Regional Impacts of Trade and Investment On Labour

Morley Gunderson*
Department of Economics
University of Toronto
Toronto, ON M5S 1V6

To understand the regional or spatial impact of trade and investment on labour, it is first necessary to understand why firms locate where they do, how this will change as a result of the liberalisation of trade and investment, and how this in turn will affect labour market outcomes. This framework also provides a convenient way of incorporating some of the new developments that have implications for the regional impact of trade and investment, including agglomeration economies, endogenous growth, social capital, regional convergence, path dependence and border effects.

Although there are a variety of ways of categorising the determinants of firm location decisions, they are grouped here into the following inter-related categories:

- access to markets as affected by such factors as distance and transportation costs;
- access to labour, raw resources and other inputs;
- agglomeration effects, clustering and increasing returns;
- endogenous growth, convergence and divergence;
- social capital;
- border effects;
- historical accident and path dependence;
- tariff and other trade barriers, and
- taxes, subsidies and other regulations.

Each of these are expanded upon in turn, with particular emphasis on the implications for labour outcomes especially wages, employment, unemployment and wage inequality.

**Access to Markets**

Firms will obviously try to locate so that they have access to the product markets where they sell their output.\(^1\) This in turn depends upon such factors as distance and transportation costs (as well as agglomeration effects and increasing returns to scale as discussed subsequently). Historically, this placed a premium on such factors as ports, waterways and natural transportation corridors. Such "natural" factors have taken on less importance given the decline of heavy industry, and its replacement by "flexible" factories linked by advanced communications and just-in-time delivery. This is especially the case for many of the new products and services associated with the information age economy, where bulk and weight play a less important role. The increased emphasis on just-in-time delivery, however, has placed a premium on proximity to the market so as to minimise delivery time.\(^2\)

Trade liberalisation essentially expands the potential market. This has two main effects with respect to labour market outcomes. First, it expands the market beyond the conventional
domestic markets within the borders of a country, and into the new external markets of the trading partner. To reduce transportation costs, firms may relocate closer to the border and hence the market of their new trading partner. Second, the new markets may give rise to a different set of demands for products and services produced in different regions compared to the demands of the domestic market. In a global economy, firms may be given product mandates to produce particular products for the world market. Obviously those firms will expand, as will the regions where those firms are located.  

Trade liberalisation brings not only export expansion but also greater import competition, and such competition obviously can have impacts on labour that differ by region. Regions with firms and industries affected by import competition may decline, especially if they are unable to restructure to meet the expanding export opportunities.  

With respect to NAFTA, the greatest regional impact of expanded product markets is expected to occur for Mexico. Prior to trade liberalisation, manufacturing production was concentrated in Mexico City, producing for the surrounding hinterland within its national borders. Economies of scale and the advantages of forward and backward linkages were generally obtained by concentrating production in that locale, given transportation costs and distance to the surrounding and captive domestic market.  

Under NAFTA, the vast markets of the U.S. were opened up. Distance and transportation costs encouraged production to shift towards the U.S. border, to cities like Monterray, Ciudad Juarez and Tijuana, and away from Mexico City. This was facilitated by the earlier and parallel expansion of the maquiladoras or special exporting processing zones which essentially import parts and components from the U.S., assemble final products using low-wage labour in Mexico, and then export the finished products back to the U.S. The maquiladoras tend to locate near the U.S. border to minimise the distance and transport costs both for the importing of parts and components and for the exporting of final products. Again, this fosters the regional relocation of production from internal centres like Mexico City to Northern border regions. This relocation is facilitated by the fact that such assembly activity is essentially "footloose", not tied to raw resources (except for the raw components from the U.S.), or specific sources of specialised, immobile, skilled labour.  

For the U.S., the relocation of economic activity towards the Mexican border is expected to be less pronounced given the smaller size of the Mexican market relative to the U.S. market. Nevertheless, such relocation has occurred and it has given rise to considerable employment expansion in U.S. border cities (Hanson 1996b, 1997b). For this reason, it is not surprising that U.S. congressional districts near the Mexican border overwhelmingly voted in favour of NAFTA, while those furthest from the border voted overwhelmingly against the agreement (Hanson 1998c).  

For Canada, the various regions can be affected somewhat differently by the export expansion and import competition associated with trade liberalisation. However, there is not likely to be a substantial shift of production to the border because the locus of production already is clustered near the border. Southern Ontario, for example, is within one day of trucking distance to three-quarters of Canadian manufacturing and over half of U.S. manufacturing (Courchene and Telmer
1998). Cities like Toronto and its surrounding "Golden Horseshoe" are within 10 hours driving time from major U.S. markets like New York, Boston, Philadelphia, Baltimore, Washington, Pittsburg, Cincinnati, Cleveland, Detroit and Chicago. The industrial heartland of the U.S. encompasses the industrial heartland of much of Canada. As aptly stated by Krugman (1991d: 2-71) "Canada is essentially closer to the United States than it is to itself ... Industrial Ontario is aptly considered by geographers to be part of a common American manufacturing belt".

In spite of this proximity, there are pressures for regional reallocations and new regional alliances. As provocatively stated by Courchene and Telmer (1998: 289): "It is time to view Canada as a series of north-south, cross border economies, with quite distinct regional structures". Vancouver increasingly looks to the Asia-Pacific market and to the surrounding Cascadia area encompassing Washington, Oregon and Alaska (Goldberg and Levi 1993).

