diagram to crystallise our thinking (Figure 13). Both panels of the diagram plot the strength of agglomeration and dispersion forces on the vertical axis. However in contrast to Figure 12, the horizontal axis plots the share of industry in the big region (US).

The left panel shows two lines, AA and DD, which illustrate how agglomeration and dispersion forces change with the concentration of industry in the US. The agglomeration-force line, AA, is rising due to circular causality (spatial concentration raises the incentive to spatially concentrate). The dispersion-force line, DD, is rising since the benefit of staying in the small region rises as more firms move to the big market due to wage congestion and local competition effects.

The locational equilibrium is at point E. This identifies the share of firms in the big market (the US) where incentives to agglomerate are just balanced by incentives to disperse. Given the US's intrinsic size advantage, it is clear that a share of half is not an equilibrium (i.e. the strength of the agglomeration force at s= ½ is 'a'; the strength of the dispersion force at s = ½ is 'b'; if the share started at ½, agglomeration forces would drive relocation until the big region's share of industry rose to s').

4.2.2 Is free trade pro- or anti- agglomeration?

The left panel was drawn for a given level of trade freeness. A critical issue for this paper is the impact that reducing trade costs have on the location of industry. This is studied in the right panel.

As discussed above, lower trade costs generally makes distance less of an issue and thus weaken both agglomeration and dispersion forces. The impact on the share of industry in the small region can go either way. If the agglomeration forces weaken more than the dispersion forces, the small region's equilibrium share rises (i.e. the US share falls). This is the case illustrated in the right panel, but plainly it could go the other way if DD fell more than AA.

As a rough rule, broad trade liberalisation in recent decades seems to have fostered a dispersal of industry, which is why the left panel depicts liberalisation as anti-agglomeration. We can see this at the global level (OECD nations are de-industrialising while the emerging economies are industrialising, see Debande, 2006), and within Europe (see the auto industry example above), and between the US and Canada. However, during the first wave of globalisation (roughly 1870 to 1914), lower trade costs were associated with a very strong agglomeration of industry in the North (especially in Britain, the US, and some West European nations) and a de-industrialisation of the South (especially India and China, see Baldwin and Martin 1999).

The NEG literature explains both outcomes with the so-called hump-shaped nature agglomeration rents, which notes that the balance of agglomeration and dispersion forces is most strongly titled toward agglomeration at <u>intermediate</u> trade costs. Consider the polar examples.

When trade is highly restricted, it is very unprofitable for firms in the core region to sell to peripheral markets. This dampens their enthusiasm for location in the "core", i.e. big region. Indeed if trade in prohibit expensive each region has to make everything it consumes, so the dispersion of industry matches the dispersion of consumers. At the other extreme of perfectly costless trade, location in the core or any other region is immaterial, so the gains from being in the core are nil. It is in between these two extremes – in other words, at intermediate trade costs – that location in the core matters most. For intermediate trade costs, clustering is both possible (since firms in the core can still sell to customers in the periphery) and profitable (since locating in the core economises on trade costs).

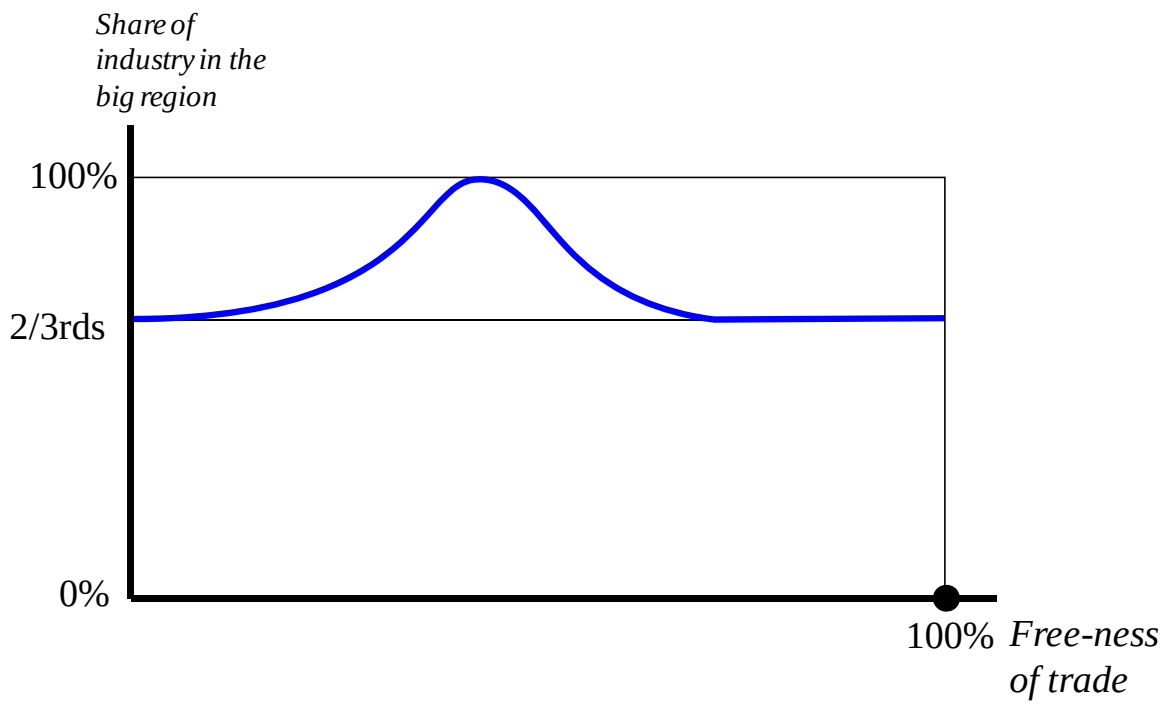


Figure 14: The hump-shape relationship between agglomeration and trade costs

This widely known feature of the NEG logic leads to the seemingly contradictory conclusion that lowering trade costs when they are high tends to produce a concentration of firms in the big region. However beyond some level of trade costs, further trade facilitation leads to dispersion away from the core.

This shown in Figure 14 for the case of a big country that is naturally big, say it has 2/3rd the world population and the small country has 1/3rd, so a neutral, or non-agglomerated location equilibrium would involve a 2/3-1/3 distribution of industry. When the free-ness of trade is zero (autarky) and 100% (cost-less trade), the neutral distribution prevails. In between, agglomeration forces tend to encourage spatial concentration in the big region. Note however, that once trade get free enough, the dispersion forces that are unrelated or positively related to trade costs take over and push the equilibrium to the un-agglomerated state.

While there is no clear empirical dividing line, many economists believe that the advanced industrialised economies are beyond the turning point. Further globalisation seems to be associated with a dispersal of manufacturing away from the big markets and towards nations with lower labour costs – especially those that are located in a way that naturally provides them excellent access to big markets. Baldwin and Krugman (2004) have used this feature to explain the changing nature of tax competition since the 1980s.

4.3 Threshold effects, trade cost/policy interactions, and hysteresis

Cluster economics presents policymakers with a set of issues that do not arise in more standard, smoother, more neoclassical models. Three of these are worth discussing in the present context (see Baldwin et al 2003 Chapter 9 for a more extensive discussion, and Chapters 12 – 18 for applications to trade policy, tax policy and subsidies policies).

When starting from a situation where industry is concentrated spatially, agglomeration forces can render small policy interventions useless, even though a large policy intervention could be effective. Firms located in a region with a large concentration of industry enjoy agglomeration economies as explained above. (In many cases, governments and labour unions in the core region attempt to ‘tax’

these agglomeration economies by charging higher taxes, demanding higher wages, higher benefits, stricter firing conditions, etc.)

