Rising trade costs?

Agglomeration and trade with endogenous transaction costs

Gilles Duranton[*,‡]
University of Toronto

Michael Storper[*,†]
London School of Economics

29 March 2006

ABSTRACT: While *transport* costs have fallen, the empirical evidence also points at rising total *trade* costs. In a model of industry location with endogenous transaction costs that seeks to replicate features from the machinery industry, we show how and under which conditions a decline in transport costs can lead to an increase in the total cost of trade. The subtle relationship between (endogenous) transport costs and the sensitivity of trade to distance is also explored.

Key words: Transaction costs, trade costs, transport costs, agglomeration, vertically linked industries.
JEL classification: D23, D24, R12.

[*]We have benefited from comments and discussions with Pol Antràs, Keith Head, Vernon Henderson, Maureen Kilkenny, Niko Matouschek, Thierry Mayer, Shin-Kun Peng, Diego Puga, Frédéric Robert-Nicoud, Jacques Thisse, Tony Venables, and seminar and conference participants in Cagliari, Marseille, and Seattle.

[‡]Department of Economics, University of Toronto, 150 Saint George Street, Toronto, Ontario, Canada M5S 3G7, gilles.duranton@utoronto.ca, http://individual.utoronto.ca/gilles/default.html. Also affiliated with the Centre for Economic Policy Research. Financial support from the Leverhulme Trust is gratefully acknowledged.

[†]Department of Geography, Houghton Street, London WC2A 2AE, United Kingdom, m.storper@lse.ac.uk. Also affiliated with the Institut d'Etudes Politiques de Paris and UCLA.

1. Introduction

The cost of shipping goods from one place to another has declined relentlessly over at least the last hundred years. Glaeser and Kohlhase (2003) document that the share of US GDP in transportation industries has been halved over the 20th century. They also show that the real dollar cost per ton-mile for US railroad shipping has decreased nearly tenfold over a century. French data exhibit a similar trend: Combes and Lafourcade (2005) document a decline of nearly 40% in the cost of road freight over 1978-1998. There is thus very little doubt that unit internal transport costs have declined very significantly. The evidence on international transport costs is arguably more difficult to assess (Anderson and van Wincoop, 2004). Hummels (1999) documents a dramatic fall in air transport costs since at least the early 1970s and a less pronounced decline in sea transport costs since 1985 (where containerisation, which improved shipping quality and efficiency, led to higher prices between 1970 and 1985). This steep decline in transport costs has led many analysts to predict the 'death of distance' (e.g., Cairncross, 2001).

But distance is not dead — not even sick. Casual evidence abounds. More systematic evidence is provided by Hillberry and Hummels (2005) who show that in the US, shipments within a six kilometre radius are three times larger than shipments beyond this radius. How can distance remain important, possibly become even more important, when transport costs decreases? As a way into this problem, let us build on the example of distance and trade costs in an industry for which the facts are well documented by Gertler (2004): the machinery industry. In this industry, the design, development, customisation, installation, start-up, servicing, and updating of complex pieces of equipment require frequent interactions between producers and end users. This is all the more so because standardised machines are increasingly being replaced by custom-made equipment, so that users must be trained to use such equipment. All these operations are time-consuming and involve numerous trips back and forth between the producer and the end user.

Purchasing such complex pieces of machinery is also fraught with difficulties because it is usually impossible to contract ex-ante on the future quality of custom-made machines. Indeed, Gertler's (2004) main argument is that the (relative) industrial decline of the US Midwest and Canada is to a large extent the consequence of the unsuccessful adoption of flexible manufacturing in the late 1980s and early 1990s. Having to modernise their equipment, manufacturing firms in the US and Canada faced a scarcity of competitive local machinery producers. Instead, German pro-

ducers were perceived to offer superior custom-made products. However, buying machinery from German producers turned out to be very difficult because of all the transaction costs mentioned above. Often, US and Canadian manufacturers ended up with machines that were well below what they expected even after incurring very large costs of installation, training, and servicing. German machine producers also ended up dissatisfied because, despite their efforts, their machines had much worse performance in North America than at home with German manufacturers. Beneath all this is the strong suggestion that the cost of trading machines across the Atlantic has increased over time rather than decreased.

Our goal in this paper is to provide a micro-founded model of endogenous trade costs able to replicate this type of development. To sharpen our intuition and to provide a guide in our choice of assumptions, we remain with the machinery industry that we model in its interactions with vertically linked industries. In the simplest version of the model, under decreasing returns final producers supply a homogenous final good using a unit of entrepreneurial labour, production labour, and a (custom-made) machine. The machine has to be bought from a machine producer. The higher the quality of the machine, the higher the marginal productivity of labour for final goods. Machine producers can produce higher quality machines at an increasing (labour) cost to themselves. The two important features are that machine quality is not fully contractible and each machine is specific to its final user. This leads to a standard hold-up problem between each final producer and her machine producer. In the absence of trade in machines between countries, final producers buy their machines locally. When we allow for international trade in machines, we also assume that the cost of attaining a given level of quality is higher between countries than within countries, because of the transport costs incurred for carrying out complex user-producer relations.

Our first result is that when machines are imported (rather than sourced locally), lower transport costs can lead to higher unit trade costs. This result arises because, as transport costs decline, exporters find it profitable to produce higher quality machines that require more interactions (i.e., more transport) between producers and users. When transport costs are sufficiently high, this quality effect more than offsets the direct effect of lower transport costs. Thus, when a strong distinction is made between *transport costs* (i.e., the 'physical' cost of a shipment) and *trade costs* (i.e., the sum of all the costs incurred to deliver a good to its user, including in this case significant back and forth exchanges between the machine producer and its end user), a decline in transport costs

need not imply a decrease in trade costs.[1] In this way, trade costs can be seen to be endogenous and may not monotonically increase with transport costs. In particular, lower transport costs may act as an incentive to produce goods for which overall transaction costs are higher and thus more costly to trade.

Our second result relates to the location of the industry. When we allow for agglomeration effects and for location to be endogenous, we find that for high transport costs, there is a unique equilibrium for which machine production is evenly spread between the two countries. By contrast, for low transport costs, multiple equilibria arise and machine production can agglomerate in one (richer) country. This is reminiscent of the standard New Economic Geography (NEG) results and in particular those of Krugman and Venables (1995); it is nonetheless original in three ways. First in our model the framework in which it is derived differs from that of the NEG.[2] Second, the framework we use is a 'realistic' description of the production process and transport costs of a specific industry, machinery. This contrasts with most of the NEG literature, which typically attempts to model a generic industry, but often borrows features from many and ends up creating an implausible hybrid. Third, many interesting extensions such as the make-or-buy decision can be readily introduced into our framework.

Our third key set of results is a more nuanced view of the sensitivity (or elasticity) of trade to distance. To obtain them, we consider that standard machines (of exogenously given quality and not subject to any hold-up problem) can be traded as well as custom-made machines. In this context, we show that under empirically plausible parameter values, lower transport costs can lead to a higher sensitivity of trade to distance. This is because when transport costs fall below a certain threshold, importers switch from standard machines, for which trade is cheap and not very sensitive to distance to custom-made machines, which are more expensive to trade and can be more distance sensitive. This resembles the evolution of some segments of the machine industry, where the rise in the use of custom-made machinery took place against a background of decreasing

[1] We explain the wedge between transport and trade costs principally via transaction costs. There are of course many other possible candidates to explain a growing wedge between trade and transport costs: policy barriers (tariffs and non-tariff barriers), local distribution costs (wholesale and retail), information costs, costs associated with the use of different currencies, and legal and regulatory costs. However none of these other costs appears to have increased. Trade barriers have declined enormously since 1950. With the development of supermarkets, large retail chains, and computerised inventory management, it seems unlikely that there have been large increases in retail and wholesale costs. Information is arguably easier to find today than in the past. Currency costs and legal and regulatory costs do not offer strong prima facie evidence of a strong increase either. By a process of elimination, the type of friction we stress is the most likely source of possibly diverging trends for trade and transport costs.

[2] Ex-ante product differentiation plays no role (and ex-post product differentiation is an outcome) and we assume decreasing returns in final production instead of increasing returns.

transport costs. These results on the sensitivity of trade to distance may also provide fresh insight into the meaning of empirical results from gravity equations and in particular those studies that have noted an increase in the sensitivity of international trade to distance (e.g., Berthelon and Freund, 2004; Combes, Mayer, and Thisse, 2006). Disdier and Head (2005) offers strong confirmation of this fact.[3]

In addition to adding to the results from the NEG, mentioned above, our model is also related to the literature on incomplete contracting in a spatial framework (Almazan, de Motta, and Titman, 2005; Helsley and Strange, 2004; Matouschek and Robert-Nicoud, 2005; Rotemberg and Saloner, 2000) and to the literature on international out-sourcing (Antràs, 2003; Grossman and Helpman, 2005; McLaren, 2000). The key difference with the location-and-incomplete-contracting literature is that we consider contracting interactions across locations (instead of within locations) but ignore how agglomeration reduces opportunism. The main difference with the literature on international out-sourcing is that we adopt a simpler contracting framework but allow for the location of activities to be fully endogenous. Finally we contribute to the emerging literature on the micro-foundations of trade costs. Coleman (2004) shows how transport, when subject to a capacity constraint, can imply significant deviations from the law of one price in the short run. Harrigan and Venables (2006) explore the importance of time costs in the determination of trade patterns, while Leamer and Storper (2001) and Storper and Venables (2004) examine the benefits of face-to-face contact. This paper is instead concerned with how incomplete contractibility affects trade, agglomeration, and productivity.