Alberta looks South to the U.S. market to export its oil and natural resources. Winnipeg increasingly regards itself as the potential Northern end of a transportation corridor through the U.S. via Interstate 29 to Mexico. Toronto and the industrial heartland of Ontario is re-orienting itself from its traditional East-West orientation to new North-South alliances. As stated by Courchene and Telmar (1998: 2), Ontario is transforming itself to the "premier economic region state within North America" -- note, they state North America, not just Canada. Montreal and other parts of Quebec are trying to revive Franco-American ties with New England, especially for the sale of its resources including hydro electric power (Konrad 1995), and to diversify if political separation occurs. The Atlantic provinces are also looking South, especially if they become isolated from the rest of Canada should Quebec separate.

**Access to Labour, Raw Resources and Other Inputs**

The regional location of firms and economic activity can be influenced by the market for inputs just as they can be affected by access to markets for outputs. Historically, this has meant that production activity especially of heavy industry was often tied to sources of raw material like iron and coal and to ports or transportation modes that could serve as access points for such resources. With the shift away from manufacturing and heavy industry and towards the service-based information economy, such ties have become less important. Business investment decisions with respect to physical capital of plant and equipment can be more "footloose" and based on factors other than access to raw resources. Labour and its embodied human capital become an increasingly important source of comparative advantage.

With respect to the labour input, opposing forces are at work in influencing location decisions. On the one hand firms can more easily locate in areas where labour costs are low. In this regard, labour costs are affected not only by wages and fringe benefits, but also by regulatory costs and productivity. While low productivity in low-wage locations often deterred investment decisions in the past, increasingly such low-wage locations are becoming high-productivity locations. This is the case, for example, with auto production in Mexico, where productivity is rapidly approaching levels that prevail in the U.S. and Canada.

On the other hand, the human capital aspect of labour has become more important, especially given the orientation to the information-age economy. The reasonably high levels of basic education in East Asia (financed in large part by refunneling their gains from export led growth
into their education infrastructure) have been important factors in their success. The human
capital concentrations (often associated with strong linkages to universities) in Silicon Valley,
Route 128 around Boston, and the Raleigh-Durham-Chapel Hill Research Triangle of North
Carolina, have contributed importantly to their success. High-performance, high value-added
workplaces are increasingly emphasized as necessary to survive in a high-wage economy.
Access to pools of skilled labour is an important factor in sustaining investment in larger urban
conglomerates that otherwise may be unattractive because of congestion externalities and high
costs of living.

Trade agreements like NAFTA will lead to even greater emphasis on the input of labour and its
embodied human capital. Multinationals are now more able to seek out areas of low labour costs
in relation to productivity, and to export from there to the higher labour cost countries given the
reductions in the tariff and non-tariff barriers to trade.\textsuperscript{11} They will also, however, seek to place
their high value-added activities and operations requiring extensive human capital into locations
with high-performance workplaces and large embodiments of human capital (Gunderson and
Verma 1994a).

Under NAFTA, the less skilled assembly type operations are more likely to be located in
Mexico, especially in the maquiladoras. Much of this, however, may simply be a diversion of
activity from low-wage Asian economies, especially given the proximity of Mexico to the U.S.\textsuperscript{12}
This is especially the case for the maquiladoras where proximity matters both for the importing
of parts and the exporting of finished products.

The higher value-added activities requiring skilled labour with large embodiments of human
capital are more likely to be located in the higher wage countries and regions within those
countries. This means, for example, that the research and development and financial and "head-
office" type functions are more likely to be done in countries like the U.S. and Canada.

In this regard, the U.S. is apt to have a greater location advantage for a number or reasons. It is
closer to Mexico in both physical distance and with respect to culture and language, especially in
border states. The proximity can be especially important for the exporting of parts for assembling
in the maquiladoras. The U.S. also has a considerable "head-start" with respect to many head
offices, and the advantage of agglomeration effects with respect to size (the importance of which
is discussed subsequently). The natural advantages of the U.S. as a hub would likely increase
further if NAFTA were extended South to include Latin American and Caribbean countries,
although Mexico could also benefit in this regard.\textsuperscript{13}

This raises the concern that Canada will mainly be the location for warehousing operations to
serve the Canadian market, with low-wage activities focussed in Mexico and high-skilled, high-
wage activities, such as research and development, concentrated in the U.S. hub. While a
possibility, it is also the case that Canadian operations can specialise and compete in the broader
market, as evidenced by such multinationals as Northern Telecom and Bombardier. The location
of their activities, however, is now much more "footloose", being able to readily shift in response
to changes in comparative advantages of different locations.
This reduced tie to inputs like natural resources, and the increased emphasis on the human capital aspect of the labour input has contributed to the increased wage inequality that has occurred since less-skilled, low-wage labour in countries like Canada and the U.S. is less in demand, while skilled labour that embodies more human capital is in greater demand. This exacerbates the growing wage inequality that is attributed in part (albeit likely in small part) to trade, as import competition into the higher-wage countries disproportionately affects low-wage labour in those countries, and export expansion helps higher-wage labour.\textsuperscript{14}

**Agglomeration Effects, Clustering and Increasing Returns**

The previous discussion referred to readily observed factors like access to output and input markets as important determinants of location and production decisions. Increasingly, however, less tangible factors have been emphasized, especially associated with interrelated concepts that have recently been developed, or at least emphasized, in economics.

