NBER WORKING PAPER SERIES

AIRPLANES AND COMPARATIVE ADVANTAGE

James Harrigan

Working Paper 11688
http://www.nber.org/papers/w11688

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2005

International Research Department, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045, james.harrigan@ny.frb.org. This paper has benefited from audience comments at Ljubljana, Illinois, Michigan, Columbia, the World Bank, NBER, CEPR, Hitotsubashi, Tokyo, City University of Hong Kong, and Hong Kong University of Science and Technology. I thank Christina Marsh and Geoffrey Barrows for excellent research assistance. The views expressed in this paper are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Airplanes and Comparative Advantage
James Harrigan
NBER Working Paper No. 11688
October 2005, Revised June 2006
JEL No. F1

ABSTRACT

Airplanes are a fast but expensive means of shipping goods, a fact which has implications for comparative advantage. The paper develops a Ricardian model with a continuum of goods which vary by weight and hence transport cost. Comparative advantage depends on relative air and surface transport costs across countries and goods, as well as stochastic productivity. In the model, countries that are far from their export markets will have low wages and tend to specialize in high value/weight products, which will be shipped on airplanes. Less remote exporters will have higher wages, and will tend to specialize in low value/weight products which will be sent by ship, train, or truck. These implications are confirmed using detailed data on U.S. imports from 1990 to 2003. Distance from the US is associated with much higher import unit values, an indication that the model identifies a quantitatively important influence on specialization and trade.

James Harrigan
International Research Department
Federal Reserve Bank-New York
33 Liberty Street
New York, NY 10045
and Columbia University and NBER
james.harrigan@ny.frb.org

1 Introduction

Countries vary in their distances from each other, and traded goods have differing physical characteristics. As a consequence, the cost of shipping goods varies dramatically by type of good and the route that it is shipped. A moments reflection suggests that these facts are probably important for understanding international trade, yet they have been widely ignored by trade economists. In this paper I focus on one aspect of this set of facts, which is that airplanes are a fast but expensive means of shipping goods.

The fact that airplanes are fast and expensive means that they will be used for shipping only when timely delivery is valuable enough to outweigh the premium that must be paid for air shipment. They will also be used disproportionately for goods that are produced far from where they are sold, since the speed advantage of airplanes over surface transport is increasing in distance. In this paper I build a simple model that illustrates some implications of these observations for specialization and wages: remote countries will have lower wages, and will specialize in lightweight goods which are air shipped. This implies that goods imported from more distant locations will have higher unit values. Using highly disaggregated data on all U.S. imports from 1990 to 2003, I show empirically that distance has a big influence on the composition of U.S imports: imports from distant trading partners have much higher unit values, and are much more likely to arrive by plane.

There is a small, recent literature that looks at some of the issues that I analyze in this paper. Limao and Venables (2002) model the interaction between specialization and trade costs, illustrating how the equilibrium pattern of specialization involves a tradeoff between comparative production costs and comparative transport costs. Deardorff (2004) elegantly shows how relative distance affects the trade pattern, arguing that local comparative advantage (defined as autarky prices in comparison to nearby countries rather than the world as a whole) is what matters in a world with trade costs. Evans and Harrigan (2005) develop a model of the demand for timeliness, and show how the pattern of US apparel imports is influenced by the interaction between relative distance and the relative value of timely delivery. Harrigan and Venables (2006) further develop microfoundations for the demand for timely delivery, and show how timeliness can lead to an incentive for agglomeration.

David Hummels has written a series of important empirical papers that directly motivated this paper. Hummels (1999) shows that ocean freight rates have not fallen on average since

World War 2, and have often risen for substantial periods. By contrast, the cost of air shipment has fallen dramatically. Chart 1 shows that these trends have continued since 1990, with the relative price of air shipping falling 40% between 1990 and 2004. Hummels (2001a) shows that shippers are willing to pay a large premium for faster delivery, a premium that has little to do with the interest cost of goods in transit[1]. Hummels (2001b) analyzes the geographical determinants of trade costs, and decomposes the negative effect of distance on trade into measured and unmeasured costs.

2 Airplanes and trade: theory

All comparative advantage trade models show how the interaction between country and product characteristics determine trade patterns. In the model developed here the relevant country characteristics are relative distance from each other and the relevant product characteristics are weight and unit labor requirements. As in the model of Eaton and Kortum (2002), in my model deterministic transport costs and stochastic labor productivity together determine a country's trade pattern. I extend the Eaton-Kortum framework by introducing differences in trade costs across goods: heavier goods cost more to ship by air. I initially simplify Eaton-Kortum by limiting the number of countries in the model to three, which is the minimum number required to make relative distance an influence on trade patterns. This simplification is useful for developing intuition and comparative statics, but is not needed for generating cross-sectional predictions, and I extend the model to the general multi-country case in section 2.3.

2.1 The three country model

In the model there are three countries, *1, 2*, and *3*, which can be thought of as "United States", "Mexico" and "China". Country *1* has a large technological advantage in a homogeneous numeraire good, so in the equilibria that I examine it specializes in this good, which it produces with a unit labor requirement of one. With *1*'s wage as the numeraire, the FOB

[1] By "the interest cost of goods in transit", I mean the financial cost of having goods in transit before they can be sold. This opportunity cost equals the value of the good × daily interest rate × days in transit.

export price of *1*'s good is also one[2]. *1* consumes the numeraire and imports from *2* and *3*. Demand for the numeraire and imports comes from a Cobb-Douglas utility function with expenditure share α on total imports.

Countries *2* and *3* are identical except for distance from *1* and the size of their labor forces. Both countries produce *x*, which they don't consume, exporting all their output to *1*, and using the export revenues to buy the numeraire from *1*. Producers in *2* and *3* face a choice of shipping mode (air or surface). Air shipment is more costly, and depends on the weight of the product being shipped. Despite its cost, air shipment may be profitable because goods shipped by air can be sold for a premium over surface shipped goods. To formalize this tradeoff, let an index $z \in [0,1]$ order goods by increasing weight (and therefore increasing value/weight, though this will be endogenous): good 0 is the lightest (computer chips), while good 1 is the heaviest (oil). Iceberg surface shipping costs τ are the same for all goods, but iceberg airfreight costs $\omega(z) > 1$ are increasing in weight and, therefore, increasing in z: good 0 is the cheapest to send by air, while good 1 is the most expensive. Furthermore, the cost of air freight is the same regardless of where the flight originates, and to make the problem interesting assume

$$\omega(z) > \tau_3 > \tau_2 > 1 \text{ for all } z. \tag{1}$$

Why would anybody pay for airfreight? The answer is, consumers like speedy delivery for some reason, so that demand is higher for the same good when it is shipped by air. Some of the reasons for such a preference are analyzed by Evans and Harrigan (2005) and Harrigan and Venables (2006), but for the purposes of this model I will simply suppose that utility is higher for goods that arrive by air. Let the set of goods shipped by air be *A,* with measure also given by *A*. Subutility for imports is

$$U(x(z)) = \int_{z \in A} a \ln x(z) dz + \int_{z \notin A} \ln x(z) dz \tag{2}$$

where $a > 1$ is the air-freight preference. The resulting demand functions are generalizations of constant-expenditure-share Cobb-Douglas:

$$x(z) = \frac{a}{aA + (1-A)} \cdot \frac{\alpha L_1}{\omega(z) p(z)} \qquad z \in A$$

[2] FOB stands for "free on board", and refers to the price of the good before transport costs are added. CIF stands for "cost, insurance, and freight", and refers to the price after transport costs have been added.

$$x(z) = \frac{1}{aA + (1-A)} \cdot \frac{\alpha L_1}{\tau p(z)} \qquad z \notin A \tag{3}$$

The relevant prices are inclusive of transport costs, which will depend on where the good is produced and perhaps on weight.

Given these demands and the structure of transport costs, the next task is to determine the equilibrium location of production. Perfect competition ensures FOB price = unit cost, but there is a choice of shipping mode and consequent CIF price paid. When buying from location *c*, consumers are willing to pay for airfreight as long as the relative marginal utility from timely delivery exceeds the relative shipping cost, or

$$\frac{\omega(z)}{\tau_c} \le a, \qquad c = 2, 3 \tag{4}$$

Since FOB production costs are the same, competition among sellers means that they will ship by air just in case this inequality is satisfied. I choose parameter values so that this never happens for country 2 and sometimes does for country 3:

$$\begin{aligned} a\tau_2 &< \omega(z) & z &\in [0,1] \\ a\tau_3 &< \omega(z) & z &\in [\bar{z},1] \\ \omega(z) &\le a\tau_3 & z &\in [0,\bar{z}] \end{aligned} \tag{5}$$

The cutoff $\bar{z}$ is an endogenous variable which is determined by the relative cost of air and surface shipping in country 3 only, given implicitly by

$$a\tau_3 = \omega(\bar{z}), \tag{6}$$

and its determination is illustrated in Figure 1. Goods $z \in [\bar{z},1]$ will never be shipped by air, regardless of where they are produced, and I call these goods heavy. Light goods, $z \in [0,\bar{z}]$, will be shipped by air if they are produced in *3*, otherwise they will be shipped by surface from *2*. The boundary between heavy and light goods will change when surface or air transport costs change, but it does not depend on comparative cost advantage, since it reflects only the decision facing a producer in one country.

Production location and shipping mode are determined jointly. Define relative surface transport costs, relative wages, and relative unit labor requirements respectively as

$$w \equiv \frac{w_2}{w_3}, \quad \tau \equiv \frac{\tau_2}{\tau_3}, \quad b(z) = \frac{b_2(z)}{b_3(z)} \tag{7}$$

For heavy goods, consumers in *1* buy from the lowest cost source, where costs are inclusive of wages and transport costs. Therefore, goods are produced in *2* if and only if

$$\tau_2 b_2(z) w_2 \leq \tau_3 b_3(z) w_3$$

or

$$\tau w b(z) \leq 1 \qquad z \in [\bar{z}, 1]. \tag{8}$$

For light goods, we know that if they are produced in *3* they'll be shipped by air, so the relevant cost comparison is between surface in *2* and air in *3*. But production cost is not the only consideration, since consumers are willing to pay more for goods shipped by air. The relevant cost comparison needs to be adjusted for this, and becomes

$$\tau_2 b_2(z) w_2 \leq \frac{\omega(z) b_3(z) w_3}{a} \qquad z \in [0, \bar{z}],$$

so production takes place in *2* if and only if

$$\frac{\tau_2 w b(z)}{\omega(z)} \leq \frac{1}{a} \qquad z \in [0, \bar{z}] \tag{9}$$

I will treat labor productivity in good z as a random variable, and I adopt the modeling strategy of Eaton and Kortum (2002). I simplify the Eaton-Kortum framework by focusing on just two countries that have identical distributions of labor productivity (the inverse of the unit labor requirement) drawn from a Fréchet distribution with parameters $T > 0$ and $\theta > 1$. With this distribution, the log of productivity has mean $\frac{\gamma + \log T}{\theta}$ and standard deviation $\frac{\pi}{\theta\sqrt{6}}$, so that smaller values of θ imply greater dispersion in productivity[3].

With random productivity, the low-cost producer is probabilistic. Adapting Eaton and Kortum's equation (8) for my purposes gives a particularly simple expression for the probability that country 2 is the supplier of heavy good z:

$$\pi_2^H(z) = \pi_2^H = \frac{(w_2\tau_2)^{-\theta}}{(w_2\tau_2)^{-\theta} + (w_3\tau_3)^{-\theta}} = \frac{1}{1 + (w\tau)^{\theta}} \qquad z \in [\bar{z}, 1] \tag{10}$$

This expression is quite intuitive: the probability that country 2 will supply any given heavy goods is decreasing in 2's relative wages and transport costs. The problem is slightly more complex for lightweight goods for two reasons. The first is that country 3's optimal shipping mode for lightweight goods is air, and the transport cost for these goods depends on weight. The second is that consumers in 1 are willing to pay a premium $a > 1$ for goods shipped by air. Using equation (9) with the Fréchet distribution for productivities implies that the probability that country 2 is the supplier of light good z is

$$\pi_2^L(z) = \frac{(\tau_2 w_2)^{-\theta}}{(\tau_2 w_2)^{-\theta} + \left(\frac{\omega(z) w_3}{a}\right)^{-\theta}} = \frac{1}{1 + (w\tau)^{\theta}\left(\frac{a\tau_3}{\omega(z)}\right)^{\theta}} \qquad z \in [0, \overline{z}) \qquad (11)$$

The term $\left(\frac{a\tau_3}{\omega(z)}\right)^{\theta}$ in equation (11) is strictly greater than one, which implies $\pi_2^L(z) < \pi_2^H$ for all $z \in [0, \overline{z})$. This result says that country *2* has a greater chance of supplying heavy goods than lightweight goods, and the lighter the good the lower the chance that *2* will be the supplier. The law of large numbers implies that for any interval of goods the average probability will be the share of goods supplied by country 2, so I'll refer to the π's from now on as market shares. The market shares for country 3 are just one minus the shares for country 2:

$$\pi_3^H = \frac{1}{1 + (w\tau)^{-\theta}}, \qquad \pi_3^L(z) = \frac{1}{1 + (w\tau)^{-\theta}\left(\frac{a\tau_3}{\omega(z)}\right)^{-\theta}} \qquad (12)$$

Figure 2 illustrates equations (10) and (11). Country *2*'s market share is increasing in the weight of the good $\omega(z)$ for all $z \in [0, \overline{z})$. In this range, if a good is supplied by country *2* it is sent by surface at a cost of τ_2 while if it is supplied by *3* it is sent by air at a cost of $\omega(z)$. For heavy goods $z \geq \overline{z}$, both countries use surface transport, and country *2* has a transport cost advantage since $\tau_3 > \tau_2$. Equation (10) implies that if wages are the same, country *2* will have a greater than 50% market share in heavy goods, but that within heavy goods *2*'s market share is constant.