Threshold effects arise since it takes a sufficiently large policy push, say a production subsidy or tax holiday, to attract firms away from the agglomeration rents available in the core. Unlike the standard neoclassical framework, one does not observe a little relocation from a little subsidy. One observes no relocation until the subsidy passes a particular threshold and then the effect can be larger than expected. The reason is that as firms start to locate away from the 'core', circular causality runs in reverse. Delocation of some firms reduces the attractiveness of the core and boosts that of the small region.

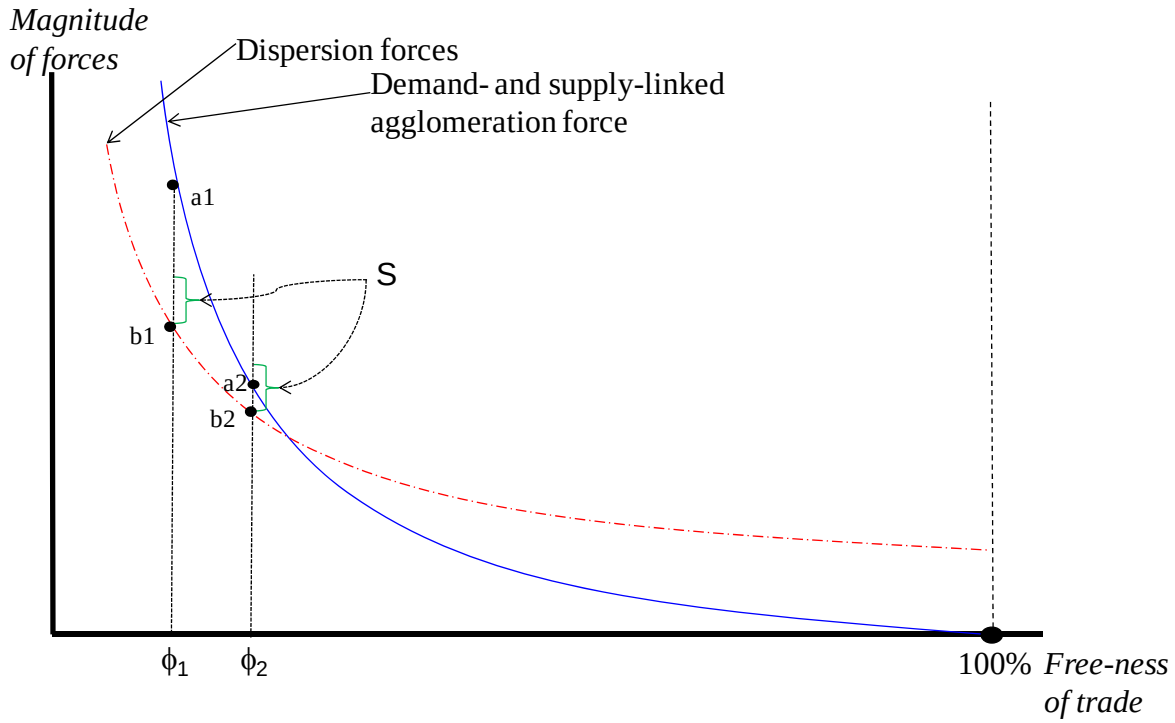


Figure 15: Trade cost and relocation policy interactions

Importantly, the level of trade costs can interact in unexpected ways with the relocation policies. Figure 15 shows an example. Consider a level of trade free-ness equal to ϕ_1, where the agglomeration forces (whose level is shown by a1) are stronger than the dispersion forces (shown by b1). At this point, a subsidy to firms that relocate to the small region will not be effective unless it is at least equal to the difference between a1 and b1. Suppose that a subsidy equal to S in offered nonetheless, but no firms relocate to take advantage of it since the subsidy doesn't offset the relative attractiveness of the big market. Things change when trade gets free.

If trade gets freer, thus narrowing the advantages of being in the big region, the same subsidy may well have an effect. In the example, a rise in trade freeness to ϕ_2 would narrow the gap between agglomeration and dispersion forces to a2 minus b2. Since the subsidy S exceeds this gap some firms would relocate in response. This means that trade facilitation programme will tend to amplify the impactfulness of Canada's pro-industry policies ranging from R&D policies to health care.

Hysteresis is the next concern. The world of real economic geography is marked by 'path dependencies'. The reason is that there are many possible places for industry to agglomerate, but once an agglomeration gets started – or for that matter starts to unwind – it can be very difficult to reverse the trend.

For example, a temporary policy that punishes firms in a particular location may lead them to depart, or may deter them from coming to the location in the first place; if trade is quite free, the policy can

even be a rather small one. To be concrete, suppose the policy change is a rise in the corporate tax in a particular region that induces firms to choose another region. When the policy mistake is reversed, and taxes are restored to their initial point, no relocation occurs. The reason is that the firms are now enjoying agglomeration economies in some other region and simply restoring the initial policy situation will not be sufficient. This property of irreversibility is called hysteresis in physics. Krugman (1991), which presents several historical cases where random events lie behind the establishment of large industrial agglomerations today, calls this the "history matters" property. Baldwin (1988) looks at the case of hysteresis due to large exchange rate fluctuations.

With these preliminaries out of the way, we turn to the main task of this section – extension of the trade-in-task framework to allow for agglomeration effects.

4.4 Adding trade in tasks: Unbundling when supply and demand linkages are important

How does production unbundling fit into the NEG? The standard NEG approach views firms as a production bundle; the range of production stages performed inside the firm is taken as immutable. Production unbundling changes this. To be specific, consider a car whose components' production must initially be spatially bundled with final assemble (due to communication costs, delays and uncertainty). When ICT advances make production unbundling possible, the components can be produced by separate factories on either side of the border. Almost surely, this will alter the spatial distribution of industry because the balance of agglomeration and dispersion forces will be different for each component.

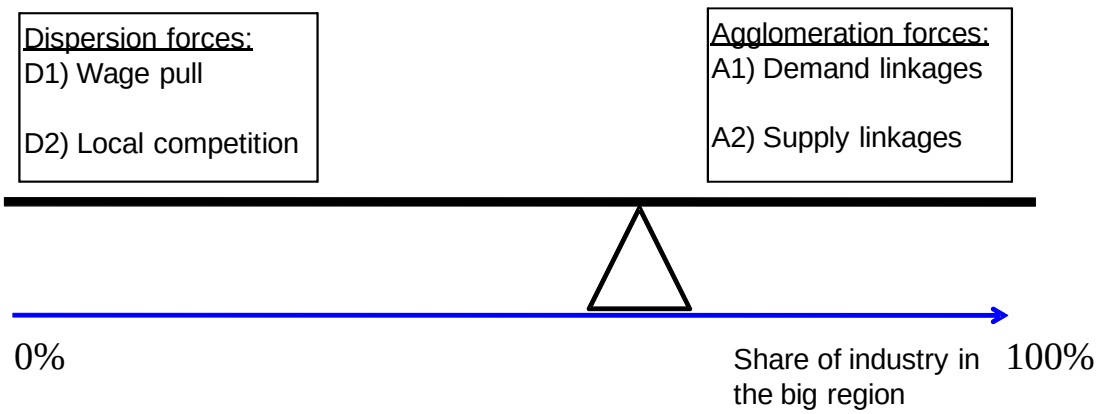


Figure 16: Location equilibrium: Fulcrum, load & effort arms, and the weight of forces

In fact there is likely to be a very clear pattern in the relocation produced by unbundling. To see this, it is useful to conceptualise the location equilibrium (i.e. the share of industry in the big market) as balancing the weight of agglomeration forces on the one hand and dispersion forces on the other (Figure 16). In this approach, the spatial division of industry is the fulcrum that balances agglomeration forces (on the right end of the lever) against the dispersion forces (on the left side of the lever). The location of the fulcrum when the lever is in balance describes the share of industry in the core region. Anything that strengthens the agglomeration force (i.e. increases the 'weight' of the agglomerate forces in our lever analogy) requires the fulcrum to shift rightward to rebalance the forces (implying an increase in the share of industry in the big region). Factors that strengthen dispersion forces will shift the fulcrum to the left (more dispersion). Of course, what really matters is the relative weigh on the two arms.