The next section highlights the general mechanism behind our argument. Section 3 presents the basic mechanics of our model. For this purpose, we first consider an economy in autarchy with exogenously set capabilities for machine production. The following section explores the model in a trade context where the location of machine production and the capabilities of the industry are endogenously determined. Section 5 extends the model to shed some light on the links between the elasticity of trade to distance and transport and trade costs. In the conclusion, we offer some possible future extensions of the model.

[3]It could be added that trade between adjacent countries remain a very large part of world trade (Leamer and Storper, 2001). All this is also consistent with more indirect evidence such as studies of agglomeration in Europe, which routinely fail to uncover much systematic changes in the location patterns of industries despite European integration (Midelfart-Knarvik, Overman, Redding, and Venables, 2000; Storper, Chen, and De Paolis, 2002).

2. Rising trade costs: The argument

To develop a general equilibrium model with detailed microeconomic foundations, theoretical consistency requires us to specify why countries trade and detailed microeconomic foundations for transaction costs. In order to keep the model tractable, fairly specific assumptions are necessary. Our modelling strategy is to accept these requirements although they reduce the scope of the model. We begin in this section with a simple sketch of how our basic approach works in a simple partial equilibrium set-up. This shows that our core argument is fairly general and relies on a more limited set of assumptions.

Consider a buyer in a country who imports one unit of a good from another country. This good is characterised by its endogenously determined quality level, K. The inverse demand of the buyer for a good of quality K is:

$$P(K) = K^{1-\epsilon}, \qquad 0 < \epsilon < 1. \tag{A}$$

The quality of the good is linear in the amount of labour used to produce it in the other country. When a good is exported, a fraction τ of the labour employed by the producer is lost in transport. The literal reading of this fact is that part of the quality 'melts' during the shipping, captured in the 'iceberg' assumption popularised by the NEG. While this interpretation may be applicable for some agricultural and other perishable goods, it is probably less appropriate for many other goods (although it should be remembered that insurance costs, among others, are ad-valorem), especially for durable goods. There are other possible reasons for the transport loss: the good may need to be adjusted to its end user; the instructions to use it need to be produced, read and even possibly discussed between the producer and the end user; and significant maintenance after installation may be needed. In all these cases, the seller will need to visit the buyer and the number of visits will increase with the quality of the good. Moreover, attaining a given level of 'quality' in a foreign country may take more labour than for one's own country.[4] Thus, imported quality is given by:

$$K = (1-\tau)L, \tag{B}$$

where L is the quantity of labour. The profit of the machine producer is $\pi = P(K) - w\,L$ where w is the wage in the exporting country. Inserting equations (A) and (B) into the profit of the exporter implies: $\pi = [(1-\tau)L]^{1-\epsilon} - w\,L$. Profit maximisation with respect to the level of quality by the

[4] Gertler (2004) insists on this last idea. He also underscores that because this quantity of labour cannot be determined in advance, foreign sourcing may turn out to be more costly than it first appears — a source of uncertainty.

exporter then implies

$$L = \left[\frac{(1-\epsilon)(1-\tau)^{1-\epsilon}}{w} \right]^{1/\epsilon}. \tag{C}$$

The unit trade cost is the labour time spent in transport τL valued at the wage in the exporting country, w. Using the above, we get after simplification:

$$TC(\tau) = \tau \, L \, w = \tau(1-\tau)^{\frac{1-\epsilon}{\epsilon}} Q, \tag{D}$$

where Q is a constellation of constants. This function is obviously bell-shaped. Starting from $\tau = 1$, lower transport costs imply first higher trade costs before leading to lower trade costs. To understand how trade costs are affected by lower transport costs, note that there are two effects. There is a direct effect of lower transport costs which multiply the value of the good being sold. There is also an indirect effect where, following a decline in transport costs, exporters find it optimal to produce higher quality goods, which in turn incur higher trade costs. When transport costs are high, the indirect effect dominates and lower transport costs lead to higher trade costs. For low transport costs, the opposite occurs.

This argument gives the thrust of the model but it leaves many questions unanswered.[5] Why is the good imported rather than produced locally? What are the underpinnings of the demand function? What is the source of the seller's market power? Why are wages fixed and how would endogenous wages affect the outcome?, and so on. The rest of the paper will propose a complete model with consistent answers to these questions in a general equilibrium framework.

3. The baseline model in autarchy

Consider a two-tier production process. There is a single homogenous final good that is used as numéraire. Each competitive final producer requires one unit of entrepreneurial labour, one machine, and uses some amount of workers' time. Each final producer operates under decreasing returns to scale.[6] After sinking her labour endowment, each entrepreneur acquires her machine from a machine producer. Workers are hired after the machines are delivered. The profit function

[5] Note further that it relies on the fact the exporters have some market power over their seller. However, it would be equivalent to assume a competitive market with decreasing returns. The only crucial elements are that transport costs should be ad-valorem and that quality should increase sufficiently when transport costs decline.

[6] What matters is the existence of a rent to be shared between the final producer and her machine provider. Decreasing returns generate such rents even for very simple market structures (i.e., competitive with homogenous goods). We conjecture that similar results could be obtained with increasing returns since they also generate rents. However increasing returns would require much more structure in the final market (e.g., product differentiation).

of final producer i is given by:

$$\pi_i = k_{i,j}^{\alpha} l_i^{\beta} - w l_i - P_{i,j} \tag{1}$$

where $\alpha + \beta < 1$ (and $(\alpha,\beta) > (0,0)$). Labour, l_i, is hired at the competitive wage w. Finally, $k_{i,j}$ denotes the quality of the machine bought from producer j at price $P_{i,j}$. After inserting the first-order condition for profit maximisation with respect to employment in equation (1), we find:

$$\pi_i = (1-\beta)\left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\beta}} k_{i,j}^{\frac{\alpha}{1-\beta}} - P_{i,j} \equiv Z_i - P_{i,j}. \tag{2}$$

In turn, machines are supplied by machine producers. Each machine producer requires one unit of entrepreneurial labour and uses some quantity of workers' time. As in final production, entrepreneurs in the machine sector must sink their labour endowment before being able to operate. For reasons that will become clear below, this sunk labour is best thought of as a training period. The production of *one* machine of quality $k_{i,j}$ by machine producer j sold to final producer i requires a sunk investment of $l_{i,j}$ units of labour such that $k_{i,j} = A l_{i,j}$. The productivity shifter A, which we refer to as the capabilities of machine producers, is temporarily taken to be exogenous. Below, it will become an endogenous variable. The profit of entrepreneur j is the sum of her operating profit across all the final producers she sells a machine to

$$\pi_j = \sum_{i \in C(j)} \pi_j(i) = \sum_{i \in C(j)} \left(P_{i,j} - \frac{w}{A} k_{i,j}\right), \tag{3}$$

where $C(j)$ denotes the set of j's buyers.

The investment made by the entrepreneur in the machine sector to produce a machine is not contractible *ex-ante*.[7] Because of customisation, each machine is specific to its prospective buyer and cannot be used by any other final producer. This specific and non-contractible investment opens the door to *ex-post* renegotiation, which in turns precludes the existence of any market where machine producers can credibly compete on prices ex-ante. Instead, each final producer is randomly matched to a machine producer. After being matched with final producer i, machine producer j makes her non-contractible investment in quality, $k_{i,j}$. After this, the two parties bargain on the price $P_{i,j}$.

[7] For consistency, we also assume that profits and employment in the two sectors cannot be contracted upon ex-ante. Otherwise it might be possible to contract on final output or some inputs and implement an efficient investment. Note that this paper takes contract incompleteness as a given and does not attempt to add anything to the debate about its micro-foundations (Hart and Moore, 1999). Furthermore, final production and machine production require their entrepreneurs to sink their labour endowment so that vertical integration is simply not an option. This assumption is relaxed in Appendix A.

Following usual practice in the incomplete contract literature (see Grossman and Hart, 1986; Hart, 1995, and their followers), we assume that the hold-up problem is resolved co-operatively ex-post. The machine is thus delivered to its buyer and the surplus is split following a Nash-bargaining solution. The ex-post surplus made by the machine producer is $P_{i,j}$.[8] The surplus made by the final producer is given by (2). If a is the 'bargaining power' of final producers, the Nash-bargaining solution is such that it maximises the following expression:

$$\underset{P_j}{\mathrm{Max}}(Z_i - P_{i,j})^a P_{i,j}^{1-a}. \tag{4}$$

This directly yields $P_{i,j} = (1-a)Z_i$ so that the profit of final producer i can be written as $\pi_i = aZ_i$ and the operating profit made by the machine producer with her machine is $\pi_j(i) = (1-a)Z_i - wk_{i,j}/A$. Expecting this *ex-post* level of operating profit with final producer i, the *ex-ante* profit-maximising investment for j is:

$$k_{i,j} = \frac{[(1-a)\alpha A]^{\frac{1-\beta}{1-\alpha-\beta}} \beta^{\frac{\beta}{1-\alpha-\beta}}}{w^{\frac{1}{1-\alpha-\beta}}}. \tag{5}$$

Note that the quality produced by machine producers increases with the efficiency shifter,A, while it decreases with the cost of labour, w. Quality also declines with the bargaining power of final producers, a.