Agglomeration effects that are external to the firm can occur whereby firms benefit by clustering around other firms. The other firms may be associated with backward linkages whereby they produce important inputs for the related firms, or forward linkages where they use the output of the related firms. The linkages reduce transportation and transaction costs associated with clustering together, although at some point they can be negative if congestion occurs or if they cannot diversify against common shocks.

The agglomeration effects of clustering together may also be associated with network or technological or other spillover externalities which are not "sold" or transacted in the market, but which nevertheless can affect the productivity of related firms and workers.\textsuperscript{15} Such may be the case, for example, with knowledge externalities often associated with high-tech developments or clustering around universities (Silicon Valley, Route 128 around Boston, and the North Carolina Research Triangle being obvious examples).\textsuperscript{16} In many cases, these beneficial effects are transacted through markets in which case any location advantage is "paid for". Ideas are patented in cases where intellectual property rights are well established. Universities can "sell their services" and industrial parks can charge higher rents for such location advantages. In other cases, however, it is extremely difficult to internalise such benefits through markets, especially the market for ideas and their implementation into practical applications.\textsuperscript{17} In such circumstances, mutually beneficial gains can be had by clustering related activities together -- "when you learn, I learn", and (.. hopefully) vice-versa.

Such clustering can be especially beneficial for job markets, for both employees and firms. This is especially the case since job matches are often made through informal contacts that are facilitated by proximity. Employees are often able to search for a better job match while working in their existing job. If laid off, their job search is facilitated if it occurs in a familiar labour market and does not entail family relocation. Employers are better able to recruit if there is a large pool of talent in the surrounding labour market. Both employees and employers may gain if there is a clustering of firms in the event of layoffs in any one firm.

Mobility of labour within a clustering may also facilitate the transmission and expansion of knowledge, as occurs for example with the constant "churning" of high-tech employees in
Silicon Valley. To the extent that the "knowledge enhancing" effect is greater than the "knowledge inhibiting" effect created by the loss of intellectual property rights associated with the high turnover of employees, then economic activity is enhanced by such clustering and employee mobility.\textsuperscript{18}

Increasing returns to scale that are internal to firms may also influence location activities as when, for example, firms consolidate their activities in one location to take advantage of the economies of scale by spreading the high fixed costs over larger output. Such economies of scale may be associated with particular plants being given world product mandates to produce a specific product for the global market.

The interaction of internal economies of scale, external agglomeration effects and transportation costs can have complicated interactive effects on location decisions.\textsuperscript{19} Reductions in transportation costs can facilitate the dispersion of production to low-cost regions because they can now more easily transport their products to the market. The low-cost region may be the peripheral area with, say, low labour and land costs, but it may also be the core area with its agglomeration economies and economies of scale even though it may have high labour and land costs. In essence, reductions in transportation costs, including those engendered by eliminating borders, can lead to a dispersal of economic activity to the peripheral regions, especially if they are closest to the new market, but it could also lead to greater concentration in the core regions, if external agglomeration and clustering effects and internal economies of scale are sufficiently large.

The relationship may also differ according to the stage of integration. At the initial stages when transportation costs are still high, activity may remain in the peripheral region or shift there to take advantage of its low labour costs. In the intermediate stage of integration, activity may shift to the core, hub region to take advantage of the internal economies of scale and external agglomeration economies. This shift may be abrupt since these effects may apply only at critical, threshold levels. Then if scale and agglomeration effects are exhausted or become negative as in the case of congestion, and wages and labour and housing costs rise in the core, then activity may disperse more to the periphery. This slow dispersal of the benefits of integration to the peripheral regions lead Puga and Venables (1997: 365) to the important policy recommendation that "a firm and credible commitment to full integration may convince peripheral regions to put up with harder times during the intermediate stages of trade liberalisation. So will transfers from the core regions, which redistribute the aggregate gains that free trade achieves during the whole integration process".

Agglomeration effects from knowledge spillovers may differ according to the typical life-cycle of an industry. They may be especially important in the early stages of innovation and commercial development when knowledge spillovers are important. They may dissipate in importance as the industry matures from innovation and development to commercial production. Agglomeration effects, however, may also enable a regional economy to "reset its innovation cycle and overcome the locational entropy of product life cycle maturity" (Gilson 1998: 15).
Endogenous Growth, Convergence and Divergence

Agglomeration effects and increasing returns underlie the concept of endogenous growth whereby diminishing returns are offset by such positive spillovers associated with agglomeration, clustering and growth itself. In essence, growth begets further growth rather than being subject to decreasing returns.

Such forces may postpone or offset the tendency towards convergence whereby growth rates would converge as the less developed regions experience increasing returns and the technology transfer from the more developed regions which in turn experience decreasing returns and the emulation of their successful technology. Indeed, increasing divergence may occur if the positive agglomeration effects and increasing returns are sustained and attract more capital, both physical and human capital -- "nothing succeeds like success" and "the rich become richer and the poor poorer". Mobility of labour can enhance such divergence through "brain drains" as the "leavers" embody considerable human capital (as well, perhaps, of unobserved characteristics such as motivation and initiative), and the "stayers" are more likely to be unskilled and unemployed. Transfer payments in the form of income support for the "stayers" can further sustain that divergence of development.