Using the lever analogy, it is easy to see that unbundling will almost surely alter the location of industry. The text in Figure 16 (which summarises the discussion in the previous section) indicates that we can focus on five key factors determining the spatial division of industry: 1) relative wages, 2) sensitivity to local competition, 3) output-selling trade costs, 4) input-buying trade costs, and 5) cost share of intermediate goods.

The initial equilibrium balanced forces for the production bundle as a whole – i.e. it is the bundle's average agglomeration and dispersion forces that mattered, where the average is across all manufacturing stages. Because the components will individually face a different balance of dispersion and agglomeration forces, the production location of some components will change. But in which direction?

While details will dominate in particular cases, the logic of trade in tasks suggests that unbundling tends to reduce concentration in the big market, but especially in upstream production stages. The key is that supply-linked agglomeration forces are systematically less important for production stages near the beginning of the value-added chain since these stages buy fewer traded inputs.

To see this, consider the stylised, 3-step supply chain in Figure 17; parts are assembled into components which are in turn combined with more parts to produce the final good.

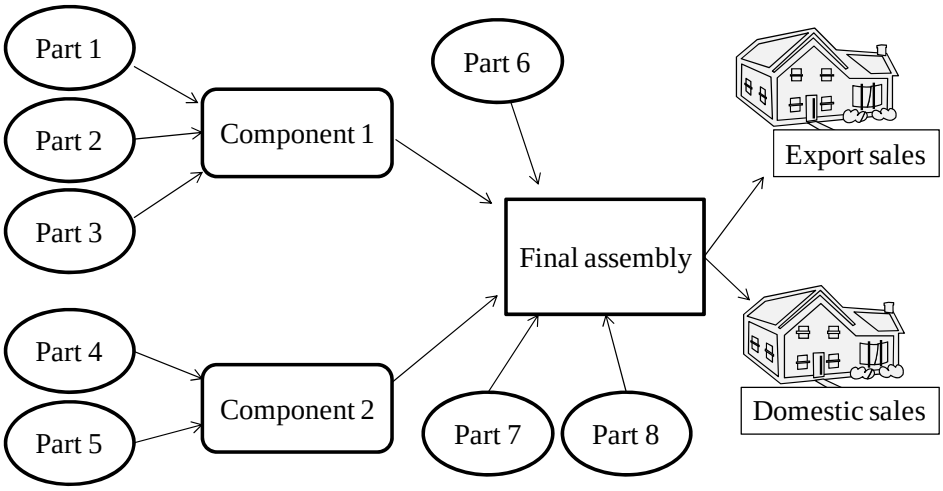


Figure 17: A stylised supply chain.

It is immediately apparent that parts, components and final assembly face different types of agglomeration forces. For all of them, demand-linked agglomeration forces matter since all of them must sell their output and thus output-selling trade costs matter. Parts, however, do not themselves buy intermediate inputs – they are, in the example, produced directly from primary factors such as capital and labour. As a consequence, the production of parts is not subject to supply-linked agglomeration forces, i.e. input-buying trade costs are irrelevant. As far as dispersion forces are concerned, the key is that the advantages of lower labour costs in the small region apply equally pre and post unbundling. Now putting these observations together with the fact that the average agglomeration force matched the average dispersion force in the pre-unbundling equilibrium, the bundled spatial equilibrium will no longer be correct for parts. The dispersion forces for parts alone will outweigh the agglomeration forces and some of the parts making will be offshored.

In Figure 16 terms, the agglomeration forces are "lighter" for parts than they were for the whole bundle, so rebalancing requires the fulcrum for the parts industry to shift leftward (i.e. the share of parts production in the big market will fall). Parts, in other words, will be outsourced and offshored to the small region because supply-linkage are weaker for the initial production stages that use fewer intermediate goods. This may not be the end of the story.

The relocation of the parts industry may have a knock-on effect due to supply linkages. The shifting location of parts makers to the small region will make the small region more attractive to components manufacturers, in particular those that have a very high intermediate-goods cost-share and those for whom input-buying trade costs are high (e.g. for the parts that are particularly expensive to transport due to fragility or weight) and whose output-selling trade costs are low (e.g. for components that are easy and cheap to ship due to their weight, fragility, etc).

In short, the greater dispersion of parts makers dampens the attractiveness of the big region to component makers. This reduction in the 'weight' of agglomeration forces for components forces a leftward shift in the fulcrum for the components industry. This in turn will feed back to weaken the demand-linked agglomeration force in the parts industry, as the component firms are the customers for parts firms.

The basic point of the knock-on effects can be rephrased as follows. As the parts must somehow get to the component-maker's facility, and the components must somehow get to the final assembly plant, the question is whether it is more efficient to ships parts across the border to component markers, or to move the component-makers facilities to the same side of the border as the parts markers.

The same logic does not apply to final assembly since it is tied to the large region's massive consumer base. The shifting of parts and components to the small region reduces the supply-linked agglomeration forces in final assembly. However, unlike the parts and components segment of the industry, final assembly continues to face unaltered demand-linked agglomeration forces arising from the location of so many final consumers in the big market. Thus there is likely to be a less than proportional shift in final assembly to the small region.

Additional factors come into play in determining the location of production after unbundling.

- The labour-intensity of the various unbundled parts and components will differ (in the initial bundled situation what mattered was the average labour-intensity). Since a typical situation in the real world (and in the NEG theoretical models) is that equilibrium labour costs are higher in the big nation, relocation to the small nation is more likely, all else equal, for labour-intensive stages of production.
- Relocation will entail higher transport costs (shipping to the big market), so relocation will tend to be more attractive for parts and components that are easily, cheaply and reliably shipped.
- Unbundling may also change scale. Bundling may have prevented some of the components from achieve their most efficient scale of production. An example of this is 'shared platforms' in the auto industry. For such components, unbundling, by allowing a single component factory to serve more than one final good producer, will raise the scale of production and result in a spatial grouping of production. The location of these large factories will involve the same balance of forces as in Figure 16.

5 POLICY ISSUES

Having explained the basic conceptual framework – in particular the extension of the trade in task approach to allow for agglomeration economics (NEG) – we turn to a number of policy implications, starting with trade facilitation

5.1 Impact of trade facilitation

The basic logic of agglomeration suggests that the US economy's size is an enormous advantage in attracting and keeping industry. But the large-market benefits of producing in the US are set against the negatives stemming from local competition and higher wage costs. Presuming that most industries are on the downhill side of Figure 14, anything that makes US-Canada trade cheaper, faster and more predictable will tend to erode the attractiveness of the US's market-size. This, in turn, would tend to promote Canada as a location for industry.

Notice that trade cost entered the equation in two ways: the cost of buying necessary intermediate goods from across the border, and the cost of selling output across the border. As the US starts with a larger concentration of both customers and input-producers, a reduction in either or both costs will reduce the US's market-size advantage and thus foster the location of industry in Canada. This is just a general prediction of Krugman's New Economic Geography approach – free trade reduces the large market's advantage. The point, however, can be augmented with considerations arising from production unbundling.

The first point to make is that production unbundling has the effect of 'putting into play' a large number of industrial jobs that were previous bundled into larger plants – plants that were in turn attracted to the US's large consumer market. More precisely, the NEG-cum-trade-in-task logic teaches us to think of unbundling as a large drop in the distance-related input-buying trade costs. Before unbundling, it would have been prohibitively expensive to try to manage the sort of international supply chain we see today. Or to put it differently, firms located the production of parts and components close by – often in the same factory complex – in order to economise on input-buying trade costs broadly defined.