To close the model I make wages endogenous. Factor market clearing requires that FOB

---

[3] In terms of the Eaton-Kortum model, I assume that both countries have the same absolute advantage parameter T_c. The constants are $\gamma = 0.577...$ and $\pi = 3.14159....$ See Eaton and Kortum

export revenue equals national income in country *2* and *3*. For both countries, FOB revenue from good *z* is the probability that it produces the good times country 1's CIF expenditure on that good, divided by the iceberg transport cost. Expenditure levels for good *z* are found by multiplying equations (3) by *p(z)*. Total expenditure on light goods is then the integral of expenditure on each good over the range $\left[0,\bar{z}\right)$, and expenditure on heavy goods is the integral over the range $\left[\bar{z},1\right]$. Using these expenditure levels and the probabilities from (10) and (11), the factor market clearing condition for country 2 becomes

$$w_2L_2 = \frac{\alpha L_1}{aA+(1-A)}\left[\int_0^{\bar{z}} \frac{\pi_2^L(z)}{\tau_2}dz + \int_{\bar{z}}^1 \frac{\pi_2^H}{\tau_2}dz\right]$$

$$= \frac{\alpha L_1}{aA+(1-A)}\left[\frac{1}{\tau_2}\int_0^{\bar{z}} \frac{1}{1+(w\tau)^\theta\left(\frac{a\tau_3}{\omega(z)}\right)^\theta}dz + \frac{1}{\tau_2}\frac{1-\bar{z}}{1+(w\tau)^\theta}\right] \qquad (13)$$

Similarly for country 3, export revenue is the sum of FOB revenue from air- and surface-shipped goods:

$$w_3L_3 = \frac{\alpha L_1}{aA+(1-A)}\left[a\int_0^{\bar{z}} \frac{\pi_3^L(z)}{\omega(z)}dz + \int_{\bar{z}}^1 \frac{\pi_3^H}{\tau_3}dz\right]$$

$$= \frac{\alpha L_1}{aA+(1-A)}\left[a\int_0^{\bar{z}} \frac{\omega(z)^{-1}}{1+(w\tau)^{-\theta}\left(\frac{a\tau_3}{\omega(z)}\right)^{-\theta}}dz + \frac{1}{\tau_3}\frac{(1-\bar{z})}{1+(w\tau)^{-\theta}}\right] \qquad (14)$$

The market clearing equations (13) and (14) along with equation (6) that defines $\bar{z}$ are three equations in the three unknowns w_2, w_3, and $\bar{z}$. With a solution to these three equations, the other endogenous variables of the model (national income and trade flows) are obtained by substitution.

The three equation system given by equations (6), (13) and (14) is highly nonlinear but fairly simple economically. Finding an analytical solution for equilibrium wages is impossible, as the integrals in (13) and (14) can not be evaluated analytically, but numerical solution is

---

(2002) for more on the Fréchet distribution and its interpretation.

straightforward using a convenient functional form for $\omega(z)$,

$$\omega(z) = \beta \tau_3 a^{1+z} \tag{15}$$

where the shift parameter β has a range of $[a^{-1}, 1]$. Recall that the condition for airfreight to be profitable for country 3 in good z is $\omega(z) \le a\tau_3$. For low values of β air freight is always profitable for country 3,

$$\beta = a^{-1} \rightarrow \qquad \omega(0) = \tau_3 \qquad \omega(1) = a\tau_3$$

while for high values it is never profitable:

$$\beta = 1 \rightarrow \qquad \omega(0) = a\tau_3 \qquad \omega(1) = a^2\tau_3$$

Substituting (15) into (6) gives the solution for $\bar{z}$:

$$\bar{z} = -\frac{\log \beta}{\log a} \in [0,1]$$

I solve the model for a numerical example, with $a = \theta = \tau_3 = 2$, $\tau_2 = 1$.

2.2 Comparative statics: wages and specialization when air freight gets cheaper

Because the model is highly stylized and solved for a numerical example, the equilibrium values of the endogenous variables are of little intrinsic interest. The interest of the model lies in its comparative static predictions about how wages and specialization change with changes in the parameters of the model. As noted in the introduction, the long-term trend is for air transport costs to decline relative to the cost of surface shipping (Chart 1). In this section, I derive the implications of such a drop in the relative cost of air shipping, which I model as a proportionate shift down in the cost of air transport (a fall in the shift parameter β).

Figure 3 shows that falling air transport costs expand the range of goods which are potentially shipped by air. The increase in $\bar{z}$ creates excess supply for country 2's labor, as some goods formerly produced in 2 are now profitable to produce in 3 and send by air. In the new equilibrium relative wages in 2 decline, and the resulting effects on market shares are illustrated in Figure 4. Country 2 increases its market share in all heavy goods, where 2's now-lower wage improves its competitiveness, and loses market share in light goods, where the lower cost of air shipping more than offsets the drop in 2's wages. Thus in equilibrium distance matters *more* to specialization rather than less when some transport costs fall, in the sense that market shares across goods are more strongly correlated with relative distance.

Equilibrium wages as a function of the cost of air shipment are illustrated in Figure 5. As expected, a fall in air freight costs (declining β) lowers the wage of *2* in both absolute and relative terms. Surprisingly, the initial effect of a decline in air freight costs on w_3 is negative. This is an instance of immiserizing technological improvement: the increased supply of goods from *3* lowers their price by more than the improvement in technology. As technology improves further, this terms of trade effect is outweighed by the efficiency gain on inframarginal goods, so w_3 increases. This result is partly an artifact of the assumption of Cobb-Douglas expenditure by country 1, and with a more elastic aggregate demand for imports the negative terms of trade effect of technological improvement would diminish.

Whatever the effect on the absolute level of wages in *3*, lower air freight costs inevitably lower real wages in *2*. This happens because *2* faces greater competition from *3* but has no use for the improved air shipping technology. The unambiguous winner is country *1*, which gets lower prices on all its imports from *2* and gets a wider range of air shipped goods from *3*. In the case where w_3 actually falls, country *1* gets more than 100% of the global welfare gain from improved technology: *1* gets both lower prices on all the goods it imports by surface and a wider selection of air shipped goods.

2.3 Model extensions: bilateral trade in a multicountry world

The model of sections 2.1 and 2.2 analyzes the competition between two countries to serve the market of a third. In formulating the model I simplified the model of Eaton and Kortum (2002) by reducing the number of countries to three and breaking symmetry, and I extended their analysis by making shipping costs endogenous and heterogeneous across goods. This is the simplest setting necessary to illustrate the underlying economic mechanisms, but is excessively simple for organizing our thoughts about how relative transport costs across countries and goods determine bilateral trade. Here I sketch an extension of my model that allows for many countries and treats them symmetrically, as in Eaton-Kortum. The purpose of the extension is to derive predictions about bilateral trade, in preparation for the data analysis in the second half of the paper.

Let there be *N* countries, and retain the simplifying assumption that they all have identical *ex ante* productivity distributions over the same continuum of goods. As in Eaton-Kortum, each country can potentially trade with any other, and transportation costs differ by country and trade route. As in the previous sections, I will focus on imports by country *1*, but I will not need to specify the details of demand in any country.

Producers in country c can choose either airfreight or surface shipment when shipping to country 1. Surface shipping costs from c to 1 are given by τ_c and airfreight costs are given by $\omega_c(z)$ [4]. Country c producers will send lighter goods by air and heavier goods by ship, with the cutoff good $\bar{z}_c$ given by equation (6), with c subscripts on τ and $\omega(z)$. Unlike what was assumed above, here I do not assume an interior solution for $\bar{z}_c$, so I do not rule out the case that c will use a single shipping mode for all exports to 1.[5] Define the premium-adjusted minimum shipping cost $t_c(z)$ as

$$t_c(z) = \min\left[\tau_c, \frac{\omega_c(z)}{a}\right]$$

In the Eaton-Kortum model, the probability that country c will supply a given market is the same for all goods (their equation (8)). In the current model, the probability varies, and will depend on $t_c(z)$ for all countries. With this modification the Eaton-Kortum logic goes through otherwise unchanged, so the probability that country c will supply good z to country 1 is

$$\pi_c(z) = \frac{\left(w_c t_c(z)\right)^{-\theta}}{\sum_{j=1}^{N}\left(w_j t_j(z)\right)^{-\theta}} = \frac{\left(w_c t_c(z)\right)^{-\theta}}{\Phi(z)} \qquad (16)$$

The summation in the denominator $\Phi(z)$ in (16) includes country 1, which reflects the fact that good z might be produced domestically rather than imported.

Closing the model requires specification of the full-employment conditions, conditions which will be analogous to (13) and (14). For my purposes here, though, it is more interesting to analyze how the market shares given by (16) vary across countries c and goods z.

For any pair of countries b and c that export to 1, (16) implies that their relative probability of exporting good z is

$$\frac{\pi_c(z)}{\pi_b(z)} = \left(\frac{w_c t(z)_c}{w_b t(z)_b}\right)^{-\theta} = \left(\frac{w_b t(z)_b}{w_c t(z)_c}\right)^{\theta} \qquad (17)$$

[4] If I were fully specifying the multicountry model, transport costs would have two subscripts, one each to denote the exporting and importing country. I omit the implicit 1 subscript here to reduce notational clutter.

[5] If $\bar{z}_c = 0$, c always uses surface shipment. If $\bar{z}_c = 1$, c ships all goods by air.

Equation (18) has the perfectly obvious implication that in head-to-head competition with *b*, *c*'s chances of success are increasing in *b*'s wages and transport costs and decreasing in *c*'s own.

How does this competition between *b* and *c* look if we compare across goods? The answer depends on how each country's optimal transport cost varies across goods. The case where *c* never finds it optimal to use air shipment (that is, $\bar{z}_c = 0$) while *b* finds it optimal for a range of lighter goods ($\bar{z}_b \in (0,1)$) is the case analyzed in sections 2.1 and 2.2, and illustrated in Figure 2. More generally, suppose that *c* is closer to *1* than *b* is, so that *b*'s air and surface transport costs are lower than *b*'s for all goods. One possible configuration of costs is illustrated in Figure 6, with relative market shares shown in Figure 7. Figure 6 is drawn on the assumption that increasing distance raises the cost of surface shipment relative to air shipment, so that $\bar{z}_c < \bar{z}_b$. This is a realistic assumption, since the speed advantage of airplanes over surface transit is increasing in distance. It is also an assumption whose implications will be validated in the empirical analysis below.

The implications for relative market shares are illustrated in Figure 7. For the lightest goods, $z \in [0, \bar{z}_c]$, country *c* will have low relative market share, since both countries ship these goods by air and *c*'s proximity advantage over *b* is not very important. For goods of intermediate weight such that *c* ships by surface and *b* ships by air, $z \in (\bar{z}_c, \bar{z}_b]$, *c*'s market share is increasing in *z*. Finally for the heaviest goods, $z \in (\bar{z}_b, 1]$, *c* has its greatest competitive advantage relative to faraway *b*.

2.4 Model extensions: relaxing some simplifying assumptions

In this subsection I discuss the implications of relaxing some of the simplifying assumptions (about technology, demand, and transport costs) that were made in the previous subsections.

So far I have assumed that surface transport costs don't depend on weight. This assumption is empirically false of course, but it is not analytically important. What matters for the mechanism of the model is that the relative cost of air to surface transport is increasing in weight. The commonplace observation that light weight goods are more likely to be shipped by air than are heavy goods is enough to validate this assumption. In terms of Figures 1, 3, and 6, the restriction is that the slope of $\tau(z)$ is less than the slope of $\omega(z)$.With this weaker assumption, the general equilibrium sorting mechanism will work exactly as described above, with nearby countries tending to specialize in heavier goods which are shipped by surface.

A key element of the model is that consumers are willing to pay a premium for air shipment. This is a reduced form assumption which can be microfounded as a demand for timely delivery, as discussed in Evans and Harrigan (2005) and Harrigan and Venables (2006). As argued in those papers, though, timely delivery can be guaranteed by proximity between supplier and customer, and Evans-Harrigan show that the interaction between proximity and the demand for timely delivery is a key competitive advantage for Mexico in serving the U.S. market. In terms of the present model, this suggests a re-specification of demand, where the premium for timely delivery a can be earned by nearby suppliers who deliver by surface transport or by faraway suppliers who ship by air. This re-specification would have two effects. First, it would guarantee that nearby countries would never use airplanes (as assumed in Sections 2.1 and 2.2). Second, it would slightly change the demand functions and market clearing conditions for nearby countries, since they would now be earning higher prices. However, specifying the model in this alternative way would have no impact at all on the key predictions, whether comparative static, across countries, or across goods.