Free-entry of entrepreneurs in final production implies that in equilibrium $\pi_i = w$. Using (2), (5), and $\pi_i = aZ_i$, we find after simplification:

$$k_{i,j} = \frac{(1-a)\alpha A}{a(1-\beta)} \tag{6}$$

and

$$w = [a(1-\beta)]^{1-\alpha-\beta} \beta^{\beta} [(1-a)\alpha A]^{\alpha}. \tag{7}$$

Using l_i as given by the first-order condition of (1) with respect to employment, the labour requirement for quality $l_{i,j} = k_{i,j}/A$, and equation (6), employment in each final producer and in the production of each machine can be derived as $l_i = \frac{\beta}{a(1-\beta)}$ and $l_{i,j} = \frac{(1-a)\alpha}{a(1-\beta)}$, respectively.

Finally, if n denotes the share of final production entrepreneurs in employment and m is the share of machine entrepreneurs, the random matching of each final producer with a machine producer implies that the expected profit of any entrepreneur in the machine sector is given by

[8] Since the investment made by the machine producer is both specific to final producer i and sunk, it must be ignored at this stage.

$E[\pi_j] = E[Card(C(j))]\pi_j(i) = \frac{n}{m}\pi_j(i)$. Assuming risk-neutrality, free-entry in machine production then yields $E[\pi_j] = \frac{n}{m}[(1-a)Z_i - wk_{i,j}/A] = w$. After normalising total employment to unity, labour market clearing writes $n + nl_i + m + nl_{i,j} = 1$. Together with the results above, these two expressions imply: $n = a(1-\beta)$ and $m = (1-a)(1-\alpha-\beta)$.

Before allowing for international trade, a few comments are in order. First note that mean income w in equation (7) is maximised when the bargaining power of final producers, i.e., their share of surplus is such that $a = \frac{1-\alpha-\beta}{1-\beta}$. This value of a (between 0 and 1) reflects a trade-off between the rents of entrepreneurs in final production and those of entrepreneurs in the machine sector. Machine producers invest efficiently when they can capture all the surplus, i.e., when $a = 0$. On the other hand, low values of a make entry in final production unattractive since a low a implies a low share of surplus for final producers. With few entrepreneurs in final production, each will have to employ a large number of workers, and this in turn will imply a very low marginal product of labour. The other key parameter in (7) is the capabilities of machine producers, A, which map directly onto wages. This is because higher capabilities lead to better machines in equilibrium and thus a higher marginal product of labour in final production and higher wages for workers and entrepreneurs in both sectors.

Since machine producers cannot capture the entire marginal returns to their investment without preventing entrepreneurs in final production from entering efficiently, the equilibrium outcome described above is inefficient. The efficient outcome would involve the ex-ante co-operative maximisation of the joint-profit by the two entrepreneurs: $\pi = k^\alpha l^\beta - wl - wk/A$. The un-constrained first-best also requires any entrepreneur in the machine sector to produce as many machines as possible. To avoid any degeneracy, it is useful to assume a limited span of control for entrepreneurs in the machine sector so that an entrepreneur cannot produce more than g machines. For simplicity, assume g to be large enough for this limit on the span of control not to be binding at the free-market equilibrium.[9] The first-best wage can then easily be derived as

$$w^{FB} = \left[\frac{g(1-\alpha-\beta)}{g+1}\right]^{1-\alpha-\beta} \alpha^\alpha \beta^\beta A^\alpha. \tag{8}$$

Simple algebra shows that this wage is higher than the equilibrium wage derived in equation (7).

The appendices shows that our model can be enriched to consider standard contracting issues such as the make-or-buy decision (Appendix A) or the role of property rights (Appendix B). The

[9]From previous algebra, it is sufficient to have $g > (1-\beta)/\alpha$.

first extension allows us to re-derive standard results (e.g., that of Grossman and Helpman, 2002) under a very different set of specifications while the second gets much closer, in modelling terms, to reference models (i.e., that of Grossman and Hart, 1986).

4. International trade in machines

A. *Transport costs and endogenous determination of* A

Consider two countries that are ex-ante similar in every respect. Assume that when a machine is produced in country 1 for a final producer in country 2, the labour employed at machine production must spend a share τ of their time travelling back and forth between the two countries. This travel time is not otherwise productive. Consequently when machines are exported, the quality production function for machinery is $k_{i(2),j(1)} = A_1(1-\tau)l_{i(2),j(1)}$ where A_1 is the capability of machine producers in country 1 while subscript $i(2)$ denotes final producer i located in country 2 and subscript $j(1)$ denotes machine producer j located in country 1.[10] Exported machines are paid in final goods, which we assume to be perfectly tradable.[11] Figure 1 represents the complete input-output structure.

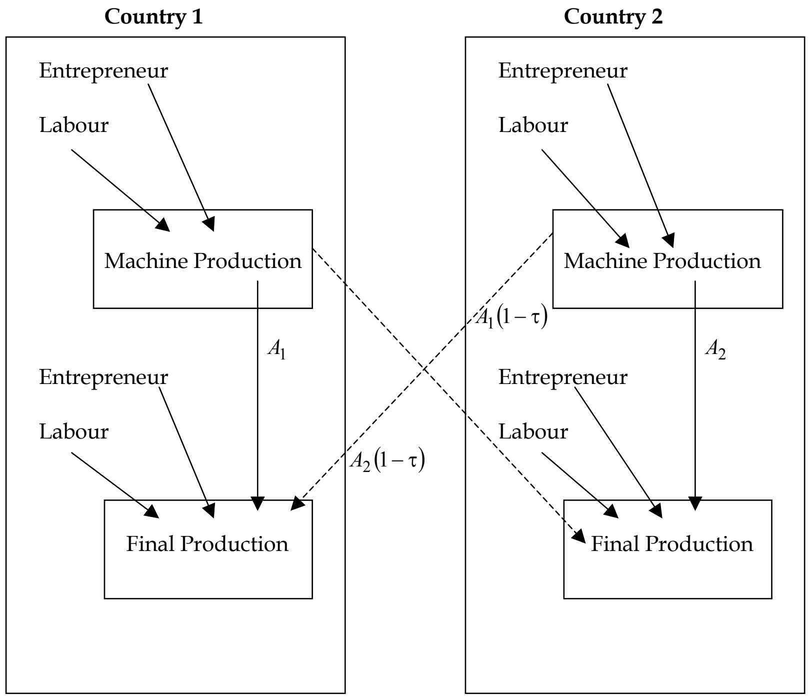
Turning to the determination of the capabilities of machine producers, A_1 and A_2, note that the empirical literature argues that both the number of firms and total employment in the sector boost productivity at a given location (Rosenthal and Strange, 2004). For simplicity and given that total employment in each country is normalised to unity, we consider that a country's share of employment in the machinery sector has a positive effect on its capabilities, A.[12] If we denote the share of employment in machinery in a country (workers and entrepreneurs) by λ, our assumption implies $A_1 = \mathcal{A}(\lambda_1)$ and $A_2 = \mathcal{A}(\lambda_2)$ with $\mathcal{A}' > 0$. For consistency with the material above, we normalise $\mathcal{A}((1-a)(1-\beta)) = A$ where $(1-a)(1-\beta)$ is the share of labour employed in machinery at the autarchy equilibrium.

[10] Note that this is formally equivalent to the standard assumption of iceberg transport costs in NEG.

[11] Intermediate goods appear empirically to be more costly to trade (Hillberry and Hummels, 2005). In any case, adding trade costs for final goods would not change the nature of the results since it would amount to adding a trade cost that *in fine* affects the revenue of the machine producers but not the quality of machines. Formally, an ad-valorem trade cost τ^f on final goods would imply that machine producers would receive a share $1-a-\tau^f$ of the surplus rather than $1-a$ and the results below would need to be modified accordingly.

[12] These agglomeration economies can accommodate a variety of microeconomic foundations. We do not develop them here since their explicit modelling would only add a layer of complication without any further insight (except for one specific case that we discuss in the conclusion). We refer instead to Duranton and Puga (2004) for a detailed exposition of the mechanics of those microeconomic foundations.

Figure 1. The production structure in open economy



B. *Symmetric equilibrium with no trade*

From the analysis above, it is easy to see that for any $\tau > 0$, there is always a stable symmetric equilibrium. This equilibrium replicates the autarchy equilibrium in both countries. The local stability of this equilibrium is straightforward because a small positive shock in machinery employment in one country would lead to only a marginal increase in local productivity, which should be offset by the discrete cost of exporting. The conditions under which this equilibrium is unique are explored below.