The second point returns to the interaction between pro-dispersion policies – like Canada's health care system, its production support, etc. – and trade costs. The example shown in Figure 15 is a rather general proposition. Indeed it is absolutely obvious once one sees that the advantages of being in the big market are eroded by lower trade costs. What this means is that the effectiveness of a particular pro-relocation policy will typically become magnified as trade costs fall. Trade facilitation, in short, can be a 'force multiplier' for Canadian industrial policy.

The third point is that the unbundling allows dispersion forces to operate more finely on the value-added chain. Tasks, in other words, will tend to migrate to nations that have the most appropriate factor prices. This should help Canada to develop more finely defined strengthens in manufacturing segments that correspond more precisely to natural comparative advantages.

The fourth point is that the crisis-linked changes in US auto sector can be thought of as a massive unbundling. The inertia created by the historical concentration of US production around Detroit is suffering a distinct blow. While politics makes it hard to close down US plants and lay off workers, the economics of the situation is forcing exactly this. The NEG-cum-trade-in-task logic provides an insightful way to view this development. Agglomeration forces are inevitably circular; spatial concentration creates forces that encourage further concentration (Figure 11). A massive disruption of the status quo will almost surely disrupt one aspect of this circular causality. The job loss and plant closings that will occur in the US will reduce the effects of path dependency, i.e. locational hysteresis. As mentioned before, economic geography outcomes are riddled with multiple equilibria. As the crisis deletes the existing spatial equilibrium, a new one must appear. In transition to the new one, small policy mistakes, even temporary ones, on either side of the border are likely to have hysteretic effects. In plain English; many plants and jobs will be reallocated in the coming years and this makes trade and industry policy far more important that it has been for a very long time. Whatever new configuration of the North American auto industry emerges after the crisis is likely to stay in place for a decade or more; auto plants do not move around like hotdog vendors.

5.2 Labour and industrial policy

One of the key dispersion forces fostering the location of industry in Canada is its low productivity-adjusted labour costs and favourable manufacturing incentives. (This paper does not address the social welfare consequences of such policies, but rather focuses on their location effects.) As discussed above, these dispersion forces are magnified by lower trade costs. This suggests that the impact of attractive labour and policies could rise as trade within North America becomes cheaper and more reliable. The same is true of R&D policies that prove attractive to component producers. This, of course, is just the reverse of the point that trade facilitation is a force multiplier for industrial policy.

It was pointed out that path-dependency, or hysteresis, is an endemic feature of the new economic geography reasoning extended to encompass trade in tasks. What does hysteresis mean for policy? The most direct implication is one of caution. Bad policies, even when they are temporary, may have long-last bad effects. Moreover, reversing the effects may be difficult and require policy reforms that are much larger than the change that lead to the initial effects. On the flip side, pro-industry policies, even if they are temporary, may have long-lasting effects on the location of industry, as mentioned above.

It is worth mentioning, however, that the logic of unbundling – especially the trade-in-task formulation of Grossman and Rossi-Hansberg – suggests a note of caution at the national level when it comes to R&D subsidies. Firms that offshore unbundled production stages often use the home-nation's technology in the foreign plants. While this is good for the companies, the benefit to the home nation is less obvious. In particular, one of the strongest arguments for forceful pro-R&D policies – tax breaks, government subsidies, competition policy exceptions and the like – is that the technology helps create 'good jobs'. However, this argument has much less force when the R&D subsidised by home nation taxpayers is used to employ workers abroad.

Formally, this aspect of offshoring using domestic technology is like exporting technology. As Samuelson (2004) points out, trade in technology does not always lead to gains from trade for both nations. In particular, the widespread existences of spillovers suggest that the 'sale' of technology via offshoring will not be properly priced from the social point of view. Ultimately, this is an empirical question. Does R&D promote a complex of competencies that boost industrial output in an unbundled world, or does it simple subsidies technology that will be used to raise the productivity of foreign workers in offshored jobs?

5.3 MFN tariff policy

The situation of Canada in North America is particular. It has a labour-cost advantage over the US, but Mexico's wage cost advantage is far greater. These bald facts, however, need to be qualified since not all labour is created equal, nor does every stage of manufacturing require the same type of labour. Roughly speaking we can classify labour into high, medium and low skilled groups. Inside North America, high-, medium- and low-skilled labour is relatively cheap in the US, Canada and Mexico respectively, at least in productivity-adjusted terms. From an old-paradigm perspective, North American integration should be expected to result in greater specialisation by the three nations into sectors that use relatively intensively, the factors with which they are relatively well endowed.

Things are more complex when we realise that trade in tasks is creating an ever finer division of factor intensities in the manufacturing sector, and when we realise that cluster economics and market size effects are important. In particular, as argued in Section 4.4, the relocation of some stages of production can alter the attractiveness of regions for other stages of production. The example presented concerned the relocation of parts to, say, Canada, arguing that this could improve the attractiveness of Canada for component manufacturers via supply-linkages. And since the component makers are the customers for parts makers, the causality is circular; having many parts producers make the location more attractive to component makers and the presents of component makers renders the locale more attractive to parts markers.

Given this logic, the attractiveness of Mexico's low-cost, but low-skilled and low-productivity labour poses a challenge to Canada's industry – including segments that are not low-skilled intensive. The Challenge comes via supply-linked agglomeration forces. This applies especially to manufacturing segments that intensively uses parts that whose production is low-skilled-labour intensive.

External trade policy can help meet this challenge. Lowering the cost of importing parts that are intensive in low-skilled labour is a direct response to the supply linkages created in Mexico by their abundance of such labour. As Robert Mundel noted decades ago, trade is a substitute for factor movements. The direct way to counter Mexico's advantage would be for Canada to 'import' low-skilled labour. Mundel's insight, however, tells us that this is not necessary. Reducing tariffs and other border costs on low-skill-intensive parts will tend to offset the attractiveness of the Mexican market and US locations near the Mexican suppliers of these goods. After all, the key to these supply linkages is the price of the input. The input's production location is relevant only in so far as it affects the price. Providing Canadian component makers with competitively price parts from third nations directly offsets the locational advantage created by the production of such parts in Mexico.

This is a line of thinking is quite consistent with the notion of effective rates of protection; lowering the tariff on imported intermediates, makes production of the downstream goods more attractive in the liberalising nation.

Because there is free trade inside NAFTA, the effectiveness of MFN tariff liberalisation depends in part on US and Mexican MFN tariffs. Taking the example of Cars, engines and engine parts (Table 3), Mexico's MFN tariffs are quite high, while those of the US are modest. In this situation, zero tariffs on engine parts in Canada can help boost its locational attractiveness to firms that use such imports. The point is that with its high MFN tariff, the price of such goods inside Mexico will tend to exceed the world price. Thus a zero MFN tariff policy could partially or entirely offset the supply-linkages in Mexico. The same is true for the comparison with the US, as long as the US maintains a higher MFN tariff on the parts.

Table 3: Canadian, US and Mexican average applied MFN tariffs (%).

	Canada	US	Mexico	Notes
Cars	6.1	2.5	50	HS 8703.2 to .9
Engines	5.2	1.3	8	HS 8407.34
(range)	(3.5 - 6.0)	(0 - 2.5)	(7-10)	
Engine parts	2.8	2	8.2	HS 8409.91
(range)	(0 - 6.0)	(0 - 2.5)	(0-10)	

Source: WTO Integrated Database.
Notes: The averages are at the HS 6-digit level.

The figures in Table 3 suggest that the current configuration of Canadian MFN applied tariffs provides a strong protection to the production of major components and the assembly of cars. This would be strengthened by further lowering of the average MFN tariff on parts used in component manufacturing, especially those subject to offshoring to Mexico.

The policy implication for this line of thinking is that the locational effects of Canada's MFN tariff depend upon the configuration of US and Mexican tariffs.