A more substantive assumption is that the air freight utility premium a is the same for all goods. Evans and Harrigan (2005) argue that there is variation in the demand for timely delivery across products, and it is this variation across goods (rather than variation in shipping costs, which they assume away) which is the source of comparative advantage for nearby countries in their paper. Allowing for heterogeneity in a across goods in the model of this paper would break the simple relationship between weight and comparative advantage. Conceptually the extension is straightforward, and I sketch it here.

Let $a(z) \geq 1$ be the utility premium for air shipment (or timely delivery) of good z, and let $\omega(z) > 1$ be the iceberg air shipment cost for z as before. Now order goods z in increasing order of the ratio $\frac{\omega(z)}{a(z)}$, and implicitly define the cutoff value $\bar{z}$ by

$$\frac{\omega_c(\bar{z})}{a(\bar{z})} = \tau_c \tag{19}$$

The optimal shipping mode choice is now quite similar to what it was before: send the good by air iff $z \leq \bar{z}$. Figures 1, 3, 6 and 7 would be unchanged, although the interpretation is slightly different, since the horizontal axis now indexes weight relative to the demand for timely delivery rather than just weight. The implications for comparative advantage are that faraway exporters will specialize in goods that are light and that have a high demand for timely delivery. Combining this extension with the extension discussed in the immediately preceding paragraphs leads to a hybrid prediction about specialization in goods where timely delivery is important: the heavy ones will be produced in nearby countries and sent by surface, and the light ones will be produced in remoter locations and sent by air.

Finally, the assumption that all countries have the same *ex ante* distribution of productivities can be relaxed somewhat. It is key to analytical tractability that all countries have the same dispersion parameter θ, but they may differ in their absolute level of productivity (what Eaton and Kortum denote T_c). The reason is that θ governs comparative advantage while T_c affects only the absolute level of wages and national income; see Eaton and Kortum (2002) for more on this.

2.5 The model's prediction for trade data

For any given level of wages, the model delivers predictions about the cross-section of goods imported by *1*. It is these predictions which will be the focus of the empirical analysis, with the United States as the importing country.

The import data used in this paper are collected by the U.S. Customs Service and reported on CD-ROM. For each year from 1990 to 2003, the raw data include information on the value, quantity (usually number or kilograms), and weight (usually in kilograms) of U.S. imports from all sources. The data also include information on tariffs, transport mode and transport fees, including total transport charges broken down by air, vessel and (implicitly) other, plus the quantity of imports that come in by air, sea, and (implicitly) land.[6] The import data are reported at the 10-digit Harmonized System (HS) level, which is extremely detailed, with over 14,000 codes in 2003.

The import data does not report prices, but since it reports both value and quantity I can calculate unit values, defined as the dollar value of imports per physical unit. Since shipping costs depend primarily on the physical characteristics of the good rather than on its value, low

[6] "other" transport modes include truck and rail, and are used exclusively on imports from Mexico and Canada.

value goods will be "heavy" in the sense of having a higher shipping cost per unit of value[7]. For example, consider shoes. Quantities of shoes are reported in import data, and the units are "number" as in "number of shoes". Expensive leather shoes from Italy and cheap canvas sneakers from China weigh about the same, but the former will have a much higher unit value. In the context of the model, Italian leather shoes are "lighter" than Chinese fabric sneakers, in the economically relevant sense that the former have lower transport costs as a share of value. It is important to keep in mind that it is meaningless to compare unit values when the units are not comparable: dollars per number of shoes is not comparable to dollars per barrel of oil or dollars per square meter of fabric.

The first prediction of the model has already been illustrated in Figures 1 and 2: country *2* will have lower market share in light-weight goods, and these light goods will be shipped by air when produced by *3*. This prediction was generalized in the discussion of the *N* country model, where it was shown that nearby countries will have higher market shares in heavy goods and faraway countries will specialize in light goods. The model's prediction can then be translated into a prediction about unit values: within a given product category, nearby countries will tend to specialize in low-value goods. High-value goods will tend to be produced in more distant locations and will be shipped by air.

The model's predictions about specialization can be expressed in two ways. The first is in terms of relative quantities: nearby countries have a higher market share in heavy goods than lightweight goods, a prediction illustrated in Figures 2 and 7. In principle one could test this prediction, but in practice there are insuperable obstacles. It would require classifying every good by its weight, which is impossible for two reasons: a majority of import records do not report weight, and when they do there is great heterogeneity in weight even within 10-digit HS codes. Even if these measurement issues were not decisive, there would be an intractable identification issue, which is that there are many reasons unrelated to weight why a country might have a high market share in a particular good. For example, Canada has a very high market share in lumber and wood products, which have relatively low value per kilo; since Canada is adjacent to the United States this would seem to support the model. But of course it would be absurd to explain trade in lumber while ignoring the fact that Canada is covered in

[7] The relationship between shipping cost and shipment value is estimated by Hummels and Skiba (2004), Table 1. They find that shipping costs increase less than proportionately with price.

trees.

Consequently, I will instead focus on the model's predictions about unit values of goods which are actually imported. For a particular group of goods, the model predicts a relationship between unit values and distance: imports from nearby countries should have lower unit values than imports from more distant countries. That is, the deviation of unit values from the group mean should be positively related to distance from the U.S.

To state this a bit more formally, suppose that a given HS code contains goods of varying weights, which we can order from lightest to heaviest. According to the model, each good within the HS code will be provided by one country, with the winner of the good-by-good competition being stochastic. Thus, a country that exports in this code must have won at least one competition. Conditional on exporting in this code, nearby countries are more likely to have won competitions among the heavier goods, and more distant countries are likely to have won in the lighter goods.

Note that this formulation of the model's prediction effectively controls for other, non-weight related determinants of specialization. This is because the prediction about the cross-section of unit values within an HS code is conditional on countries exporting in that category at all.

3 Airplanes and trade: empirical evidence

I aggregate the 10-digit import data for analysis in various ways. For most of the descriptive charts and tables, I work with a broad aggregation scheme that updates Leamer's (1984) classification, which is reported in Table 1. For the regression analysis, I work with the individual 10-digit HS categories.

3.1 Data description

Table 1 illustrates the great heterogeneity in the prevalence of air freight for U.S. imports, as well as some important changes over the sample. Many products come entirely or nearly entirely by surface transport (oil, iron and steel, road vehicles) while others come primarily by air (computers, telecommunications equipment, cameras, medicine). Scanning the list of products and their associated air shipment shares hints at the importance of value to weight and the demand for timely delivery in determining shipment mode. Charts 2 to 8 illustrate the variation in air freight across regions and goods (the regional aggregates are defined in Table 2, while the product aggregates correspond to the headings in Table 1). Chart 2 shows that about a quarter of US (non-oil) imports arrived by air in 2003, up from 20% in 1990 (for brevity, in what follows I'll call the proportion of imports that arrive by air "air share"). Excluding NAFTA, the non-oil air share was 35% in 2003. Chart 3 shows that this average conceals great regional variation, which is related to distance: essentially no imports come by air from Mexico and Canada, while Europe's air share is almost half by 2003, up from under 40% in 1990. East Asia's air share increased by about half from over the sample, from 20 to 30%. The airshare from the Caribbean and South America was about one-fifth over the sample. Chart 4 shows that air shipment is particularly high in labor-intensive manufactures, machinery, and chemicals (as Table 1 shows, the capital-intensive aggregate is mainly steel and other metals, which are very heavy). The biggest increase in air share came in chemicals, which (as Table 1 reveals) is accounted for by the increasing importance of pharmaceuticals, which had a 65% air share in 2003.

The remaining Charts 5 through 8 show the evolution of air share by major regions and product aggregates. The most notable fact about Chart 5 is the sharp increase in machinery's air share from East Asia, to levels similar to Europe's by 2003. Chart 6 shows a drop in the air share for labor intensive manufactures (the biggest component of which is apparel) from the Caribbean and Western and South Asia. This has to do with the phenomenon documented by Evans and Harrigan (2005): as apparel production moved to Mexico during the 1990s, the shift was concentrated in goods where timely delivery is important. Essentially, U.S. apparel retailers who wanted timely delivery replaced air shipments from South Asia and the Caribbean for surface shipments from next-door Mexico.

Heavy capital-intensive goods have not shown any increase in air-share from any source since 1990 (Chart 7). Finally, the shift toward air shipment in chemicals from major suppliers was world-wide, with the exception of nearby Caribbean suppliers (Chart 8).

Tables 3 and 4 illustrate how weighted average transport costs vary by region and product category. Not surprisingly, the products that have the highest air share, Machinery and Chemicals, have the lowest air freight costs. In some of the categories, the average transport cost is lower for air than overall, which of course reflects selection: very low value/weight items, which cost a lot to move even by ship, don't get put on planes.

3.2 Statistical results

The theory model of section 2 makes a number of predictions. The one I focus on here concerns the price of imports across source countries: the model predicts that imports from faraway countries will have higher f.o.b. prices per unit than goods shipped from nearby countries. Statistically, I investigate this by looking at variation in unit values across exporters within 10-digit HS categories. The econometric model I use is

$$v_{ict} = \alpha_{it} + \beta_t d_c + other\ controls + error \qquad (20)$$

where

v_{ict} = log unit value of imports of product *i* from country *c* in year *t*

α_{it} = fixed effect for 10-digit HS code *i* in year *t*

d_c = distance of *c* from United States

Note that import values are measured f.o.b, so they do not include transport charges. The model predicts $\beta_t > 0$ in equation (20): across exporters within a 10-digit commodity category, more distant exporters will sell products with higher unit values, controlling for other observable country-specific factors which might affect unit values. When the units are kilograms, then the prediction for unit values is a prediction about the value-weight ratio.

The fact that equation (20) uses only cross-exporter variation within each 10-digit HS code is the key modeling choice of the data analysis. The decisive advantage of using a within-product estimator is that it controls for which goods a country exports: if a country does not export product *i* to the US, then that country's distance from the US is (appropriately) irrelevant to the effect of distance on unit values within product *i*. Product fixed effects also control for differences in physical characteristics of products, making it possible to meaningfully pool information from microchips and potato chips.

The main disadvantages of using product fixed effects is that it prevents the data from saying anything about cross-product specialization, which is a central prediction of the model. To the extent that weight determines specialization across products, estimating equation (20) will miss such effects. As a consequence, the within-product specialization captured by equation (20) will be a lower bound on the importance of weight as a source of comparative advantage.

I measure distance by five indicator variables:

1. adjacent to the US (Mexico and Canada).

2. between 1 and 4,000 kilometers (Caribbean islands and the northern coast of South America).
3. between 4,000km and 7,800km (Europe west of Russia, most of South America, a few countries on the West Coast of Africa)
4. between 7,800km and 14,000km (most of Asia and Africa, the Middle East, and, Argentina/Chile)
5. over 14,000km (Australia/New Zealand, Thailand, Indonesia, Malaysia/Singapore)

In some of the regressions, I aggregate the distance classes into near (less than 4,000km) and far. I also include a dummy for if a country is landlocked.

There are many other factors that could affect unit values, and I control for some of these. *Other controls* include

1. trade cost variables (shipping cost and tariff, both measured as *ad valorem* percentages), which should have negative signs to the extent that trade costs are borne by producers.
2. macro indicators of comparative advantage (log aggregate real GDP per worker and log overall price level, both measured relative to the US, from the Penn-World Tables). My model is silent on how these aggregate measures might affect prices, but if more advanced countries specialize in more advanced and/or higher quality goods, we would expect positive effects of these variables on log unit values. Evidence of such effects is reported by Schott (2004), Hallak (2006) and Hummels and Klenow (2005).

Tables 5, 6, 7, and 8 report the results of estimating equation (20)[8]. For each year, log unit value is regressed on the controls as well as fixed effects for 10-digit HS codes. In the interest of reducing the quantity of numbers presented, I report results for only four selected years (1990, 1995, 2000, and 2003), although all regressions were estimated on all 14 years from 1990 to 2003 (complete results available on request). Each column shows results for a single year's regression, with *t*-statistics in *italics*.

The specifications in Tables 5 to 8 differ in the definition of the dependent variable and the scope of the sample. Tables 5 and 7 use the broadest definition of unit value, and include all observations for which units are reported, whether those units are number, barrels, dozens, kilos,

[8] All regressions are estimated by OLS, with product fixed effects. Note that standard errors are not adjusted for clustering by exporting country, despite the fact that cross-product correlation of errors within an exporter is *a priori* plausible. For an explanation, see the Appendix.

or something else. Tables 6 and 8 include only observations for which weight in kilos is reported, so the unit value in the Tables 6 and 8 regressions is precisely the value-weight ratio for all of the observations. Tables 5 and 6 include all available observations, while Tables 7 and 8 restrict the sample in two ways. First, very small and potentially erratic observations are eliminated by dropping all import records of less than $10,000, and/or that report just one unit. Secondly, to focus attention on manufactured goods, Tables 7 and 8 include only HS codes that belong to SITC 6 (manufactured goods), 7 (machinery and transport equipment), and 8 (miscellaneous manufactures). The differences in the coverage of the Tables are summarized here:

	definition of dependent variable	
	log unit value	log value/weight
all available observations	Table 5	Table 6
exclude small & non-manufactured obs	Table 7	Table 8

Tables 5 and 6 show that the effect of distance on unit values is large, robust, and statistically significant. The first column for each year in Tables 5 and 6 has a single indicator for distance greater than 4000km from the United States. As the first row of Table 5 shows, for the full sample unit values are between 19 and 37 percent higher when they come from more distant locations. The effect is even larger when the sample is restricted to observations with units in kilos, with the distance effect between 35 and 55 percent (first row, Table 6). The second four rows of Tables 5 and 6 break down distance into a larger number of categories, with Mexico/Canada as the excluded category. The effect of being less than 4000km but not adjacent to the US is positive and significant in each year in Table 6, and for all but one year in Table 5, but the effect is fairly small. The effect of being more than 4000km from the US is much larger, though it is not monotonic in distance, with larger effects in the 4000-7800km range than in more distant categories. The additional effect of being landlocked is also large, ranging between 15 and 48 percent across specifications in Tables 5 and 6.