C. *Asymmetric equilibrium with trade*

We now explore the asymmetric configuration where all machinery production takes place in country 1, which is also active in final production.[13] As a result of agglomeration economies, machine producers in country 1 have greater capabilities than machine producers in country 2: $A_1 = \mathcal{A}(\lambda_1) > A_2 = \mathcal{A}(0)$.

[13]Formally, the condition for this country not to be fully specialised will turn out to be $(1-\alpha)(1-\beta) < 0.5$. Given that β is the share of labour in final production for which 0.5 is possibly a lower bound, this condition is very unlikely to be binding empirically.

Equilibrium wage and machine quality in country 1 can be derived as in equations (1)-(7) to obtain:

$$k_{i(1),j(1)} = \frac{(1-a)\alpha A_1}{a(1-\beta)} \tag{9}$$

and

$$w_1 = [a(1-\beta)]^{1-\alpha-\beta}\beta^{\beta}[(1-a)\alpha A_1]^{\alpha}. \tag{10}$$

Because of constant returns to scale in aggregate production, these results are the same as for the closed economy, except of course for the shares of employment in the different occupations and, as a result, the higher productivity shifter in machine production. These two quantities are derived below.

Turning to country 2, profit maximisation by final producers implies:

$$\pi_{i(2)} = (1-\beta)\left(\frac{\beta}{w_2}\right)^{\frac{\beta}{1-\beta}} k_{i(2),j(1)}^{\frac{\alpha}{1-\beta}} - P_{i(2),j(1)} \equiv Z_{i(2)} - P_{i(2),j(1)}. \tag{11}$$

Ex-post bargaining between final producers and their machine providers leads to $P_{i(2),j(1)} = (1-a)Z_{i(2)}$. With the quality production function for (exported) machinery being $k_{i(2),j(1)} = A_1(1-\tau)l_{i(2),j(1)}$, profit maximisation by exporting machine producers in country 1 implies:

$$k_{i(2),j(1)} = \left[\frac{(1-a)\alpha A_1(1-\tau)}{w_1}\right]^{\frac{1-\beta}{1-\alpha-\beta}} \left(\frac{\beta}{w_2}\right)^{\frac{\beta}{1-\alpha-\beta}}. \tag{12}$$

Free-entry in final production in country 2 then implies that labour should be indifferent between the different occupations: $\pi_{i(2)} = w_2$. After simplification we obtain:

$$k_{i(2),j(1)} = \frac{(1-a)\alpha(1-\tau)^{\frac{1}{1-\alpha}}A_1}{a(1-\beta)} = (1-\tau)^{\frac{1}{1-\alpha}}k_{i(1),j(1)} \tag{13}$$

and

$$w_2 = (1-\tau)^{\frac{\alpha}{1-\alpha}}[a(1-\beta)]^{1-\alpha-\beta}\beta^{\beta}\left[(1-a)\alpha A_1\right]^{\alpha} = (1-\tau)^{\frac{\alpha}{1-\alpha}}w_1. \tag{14}$$

Machine producers must customise the machines for their customers, install them, and train their workers to use them. When machines are exported, all these operations involve a large amount of travelling back and forth between the two countries. As a consequence, quality is more costly to produce for export and, at a given price, exported machines will be lower quality than those sold at home. In turn, this lower quality has a negative effect on the marginal productivity of the labour employed in final production in the other country. However, equation (14) also implies that

lower transport costs lead to better quality machines for exports and higher wages in the importing country.[14]

The first effect is consistent with the description given by Gertler (2004) of the failure of Canadian and American manufacturers to adopt efficiently the latest generation of machinery mostly imported from Germany. According to Gertler (2004), there was a very large cost of distance, making it very costly for machine producers to gauge the needs of their clients and then to fine-tune properly their equipment after delivery. Furthermore, in case of a problem, the machines took very long to be serviced because of the distance.[15]

After replacement and using the first-order conditions for profit maximisation in final production, employment in final production is $l_1 = l_2 = \frac{\beta}{a(1-\beta)}$. Using (13), (6) and the production function for machines, simple algebra yields $l_{i(2),j(1)} = (1-\tau)^{\frac{\alpha}{1-\alpha}} l_{i(1),j(1)}$ and $l_{i(1),j(1)} = \frac{(1-a)\alpha}{a(1-\beta)}$. Free-entry in machine production in country 1 implies that the expected income of entrepreneurs in this sector is equal to the wage: $\frac{n_1}{m_1}\pi_{j(1)}(i(1)) + \frac{n_2}{m_1}\pi_{j(1)}(i(2)) = w_1$. Together with the labour market clearing equations in both countries, $n_1 + n_1 l_{i(1)} + m_1 + n_1 l_{i(1),j(1)} + n_2 l_{i(2),j(1)}$ and $n_2 + n_2 l_{i(2)} = 1$, we can determine total employment in machine production in country 1:

$$\lambda_1 \equiv m_1 + n_1 l_{i(1),j(1)} + n_2 l_{i(2),j(1)} = (1-a)(1-\beta)\left[1 + (1-\tau)^{\frac{\alpha}{1-\alpha}}\right] \tag{15}$$

The unit trade cost for each exported machine can finally be computed as the time spent in transport by labour employed in the production of a machine, $\tau l_{i(2),j(1)}$ valued at the wage in the country that produces it, w_1. After simplification, this trade cost is equal to:

$$TC(\tau) = \tau(1-\tau)^{\frac{\alpha}{1-\alpha}} Q\mathcal{A}\left((1-a)(1-\beta)\left[1 + (1-\tau)^{\frac{\alpha}{1-\alpha}}\right]\right)^{\alpha} \tag{16}$$

where $Q \equiv \beta^{\beta}\left[(1-a)\alpha\right]^{1+\alpha}\left[a(1-\beta)\right]^{-\alpha-\beta}$ is a constant.

The product of the first two terms of $TC(\tau)$ in (16) is obviously bell-shaped and reaches a maximum for $\tau = 1-\alpha$. With $\mathcal{A}' > 0$, the last term of $TC(\tau)$ is decreasing with τ. This ensures

[14] Note that this contrasts with the Alchian and Allen (1964) result according to which it is the 'good apples' that are shipped away while the 'bad apples' are sold locally. The key difference between these two results is that Alchian and Allen (1964) consider an additive transport cost independent from the value of the goods being shipped. We consider instead an ad-valorem transport cost. Furthermore, in Alchian and Allen (1964), qualities are exogenously determined, while they are optimally chosen in our model.

[15] Gertler (2004) actually goes further and argues that transaction costs are not driven solely by τ, a pure transport cost parameter. Instead, cultural and language differences made the co-ordination of machine producers with their North-American customers much more difficult than it was with their local customers (which is consistent with the robustness of the common language dummy in gravity equations). Gertler (2004) also insists on institutional differences. German machines were designed to be used by workers who maintain them lavishly and make a heavy personal investment to understand the details of their workings. North-American workers are typically reluctant to make such long-term investment because of their much higher job turn-over.

that $TC(\tau)$ is decreasing at least over $[1-\alpha,1]$. If $\mathcal{A}'((1-a)(1-\beta))$ is not infinite, it is also easy to show that $TC(\tau)$ is increasing with τ in the neighbourhood of zero. All this implies that in a region of high transport costs, their decline leads to higher trade costs, while in a region of low transport costs, their decline leads to lower trade costs. It should also be noted that lower transport costs imply an increase in the value of trade.[16]

It can be verified that this equilibrium is stable as long as no machine producer finds it profitable to enter in country 2. Simple calculations show that this holds provided τ is below a sustain point τ_{sust} such that:

$$(1-\tau_{sust})^{\frac{\alpha}{1-\alpha}}\mathcal{A}\left((1-a)(1-\beta)\left[1+(1-\tau_{sust})^{\frac{\alpha}{1-\alpha}}\right]\right)^{\alpha}=\mathcal{A}(0). \qquad (17)$$