5.4 Unbundling and rules of origin

Any nation that applies preferential tariffs must have rules of origin to guide customs officials. These rules, however, have de facto been used to influence the location of industry. The traditional view is

that strict rules of origin foster the production of upstream intermediate goods. The trade-in-task framework, suitably extended to allow for cluster economics can provide some new insights.

Unbundling has and will continue to alter the politics and economics of rules of origin. In a nutshell, rules of origin are a way of bundling together the tariff protection enjoyed by upstream and downstream producers. In so far as unbundling further fragments a sector, it tends to erode the coalitions backing tariff-protection bundling (i.e. strict rules of origin). To explain this reasoning more carefully, it is necessary to provide some background on the economic effects and political economy of rules of origin (ROOs).

5.4.1 The basic economic effects of ROOs

Rules of origin (ROOs) put constraints on the inputs a firm in the downstream industry can use in making products that qualify for preferential tariff treatment. In the world of modern manufacturing, this means ROOs are, to a large extent, controls on the source and nature of intermediate inputs. Any serious analysis of ROOs must therefore distinguish between final goods and intermediate inputs (although of course the 'final good' itself may be a component rather than a consumer good). In most cases – and certainly in cases where unbundling is applicable – both the final and intermediate goods are traded. This means ROOs will interact with tariffs in intricate ways.

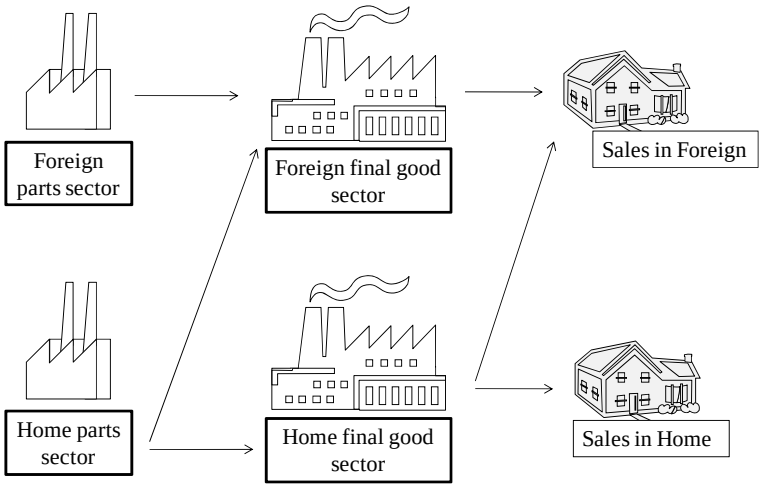


Figure 18: A simplified supply chain.

To keep the analysis concrete and reduce complexities to the bare minimum, we work with a stylized supply chain shown in Figure 18. We have two nations, Home and Foreign that compete in both parts and final goods. As drawn, Home is an importer of parts and an exporter of final goods. This pattern of trade does matter, but we can introduce the basic linkages most simply in this setting.

The economic linkages between the parts and final goods sector are twofold. First, the demand for parts is a derived demand, i.e. it depends upon the production level of the final good. In the case at hand, where Home does not export parts, the demand for Home parts is linked directly to Home production of the final good. This is a positive link in the sense that anything that boost Home final goods production will also tend to boost the fortunes of the Home parts sector.

The second link comes from the cost of parts. Since parts are an input to final goods, the supply side of the final goods sector is linked to the equilibrium price of parts. This is a negative link; policies such as tariffs and ROOs that raise the price of parts are good for the parts sector but bad for the final goods sector.

In short, the parts sector is linked to the final goods sector via demand and the final goods sector is linked to the parts sector via the supply side.

These links continue to exist, but are dampened, when the parts and or final goods sectors switch trade status. The final-to-parts link via demand now depends upon local and export sales, not just local sales. The parts-to-final link via supply now depends upon local and export sales of parts. With these basic points in hand, we turn to the politics of ROOs.

5.4.2 The basic political economy of ROOs

The starting point of any discussion of ROOs should be that ROOs will – absent some outside arrangement – typically find the interest of downstream producers in conflict with the interest of upstream producers. Of course, this is often not the case in the real world, so will shall have to address such 'outside arrangements', but we start by establish the baseline conflict over ROOs.

The point is that if ROOs are binding, and the downstream preferences tempting enough, the ROO will induce downstream firms to switch to higher cost inputs in order to qualify their good for duty free treatment. Typically, this will boost profits and output of the upstream firms, but lower those of downstream firms. Or to put it differently, strict rules of origin for a particular good tends to reduce final good's effective rate of protection. For the upstream firms and their workers, strict rules of origin are like a government mandated restriction on competition from third-nation competitors. Given the choice, and again ignoring outside arrangements, downstream producers would like the tariff preferences without the ROO-imposed restriction on inputs, while the upstream producers only profit from the downstream tariff preferences, if ROOs throw more business their way.

As mentioned, one often observes upstream and downstream firms lobbying in cooperation for strict rules of origin – the textile and clothing sectors being a classic example. These ROO coalitions must be based on a broader agreement since ROOs in isolation are a divisive factor. Consider some examples.

One situation where an outside arrangement is possible occurs when the ROO helps cement a coalition for including a particular sector in the liberalisation. Ravenhill (2008 p.194) provides the example of how ROOs for tomato ketchup switched from CUSFTA to NAFTA. Under CUSFTA, tomato ketchup was considered as originating in CUSFTA when made from tomato paste from third nations. This was convenient since US ketchup producers imported a great deal of tomato paste from Chile. Under NAFTA, however, ketchup cannot get duty-free treatment if it is made from third-nation tomato paste. The effect was that the US ketchup makers switched from Chilean tomato paste to Mexican tomato paste. While this raised US ketchup makers' costs somewhat, it had the effect of turning Mexican tomato paste producers into big fans of including ketchup in Mexico's list of products to be granted duty free status. The gain to US ketchup makers was the duty-free market access to Mexico, while the cost was a higher priced input. The gain to Mexican firms was a new export market for their tomato paste, while the cost was greater competition in the downstream good, ketchup. This is an example of a cross-border coalition, but it could also work inside one nation if the ROOs helped the parts makers expand exports within the FTA.

Another reason producers of a particular product may support strict rules of origin is to raise their rivals cost. For example, consider a product where Canada imposes a zero MFN tariff on certain parts while the US has prohibitive tariffs on the same parts. In this case, a very strict rule of origin will raise the Canadian firms' costs by forcing them to switch to higher priced parts made within the FTA. The ROO has no impact on US firms as they are already buying expensive parts locally. In this case, the strict rule of origin is offsetting some of the competitive pressure that the US firm faces from Canada when the US tariff on the downstream good is removed.[5] In this case, strict ROOs could reduce the opposition of US downstream firms to having their sector included in the FTA. US downstream

[5] See the excellent account in Destler (2003).

producers may still lose from the tariff removal but since the strict ROO raises their rivals' costs more than their own, the ROOs may dampen to an acceptable level the sector's opposition.

A third case, and the one most directly affected by unbundling, is when the parts and final goods are made by the same company, so the differences are internalised.

5.4.3 Unbundling and ROOs

The basic impact of unbundling is to make the ROO coalitions more difficult to manage. As production stages are separated spatially, especially when their ownership is also separated, the intrinsic conflicts between parts makers and parts buyers become more problematic – especially when the outsources parts are moved outside the nation. It may be the case the ROOs are 'saving industry jobs', but whose? As unbundling and spatial dispersion of upstream manufacturing proceeds, the nationalistic argument for ROOs tends to get blurred. Moreover, if unbundling results in a multiplication of firms, it will make political organisation more difficult.