The estimated effects of distance are invariably larger in Table 6 than in Table 5. This discrepancy is supportive of the model's predictions, since the dependent variable in Table 6 (log value/kilo), is more closely connected to the theory than the dependent variable in Table 6 (log unit value). To the extent that different units within a product have different weights (which they often do), one would expect a weaker connection between unit value and distance.

The non-monotonicity of the distance effect on unit values probably reflects imperfectly measured country characteristics that are correlated with distance, since the 4000-7800 range includes many of the most developed countries. The importance of development in affecting unit values was found in Schott (2004), Hallak (2006) and Hummels and Klenow (2005), and is confirmed here: a higher aggregate price level (which is associated with development) raises unit values with a large and significant elasticity, between 0.4 and 0.8, in every regression in Tables 5 and 6. The effect of aggregate productivity is inconsistent across specifications, but this merely reflects the very high correlation between aggregate productivity and price level.

Although the tariff and transport cost effects are not the focus of the paper, it is interesting that they are consistently estimated to be small, negative and mostly statistically significant (ranging from 0 to -0.016 across specifications). These negative effects are consistent with the US being a large market for most exporters, and are suggestive of a terms of trade gain from protection.

Tables 7 and 8 repeat the specifications of Tables 5 and 6 for a narrower sample, excluding non-manufacturing imports as well as very small observations (value of less than $10,000, and/or just one physical unit). The results are largely consistent with Tables 5 and 6, which confirms that the overall results are not driven by a small number of observations or by non-manufacturing SITC categories.

As noted above, the non-monotonicity of the distance effect is potentially troublesome, since it suggests that unmeasured country characteristics that are correlated with distance may be driving the estimated effect of distance on import unit values. To address this possibility, I first estimated a model with both exporter and product fixed effects:

$$v_{ict} = \alpha_{it} + \lambda_{ct} + other\ controls + error \qquad (21)$$

In this specification, the only "*other controls*" are the trade cost variables (shipping cost and tariff, both measured as *ad valorem* percentages), since these are the only controls that are not country-specific. Thus, the effect of all exporter-specific characteristics on log unit value, including distance from the US, are captured by the λ_{ct}'s in (21). I then regressed the estimated exporter fixed effects on log aggregate price level *p* and a measure of distance:

$$\hat{\lambda}_{ct} = \alpha_t p_{ct} + \beta_t d_c + error \qquad (22)$$

In equation (22), β answers the following question: "are the country-average differences in log unit value correlated with distance, after controlling for the level of development?" In (22) distance is measured by an indicator equal to 1 for distances greater than 4000km from the U.S., and also as log kilometers. To control for the first-stage estimation error in the $\hat{\lambda}_{ct}$'s, equation (22) was estimated by weighted least squares with weights equal to the reciprocal of the standard errors of the $\hat{\lambda}_{ct}$'s.

Equations (21) and (22) are estimated on the same four samples as reported in Tables 5 through 8, and the results are reported in Table 9. The table shows that there is a strong and consistent relationship between the exporter fixed effects and the distance indicator, with some evidence that the effect has grown stronger over time: in 1990 the effect of distance > 4000km raises import unit values between 25 and 34 percent, while in 2003 the effect is between 43 and 57 percent. In contrast, the correlation between distance measured continuously and the exporter fixed effects is essentially zero. This zero correlation is not particularly surprising in light of the theory: the sorting mechanism in the model is not continuous in distance either. Finally, the last four rows of Table 9 confirm the results of Schott (2004), Hallak (2006) and Hummels-Klenow (2005): there is a strong, robust relationship between import unit values and the level of development, which is proxied here by the aggregate price level[9].

A final empirical question concerns the choice between air and surface shipment by exporters. According to the model, the mechanism behind the specialization documented in Tables 5 through 8 is that remote exporters are more likely to ship goods by air. An alternative, though not competing, hypothesis is that high-quality goods are more likely to be shipped by air. To address this question, I estimate a discrete choice model for whether or not goods are shipped by air. Because of inherent differences in the physical characteristics of goods (bushels of wheat vs. auto parts), it is important to estimate such a model using only across-country variation within products. Defining the indicator variable $a_{ic} = 1$ if product i is shipped by air from exporter c, I estimate the following conditional fixed effects logit model for each year:

[9] Very similar results are obtained when log aggregate productivity is used as a proxy for development instead of log aggregate price level.

$$\Pr\left(a_{ict}=1\,|\,data\right)=\frac{\exp\left(\alpha_{it}+\beta_t d_c+other\ controls\right)}{1+\exp\left(\alpha_{it}+\beta_t d_c+other\ controls\right)} \tag{23}$$

where the notation is the same as for equation (20)[10]. Estimation is by maximum likelihood, and the results are reported in Table 10. Because of the product fixed effects α_{it}, products where all exporters choose the same mode (air or non-air) contribute nothing to the log likelihood function and so have no influence on the parameter estimates. Intuitively, if there is no variation within a product across exporters then that product can't tell us anything about how exporter characteristics affect the choice of shipping mode. The final line of Table 10 indicates that just over half of the products have some variation in shipping mode, and can thus be used in identifying the parameters of equation (23).

The parameter estimates reported in the top half of Table 10 are unsurprising, with distance and indicators of development both having positive and statistically significant effects on the probability of air shipment. The bottom panel of the table quantifies the importance of the estimated parameters by reporting the average estimated odds ratios implied by the model[11]. The odds ratios are the ratios of predicted probabilities under different configurations of the distance dummies. For example, the first number in bottom panel says that, in 1990, the average probability of air shipment from an exporter more than 4,000km from the U.S. was 4.9 times the probability of air shipment from an exporter less than 4,000km from the U.S. As expected, all the odds ratios are much bigger than one. The second row shows that even relatively nearby exporters were 7 times as likely to ship by air in 2003 as were Canada and Mexico. As was seen repeatedly in Tables 5 through 9, the largest distance effect is found for exporters between 4,000 and 7,800km from the US: goods are around 20 times more likely to be exported by air from these countries than from NAFTA. Curiously, the probability of air shipment falls off for the most extreme distances, though it remains much larger than the probability of air shipment from NAFTA.

[10] For a small number of product-exporter-year observations, the share of goods that arrive by air is neither zero nor one. In these cases I define a_{ict} = 1 if the share that arrives by air exceeds 0.5. The results are not sensitive at all to including this small number of observations.

[11] Because the distance indicators are not continuous, the derivative of the probability with respect to distance is not defined. As a consequence, marginal effects are not defined, which is why I report average odds ratios to quantify the effect of distance.

3.3 Summarizing the evidence

The goal of the statistical models of this section is to test the model's prediction that remote exporters will specialize in light-weight goods. Because it is not possible to adequately control for non-distance related determinants of specialization, the methodology has focused on within-category variation in import unit values. Thus the statistical model of equation (20) asks the question: if a country exports a good to the U.S., is the unit value of that good related to distance? The answer is emphatically yes, as shown in Tables 5 through 9. If we focus on the final year of the sample, we find that exports that arrive from destinations more than 4000km from the U.S. (that is, from sources other than Mexico, Canada, and the Caribbean) have import unit values between 31 and 50 percent higher than those from nearby sources. Looking just at observations where I can compute value per kilo, which is the measure of unit value most closely connected to the theory model, I find even larger effects, between 48 and 57 percent depending on the specification.

A related finding is that the effect of distance on unit values is non-monotonic, with the effect seemingly peaking in the distance range of 4000-7800km. Since this distance category includes Europe, the large estimated coefficients may be conflating the effect of distance and Europe's comparative advantage in producing high-quality goods. This is an important caveat for interpreting the size of the distance effect but does not overturn the strong relationship between distance and import unit value.

The results of Tables 5 through 9 confirm the importance of distance on unit values, but they don't say anything about the role of shipment mode choice. Table 10 fills in this gap, with the unsurprising finding that air shipment is strongly related to distance. Thus, we can conclude that the findings of Tables 5 through 9 are driven at least in part by the mechanism studied in the model: remote exporters specialize in lightweight goods which are shipped by air.

Lastly, the results confirm the results of Schott (2004), Hallak (2006), and Hummels-Klenow (2005) that there is an important relationship between import unit values and the level of development, which probably reflects a comparative advantage that rich countries have in high-quality goods.

4 Conclusion

This paper has focused on the interaction between trade, distance, product characteristics, and the choice of shipping mode. In the theory model, I showed how the existence of airplanes implies that countries that are far from their major export markets will have a comparative advantage in lightweight goods. This prediction is strongly supported by the data.

In the empirical section, I documented the heterogeneity across regions and goods of the prevalence of air shipment in US imports. The statistical analysis uncovered a large and robust relationship between distance and unit values: U.S. imports from remote suppliers have unit values on the order of a third higher than those from nearby countries. While strong, the econometric results probably understate the importance of relative distance on US trading patterns, since the analysis uses only cross-exporter variation within narrowly defined product categories. Cross-product specialization on the basis of weight is surely important as well, as indicated by the descriptive data analysis.

Distance is not dead, and the theory and empirical results of this paper suggest that it will not be expiring any time soon. The fall in the relative cost of air shipment implies that relative distance may become even more important in determining comparative advantage, as nearby countries increasingly trade heavy goods with each other while trading lighter goods with their more distant trading partners.

References

Deardorff, Alan, 2004, "Local Comparative Advantage: Trade Costs and the Pattern of Trade", University of Michigan RSIE Working paper no. 500.

Eaton, Jonathan and Samuel Kortum, 2002, "Technology, Geography, and Trade", *Econometrica 70(5)*: 1741-1779 (September).

Evans, Carolyn E., and James Harrigan, 2005, "Distance, Time, and Specialization: Lean Retailing in General Equilibrium" *American Economic Review 95(1)*: 292-313 (March).

Hallak, Juan Carlos, 2006, "Product Quality and the Direction of Trade", *Journal of International Economic*s 68 (1), pp. 238-265.

Harrigan, James, and A.J. Venables, 2006, "Timeliness and Agglomeration", *Journal of Urban Economics 59*: 300-316 (March).

Hummels, David, 1999, "Have International Transportation Costs Declined?", http://www.mgmt.purdue.edu/faculty/hummelsd/

Hummels, David, 2001a, "Time as a Trade Barrier", http://www.mgmt.purdue.edu/faculty/hummelsd/

Hummels, David, 2001b, "Toward a Geography of Trade Costs", http://www.mgmt.purdue.edu/faculty/hummelsd/

Hummels, David, and Peter Klenow, 2005, "The Variety and Quality of a Nation's Exports", *American Economic Review 95(3)*: 704-723 (June).

Hummels, David, and Alexandre Skiba, 2004, "Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture", *Journal of Political Economy 112(6)*: 1384-1402 (December).

Leamer, Edward, 1984, *Sources of International Comparative Advantage*, Cambridge, MA: MIT Press.

Limão, Nuno, and A.J. Venables, 2002, "Geographical Disadvantage: a Hecksher-Ohlin-Von Thunen Model of International Specialisation," *Journal of International Economics 58(2)*: 239-63 (December).

Redding, Stephen, and A.J. Venables, 2004, "Economic Geography and International Inequality", *Journal of International Economics 62(1)*: 53-82.

Schott, Peter K., 2004, "Across-product versus within-product specialization in international trade", *Quarterly Journal of Economics119(2):*647-678 (May).