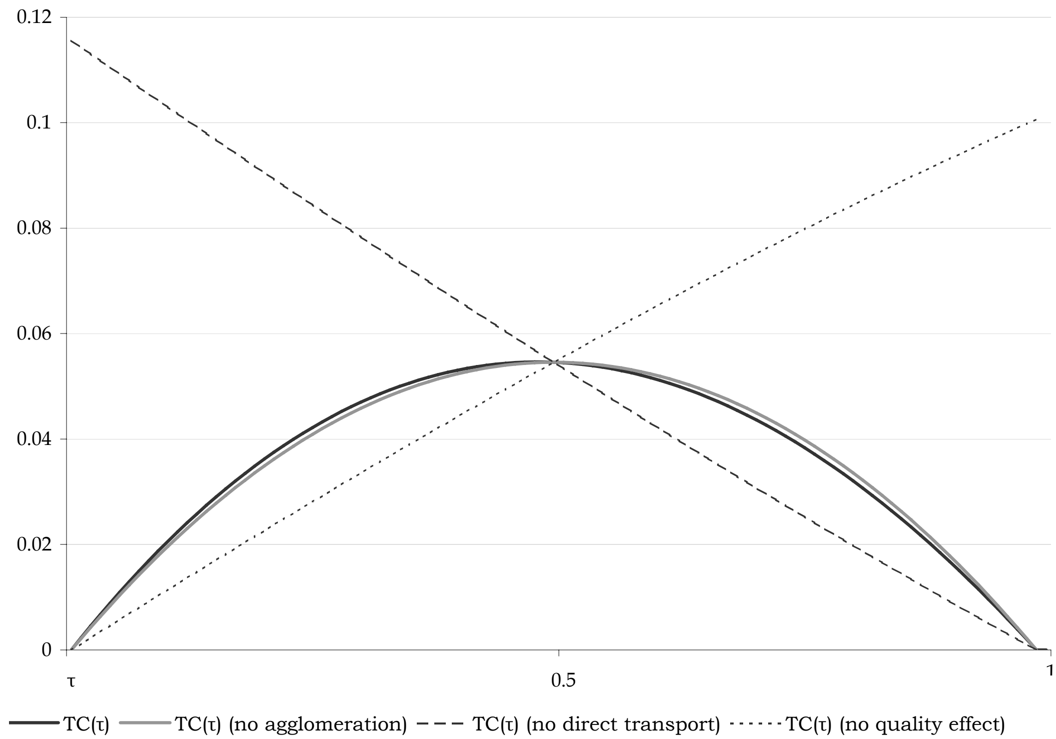
Since the term on the left hand-side is decreasing in τ, it is easy to see that τ_{sust} is unique. It is also easy to show that because (i) $\mathcal{A}'>0$ and (ii) trade costs are ad-valorem, there is no asymmetric equilibrium with active machine producers in both countries.

In addition, note that a decline in transport costs implies higher trade costs only when transport costs are high, so that it may be the case that for this parameter region, the asymmetric equilibrium is not stable. A sufficient condition for the asymmetric equilibrium to be stable in a region of increasing trade costs is when $\tau = 1-\alpha$, value for which we know that trade costs increase with falling transport costs. This sufficient condition reduces to: $\mathcal{A}(0) < \alpha^{\alpha/(1-\alpha)}\mathcal{A}\left((1-a)(1-\beta)\left(1+\alpha^{\frac{\alpha}{1-\alpha}}\right)\right)$.

In the determination of unit trade costs, the transport cost parameter, τ, intervenes three times (equation 16). First, it intervenes directly since a fraction τ of the machine producer's labour is lost in transport. Through this direct effect, lower transport costs push towards lower trade costs. This direct effect is counterbalanced by two indirect effects. When transport costs decrease, a lesser fraction of labour is lost in transport so that it is optimal for the producer to produce higher quality machines for export. This indirect effect dominates the direct effect for high trade costs while it is dominated for low trade costs. There is another indirect effect whereby lower transport costs and the subsequent increase in quality leads to higher employment in machinery. In turn, because of agglomeration economies, this leads to higher wages in the exporting economy. Since

[16] The volume of trade at the asymmetric equilibrium is constant by assumption since each final producer in country 2 is restricted to import only one machine. With an elastic demand for machines (caused for instance by heterogeneous capabilities of entrepreneurs in final production), it would be straightforward to extend the model and get increasing volumes of trade together with higher trade costs as consequences of lower transport costs. See next section for more discussion of this issue.

Figure 2. Evolution of unit trade costs as a function of transport costs



unit transport costs are paid in the labour of the exporting country, this agglomeration effect also contributes to higher trade costs as a consequence of lower transport costs.

The plain black curve in Figure 2 illustrates these results by plotting $TC(\tau)$ for the following parameter values: $\mathcal{A}(\lambda) = \lambda^{0.4}$, $a = 0.5$, $\alpha = 0.5$ and $\beta = 0.4$. Note that for these parameter values $\mathcal{A}(0) = 0$ so that the asymmetric equilibrium is stable for any $0 \leq \tau < 1$. The dashed curve in the same figure illustrates what the evolution of trade costs as a function of transport costs would be if there was no direct effect of transport (i.e., the first τ in equation 16 is set to 0.5 throughout). The dotted curve plots the evolution of trade costs when the indirect quality effect is conditioned out (i.e., the second τ in equation (16) is set to 0.5 throughout). Finally, the plain grey curve illustrates what the evolution of trade costs would be in absence of agglomeration effects (i.e., the third τ in equation (16) is set to 0.5 throughout).

Turning finally to the welfare analysis, note that the inefficiencies associated with incomplete contracting, which are discussed above, are still present in the open economy case. Our agglomeration mechanism leads to another possible source of inefficiencies. For low enough trade costs, the symmetric equilibrium is dominated in a Pareto sense by the asymmetric equilibrium. This

is because of a standard co-ordination failure where no-one wants to deviate from the symmetric equilibrium although the asymmetric equilibrium would make machine production sufficiently more efficient to offset trade costs in the country that specialises in final production. On the other hand, the asymmetric equilibrium cannot be Pareto-dominated because a country always gains from the agglomeration of machine production. However, this asymmetric equilibrium may not maximise total output. This can occur when $\mathcal{A}(.)$ is very concave with say $\mathcal{A}(0) = 0$ and $\mathcal{A}'$ very small for $\lambda > (1-a)(1-b)$. In this case, with $\mathcal{A}'$ very small, the gains from agglomeration are very small and can be more than offset by the trade costs even though the asymmetric situation is always in equilibrium (since $\mathcal{A}(0) = 0$).

D. *Discussion*

It should be clear at this stage that the detailed specification for transport costs matters. We chose an ad-valorem specification rather than a fixed amount per unit. This choice was dictated by the greater realism of this assumption for the machinery industry. As already noted, our specification is equivalent to the iceberg transport cost assumption, thus making comparisons with existing results easier. An additive specification for transport costs may lead to similar results provided two conditions are met: (i) lower transport costs lead to a strong enough positive quality effect and (ii) higher quality is more costly to trade. In this case, the increase in trade costs caused by higher quality dominates the direct effect of lower transport costs.

The other assumptions are far less central to our main result. For instance, our model is rooted in an incomplete contracting framework. We believe this is an important feature of the industry we model. The rent-splitting that occurs because of incomplete contracts also introduces some very tractable proportionalities, which allow us to solve the model easily and obtain closed-form solutions. However incomplete contracting is not a necessary condition for a non-monotonic evolution of the trade costs. To see this, note that the first-best quality of exported machines is given by the maximisation of the joint surplus of the two firms and is such that $k^{FB}_{i(2),j(1)} = k_{i(2),j(1)}/(1-a)^{\frac{1-\beta}{1-\alpha-\beta}}$ where $k_{i(2),j(1)}$ is the equilibrium quality given by (13). This implies that first-best employment to produce an exported machine, $l^{FB}_{i(2),j(1)}$, will contain a term in $1-\tau$. Hence the trade costs, $\tau l^{FB}_{i(2),j(1)} w_1$, will contain a term in τ multiplied by a term in $1-\tau$ and some wage/agglomeration term affected negatively by τ. Consequently, trade costs should initially increase when transport costs decline from 1 and decrease when transport costs approach 0.

Empirically, our model appears to capture rather well some aspects of the evolution of trade costs in the machinery industry. Whether trade costs have also increased in other industries is still an open question.[17] This is because trade costs are extremely hard to measure directly (Anderson and van Wincoop, 2004). There is nonetheless some suggestive empirical evidence that trade costs have indeed increased in the last thirty years. For instance, in his estimation of the trade cost parameter of a New Economic Geography model, Hanson (2005) finds that internal US trade costs increased between 1980 and 1990.[18]

More speculatively, a rise in trade costs could provide an answer to some problems in the existing literature on trade and location, where the NEG figures prominently. These NEG models, albeit very stylised, provide a compelling story about the determinants of industrial agglomeration at large spatial scales. The comparative statics of those models with respect to trade costs are now heavily used for policy and predictive purposes. The usual thought experiment is to consider the effects of a decline in trade costs (which are not distinguished from transport costs) on the location of economic activities. Yet, the strong conclusions of these models have not received much convincing empirical backing (see Head and Mayer, 2004, for a recent and very thorough review). This weak empirical performance is increasingly interpreted as a failure of these models. This paper suggests a different interpretation. It may not be the basic reasoning of NEG models that is flawed, but the assumption that in the real world, declining transport costs have actually brought about a powerful reduction in total trade costs across the board.

[17] To take an example radically different from the machinery industry, the wine industry also appears to support a similar argument about rising trade costs. The bulk of wine consumption in Europe used to be low quality wine transported directly in containers. The trade costs for low quality wines were probably quite low. Today, wine quality is much higher and bottles rather than containers need to be shipped. This makes the shipping of wine much more difficult, all the more so since the temperature and storage conditions must be closely monitored. On the consumer front, the number of varieties has also considerably increased making the choice of wine much more difficult, etc. As in the machine industry, trade costs for wine are also likely to have increased following an increase in product quality. Still another example may be found in the clothing industry. The Spanish firm Zara now has stores in many cities around the world. Their policy is to turn over about one third of the product mix every ten days, and to do so differently from one store to another, depending on the local market. As a result, their trade costs to market outlets are relatively high. Moreover, in order to innovate in product mix so rapidly and continuously, the firm has agglomerated much of its supply chain in the Basque region of Spain, as a way of facilitating rapid decisions based on face-to-face contact, reputation and cooperation with suppliers. Product quality, in the sense of variety, would seem to have been raised by the effect of lower transport costs, generating higher downstream trade costs and agglomeration upstream.