Another aspect of this has to do with the difference the time of choosing tariff preferences (i.e. which sectors to excluded from the FTA) and ROOs. When NAFTA was set up, the two were chosen simultaneously. This allowed the formation of the sort of within-nation and across-nation coalitions favouring the bundling of up and downstream protection. When rules of origin are revised, however, the inclusion of particular sectors must be taken a fixed. Here it should be noted that downstream producers almost always have an interest in getting their inputs for less, i.e. for favouring weaker rules of origin. The raising-rivals'-cost example does not fit this, but the others do.

All this suggests that the political economy supporting strict rules of origin will erode as unbundling proceeds, especially in sectors where unbundling is proceeding rapidly. An extreme version of this can be seen in the electronics industry where production unbundling and offshoring has reached its greatest extent.

In the electronics sector, the international supply chain is so complex and geographically fragmented, that rules of origin are very expensive to document. It is also very hard to work out which firms would benefit from a strict rule of origin because the input makers are themselves outsourcing many parts. Given the near impossibility of harmonising rules of origin between the US and the EU, the reaction was to make the rules of origin irrelevant by switching to global duty free trade in parts and components with the 1997 Information Technology Agreement (ITA). This was a WTO-orchestrated deal that saw nations accounting for more than 80% of global production switching to zero MFN bound rates on an extensive list of goods (mostly parts and components; many final goods were excluded, see Mann and Liu, 2009 for details). What are the lessons for NAFTA?

It seems likely that a more modest version of this dynamic will play out in NAFTA. Unbundling will render strict rules of origin both more difficult to support politically and more costly administratively. The first point follows on the coalition reasoning laid out above. The second point can be seen by contrasting the difficulty of determining the origin of a Model T Ford (where almost every manufactured input was made in Michigan) and a modern Ford car (which may involve thousands of imported parts and components). Moreover, the growing complexity of the international supply chain will make it every more difficult to ascertain the true winners from strict rules of origin, and in any case unbundling will mean that more of the winners are located in foreign nations.

There is a close analogy between the logic that unbundling will unravel the US's big-market advantage starting from the beginning of the supply chain and moving forward. As the upstream suppliers become dispersed and fragmented, the winners from strict ROOs or origin will become less clear and there will be pressure from downstream producers to allow cheaper, third-nation parts and components to be used.

As discussed in the section on MFN tariffs, unbundling also triggers a dynamic that forces down MFN tariffs on inputs that are relatively far up the supply chain, but this dynamic continues to work its way down the supply chain, right up to but not including final assembly. An extreme example of this can be seen in East Asia where the unilateral liberalisation of parts and components in electrical and mechanical machinery (HD 84 and 85) is almost complete.

The relevance of the MFN tariffs observation is that this trend undermines the effectiveness of strict rules of origin. Firms, after all, always have the option of paying the MFN tariff and ignoring the ROOs, so if MFN tariffs get low enough – thus rendering the preferences too small to make the ROOs worth complying with – firms will simply ignore the rules of origin. Again, this has happened in East Asia, where the utilisation rate in the ASEAN FTA rarely exceeds 10% (Baldwin 2006).

As far as Canadian policy is concerned, the idea would be to use this insight to get a head of the game – to create some locational hysteresis in industries were tariffs and rules of origin still matter in advance of their phase out.

5.5 Product standards and unbundling

Citizens expect their governments to impose health, safety and environmental standard on the goods they buy. As intermediate inputs are an essential element in many final goods, it is also natural to impose standards on upstream products as well as consumer products. But product standard also play a protection role (Baldwin 2000). Unbundling and the Figure 16 logic suggest that pressures for this sort of protection will unwind from the beginning of the value added chain and moving forward. There is a close analogy with the logic concerning ROOs and MFN tariffs.

Take the example of the auto industry. Here NAFTA standards make it harder for European and Asia producers to sell in the North American market (just and European and Asia standard do reciprocally). But standards for which goods? When much more of manufacturing was bundled in a single firm, it was clear who would benefit from a particular idiosyncratic standard, just as it was clear who would benefit from strict rules of origin.

As manufacturing becomes unbundled and geographical dispersed, especially when parts production is both outsourced (produced by a different company) and offshored (produced in a differ nation), firms that ‘won’ from the protection provided by idiosyncratic standards may find themselves turned into losers. That is, as unbundling turns large companies into buyers of parts, there will be increasing pressure to lower costs by adopting international standards. This basic logic parallels that concerning the difficulties of maintaining coalitions in favours of strict rules of origin.

As far as policy is concerned, the usefulness of this insight is to avoid developing industry that will not be viable once international norms are adopted, especially in upstream stages of the value added chain.

There is a second logic that suggests unbundling will favour of the adoption of international standards. As discussed above, unbundling may allow certain segments of the production chain to achieve the minimum efficient scale that were not possible when they were tied to an individual downstream firm. That is, the unbundled firm can, by selling to more than one downstream firm, achieve greater economies of scale. But once firms start selling to more than one final good producer, they may face the problem of multiple standards. Since lowering these costs is likely to be in the interest of buyers and the sellers, this aspect of unbundling may foster the elimination of standards-based protection. To put it differently, overlapping standards becomes more of a problem when the supply chain gets unbundled and dispersed around the world. For policy makers, this suggests that efforts be made to advance the internationalisation of industrial standards.

6 FUTURE STEPS

Globalisation will continue. The world's capital market, international banking in particular, is likely to see a round of rising regulation that will dampen some of its globalisation, but other than this, recent talk of de-globalisation is exaggerated. The globalisation of manufacturing will surely continue, and the globalisation of services production is just starting. Two of the forces driving this globalisation are the rising competitiveness of emerging economies producers (China, India, etc.) and the advancing sophistication and falling cost of communication and information management systems.

On the rise of the emerging market manufactures little needs to be add to the voluminous discussion; suffice it to say that a growth take-off has begun in these nations. While they may find growth gets harder as the approach the productivity frontier, it looks certain that they will at least reach the output per person of nations like Korea or Taiwan. That would mean a fourfold increase in output, with a more than proportional rise in manufacturing output. Such a large increase in selling and buying power will greatly magnify the pull and push factors driving globalisation. The world's economic landscape will surely continue to flatten as far as manufacturing is concerned.

On the advance of communications technology, there is no end of the ways things could improve. If today's most advanced teleconferencing technology (large screens, multiple cameras and microphones, etc.) became as cheap and widespread as online telephony is now, trade in services would be revolutionised; the need for face-to-face meetings would be greatly diminished. The unbundling of the service sector has only just begun, hindered as it is, by the facts that it is still very expensive to move people around the world (falling airfares are offset by rising opportunity costs of time) and it is still necessary for many service producers to interact in person, at least periodically.

Of course, all this could be subverted if the world falls into a protectionist spiral as it did in the 1930s; globalisation can and has been reversed many times over the centuries (Findlay and O'Rourke 2007). But absent this, globalisation will continue to spatially unbundle economic relationships. The price of goods, technology and labour will increasingly switch from being set in set in local markets to being set in global markets. As part of the unbundling, the broad classifications such as 'skilled labour' will dissolve into a myriad of specialities, each with its own price.

6.1 New paradigm and trade in services

The rising tradability of services has given birth to some very popular but very confused thinking by non-specialists. One of the world's leading macroeconomic policy specialists, Alan Blinder, has recently been doing a brisk stroll thorough trade theory, making dire predictions – "I estimated that 30 million to 40 million U.S. jobs are potentially offshoreable." This attracts a great deal of media attention. Unfortunately, it seems Blinder's training in trade theory stopped before Paul Krugman led the 'new trade theory' revolution in the 1980s.