Figure 1 - the air shipping decision

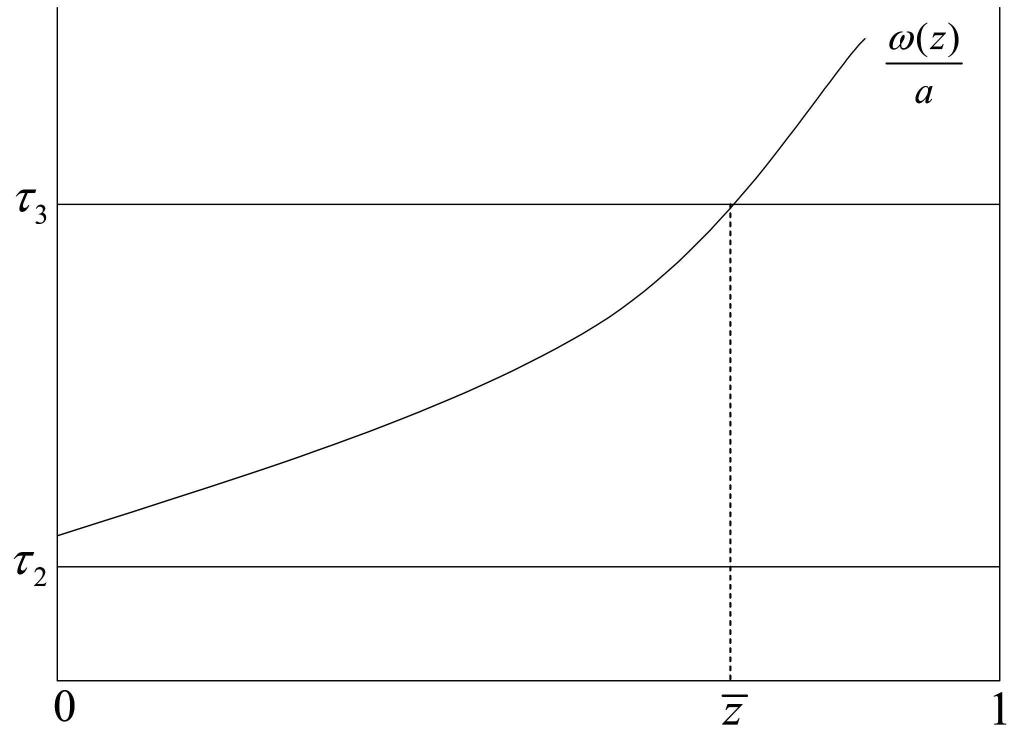


Figure 2 - Market shares for country 2

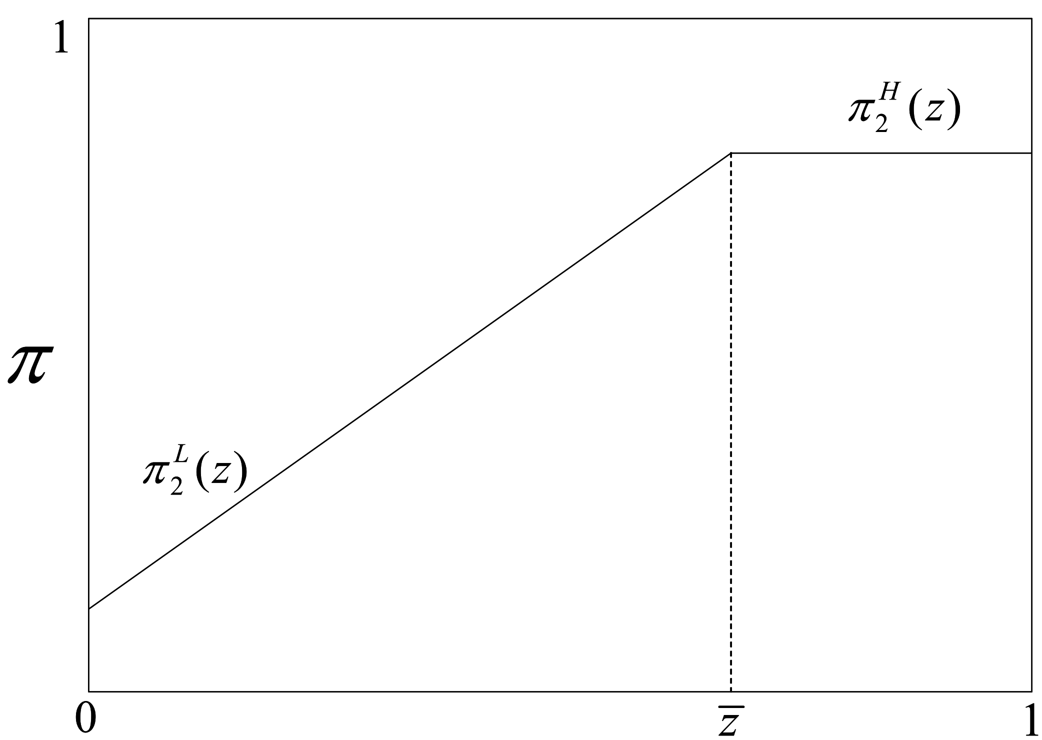


Figure 3 - Change in $\bar{z}$ when air freight costs fall

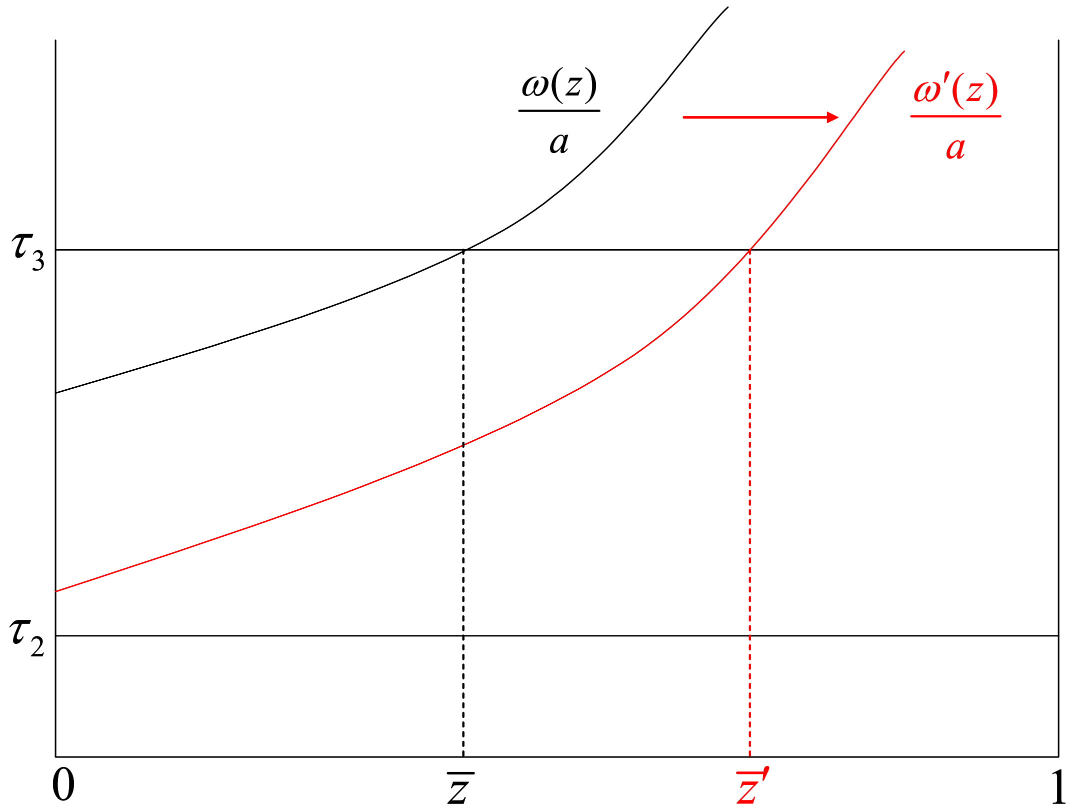


Figure 4 - Change in market shares when air freight costs fall

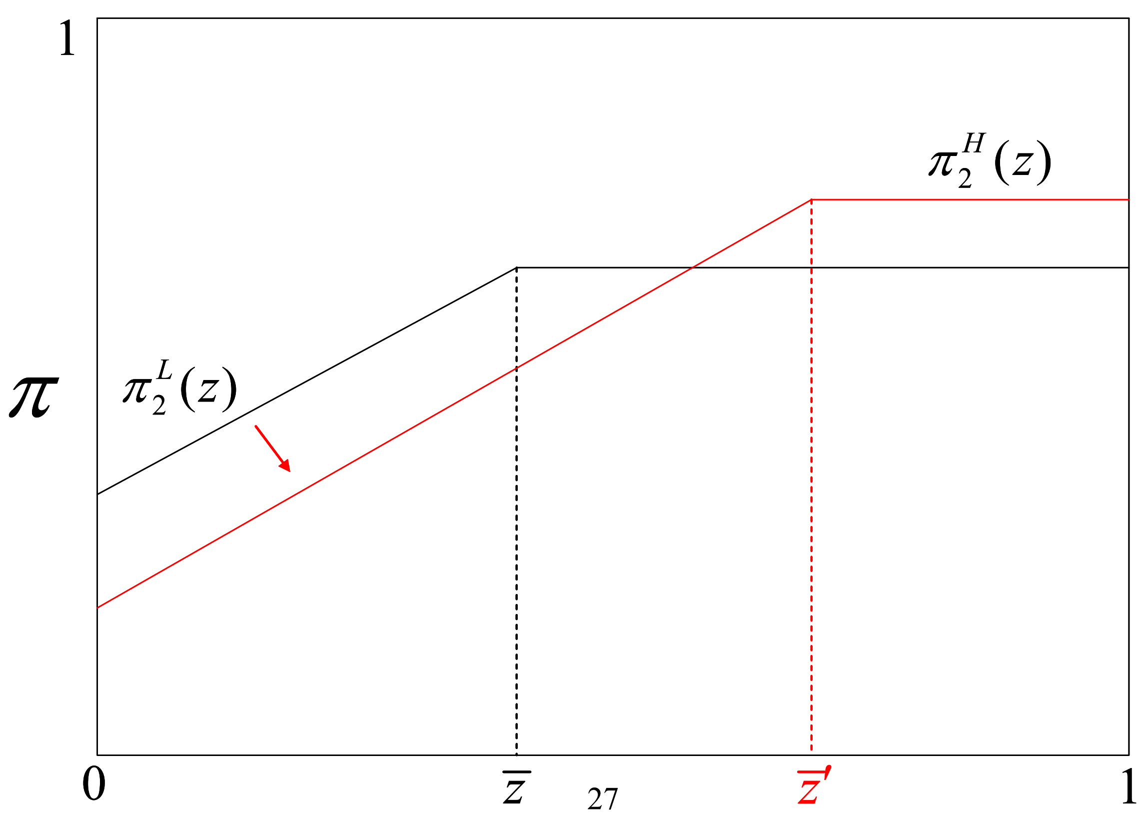
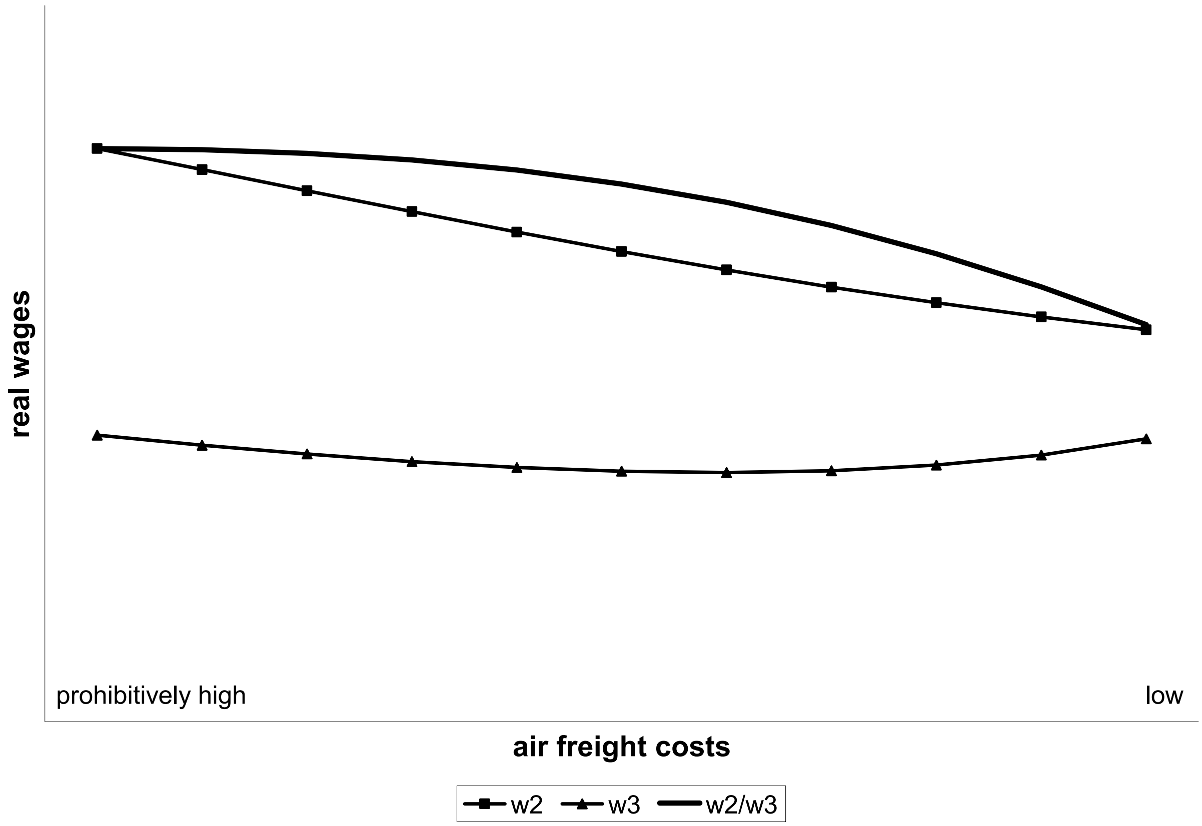


Figure 5- Wages as a function of air freight costs



Notes to Figure 5: Illustrates equilibrium wages as a function of air freight costs for a numerical example, with air freight costs varying from prohibitive (left axis) to low enough so that country 3 always uses air freight (right axis). Equilibria are the solutions to equations (6), (13), and (14) as the shift parameter β falls from 1 to a^{-1}. Parameter values are $a = \theta = \tau_3 = 2$, $\tau_2 = 1$, $L_2 = L_3$, the same as used to generate Figures 1 to 4 .