[18] Trade costs can also be measured using gravity equations. As shown by Anderson and van Wincoop (2003), the coefficient on the log distance in 'augmented' gravity equations can be interpreted as a trade costs parameter in monopolistic competition models with iceberg trade costs. However it is only one component of the trade costs — the elasticity of trade to distance. The other parameters of the trade cost (16), including τ, will appear in the constant of the same gravity regression. With trade flows being equal to the number of machines times their price, a number of parameters related to quantities will also enter the constant of the gravity estimation making it extremely hard to isolate the part that belong to trade costs (see Anderson and van Wincoop, 2004, for a complete discussion of this issue).

5. The sensitivity of trade to distance with standard and custom-made machines

A. *Standard and custom-made machines*

Thus far an important feature of the machinery industry has been has been left aside: machines need not be custom-made, as end users may instead opt for standard equipment. Simple pieces of equipment like personal computers or basic microscopes are typically produced in large series and do not require numerous movements of personnel between their producer and their end users. Others, such as commercial or military aircraft, nearly always do (pilots, crews, mechanics need to be thoroughly trained to any new model). Between these extremes, end users often decide whether to buy something 'off the shelf' or have it custom-made. By considering this issue, we can now derive a number of additional insights about the sensitivity of trade to distance.

For simplicity, assume that final producers in country 2 need to import their (standard or custom-made) machines from country 1 where machine producers have capabilities A_1 for custom-made machines and $\underline{A}_1$ for standard machines.[19] Custom-made machines are produced and traded as previously. Standard machines, of quality normalised to unity, can then be produced using $1/\underline{A}_1$ units of labour. With a competitive production of standard machines, simple algebra shows that assuming:

$$a^{1-\alpha-\beta}\left[(1-a)\alpha A_1\right]^{\alpha} > (1-\beta)^{\alpha}\left(\frac{\underline{A}_1}{1+\underline{A}_1}\right)^{1-\beta} \tag{18}$$

ensures that custom-made machines (rather than standard machines) are sold to final producers in country 1. Note that this condition will also be sufficient to guarantee that final producers in country 2 buy custom-made machines when $\tau = 0$. This condition requires the capabilities of custom-made machine producers (measured by A_1) to be sufficiently larger than those of standard machine producers (measured by $\underline{A}_1$) to compensate for the transaction costs associated with custom-made machines (which depend on a).

Standard machines can be traded between the two countries at a cost τ. For consistency with custom-made machines, this cost is also measured in country 1's labour. From the results below, it is easy to see that very similar results would be obtained if trade costs were paid in numéraire.

[19] In other words, we rule out the no-trade equilibrium and assume exogenous capabilities for machine producers. Allowing for machines to be produced in country 2 would only complicate the analysis and for us to consider an extra condition making sure that domestic sourcing does not occur in equilibrium. As seen above, we can easily make sure that this type of condition is not binding (for instance by assuming $\mathcal{A}(0) = 0$). Exogenous capabilities lead us to leave aside the (agglomeration) effect of changing employment in machine production in country 1, in particular at the time final producers from country 2 switch from one type to another. This is arguably a secondary issue not worth the extra cost in terms of non-linear algebra.

When importing a standard machine, final producer i in country 2 earns a profit equal to

$$\pi^S_{i(2)} = l_2^\beta - w_2 l_2 - w_1 \left(\frac{1}{\underline{A_1}} + \tau \right). \tag{19}$$

Profit maximisation in final production and free entry in final production in country 2 (i.e., $\pi^S_{i(2)} = w_2$) then imply that the wage in country 2 when standard machines are imported solves:

$$(1-\beta) \left(\frac{\beta}{w_2^S} \right)^{\frac{\beta}{1-\beta}} - w_2^S - w_1 \left(\frac{1}{\underline{A_1}} + \tau \right) = 0, \tag{20}$$

where w_1 is given by (10). Straightforward applications of the implicit function theorem imply that

$$\frac{\partial w_2^S}{\partial \tau} < 0 \qquad \text{and} \qquad \frac{\partial^2 w_2^S}{\partial \tau^2} < 0. \tag{21}$$

As in the case of custom-made machines, the wage in the country that imports standard machines declines with transport costs. This is because transport costs are ultimately paid by the importing country. Equations (21) and direct inspection of (20) guarantee that there is a unique positive solution to (20) in w_2 for any τ between 0 and 1.

From equations (11)-(12) and the fact that $P_{i(2),j(1)} = (1-a)Z_{i(2)}$, the profit of final producer i when importing a custom-made machine is

$$\pi^C_{i(2)} = a(1-\beta) \left(\frac{\beta}{w_2} \right)^{\frac{\beta}{1-\alpha-\beta}} \left[\frac{(1-a)\alpha A_1 (1-\tau)}{w_1} \right]^{\frac{\alpha}{1-\alpha-\beta}}. \tag{22}$$

When all other final producers in country 2 import standard machines, it pays off to start importing custom-made machines when: $\pi^C_{i(2)} > w_2^S$. Using equation (22), simple algebra shows that this condition boils down to

$$w_1 (1-\tau)^{\frac{\alpha}{1-\alpha}} > w_2^S, \tag{23}$$

where w_1 is given by equation (10) while w_2 is the positive solution to equation (20). To compare the two sides of this inequality note first that both sides decrease with τ. Then note also that when $\alpha \geq 0.5$, the *lhs* of (23) is convex. Because $w_2^S > 0$ when $\tau = 1$ and equation (18), there is a unique threshold $\hat{\tau}$ such that custom-made machines are imported when $\tau \leq \hat{\tau}$ and standard machines are imported otherwise. When $\alpha < 0.5$, the *lhs* of (23) is concave so that there could be any number of values for τ such that the two sides of (23) are equal. However, equation (18) also implies the existence of a threshold $\hat{\tau}$ below which equation (23) is always satisfied. Some arduous algebra then shows that imposing a condition slightly stronger than $w_2^S > 0$ when $\tau = 1$ guarantees that this threshold is unique.[20]

[20] This condition, $2 - \beta + (2 - \beta - 3/w_2)(\beta/w_2)^{1/(1-\beta)} > 0$ when $\tau = 1$, guarantees that the third derivative of w_2^S in τ is positive, which then leads to the uniqueness of $\hat{\tau}$.

These results show that when transports costs are small enough country 2 imports custom-made machines from country 1 while for large enough transport costs it imports standard machines. Assuming a further mild regularity condition guarantees that there is a unique value $\hat{\tau}$ below which only custom-made machines are traded and above which only standard machines are traded. This result lends support to those derived previously. Lower transport costs lead to the trade of increasingly sophisticated goods: there is first a switch from standard to custom-made machines and then an increase in the quality of custom-made machines.

B. ***The sensitivity of trade to distance***

To assess the effect of these changes on the sensitivity of trade to distance, we need to give a distance interpretation to our transport cost parameter, τ. Consistent with standard specifications in the literature, we assume that $1-\tau = \frac{1}{t\, d_{1,2}}$ where t is a transport cost parameter per unit of distance and $d_{1,2}$ is the distance separating the two countries. This relation between distance and the transport cost may seem arbitrary. A first response is that elevating distance to some power different from unity would only multiply to the two key elasticities computed below in expressions (25) and (27) by this power (and it would also elevate the distance variable by the same power in the latter equation). Hence, the comparison of these two elasticities of trade to distance would be unchanged (up to a transformation of the distance). Second, the objective here is to show that transport costs and the elasticity of trade to distance may not be monotonically related. For this, only a (realistic) counter-example is needed.

When only custom-made machines are traded (i.e., when $\tau \leq \hat{\tau}$), imports from country 1 to country 2 are given by

$$x^C_{2,1} = n_2\, P_{2,1}. \tag{24}$$

As shown above, the number of machines imported is given by $n_2 = \frac{1}{1+l_2}$ with $l_2 = \frac{\beta}{a(1-\beta)}$ and the price of machines is equal to $(1-\tau)^{\frac{\alpha}{1-\alpha}} \beta^\beta \left[(1-a)\alpha\right]^{1+\alpha} \left[a(1-\beta)\right]^{-\alpha-\beta} A_1^\alpha$. Simple algebra then shows that the elasticity of trade with respect to distance is given by:

$$\zeta^C(d_{1,2}) = -\frac{\frac{\partial x^C_{2,1}}{\partial d_{1,2}}}{\frac{x^C_{1,2}}{d_{1,2}}} = \frac{\alpha}{1-\alpha}. \tag{25}$$

Note that this elasticity is also the coefficient that would be estimated in a simple gravity equation (regressing the log of the exports on the log of distance).