Krugman's contribution, which was rewarded with a Nobel Prize in 2008, was to crystallise the profession's thinking on two-way trade in similar goods. This was a revolution since the pre-Krugman received wisdom assumed away such trade, or misunderstood its importance.[6] Trade economists when Blinder took his graduate courses, took it as an article of faith that trade flows are caused by macro-level differences between nations – for example, national differences between the cost of capital versus labour. Nations that had relatively low labour costs exported relatively labour intensive goods to nations where labour was relatively expensive; that is the traditional view that Blinder is embracing.

[6] Harvard professor Richard Cooper, a doyen of international economics in the pre-Krugman era, notes the rapid rise in two-way trade among similar nations. He asserts that this sort of trade is not welfare improve and concludes that it should be taxed since the rising trade flows made it difficult to maintain the fixed exchange rate system which he considered as the really important thing from a welfare and policy perspective. See Cooper (1968, p.273).

What Krugman (especially Krugman 1979, 1980) showed was the one does not need macro-level differences to generate trade. Firm-level differences would do. In a world of differentiated products (and services are a good example of this), scale economies can create competitiveness even when the macro-level variables are the same across nations. Krugman, a pure theorist at the time, assumed that nation's were identical in every aspect in order focus on the novel element in his theory (and to shock the 'trade is caused by national differences' traditionalists). His insight, however, extends effortlessly to nations that also have macro-level differences, like the US and India.

Blinders approach is simplistic and thus easy to explain – a fact that accounts for much of its allure as well as its shortcomings. Step 1 is to note that Indian wages are a fraction of US wages. Step 1a is to implicitly assume that Indians' productivity-adjusted wages are below those of US service sector workers since their productivity gap with US workers is not as large as their wage gap, at least in tradable services. Step 2, and this is where Blinder focused his efforts, is to note that advancing ICT makes many more services tradable. The key characteristic, Blinder claims, is the ease with which the service can be delivered to the end-user electronically over long distances. Step 3 (the critical unstated assumption, if not by Blinder, at least by the media reporting his results) is that the new trade in services will obey the pre-Krugman trade paradigm – it will largely be one-way trade. Nations with relatively low labour costs (read 'India') will export relatively labour intensive goods (read 'tradable services') to nations where labour was relatively expensive (read 'the US'). This last step is factually incorrect, as recent work by Amiti and Wei (2005) has shown.

Amiti and Wei note: "Like trade in goods, trade in services is a two-way street. Most countries receive outsourcing of services from other countries as well as outsource to other countries." The US, as it turns out, is a net 'insourcer'; the world sends more service sector jobs to the US, than the US sends to the world, where the jobs under discussion involve computing (which includes computer software designs) and other business services (which include accounting and other back-office operations). Figure 19 shows the facts for the 1980 to 2003 period. What we see that that Blinder is right about the US importing an ever growing range of commercial services – or as he was say, the third industrial revolution has resulted in offshoring ever more service sector jobs. However, the US is also 'insourcing' and ever growing number of service sector jobs via its growing exporters. The startling fact is that not only is the trade not a one-way ticket to job destruction, the US is actually running a surplus.

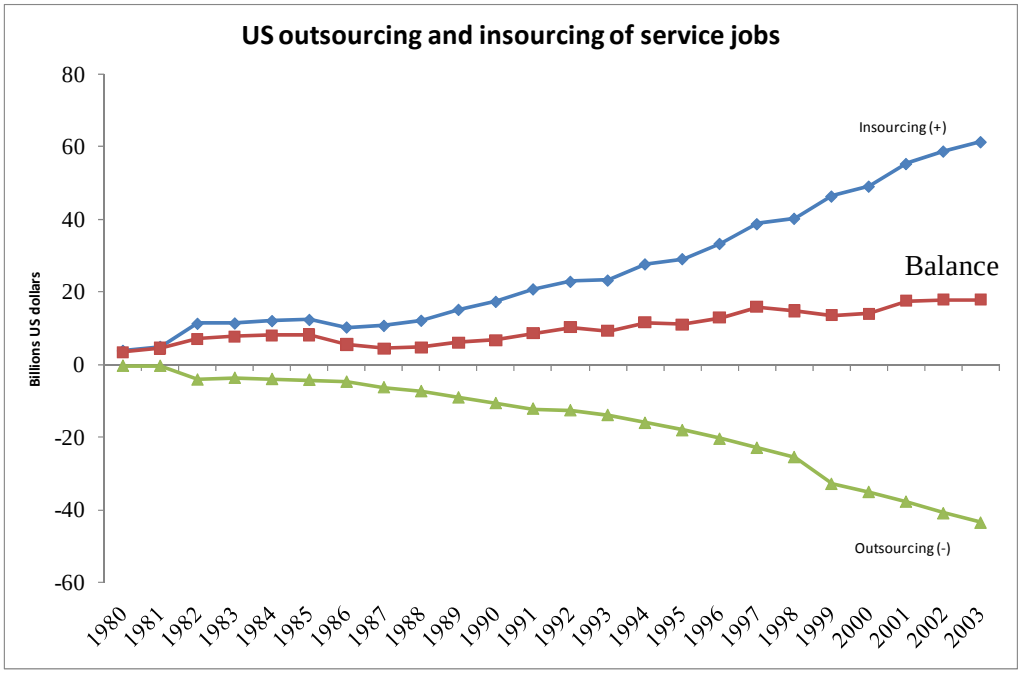


Figure 19: Blinders' blinders – service outsourcing is a two-way street.

Source: Author's manipulation of data from Amiti and Wei (2005), originally from IMF sources.

None of this should be unexpected. As North American manufactures have known for decades, falling trade costs create new opportunities and new challenges. To apply Blinder's logic to Canada's auto sector in the early 1960s would have been to claim that since US industry was so much more productivity than Canadian industry at the time, freer trade between the two would make all Canadian auto jobs "lose-able". When trade did open up, many jobs were lost, but many were created. As it turned out, micro-level factors allowed some Canadian firms to thrive while others floundered. Surely the same sort of thing will happen in services, as trade barriers in that sector fall with advancing ICT.

6.2 Future research and data gathering

One very useful step that governments could make would be to refine the data collection procedures. Government statistical collection procedures were set up to track the post-war industrial boom when jobs were associated with particular firms and particular firms were associated with particular sectors. Now, jobs are associated with particular tasks and tasks are increasing reallocated across firms (outsourcing) and across nations (offshoring), yet we do not have the data necessary to track the trends. One minor, but well known example of this is the way that unbundling of some service jobs has exaggerated the recorded decline of manufacturing jobs.

A second useful step would be to gather more detailed evidence on what caused the bundling of manufacturing stages in the first place. Was it the need for faces-to-face interactions, or the nature of the products? Until economist know a lot more about the 'glue' that produced bundling, they will not be very good at predicting future unbundling.

A third step would to gather much more detailed information on the unbundled supply chains, at least in key industries. Standard input-output tables are far to aggregate to capture the complexities of modern manufacturing.

REFERENCES

Amiti, M. and S.J. Wei (2005), “Fear of Service Outsourcing: Is it Justified?”, Economic Policy, 20, pp. 308-348.

Amiti, Mary and Shang-Jin Wei (2005). “Service Outsourcing: Is the Fear Justified?” Economic Policy,

Ando, M. and Kimura, F. (2005). The formation of international production and distribution networks in East Asia. In T. Ito and A. Rose (Eds.), International trade (NBER-East Asia seminar on economics, volume 14), Chicago: The University of Chicago Press. First version, NBER Working Paper 10167.

Baldwin, R. and P. Martin (1999), “Two waves of globalisation: Superficial similarities and fundamental differences,” Chapter 1, pp 3-59, in H.Siebert (ed.), Globalisation and Labour, J.C.B. Mohr for Kiel Institute of World Economics, Tubingen, 1999.

Baldwin, Richard (1988). "Hyteresis in Import Prices: The Beachhead Effect," American Economic Review, American Economic Association, vol. 78(4), pages 773-85.