In such circumstances, trade liberalisation and increased capital flows could foster convergence or divergence, depending upon the circumstances. It may foster convergence, if the low-cost regions are able to take advantage of the new opportunities and use their low-costs as a source of comparative advantage, and if capital is attracted to such regions. In contrast, if the positive agglomeration and spillover effects dominate, then further divergence can occur.

As indicated previously, this can give rise to a U-shaped relationship between regional convergence and trade liberalisation: convergence in the initial stages as economic activity is immediately attracted to the poorer community to take advantage of their low labour costs; divergence in the intermediate stages as economic activity moves to the more developed region to take advantage of agglomeration and scale effects; and convergence in the later stages as economic activity moves from the congested, high-cost regions back to the less developed regions.

Social Capital

Benefits from clustering of activities may also be associated with the social capital that may exist in such clustering. Social capital is generally associated with intangibles like informal networks and safety nets, family and community connections, common value systems and norms, and information spillovers and other neighbourhood effects. Unlike physical and human capital that is produced and appropriated by firms and individuals, social capital is generally the byproduct of other activities and it "resides" in the community. While it is generally regarded as producing positive effects on members of the community, there can be a darker side as for example with "old-boy networks", or social ostracism and intolerance, or discrimination against "outsiders", or self-perpetuating ghettos. The location of economic activity obviously may be influenced by such social capital since it can have very real effects on economic activity, even though social capital itself is often intangible.
While social capital is often an unintended byproduct of other activities, it can be influenced -- indeed possibly "engineered" -- by public actions. The rise of the secondary school system in the U.S. in the period 1910 to 1940 is attributed in part to the desire of older generations in rural communities to educate their youth as a form of inter-generational social insurance (Goldin and Katz forthcoming). Creating "neighbourhoods" in otherwise large amorphous cities can be a conscious policy to reduce crime and foster community involvement.

Clearly, less tangible factors such as agglomeration and clustering effects, economies of scale, and social capital can have important impacts on the regional location of economic activity. What is less clear, however, is how these effects change as a result of trade liberalisation such as NAFTA.

Such trade liberalisation can lead to new clusters of economic activity, with their forward and backward linkages and agglomeration effects. This was discussed previously with respect to the relocation of economic activity from Mexico City north towards the U.S. border, including the maquiladoras with their obvious forward and backward linkages. This did not have to occur, however, if the economies of scale and agglomeration could have outweighed the reduction of transportation costs. The relocation towards the border, however, suggests that the economies of scale and agglomeration were not that great, or they were offset by the negative congestion and other pollution effects associated with concentration in Mexico City, or that transportation costs to the border were substantial given the distance to Mexico City and the absence of a well-developed transportation infrastructure.24

Such a reallocation of economic activity away from core regions and cities is unlikely to occur in Canada given that most activity was already clustered along the border, and there is a reasonably well-developed transportation structure integrated with the U.S. transportation network. As well, positive agglomeration and scale economies in cities like Toronto and its surrounding Golden Horseshoe are less likely to be offset by negative congestion and pollution externalities as was the case with Mexico City.

Nevertheless, as discussed previously, new North-South regional alignments may occur. There is some concern that the clustering may take the form of low-wage assembling being done in Mexico, high-wage, high-value added jobs including research and development clustering in the U.S., and warehousing operations being relegated to Canada. This later concern, however, should be mitigated by the fact that Canada should be able to develop world product mandates in particular items.

The "production" of social capital may be affected in an indeterminate fashion as the result of increased trade. It may dissipate somewhat to the extent that internal networks, contacts and value systems are displaced by external networks and contacts, and as internal homogeneity is displaced by greater heterogeneity. The concern that capital has no "heart" and multinationals have no specific country or community loyalty mirrors this concern.

In contrast, the production of social capital may be enhanced by the greater exposure to other norms and values that occurs when countries trade more with one another. It may also be enhanced by the information-sharing networks that are established by internationally dispersed
ethnic groups (Rauch and Casella 1998). Employee loyalty and commitment to the organisation have been "imported" into North America, in part because of the perceived success of these intangible factors in Japanese workplaces. The same applies to employee participation schemes - a workplace practice that can be regarded as an attempt to generate social capital at the workplace.

**Border Effects**

"Border effects", which can be influential in the location of economic activity, may also reflect the social capital that is internal to specific borders. Empirical evidence indicates that inter-provincial trade flows between Canadian provinces are approximately 20 times larger than are trade flows between Canadian provinces and U.S. states, even after adjusting for differences in distance and size of the markets. This magnitude is much larger than that which would be generated by tariffs between countries. In essence, even after adjusting for differences in the factors that should affect trade flows, Canadians seem to prefer domestic internal trade within their own borders compared to external trade across borders with the U.S. Helliwell (1998) cites evidence indicating that such internal preferences for domestic transactions within borders also applies to portfolio investments, in spite of the fact that diversification against risk can be had through international investments. Helliwell (1998: 71) concludes: "For securities, just as for goods, the perceived advantages of dealing within familiar and trusted networks, institutions, and markets are greater than the potential gains from further international diversification". The importance of familiarity and networks is further illustrated by the fact that immigration enhances trade with the country of origin of the immigrants in Canada (Head and Reis 1998).