When only standard machines are traded (i.e., when $\tau > \hat{\tau}$), imports from country 1 to country 2 are given by

$$x^S_{2,1} = n_2\, P_{2,1} = w_1 \frac{\frac{1}{\underline{A}_1} + 1 - \frac{1}{t\, d_{1,2}}}{1 + (\frac{\beta}{w_2})^{1/(1-\beta)}}, \tag{26}$$

where w_2 is the positive solution to the implicit equation (20) and w_1 is given by (10). This implies that the elasticity of exports to distance cannot be expressed analytically in the general case. However, inspection of (25) and (26) shows that $\zeta^C(d_{1,2})$ will in general differ from $\zeta^S(d_{1,2})$. This implies that the sensitivity of trade to distance will differ depending on whether custom-made or standard machines are traded.

Further insights can be gained by assuming the share of labour, β, is equal to 0.5 and normalising $\underline{A}_1$ to unity. Calculations then show that

$$\zeta^S(d_{1,2}) = \frac{1}{2t\, d_{1,2} - 1} - \frac{8\, w_1^2 (t\, d_{1,2} - 1)}{t^2\, d_{1,2}^2 + 4\, w_1^2 (t\, d_{1,2} - 1)^2} \tag{27}$$

This elasticity decreases with w_1. It is thus highest for the lowest levels of w_1. With $\beta = 0.5$, condition (23) imposes at least $w_1 > 0.5$. Assuming the lowest admissible value for w_1 and $\alpha = 1/3$ (a realistic value for the share of capital), a straightforward comparison between (25) and (27) shows that $\zeta^C(d_{1,2}) > \zeta^S(d_{1,2})$ provided $t\, d_{1,2} > 1.174$. Put differently, the sensitivity of trade to distance increases when custom-made machines are imported instead of standard machine provided trade costs represent at least 21% of the cost of production of standard machines. This value is arguably a lower bound for trade costs (Anderson and van Wincoop, 2004). When $w_1 = 1$ (i.e., one unit of labour can produce a standard machine of unit quality or else add one unit of quality to a custom-made machine), trade costs need to be only 5% or more of production costs for standard machines for the switch to custom-made machines to increase the sensitivity of trade to distance.

Note also that ζ^S varies non-monotonically with $d_{1,2}$ (or equivalently with t). The derivation of (27) shows that as distance increases, the sensitivity of trade to distance first decreases and then increases. For $w_1 = 0.5$, when trade costs are below 130% of production costs a rise in distance leads to a lower elasticity of trade to distance. When trade costs are above 130% of production costs, ζ^S increases with $d_{1,2}$. For $w_1 = 1$, the elasticity of trade to distance first falls as distance increases when trade costs are below 38% of production costs and rises above this threshold.

C. *Discussion*

To understand these results, recall first that for *custom-made machines*, trade costs are a fraction of their value. In turn, trade costs affect their quality (and thus their price). The value of trade in equilibrium is then proportional to $(1-\tau)^{\alpha/(1-\alpha)}$. With distance being related to trade costs in such a way that what is left after trade costs is inversely proportional to distance, imports from country 1 are thus proportional to $d_{1,2}^{-\alpha/(1-\alpha)}$, which obviously leads to a constant elasticity of trade to distance equal to $\alpha/(1-\alpha)$. Taking a realistic value of $1/3$ for the share of machinery in final production yields an elasticity of trade to distance of $1/2$. This is below most existing empirical estimates but remember that the number of traded machines remains the same regardless of transport costs (as a result of each final producer buying exactly one machine and the number of final producers being independent of transport costs in equilibrium). Allowing for an elastic demand for machines would of course increase this coefficient (but at the cost of a much greater analytical complexity).[21]

When *standard machines* are traded, the elasticity of trade to distance depends on a negative price effect and a positive quantity effect. As shown by the numerator of equation (26), a larger distance leads to higher prices and thus a higher value of imports, all else equal. As shown by the denominator of equation (26), a larger distance also leads to fewer machines being imported and thus a lower volume of imports, all else equal. This second effect does not always dominate the first one so that the elasticity of trade to distance may be positive or negative.

To shed light on the variations of this elasticity, note first that an increase in distance implies a larger trade cost component in the price. The elasticity of prices to distance thus increases with distance (formally: $-(\partial P/\partial d)/(P/d) = -1/(td\,P)$ which increases with d). On the other hand, when $\beta = 0.5$, the elasticity of quantities to distance is decreasing with distance because the effect of trade costs over the number of final producers becomes negligible for large distance. Hence, these two effects vary in opposite directions. It turns out that for short distances, the variation of the price effect is dominated by that of the quantity effect. The opposite holds for large distances. Hence the elasticity of trade to distance decreases and then increases. As shown by the numerical examples above, the elasticity of trade to distance can increase following lower transport costs for reasonable parameter values.

[21] Models of monopolistic competition also typically generate a constant elasticity of trade to distance (Anderson and van Wincoop, 2003). However a different margin is at stake in those models. Here quantities are fixed whereas quality varies with trade costs. In monopolistic competition models, quality is fixed whereas quantities vary.

It is also interesting to note that both the price and quantity effects on ζ^S become small as distance increase. As shown by equation (27), the elasticity of trade in standard machines to distance converges to zero when distance gets arbitrarily large. Since the elasticity of trade in custom-made machines to distance is constant, for a large enough distance: $\zeta^C > \zeta^S$. The numerical examples above show that for reasonable values of the parameters, the level of trade costs for standard machines at the time of the switch to custom-made machines need not be unrealistically large.

To summarise, lower transport costs (i.e., a lower t) imply a decrease and then an increase in the sensitivity of trade to distance when standard machines are still traded. A further decline in transport costs then triggers a switch to custom-made machines for which the sensitivity to distance is also likely to be greater.

This finding can be interpreted in light of the empirical literature on the sensitivity of trade to distance. First, when standard machines are traded, standard gravity regressions will be misspecified since distance does not appear multiplicatively to determine trade flows. This is a well-known problem (see Anderson and van Wincoop, 2004). More subtly, income (or more accurately earnings) in the machine importing country is determined simultaneously with trade costs. This points at a possible endogeneity bias. Leaving these two technical issues aside, note that the results above are consistent with the well documented increase in the sensitivity of trade to distance (Disdier and Head, 2005; Combes *et al.*, 2006), as well as the more detailed findings of Berthelon and Freund (2004) that this evolution is not the outcome of the composition of international trade having shifted towards more distance-sensitive industries but one within industries.[22]

6. Conclusion

This paper proposes a model of vertically linked industries in which (i) the quality of inputs is not contractible and (ii) providing a given level of quality to suppliers becomes more costly with distance. Lower transport costs imply that higher quality inputs are traded in equilibrium, with the effect of this higher quality that trade costs increase. An extension of the baseline model also shows how the sensitivity of trade to distance can also increase following lower transport costs.

[22] Another interesting result of Berthelon and Freund (2004) is that the increase in the sensitivity to distance is not apparently related to the type of good being traded. A similar result occurs in the model here. This is because trade flows must be balanced. Consequently, trade in final goods (imported by country 1) will show the same effects as the trade in machines (imported by country 2). Whether this result holds in a multi-country and multi-sector sector framework is beyond the scope of this paper.

Three key extensions of this model can be envisaged in further work. First, the productivity shifter in machine production, A, needs to be made endogenous. As argued above, adding an extraneous source of agglomeration economies would merely lead to an extra layer of complexity. It would be more interesting to make A endogenous within the current framework and use more intensively the features of our incomplete contracting framework. For instance, we could allow for a thicker local market to reduce opportunistic behaviour and increase efficiency as in Helsley and Strange (2004). A second important extension would be to introduce some heterogeneity to make the equilibria less extreme than the stark no-trade vs. complete-agglomeration configurations. A natural way to do it would be to introduce an idiosyncratic value for each possible match between machine and final producers. Despite high transport costs, one would still expect some final producer to get a much better match abroad than at home and thus import their machinery while others would source locally. At some critical level of transport costs, this 'interior' equilibrium may become unstable and machine production would then agglomerate in one country. These results would then be closer to those of Krugman and Venables (1995).[23] The third extension would be to consider more than two goods and two countries. The multi-country extension is important because a two country framework ignores relative locations and thus does not allow us to deal with possible trade diversions and changes in industry location triggered by changes in transport costs.

[23]More formally, in the model presented here the symmetric equilibrium does not undergo a bifurcation. It is conjectured that introducing some heterogeneity in matches may imply the occurrence of subcritical bifurcations with the symmetric equilibrium becoming unstable.

References

Alchian, Armen A. and William R. Allen. 1964. *University Economics*. Belmont, CA: Wadsworth.

Almazan, Andres, Adolfo de Motta, and Sheridan Titman. 2005. Firm location and the creation and utilization of human capital. Processed, University of Texas at Austin.

Anderson, James E. and Eric van Wincoop. 2003. Gravity with gravitas: A solution to the border puzzle. *American Economic Review* 93(1):170–192.

Anderson, James E. and Eric van Wincoop. 2004. Trade costs. *Journal of Economic Literature* 42(3):691–751.

Antràs, Pol. 2003. Firms, contracts, and trade structure. *Quarterly Journal of Economics* 118(4):1374–1418.

Berthelon, Matias and Caroline L. Freund. 2004. On the conservation of distance in international trade. Policy Research Working Paper 3293, World Bank.