Baldwin, Richard (2000) Regulatory Protectionism, Developing Nations and a Two-Tier World Trading System, Brookings Trade Forum, edited by Susan Collins and Dani Rodrik, The Brookings Institution, Washington DC, 237-293.

Baldwin, Richard (2006).“Managing the Noodle Bowl: The fragility of East Asian Regionalism”, CEPR DP 5561. hei.unige.ch/~baldwin/PapersBooks/ManagingNoodles_CEPR5561.pdf

Baldwin, Richard and Frederic Robert-Nicoud (2007). "Offshoring: General Equilibrium Effects on Wages, Production and Trade," NBER Working Papers 12991.

Baldwin, Richard and Krugman, Paul (2004). "Agglomeration, integration and tax harmonisation," European Economic Review, vol. 48(1), pages 1-23.

Blinder, Alan (2009). “How Many U.S. Jobs Might Be Offshorable,” World Economy, 2009, forthcoming.

Bordenave, Gérard and Yannick Lung (1996). “New Spatial Configurations in the European Automobile Industry,” European Urban and Regional Studies 1996; 3; 305-321.

Cooper, R. (1968). The economics of interdependence. New York: McGraw-Hill.

Dallas Fed (2002). “Maquiladora Industry: Past, Present and Future.” Federal Reserve Bank of Dallas, El Paso Branch, Issue 2.

Deardorff, A. V. (1998a). Fragmentation in Simple Trade Models, RSIE Discussion Paper 422, University of Michigan, January 8, 1998. www.spp.umich.edu/rsie/workingpapers/wp.html

Deardorff, A. V. (1998b). Fragmentation across cones, RSIE Discussion Paper 427, Discussion Paper No. 427, August 7, 1998. www.spp.umich.edu/rsie/workingpapers/wp.html

Debande, Olivier (2009). "De-industrialisation," EIB Papers 3/2006, European Investment Bank, Economic and Financial Studies.

Destler, I. M. (2003). “ US Trade Politics and Rules of Origin: Notes Toward A Paper,” Presentation to Workshop, 24 May 2003, Paris. http://www.publicpolicy.umd.edu/faculty/destler/ROA.pdf

Dixit, A. and G. M. Grossman, (1982). Trade and Protection with Multi-Stage Production, Review of Economic Studies 49(4), 583-594.

Feenstra, R.C. and G.H. Hanson (1996), “Globalization, Outsourcing, and Wage Inequality”, American Economic Review, 86, pp. 240-45.

Findlay, Ronald and Kevin O'Rourke (2009). Power and Plenty: Trade, War, and the World Economy in the Second Millennium, Princeton University Press.

Greenaway, D. and D. Nelson (2001). “Globalisation and Labour Markets: Literature Review and Synthesis,” Research Paper 2001/29, Nottingham.

Grossman, G. and E. Rossi-Hansberg (2006a). “The Rise of Offshoring: It’s Not Wine for Cloth Anymore,” July 2006. Paper presented at Kansas Fed’s Jackson Hole conference for Central Bankers. http://www.kc.frb.org/

Grossman, G. and E. Rossi-Hansberg (2006a). “The Rise of Offshoring: It’s Not Wine for Cloth Anymore,” July 2006. Paper presented at Kansas Fed’s Jackson Hole conference for Central Bankers. http://www.kc.frb.org/

Grossman, G. and E. Rossi-Hansberg (2006b). “Trading Tasks: A Simple Theory of Offshoring,” August 2006. PDF file. www.princeton.edu/~grossman/offshoring.pdf

Grossman, G. and E. Rossi-Hansberg (2006b). “Trading Tasks: A Simple Theory of Offshoring,” August 2006. PDF file. www.princeton.edu/~grossman/offshoring.pdf

Hamaguchi, Mobuaki (2008). “Making sense of the timeliness of transportation in economic integration,” in East Asia’s economic integration: Progress and benefit, D. Hiratsuka and Fukunari Kimura (eds), Palgrave-Macmillan, IDE-JETRO, London.

Hanson, Gordon H., R. Mataloni, M. Slaughter (2005). “Vertical Production Networks in Multinational Firms,” Review of Economics and Statistics, Vol. 87, No. 4, Pages 664-678.

Head, K., and T. Mayer (2004). “The Empirics of Agglomeration and Trade", in the Handbook of Regional and Urban Economics Vol. 4 edited by V. Henderson and J.F. Thisse, North-Holland.

Hummels, D. (2007). “Transportation Costs and International Trade in the Second Era of Globalization,” Journal of Economic Perspectives, Volume 21, Number 3, pp. 131-154.

Jones, R. (1965). The structure of simple general equilibrium models, Journal of Political Economy, 73(6), 557-572.

Jones, R. and H. Kierzkowski (2000). A Framework for Fragmentation, Tinbergen Institute Discussion Paper, TI 2000-056/2.

Jones, R. W. and H. Kierzkowski (1990). The Role of Services in Production and International Trade: A Theoretical Framework, in R. Jones and A. Krueger (eds.), The Political Economy of International Trade, Oxford: Basil Blackwell.

Jones, R. W. and H. Kierzkowski (1998). Globalization and the Consequences of International Fragmentation, forthcoming in R. Dornbusch, G. Calvo and M. Obsfeld (eds.), Money, Factor Mobility and Trade: The Festschrift in Honor of Robert A. Mundell, Cambridge, MA: MIT Press.

Jones, R. W. and S. Marjit (1992). International trade and endogenous production structures, in W. Neuefrind and R. Riezman (eds), Economic Theory and International Trade, Springer-Verlag.

Jones, R. W., H. Kierzkowski and G. Leonard (2002). Fragmentation and intra-industry trade, in P. Lloyd and H. Lee (eds), Frontiers of research in intra-industry trade, Palgrave Macmillian.

Krugman, Paul R., 1979. "Increasing returns, monopolistic competition, and international trade," Journal of International Economics, Elsevier, vol. 9(4), pages 469-479, Krugman, P. (1991), “Increasing Returns and Economic Geography”, Journal of Political Economy 99, 483-499.

Krugman, Paul, 1980. "Scale Economies, Product Differentiation, and the Pattern of Trade," American Economic Review, American Economic Association, vol. 70(5), pages 950-59, December.

Mann, Catherine and Xuepeng Liu (2009). “The Information Technology Agreement: Sui Generis or Model Stepping Stone?” in Multilateralising Regionalism, Richard Baldwin and Patrick Low (eds), Cambridge University Press, Cambridge.

Markusen, James (2006). Modeling the offshoring of white-collar services: from comparative advantage to the new theories of trade and fdi, in S. L. Brainard and S. Collins (eds.), Brookings Trade Forum 2005: Offshoring White-Collar Work, Washington: the Brookings Institution, 1-34.

Pritchett, Lant, 1997. "Divergence, Big Time," Journal of Economic Perspectives, American Economic Association, vol. 11(3), pages 3-17, Summer.

Puga, Diego (1997). ‘The rise and fall of regional inequalities’, European Economic Review 43(2), February 1999: 303–334.

Ravenhill, John (2008). Global Political Economy, Oxford University Press.

Samuelson, Paul (2004). “Where Ricardo and Mill Rebut and Confirm Arguments of Mainstream Economists Supporting Globalization.” Journal of Economic Perspectives. Summer, 18:3, pp. 135–46.

Sturgeon, T., Olga Memedovic, Johannes Van Biesebroeck and Gary Gereffi (2008). “Globalisation of the automotive industry: main features and trends,” Int. J. Technological Learning, Innovation and Development.

Venables, A. J. (1999). Fragmentation and multinational production, European Economic Review 43, 935-945.

WTO (2008), World Trade Report 2008, World Trade Organisation, Geneva.

Yi, K. (2003). “Can Vertical Specialization Explain the Growth of World Trade?,” Journal of Political Economy.