Pure border effects should dissipate under trade liberalisation. In essence, trade liberalisation is likely to lead to a reallocation of trade from within the national borders of countries to across the borders with the new trading partners, in response to the new market opportunities. Differential effects by regions are likely to occur reflecting the different comparative advantages of the different regions.

Helliwell (1998) does find a large reduction in the border effects after the Canada-U.S. Free Trade Agreement of 1989. Specifically, internal trade flows amongst provinces within Canada were reduced from being 20 times more prominent than external flows to the U.S. in 1990, to 12 times more prominent by 1993, after the agreement. The impact seems to have been one-time since the ratio levelled of at around 12 after 1993. While the agreement appears to have resulted in a large reallocation from internal (East-West) trade, to external (North-South) trade, as Helliwell emphasizes, the preference for internal trade remains strong, even amongst equally situated trading partners (i.e., after controlling for differences in distance and the size of the market amongst provinces and states).

**TABLE 1 Ratio Of Internal (Inter-provincial) Trade Relative To External (Provincial-state Trade), 1988-1996**

<table>
<thead>
<tr>
<th>Year</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 1 Ratio Of Internal (Inter-provincial) Trade Relative To External (Provincial-state Trade), 1988-1996**
<table>
<thead>
<tr>
<th>Province</th>
<th>Pre-FTA&lt;sup&gt;2&lt;/sup&gt; 1988-1990</th>
<th>Post FTA&lt;sup&gt;2&lt;/sup&gt; 1991-1996</th>
<th>% Change&lt;sup&gt;3&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>17.7</td>
<td>13.6</td>
<td>-23</td>
</tr>
<tr>
<td>Newfoundland</td>
<td>9.8</td>
<td>10.6</td>
<td>8</td>
</tr>
<tr>
<td>New Brunswick</td>
<td>15.3</td>
<td>12.2</td>
<td>-20</td>
</tr>
<tr>
<td>Nova Scotia</td>
<td>29.1</td>
<td>21.9</td>
<td>-25</td>
</tr>
<tr>
<td>Prince Edward Island</td>
<td>18.8</td>
<td>20.6</td>
<td>10</td>
</tr>
<tr>
<td>Quebec</td>
<td>17.4</td>
<td>11.8</td>
<td>-32</td>
</tr>
<tr>
<td>Ontario</td>
<td>24.3</td>
<td>16.3</td>
<td>-33</td>
</tr>
<tr>
<td>Manitoba</td>
<td>15.9</td>
<td>11.1</td>
<td>-30</td>
</tr>
<tr>
<td>Saskatchewan</td>
<td>15.7</td>
<td>10.0</td>
<td>-36</td>
</tr>
<tr>
<td>Alberta</td>
<td>8.7</td>
<td>7.0</td>
<td>-20</td>
</tr>
<tr>
<td>British Columbia</td>
<td>9.4</td>
<td>6.4</td>
<td>-32</td>
</tr>
</tbody>
</table>

Notes: 1. After controlling for differences in distance and size of markets.

2. Even though the FTA came into effect in 1989, the period 1998-90 is labelled here as pre-FTA and the period 1991-96 as post-FTA because its effect would not likely show up for at least a year. Helliwell (1998, p. 3) refers to the post-1990 period as reflecting the increased bilateral trade between Canada and the US.

3. Post minus Pre divided by Pre times 100.

Source: Calculated as the simple average of the border effects for each of the years given in Table 2-3 in Helliwell (1998).

Table 1, based on the estimates of Helliwell (1998), summarises the importance of internal trade relative to external trade for the different regions of Canada, illustrating how that relative importance changed after the FTA. The large positive numbers in both the pre- and post-FTA period indicates that all provinces relied much more on internal inter-provincial trade compared to external trade with U.S. states. The reliance in internal trade with other provinces compared to the U.S. was particularly high (i.e., above the national average) in Nova Scotia, Ontario and Prince Edward Island. Reliance on internal trade relative to external trade was below the national average in all other provinces, but especially in Alberta and British Columbia.

The last column of Table 1 indicates the substantial restructuring away from internal, East-West inter-provincial trade towards external North-South U.S. trade after the FTA, with a negative entry indicating a reduction in the ratio of internal compared to external trade. That restructuring from East-West to North-South occurred in all provinces except Newfoundland and Prince Edward Island, which became slightly more reliant on internal inter-provincial trade. There was considerable variation in the restructuring, however, being particularly dramatic in Ontario, Quebec, British Columbia and Saskatchewan. The post-FTA period is noticeably characterised by the four Western provinces being most integrated with the U.S. in the sense of having the lowest ratio of East-West trade relative to North-South trade with the U.S. Most of the restructuring seems to have gone on over a three-year period from 1991 to 1993, having
plateaued since that time. While the FTA appears to have led to a substantial restructuring from East-West inter-provincial trade to North-South trade with the U.S., the fact remains that inter-provincial trade is still relatively much more important in Canada.