Cairncross, Frances. 2001. *The Death of Distance: 2.0 How the Communications Revolution will Change our Lives*. London: Texere.

Coleman, Andrew. 2004. Storage, slow transport, and the law of one price: Evidence from the nineteenth century US corn market. Processed, University of Michigan.

Combes, Pierre-Philippe and Miren Lafourcade. 2005. Transport costs: Measures, determinants, and regional policy implications for France. *Journal of Economic Geography* 5(3):319–349.

Combes, Pierre-Philippe, Thierry Mayer, and Jacques Thisse. 2006. *Economic Geography*. Princeton (NJ): Princeton University Press. Forthcoming.

Disdier, Anne-Célia and Keith Head. 2005. The puzzling persistence of the distance effect on bilateral trade. Processed, University of British Columbia.

Duranton, Gilles and Diego Puga. 2004. Micro-foundations of urban agglomeration economies. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland, 2063–2117.

Gertler, Meric S. 2004. *Manufacturing Culture: The institutional geography of industrial practice*. Oxford: Oxford University Press.

Glaeser, Edward L. and Janet E. Kohlhase. 2003. Cities, regions and the decline of transport costs. *Papers in Regional Science* 83(1):197–228.

Grossman, Gene M. and Elhanan Helpman. 2002. Integration versus outsourcing in industry equilibrium. *Quarterly Journal of Economics* 117(1):85–120.

Grossman, Gene M. and Elhanan Helpman. 2005. Outsourcing in a global economy. *Review of Economic Studies* 72(1):135–160.

Grossman, Sanford J. and Oliver D. Hart. 1986. The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy* 94(4):691–719.

Hanson, Gordon H. 2005. Market potential, increasing returns, and geographic concentration. *Journal of International Economics* 67(1):1–24.

Harrigan, James and Anthony J. Venables. 2006. Timeliness and agglomeration. *Journal of Urban Economics* (forthcoming).

Hart, Oliver. 1995. *Firms, Contracts, and Financial Structure*. Oxford: Oxford University Press.

Hart, Oliver and John Moore. 1999. Foundations of incomplete contracts. *Review of Economic Studies* 66(1):115–138.

Head, Keith and Thierry Mayer. 2004. The empirics of agglomeration and trade. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland, 2609–2669.

Helsley, Robert W. and William C. Strange. 2004. Agglomeration, opportunism, and the organization of production. Processed, University of Toronto.

Hillberry, Russell and David Hummels. 2005. Trade responses to geographic frictions: A decomposition using micro-data. Processed, Purdue University.

Hummels, David. 1999. Have international transportation costs declined? Processed, Purdue University.

Krugman, Paul R. and Anthony J. Venables. 1995. Globalization and the inequality of nations. *Quarterly Journal of Economics* 110(4):857–880.

Leamer, Edward E. and Michael Storper. 2001. The economic geography of the Internet age. *Journal of International Business Studies* 32(4):641–665.

Matouschek, Niko and Frédéric Robert-Nicoud. 2005. The role of human capital investments in the location decision of firms. *Regional Science and Urban Economics* 35(5):570–583.

McLaren, John. 2000. 'Globalization' and vertical structure. *American Economic Review* 90(5):1239–1254.

Midelfart-Knarvik, Karen Helene, Henry G. Overman, Stephen J. Redding, and Anthony J. Venables. 2000. The location of European industry. Economic Papers 142, European Commission Directorate-General for Economic and Financial Affairs.

Rosenthal, Stuart S. and William C. Strange. 2004. Evidence on the nature and sources of agglomeration economies. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland, 2119–2171.

Rotemberg, Julio J. and Garth Saloner. 2000. Competition and human capital accumulation: a theory of interregional specialization and trade. *Regional Science and Urban Economics* 30(4):373–404.

Storper, Michael, Yun-chung Chen, and Fernando De Paolis. 2002. Trade and the location of industries in the OECD and European Union. *Journal of Economic Geography* 2(2):73–107.

Storper, Michael and Anthony J. Venables. 2004. Buzz: Face-to-face contact and the urban economy. *Journal of Economic Geography* 4(4):351–370.

Appendix A. Allowing for a make-or-buy decision

To show that our simple framework could easily be extended in a number of directions and that our contracting structure could be used to explore standard contracting issues, in this appendix we explore briefly the make-or-buy decision. More specifically, we consider the case where final producers can be integrated and produce their own machine, albeit less efficiently than 'specialist' machine producers.[24] The assumption that non-integrated machine producers have higher capabilities is justified by the fact that the labour endowments of entrepreneurs in the machine sector have been sunk in specialist training (i.e., there are benefits to specialisation). Alternatively, superior efficiency of specialists could be due to higher bureaucratic costs associated with vertical integration. Formally, we assume that final producers can also produce their own machine according to $k_i = A_f l_{i,i}$ with $A_f < A$.

The profit of integrated final producers (i.e., those who decide to make their own machine) is then:

$$\pi_i = k_{i,i}^{\alpha} l_i^{\beta} - w l_i - \frac{w}{A_f} k_{i,i}. \tag{A 1}$$

The two first-order conditions for profit maximisation with respect to $k_{i,i}$ and l_i imply that equilibrium profit is given by

$$\pi_i = (1-\alpha-\beta)\frac{(\alpha A_f)^{\frac{\alpha}{1-\alpha-\beta}} \beta^{\frac{\beta}{1-\alpha-\beta}}}{w^{\frac{\alpha+\beta}{1-\alpha-\beta}}}. \tag{A 2}$$

When all final producers are integrated and produce their own machine, free-entry (i.e., $\pi_i = w$) implies that the equilibrium wage is:

$$w = (1-\alpha-\beta)^{1-\alpha-\beta} \alpha^{\alpha} \beta^{\beta} A_f^{\alpha}. \tag{A 3}$$

Using equations (7) and (A 3), it is easy to verify that if

$$A_f > \left[\frac{a(1-\beta)}{1-\alpha-\beta}\right]^{\frac{1-\alpha-\beta}{\alpha}} (1-a)A, \tag{A 4}$$

the unique equilibrium involves only integrated producers whereas in the opposite case all machines are produced by specialised machine producers. When the two sides of equation (A 4) are equal, there is a continuum of equilibria with machines being produced by both integrated final

[24] Note that we only consider the case where final producers can integrate 'upstream'. Allowing machine producers to integrate 'downstream' by engaging in final production with a lower efficiency than specialised final producers, albeit straightforward, would only lead to a similar trade-off within a more complicated choice structure. We ignore this extension here.

producers and machine producers. The stability of these equilibria can be readily checked. When (A 4) is satisfied, no final producer finds it profitable to out-source the production of her machine. Conversely, when the inequality is reversed, no final producer finds it profitable to produce her own machine.[25] Equation (A 4) highlights a trade-off between integration, which saves on set-up and transaction costs, and out-sourcing, which allows access to specialist producers with greater capabilities.

Despite significant modelling differences, these results remind us of Grossman and Helpman (2002), who consider differentiated final producers operating under increasing returns. In their approach, there is an intermediate input which can be either of high or low quality. By contrast, in the framework proposed here, the final good is homogeneous and produced under decreasing returns, while there is a continuum of possible qualities for the machines. However, like Grossman and Helpman (2002), we find that the equilibrium market structure is characterised by pervasive out-sourcing when the difference in capabilities between specialised machine producers and integrated final producers is large and integration when this difference is small. In both,the equilibrium is unique, so that integrated and non-integrated final producers do not co-exist except in knife-edge cases.

Appendix B. Property rights: a sketch of extension

Note that there is no role for property rights in the model exposed in Section 3 since there are no productive assets that can be traded ex-ante. This contrasts with the theory of firm developed by Grossman and Hart (1986) and their followers. In our model, as in Grossman and Helpman (2002), the difference between integrated firms and out-sourcing depends on the identity of the investor — the machine producer or the final producer — not on who holds property rights. Consistent with this, the reserve option of machine producers is always zero in equation (4). This differs from Grossman and Hart (1986), where property rights play a role ex-post (by determining reserve options) and thus affect the ex-ante incentives to invest.

It would be however feasible to extend our model and allow property rights to matter. For property rights to play a role, we could first assume that a machine — before being used for final production — could belong to one or the other party. To avoid a trivial optimal allocation of

[25]Note that when the bargaining power is such that it maximises wages ($a = \frac{1-\alpha-\beta}{1-\beta}$), condition (A 4) boils down to $A_f > \frac{\alpha}{1-\beta} A$.

property rights we could then assume that final producers make an ex-ante investment (e.g., they hire their labour before getting a machine). Finally, machines should have some value to other potential users rather than be entirely specific to their original buyer.

Under such assumptions, the reserve option of machine producers and final producers would depend on the allocation of property rights. When machine producers hold the property right over the freshly-built machine, they can walk out of the bargaining with a machine that has some resale value whereas they retain nothing in absence of a property right. Conversely, final producers are able to retain the machine when they formally own it, whereas they end up with nothing in absence of property right.

As in Grossman and Hart (1986), the optimal allocation of property rights is then non-trivial. When dis-integration is the optimal structure, the results would be very similar to those of the (simpler) model developed here.