Figure 6- the air shipping decision in the *N* country model

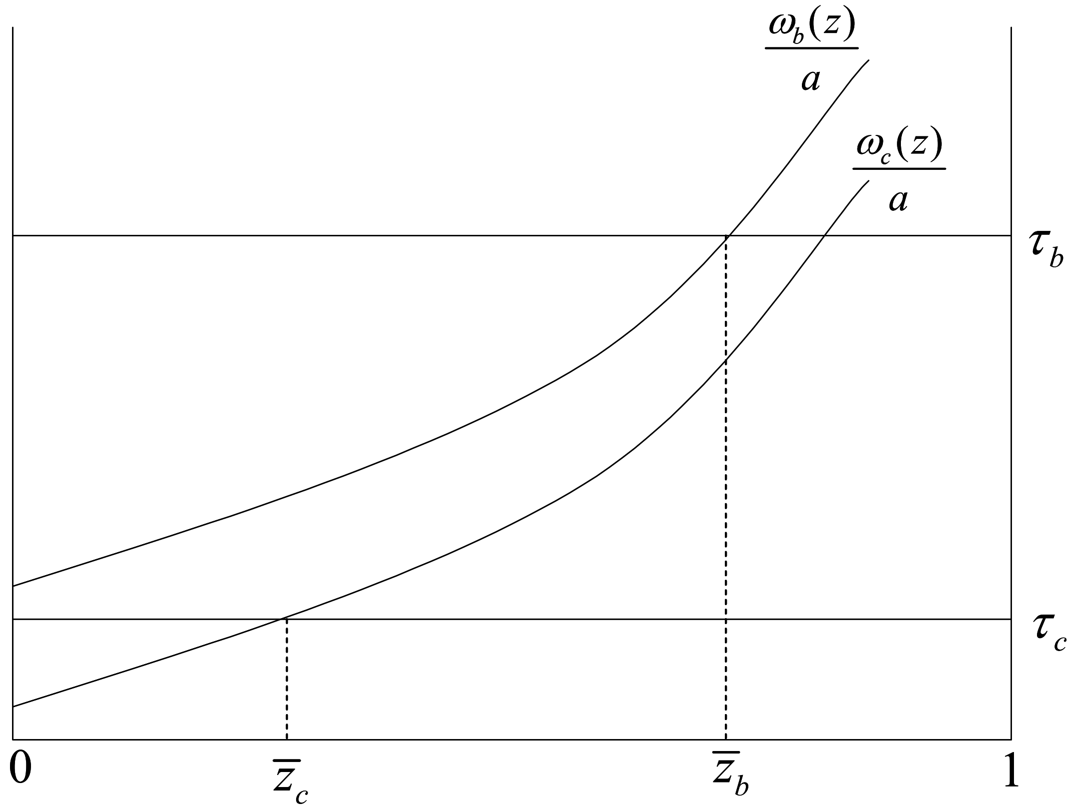


Figure 7- relative market shares

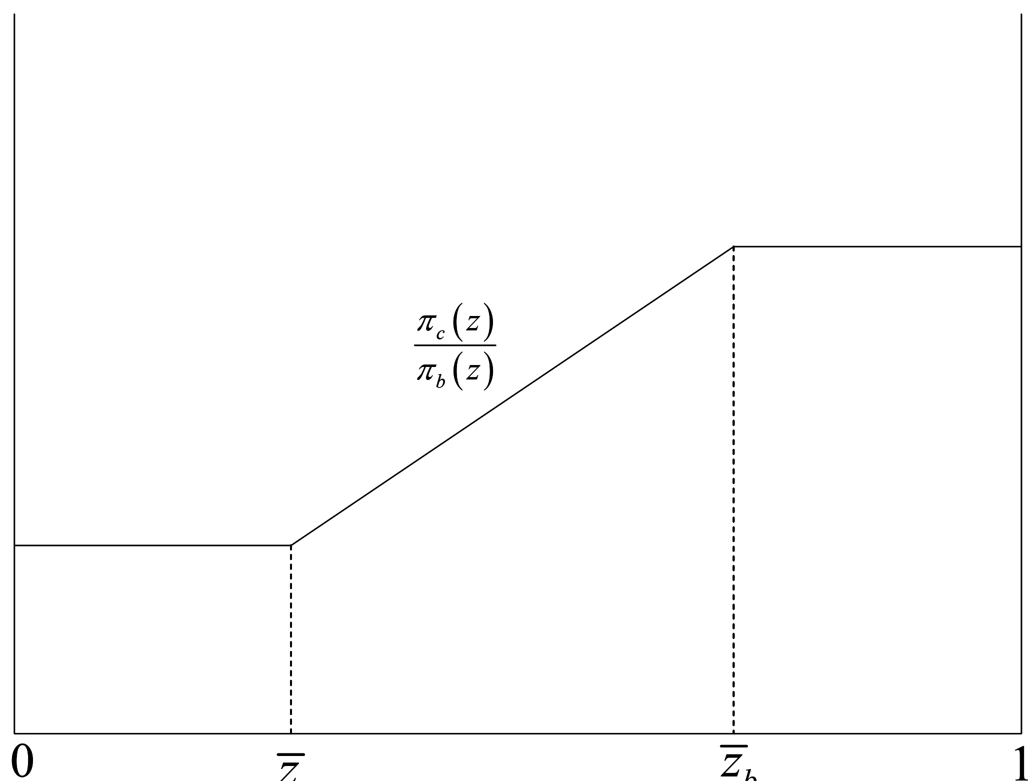
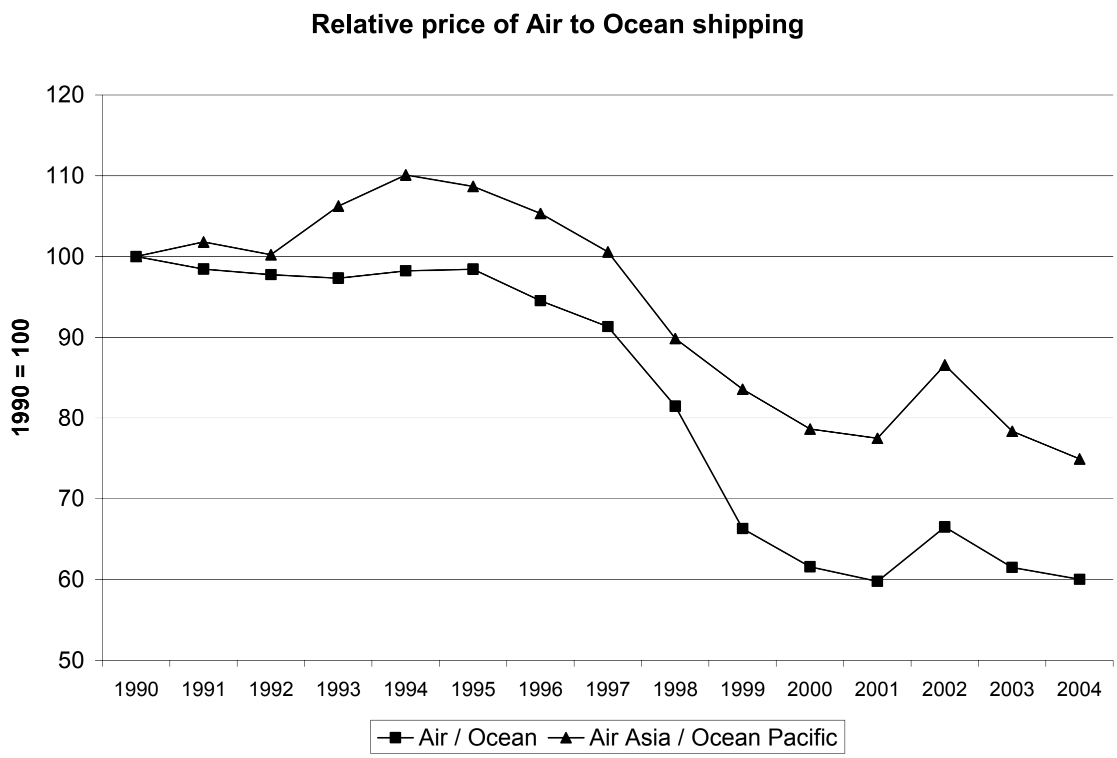


Chart 1



Notes to Chart 1: Data are price indices for U.S. imports of air freight and ocean liner shipping services from the Bureau of Labor Statistics, www.bls.gov/mxp. The "Air/Ocean" series divides all US imports of air freight services by all imports of ocean liner services, while the "Air Asia / Ocean Pacific" series divides the index for air freight imports from Asia by the index for ocean liner imports from the Pacific region.