Historical Accident and Path Dependence

Considerable anecdotal evidence suggests that the location of economic activity may be influenced substantially by historical accident. Once that initial location decision is made, however, it establishes a path dependence that is self-perpetuating and difficult to dislodge. The path dependence can arise because of such factors as the forward and backward linkages and the agglomeration and scale effects as previously discussed. In essence, the subsequent time path of the economic activity becomes influenced by economic considerations, albeit the initial conditions were established more by historical accident.28

Of course, historical accidents likely would not flourish into a self-perpetuating chain of subsequent activity if basic conditions sustaining that activity are not present. For every historical accident that leads to a success story there are likely hundreds that do not -- and only the success stories are around to be told. There is invariably a selection bias in the reporting of successes compared to failures.

Nevertheless, historical accident, often associated with a single individual, can be an important determinant of initial location decisions, and such initial conditions can set up a self-perpetuating path dependence. Krugman (1991d) colourfully describes many such cases, often arising from the acts of a single individual: Catherine Evans and the carpet industry around Dalton Georgia; John Adams Dagyr and the shoe industry in Massachusetts; a local individual who invented "filled gold" and the jewellery industry in Providence Rhode Island. Modern high-tech examples also exist: Fred Terman, the vice-president of Stanford, and his coalition with Hewlett-Packard and the provision of university land for a research park that subsequently mushroomed into Silicon Valley; Karl Compton, the president of MIT, and his encouragement of MIT faculty to form entrepreneurial links with private venture capital, which facilitated Route 128 around Boston. These later examples are not so much historical accident, but rather are conscious efforts of visionary individuals to establish a mutually beneficial path dependence, in this case between universities and the high-tech private sector. In that vein, the Research Triangle around Chapel Hill, Raleigh and Durham North Carolina was consciously established with state funds in direct emulation of these earlier success stories.

While these initial events spawned a self-perpetuating clustering of economic activity, such clusterings can dissipate. Path dependence need not be forever. As Krugman (1998) points out, Akron Ohio, once the tire-capital of the world, had over one hundred tire manufacturing firms; it now does not have a single major producer.

Trade liberalisation can affect economic activity established through historical accident and path dependence, albeit not in a unidirectional fashion. Such activity is economically viable in its current location prior to trade liberalisation, even if economic factors may not have played a crucial role in its initially being established. To the extent that its economic viability depends upon the agglomeration effects and scale economies that fostered the path dependence, then such
effects are likely to be further fostered by the expanded market associated with trade liberalisation. The clusters are now serving an even larger market. It is possible, however, that they could be challenged by new clusters that could be better situated to serve the new markets, and it is possible that they could be challenged by clusters in the new markets themselves. Silicon Valley and high-tech clusters around Austin Texas may be better situated to serve the Mexican market then are the more northern high-tech clusters. Northern Mexico could become a nucleus for automobile assembly and parts production, especially if trade agreements expand south to Latin America. In this vein, Canada seems to be more marginally situated, albeit obviously linked to the Northern U.S.

As indicated previously, when such changes in clustering occur they tend to be abrupt, often involving a complete shift or tipping to a new location, given that the agglomeration spillovers and scale economies may only be able to sustain one location. Multiple equilibria, however, are possible if the market can sustain two or more clusters.

**Tariff and Other Barriers**

The geographic location of economic activity obviously can be affected by tariffs and other barriers to trade. In Canada, for example, much of the East-West orientation of trade resulted from the tariffs established under the National Economic Policy of 1878. Such a policy redirected trade from its more natural North-South orientation based on the large U.S. markets. The higher transportation costs of redirecting trade in an East-West direction was offset in part by the construction of a subsidised railroad, also as part of the National Economic Policy. These twin policies led to a proliferation of small scale firms, often branch plants of U.S. multinationals, set up in Canada to "jump the tariff walls". If the multinationals could not export to Canada, they would produce in Canada, even at an inefficiently small scale.

Obviously, the reduction in tariffs and other non-tariff barriers to trade should lead to a reallocation of trade from its more artificially sustained East-West orientation to its more natural North-South orientation. And this has occurred in response to the FTA as discussed previously, although the inter-provincial trade still remains substantially more important than provincial-state trade with the U.S. As well, consolidation of the inefficiently small branch plants should occur to obtain the economies of scale necessary to compete in the larger market. These phenomenon likely account for much of the industrial restructuring, and associated labour adjustment, that has occurred in recent years.

**Taxes, Subsidies and Other Regulations**

The location of economic activity can also be affected by taxes, subsidies and other regulations. High taxes and extensive regulations can repel business, while subsidies in various form can attract business. As pointed out by Krugman (1991d) such "boosterism" was important in such diverse ventures as establishing the rubber industry in Akron Ohio, and the Research Triangle of North Carolina.

The actions can be more indirect, as was the case with the so-called "Michilin Bill" of 1979, when the government of Nova Scotia hastily modified its labour law at the same time as it
wanted to attract Michelin tire plants. The modification essentially required that a union must organise all of the employees of all inter-dependent plants of the same company at the same time and be certified as a single bargaining unit -- a requirement that made certification extremely difficult. The regionally extended benefits under the former unemployment insurance system of Canada likely deterred the outflow of capital, and certainly labour, out of high unemployment regions like the Atlantic provinces. Political jurisdictions can also try to attract investment by appearing to be more "open for business" and providing an economic climate that is conducive to investment. They may do this, for example, by trying to reduce the regulations and red tape that is involved in conducting business in a province, as is the case with Ontario's recent Red Tape Commission (Sheehan 1997).

Trade agreements like the FTA and NAFTA are likely to have a complicated set of effects on such regulations and hence on their impact on the location of economic activity. Changes may come from regulations within the trade agreements, such as the prohibitions on energy price fixing policies like those contained in the National Energy Policy, which provided artificially cheap energy to Central Canada, at the expense of the energy producing Western provinces (Whalley 1987). Regional development policies could be under pressure to dissipate to the extent that they contained elements that could be interpreted as unfair subsidies. Free trade will also put pressure on provincial governments to harmonise their internal standards and policies so as to facilitate internal competitiveness. In the labour area, such policies include professional occupational licensing, government licensing of trades, preferential hiring, income security programs, employment standards, education standards and language, all of which can have elements that impede the efficient internal allocation of labour, especially by providing barriers to internal labour mobility. The recently negotiated Internal Free Trade Agreement of 1994 was a step in the direction of eliminating barriers to the flow of goods and services across the country. The intent was to "encourage firms to grow and expand their capabilities, and this in turn will increase their ability to tackle international markets" (Daly 1994: 10).

In general, trade liberalisation should lead to the dissipation of regulations and taxes on business. This is so, because firms are now more able to relocate to countries and jurisdictions within countries that minimise such costly regulations. Firms can now relocate and export back into the higher cost countries, given the reductions in the tariff and non-tariff barriers to trade. In these circumstances, political jurisdictions are more likely to compete for such investment and the jobs associated with that investment by reducing their regulations and business taxes -- the "imploding" of the tax base, to paraphrase Ross Perot (Orme 1996). The pressure is to harmonise such regulations, with the possibility that the harmonisation will be downward "to the lowest common denominator". The negative side of this process is referred to as the "social dumping" or "regulatory meltdown" that will lead to a "race to the bottom" in terms of social policies and legislative and regulatory initiatives. The concern is that the maquiladoras of Mexico and the "right-to-work" states of the U.S. South, may be the new greenfield sites for Canada. The concern is exacerbated by the fact that Canada tends to have a fairly extensive degree of regulation in various dimensions of the labour area: laws that govern the certification and rights of unions; labour standards pertaining to such factors as minimum wages, hours of work and overtime, maternity leave, and terminations; pay and employment equity legislation; and health and safety, workers' compensation and reasonable accommodation requirements for disabled persons.
While downward harmonisation can occur, it need not occur, and it need not have the negative connotations with which it is typically associated. Regulations and laws can have benefits as well as costs, and those that have positive effects on efficiency and competitiveness will not dissipate -- indeed they will thrive. Workers' compensation, for example, saves on the costs of the tort liability system since workers effectively give up the right to sue their employer in return for such "no-fault" insurance. Legislation requiring advance notice in the case of plant closings or mass layoffs may facilitate new job matches that can benefit employers as well as employees. Social safety nets and labour adjustment assistance can reduce worker resistance to otherwise efficient changes, including trade liberalisation. Taxes can be effectively used for public infrastructures that can enhance a country's competitiveness. There is general recognition that one of the preconditions for investment and growth is well established property rights and enforceable contracts -- preconditions that generally require public resources to establish and enforce. Canada's public health care system may well be more cost effective than the private system of the U.S.35

In such circumstances, the threat of interjurisdictional competition for investment simply compels jurisdictions to pay more attention to the cost consequences of their regulatory and legislative decisions. It need not lead to the dissipation of such regulations -- except for those for which the costs are excessive relative to the benefits. In essence, such interjurisdictional competition puts pressure on governments to deliver their services in a more efficient and cost effective fashion.

In fact, with some notable exceptions, the regulations that are most at risk of dissipation under trade liberalisation are those that do not have positive feedback effects on efficiency and competitiveness, but that are socially wasteful and protect the "rents" and special positions of particular interest groups. An important exception pertains to pure distributional or equity oriented policies designed to help the most disadvantaged. If these do not have positive feedback effects on efficiency (e.g., by providing a safety net that reduces resistance to otherwise efficient changes) then such programs will be difficult to sustain. The problem is exacerbated by the fact that equity oriented policies may be under pressure to dissipate just when there is more need for such policies to deal with the adjustment consequences and growing wage inequality that is associated with trade liberalisation and a growing emphasis on market forces.

It may be harder, for example, for a country like Canada to sustain high taxes to pay for purely redistributive programs to provide a more egalitarian society if business can now more easily move to other countries and export back into Canada. It may be harder for a particular provincial government to provide more liberal social services if it is no longer necessary to have a branch plant in every province, and the location decision of the new conglomerate is influenced by taxes or regulatory costs. It may be harder for the federal government to pay for regional development policies if their tax base is threatened by the potential exodus of capital, including human capital through the brain drain if it is highly taxed.

The problem is exacerbated even further by the fact that if it is not possible to tax the mobile factors of production (e.g., capital and professional labour) then there may be more pressure to tax the immobile factors of production -- labour and especially unskilled labour. This can further
exacerbate the inequality that is already enhanced by such forces as technological change and trade liberalisation.