Chart 2

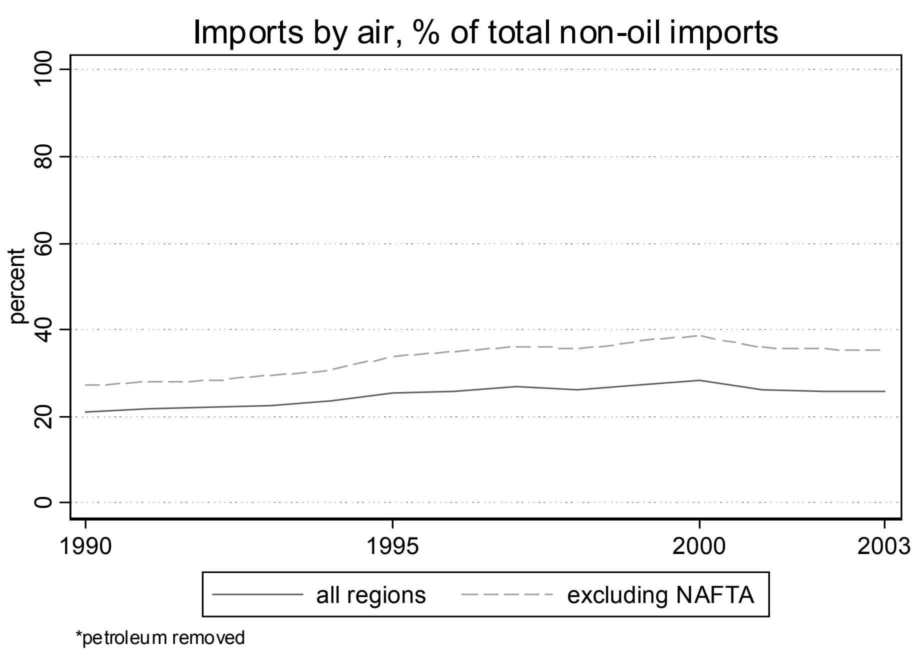


Chart 3

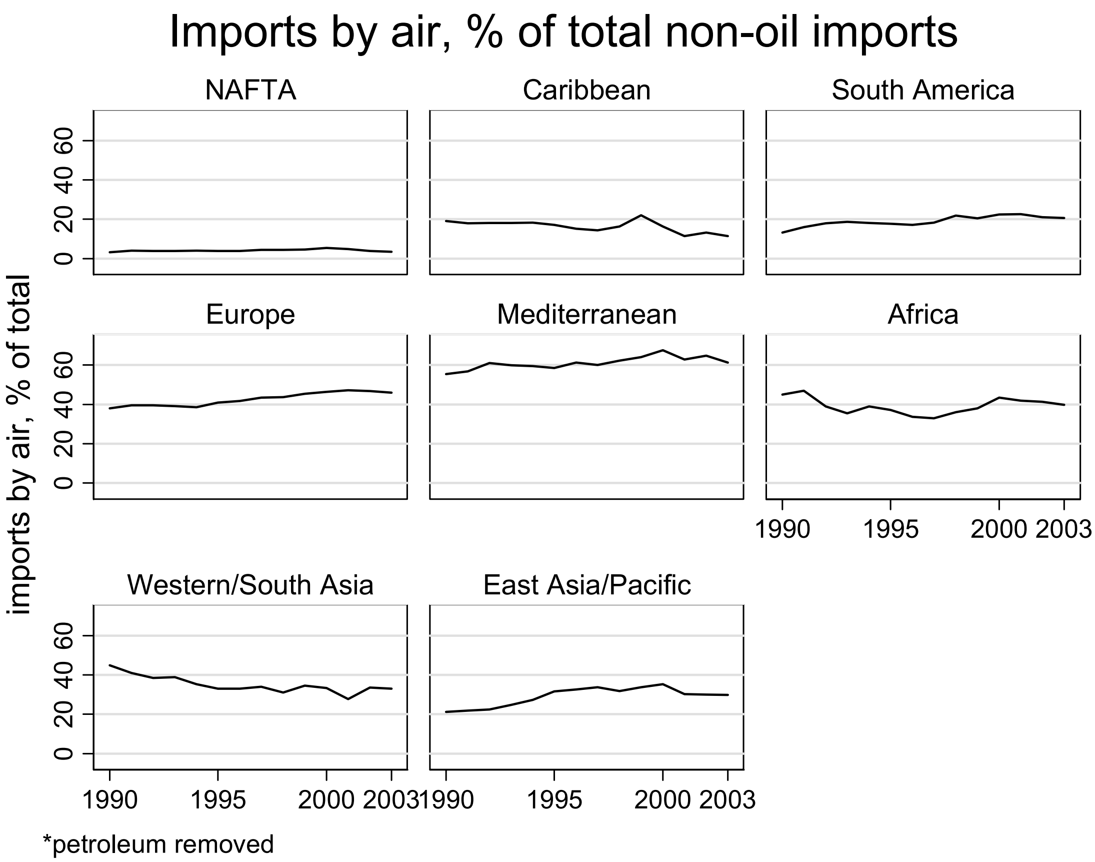


Chart 4

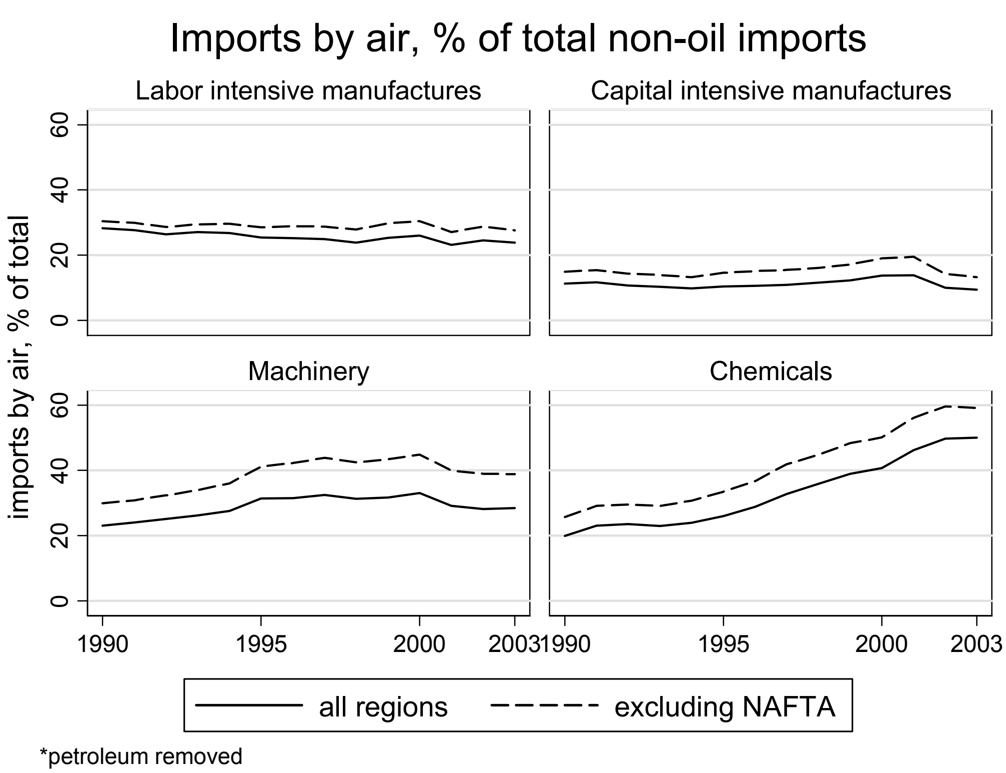
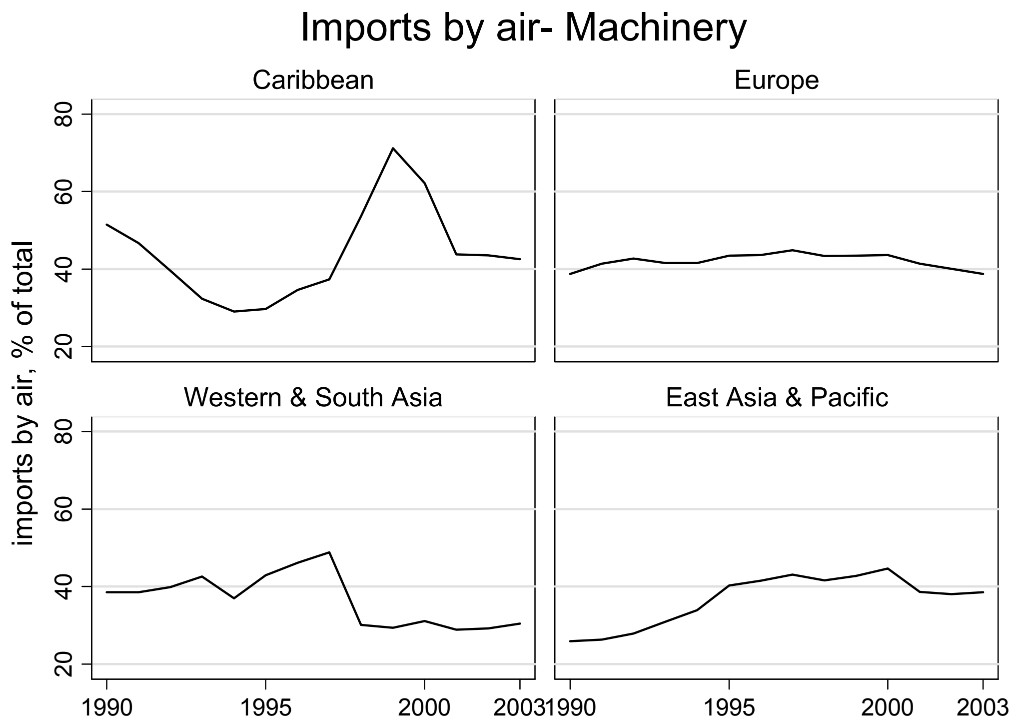
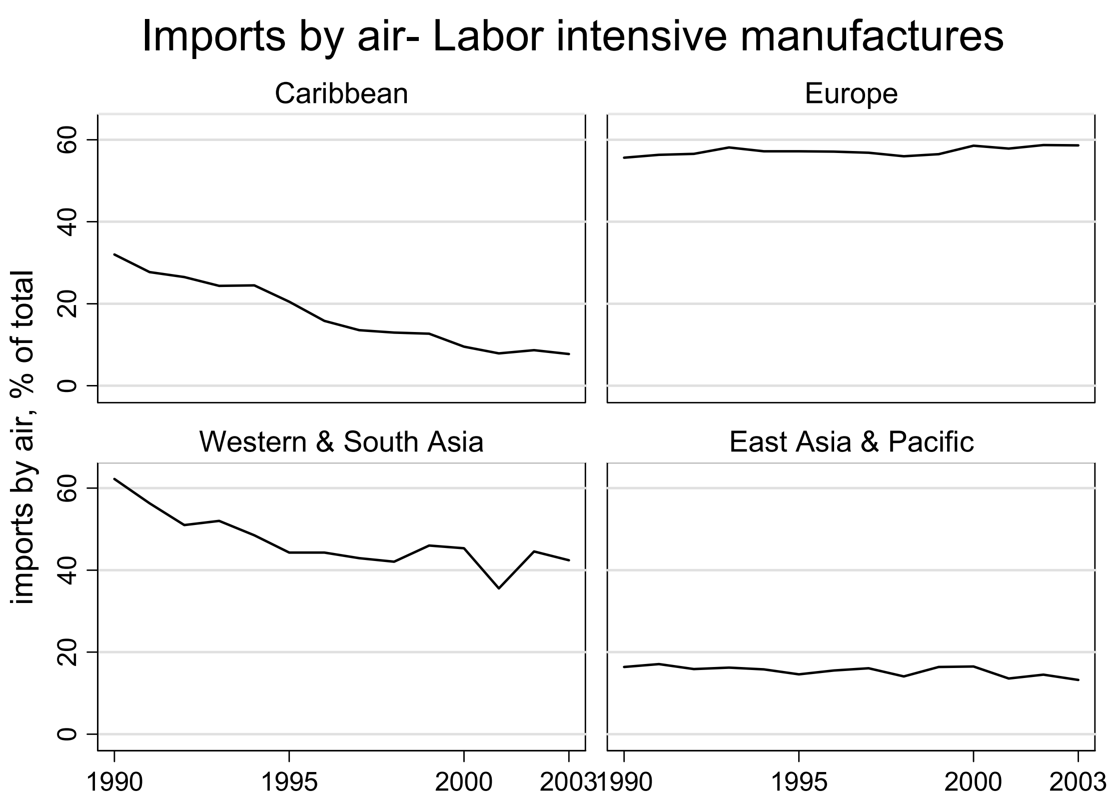