In essence, the efficiency gains from trade liberalisation enhance the size of the economic pie, but they do not guarantee that the pie will be more equitably distributed. In fact, as indicated previously, trade liberalisation has contributed to the growing wage inequality that is exacerbated even more by such forces as technological change. Furthermore, trade liberalisation and the associated interjurisdictional competition for investment and the jobs associated with that investment, somewhat ties the hands of governments to deal with the pure distributional issues unless they have positive feedback effects on efficiency and competitiveness. It may be cheaper to pay for a social safety net than prisons, but if not, then sustaining redistributive policy initiatives will be difficult. This is likely to be the real policy challenge associated with trade liberalisation.

Unfortunately, we have little empirical evidence on the extent to which costly social policies and regulations deter investment, or the extent to which positive subsidies and incentives attract investment, or the extent to which trade and openness in turn inhibit the ability of a country to develop and sustain such social programs.

The policy challenge is exacerbated by the fact that such redistributive, equity issues are best dealt with at more aggregate levels of government so as to minimise escaping the taxes and regulations. Yet most pressures are for decentralisation, and for dissipation of the nation state (Ohmae 1996), although there is also some pressure for global initiatives -- a bifurcation into both global and local initiatives termed "glocalisation" (Courchene 1995). For NAFTA, however, supranational institutions, at least at this stage, will not likely be a substitute for the loss of a national presence since the parties are not likely to hand over tax or regulatory powers to a supranational regulatory body. Requiring poorer countries to raise their regulations to the levels of those of the richer countries as a precondition for entering into a free trade agreement is likely to simply be thinly veiled protectionism.

Solutions will not be easy, and they require information on such factors as the impact of trade on labour market inequality, and the extent to which pressure towards the harmonisation of policy initiatives will have more of an impact on "good" regulations or "bad" regulations. It is also important to have information on the impact of pure equity oriented initiatives as well as the effect of changes in such initiatives on society at large.

Conclusion

Trade liberalisation, such as that which is embodied in agreements like the FTA and NAFTA, can have important impacts on the spatial allocation of economic activity across regions within a country. This, in turn, can have important impacts on the allocation of labour and the associated labour market outcomes.

In general, agreements like NAFTA will shift the locus of economic activity towards the border regions to take advantage of the lower transportation and communication costs associated with the new market opportunities. However, if the internal economies of scale and positive effects of
agglomeration, clustering, social capital and history and path dependence are sufficiently large (and cannot be replicated in the border regions) then the former core of economic activity could become even more dominant since the new larger market enables them to take even more advantage of their positive internal economies of scale and external economies of agglomeration and clustering. The end result depends upon the relative importance of such factors as transportation costs, proximity to markets for outputs and inputs, and agglomeration and clustering effects.

After NAFTA, such reallocation in Mexico has occurred from the interior around Mexico City and towards the border with the U.S. Apparently, the "pull" of the border is greater than the retention power of the positive scale economies and agglomeration effects that were associated with the former clustering of economic activity around Mexico City, especially since the later were already being offset by negative congestion and pollution externalities.

In the U.S., such reallocation of economic activity has also occurred towards the Southern border regions, although the relative effect has not been as large given the large size of the internal U.S. economy relative to the Mexican economy.

In Canada, economic activity was already clustered near the border, and while internal trade remains dominant, there has been a substantial shift from the traditional East-West internal inter-provincial trade towards North-South external trade with the U.S. This re-orientation has occurred across most regions within Canada, although the effect is most pronounced in Ontario, Quebec, Saskatchewan and British Columbia. The post-FTA period is noticeably characterised by the greatest integration with the U.S. occurring in the four Western provinces.

The fact that economic activity in Canada did not have to reallocate towards the border is fortuitous since it likely reduced the adjustment consequences that otherwise would have occurred. Substantial restructuring was already emanating from a wide range of inter-related forces including technological change, globalisation, international competition, the recession of the early 1990s, and the shift from manufacturing to services and to the knowledge-based economy. Adding a geographic dimension associated with a reallocation towards the border would have exacerbated those pressures. The industrial heartland of Ontario, being already near the U.S. border, was able to absorb the loss of its privileged East-West trade that was built up under the protective tariff, and reallocate towards the newly available North-South trade opportunities. It was able to do so because of its proximity to the border, its internal economies of scale and its external economies of agglomeration, clustering and social capital that enables it to compete in the larger market south of the border.

In addition to encouraging the restructuring away from East-West and towards North-South trade, the FTA and NAFTA will likely have an indirect effect on regional development through its effects on regulations and policy initiatives. Stronger pressure now exists to internally harmonise regulations and policy initiatives so as to enhance internal competitiveness as a precondition for external competitiveness. Indeed, this is already being manifest in such forms as the Internal Free Trade Agreement of 1994. The impetus also exists, especially under the threat of capital mobility, for external harmonisation of regulations and policies with those of the trading partners.
Contrary to general perceptions, however, such inter-jurisdictional competition for investment and the jobs associated with that investment, need not always be negative. Regulations that have positive effects on efficiency and competitiveness will thrive rather than dissipate, with inefficient rent seeking regulations being under the most pressure to dissipate. The important exception is pure equity oriented distributional policies that will be harder to sustain, at the very point in time when they will be needed to deal with the adjustment consequences of trade liberalisation and other inter-related pressures. This is the key policy challenge that is arising under trade liberalisation.

References


Daly, J. 1994. "Open for Business". *Challenges*, Fall.