Chart 5



*petroleum removed

Chart 6



*petroleum removed

Chart 7

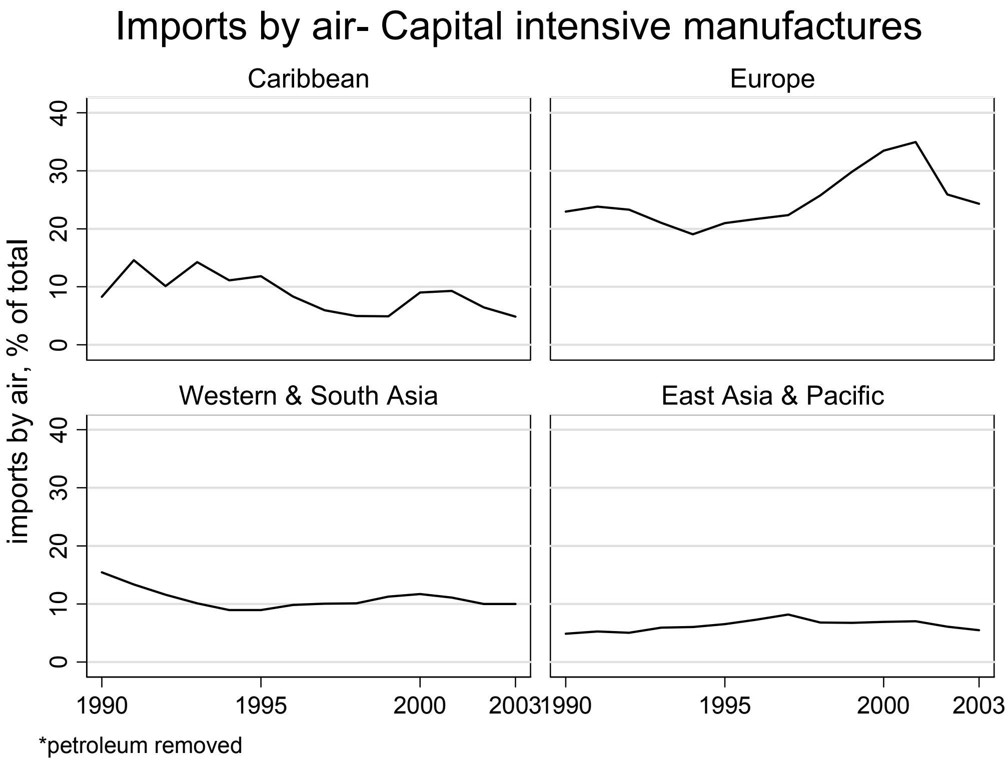


Chart 8

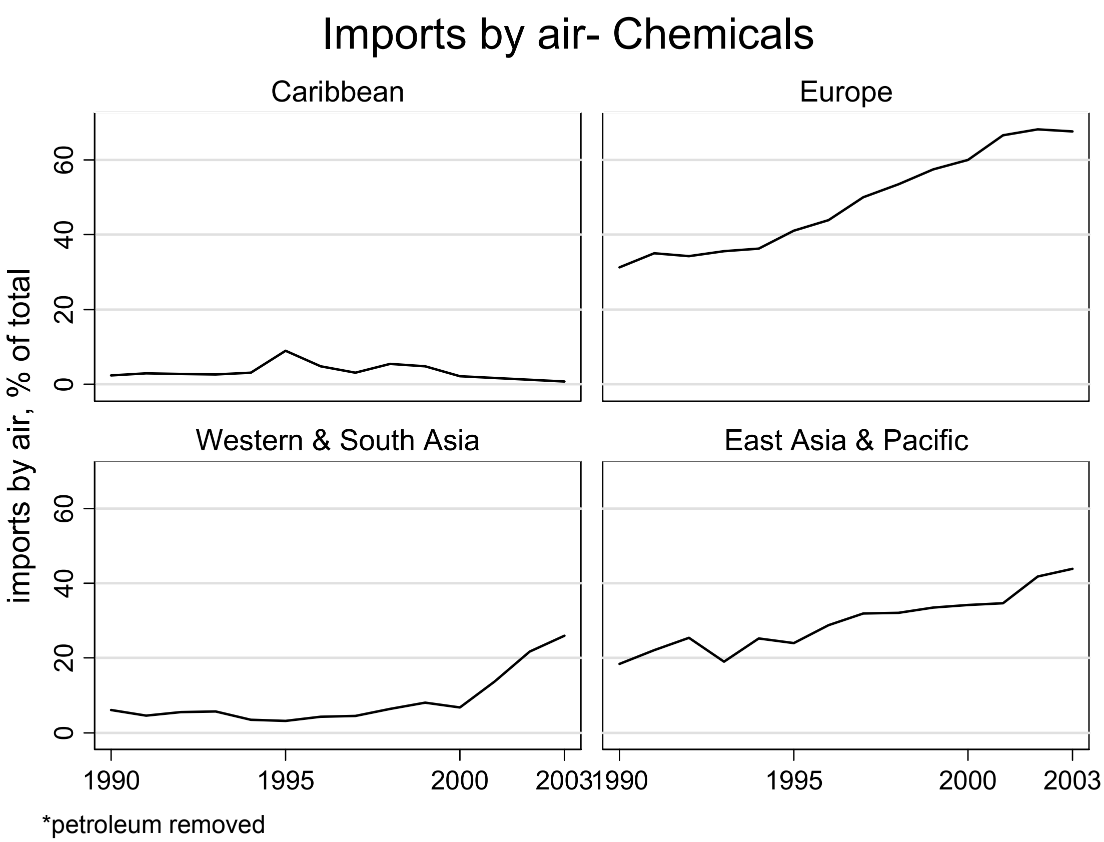


Table 1 Imports by product and percent air shipped, 1990 and 2003

Share of total 1990	Share of total 2003	SITC	% of category 1990	Imports by air, % of total 1990	% of category 2003	Imports by air, % of total 2003	SITC description
12.2	**8.9**	**Petroleum**					
		33	100.0	2.4	100.0	2.9	Petroleum, petroleum products and related materials
2.2	**3.3**	**Other fuel & raw materials**					
		28	35.0	9.1	7.5	13.6	Metalliferous ores and metal scrap
		34	30.6	1.4	74.3	7.6	Gas, natural and manufactured
		23	11.2	6.8	4.5	10.0	Crude rubber
		27	10.6	19.4	5.7	19.6	Crude fertilizers and crude minerals
		26	5.5	7.7	1.6	29.1	Textile fibres
		35	4.3	0.0	3.4	0.0	Electric current
		32	2.7	1.7	2.9	0.3	Coal, coke and briquettes
3.4	**2.8**	**Forest products**					
		64	51.2	23.8	43.5	19.5	Paper, paperboard and articles thereof
		24	18.9	4.9	21.2	4.0	Cork and wood
		25	17.3	12.3	7.6	9.5	Pulp and waste paper
		63	12.6	18.3	27.7	13.0	Cork and wood manufactures (excluding furniture)
5.7	**4.7**	**Animal and vegetable products**					
		05	19.8	7.4	19.5	5.8	Vegetables and fruit
		03	18.5	31.2	17.4	26.6	Fish, crustaceans, molluscs and aquatic invertebrates
		11	12.7	4.7	17.9	2.7	Beverages
		07	12.0	6.2	9.0	8.2	Coffee, tea, cocoa, spices, and manufactures thereof
		01	10.5	15.0	7.6	14.7	Meat and meat preparations
		29	4.4	36.8	4.8	40.6	Crude animal and vegetable materials, n.e.s.
		06	4.3	3.0	3.7	3.9	Sugars, sugar preparations and honey
		0	4.3	86.6	2.8	83.8	Live animals other than animals of division 03
		04	3.2	5.6	5.6	4.0	Cereals and cereal preparations
		12	2.3	16.2	2.2	14.1	Tobacco and tobacco manufactures

Table 1 Imports by product and percent air shipped, 1990 and 2003

		42	2.3	3.7	2.3	6.6	Fixed vegetable fats and oils, crude, refined or fractionated
		02	1.7	18.9	1.9	14.2	Dairy products and birds' eggs
		09	1.4	6.9	3.2	7.9	Miscellaneous edible products and preparations
		08	1.2	9.6	1.2	5.2	Feeding stuff for animals (not including unmilled cereals)
		22	0.7	11.8	0.5	10.8	Oil-seeds and oleaginous fruits
		21	0.6	44.2	0.2	61.1	Hides, skins and furskins, raw
		43	0.2	3.1	0.3	10.3	Animal or vegetable fats and oils
		41	0.1	13.4	0.1	18.7	Animal oils and fats
15.3	**15.9**	**Labor intensive manufactures**					
		84	33.3	56.5	31.4	47.5	Articles of apparel and clothing accessories
		89	32.9	47.0	32.6	46.4	Miscellaneous manufactured articles, n.e.s.
		85	12.8	46.3	7.7	50.1	Footwear
		66	11.6	28.9	13.7	24.5	Non-metallic mineral manufactures, n.e.s.
		82	6.6	12.8	12.2	12.3	Furniture, and parts thereof
		83	2.9	60.9	2.3	58.2	Travel goods, handbags and similar containers
8.1	**7.0**	**Capital intensive manufactures**					
		67	24.5	3.3	14.5	7.1	Iron and steel
		68	23.4	17.0	19.4	25.2	Non-ferrous metals
		69	22.2	24.2	28.5	29.4	Manufactures of metals, n.e.s.
		65	15.8	45.8	19.7	46.3	Textile yarn, fabrics, made-up articles, n.e.s
		62	8.7	14.6	9.7	27.0	Rubber manufactures, n.e.s.
		81	3.1	18.4	6.9	20.0	Prefabricated buildings, lighting & plumbing fixtures
		61	2.2	59.1	1.3	65.3	Leather, leather manufactures, n.e.s., and dressed furskins
45.2	**45.0**	**Machinery**					
		78	34.2	14.7	31.1	19.4	Road vehicles (including air-cushion vehicles)
		77	15.0	55.3	14.6	60.9	Electrical machinery, apparatus and appliances, n.e.s., and parts
		75	12.3	73.6	14.5	74.5	Office machines and automatic data-processing machines
		76	9.7	57.6	12.7	74.5	Telecommunications and sound-recording/reproducing apparatus
		74	6.4	31.5	6.8	34.7	General industrial machinery and equipment and parts, n.e.s.
		72	5.9	30.2	3.7	35.7	Machinery specialized for particular industries
		71	5.8	40.1	5.7	43.9	Power-generating machinery and equipment

Table 1 Imports by product and percent air shipped, 1990 and 2003

		79	3.3	42.0	3.6	42.8	Other transport equipment
		88	3.0	68.2	2.1	73.7	Photographic apparatus, optical goods, watches
		87	2.7	64.0	4.2	75.5	Professional, scientific and controlling instruments, n.e.s.
		73	1.7	28.9	1.0	41.4	Metalworking machinery
4.5	**8.3**	**Chemicals**					
		51	32.7	23.4	33.5	32.2	Organic chemicals
		52	14.0	11.1	7.1	16.8	Inorganic chemicals
		54	11.7	54.6	31.2	65.0	Medicinal and pharmaceutical products
		59	9.2	10.9	6.7	21.6	Chemical materials and products, n.e.s.
		57	8.9	7.2	7.1	19.5	Plastics in primary forms
		58	8.0	24.2	4.6	27.2	Plastics in non-primary forms
		53	5.9	10.4	2.4	18.4	Dyeing, tanning and colouring materials
		55	5.3	28.3	5.4	22.7	Essential oils and resinoids and perfume; cleanser
		56	4.2	8.7	2.0	6.9	Fertilizers (other than those of group 272)

Table 2 Country categories

distance	country	region	country	region
0 km from USA	Canada	NAFTA	Mexico	NAFTA
1-4000 km from USA	Bahamas	Caribbean	Barbados	Caribbean
	Belize	Caribbean	Costa.Rica	Caribbean
	Dominican.Rep.	Caribbean	El.Salvador	Caribbean
	Guatemala	Caribbean	Haiti	Caribbean
	Honduras	Caribbean	Jamaica	Caribbean
	Nicaragua	Caribbean	Panama	Caribbean
	TrinidadTobago	Caribbean	Colombia	South America
	Venezuela	South America		
4000-7800 km from USA	Bolivia	South America	Brazil	South America
	Ecuador	South America	Guyana	South America
	Paraguay	South America	Peru	South America
	Suriname	South America	Austria	Europe
	Belgium-Lux	Europe	Czechoslovakia	Europe
	Denmark	Europe	Finland	Europe
	France	Europe	Germany	Europe
	Hungary	Europe	Iceland	Europe
	Ireland	Europe	Italy	Europe
	Netherlands	Europe	Norway	Europe
	Poland	Europe	Portugal	Europe
	Spain	Europe	Sweden	Europe
	Switzerland	Europe	United.Kingdom	Europe
	Yugoslavia	Europe	Algeria	Mediterranean
	Malta	Mediterranean	Morocco	Mediterranean
	Tunisia	Mediterranean	Gambia	Africa
	Guinea	Africa	Guinea.Bissau	Africa
	Liberia	Africa	Mali	Africa
	Mauritania	Africa	Senegal	Africa
	Sierra.Leone	Africa		

Table 2 Country categories, continued

distance	country	region	country	region
7800-14000 km from USA	Argentina	South America	Chile	South America
	Uruguay	South America	Bulgaria	Europe
	Romania	Europe	Russia	Europe
	Cyprus	Mediterranean	Egypt	Mediterranean
	Greece	Mediterranean	Israel	Mediterranean
	Syria	Mediterranean	Turkey	Mediterranean
	Angola	Africa	Benin	Africa
	Burkina Faso	Africa	Burundi	Africa
	Cameroon	Africa	CentAfrRepublic	Africa
	Chad	Africa	Comoros	Africa
	Congo	Africa	Côte d'Ivoire	Africa
	Djibouti	Africa	Ethiopia	Africa
	Gabon	Africa	Ghana	Africa
	Kenya	Africa	Malawi	Africa
	Mozambique	Africa	Niger	Africa
	Nigeria	Africa	Rwanda	Africa
	Somalia	Africa	South.Africa	Africa
	Sudan	Africa	Tanzania	Africa
	Togo	Africa	Uganda	Africa
	Zaire	Africa	Zambia	Africa
	Zimbabwe	Africa	Afghanistan	W/S Asia
	Bahrain	W/S Asia	Bangladesh	W/S Asia
	Bhutan	W/S Asia	India	W/S Asia
	Iran	W/S Asia	Iraq	W/S Asia
	Jordan	W/S Asia	Kuwait	W/S Asia
	Mongolia	W/S Asia	Myanmar	W/S Asia
	Nepal	W/S Asia	Oman	W/S Asia
	Pakistan	W/S Asia	Qatar	W/S Asia
	Saudi.Arabia	W/S Asia	UAE	W/S Asia
	Yemen	W/S Asia	China	E Asia/Pacific
	Fiji	E Asia/Pacific	Hong.Kong	E Asia/Pacific
	Japan	E Asia/Pacific	Korea.RP.(S)	E Asia/Pacific
	Laos	E Asia/Pacific	Phillipines	E Asia/Pacific
	Solomon.Islands	E Asia/Pacific	Taiwan	E Asia/Pacific
over 14000 km from USA	Madagascar	Africa	Mauritius	Africa
	Seychelles	Africa	Reunion	W/S Asia
	Sri Lanka	W/S Asia	Australia	E Asia/Pacific
	Indonesia	E Asia/Pacific	Malaysia	E Asia/Pacific
	New Zealand	E Asia/Pacific	PapuaNGuinea	E Asia/Pacific
	Singapore	E Asia/Pacific	Thailand	E Asia/Pacific

Table 3- Transport costs by region, 2003

	transport cost, % of import value	air freight cost, % of air value
NAFTA	1.50	5.17
Caribbean	2.34	6.47
South America	9.17	7.04
Europe	4.45	4.96
Mediterranean	5.09	10.18
Africa	7.02	14.57
Western/South Asia	7.12	15.38
East Asia/Pacific	6.17	12.76

Table 4- Transport costs by product, 2003

	transport cost, % of import value	air freight cost, % of air value
Petroleum	5.00	22.37
Other fuel & raw materials	4.74	3.76
Forest products	6.44	20.88
Animal and vegetable products	7.30	23.77
Labor intensive manufactures	5.71	4.43
Capital intensive manufactures	5.48	6.97
Machinery	1.97	2.37
Chemicals	2.73	1.04

Table 5 - Regression of U.S. import unit values on distance and other controls, all available observations

	1990		1995		2000		2003	
more than 4000km	0.260		0.190		0.371		0.362	
	21.6		*15.3*		*32.8*		*31.9*	
1-4000km		0.056		-0.148		0.092		0.047
		2.4		*-6.6*		*4.2*		*2.2*
4000-7800km		0.446		0.390		0.517		0.530
		31.4		*27.4*		*40.4*		*41.3*
7800-14,000km		-0.023		-0.133		0.139		0.078
		-1.5		*-8.9*		*9.8*		*5.5*
more than 14,000km		0.239		0.074		0.355		0.290
		12.6		*4.3*		*20.6*		*17.0*
log Y/L	-0.051	-0.130	-0.056	-0.202	0.200	0.031	0.276	0.081
	-4.0	*-9.6*	*-4.7*	*-16.1*	*18.7*	*2.5*	*25.7*	*6.6*
log price level	0.829	0.760	0.769	0.768	0.403	0.485	0.364	0.445
	46.0	*39.9*	*52.3*	*50.3*	*28.7*	*31.6*	*25.7*	*29.0*
landlocked	0.294	0.178	0.322	0.154	0.337	0.217	0.443	0.291
	14.8	*9.0*	16.96	*8.2*	19.41	*12.4*	25	*16.3*
tariff	-0.014	-0.013	-0.011	-0.013	-0.001	-0.002	-0.003	-0.002
	-18.5	*-18.1*	*-8.7*	*-9.9*	*-0.8*	*-1.5*	*-3.0*	*-1.9*
transport cost	-0.012	-0.012	-0.004	-0.004	-0.013	-0.013	-0.011	-0.011
	-25.4	*-25.9*	*-18.2*	*-18.3*	*-33.5*	*-32.7*	*-30.1*	*-29.4*
R^2 within	0.145	0.168	0.135	0.162	0.119	0.131	0.132	0.149
N	88,984		108,837		121,830		127,602	
HS codes	11,815		13,131		13,788		14,103	

Notes to Table 5: Estimates of equation 21 in the text. For each year, log U.S. import unit value is regressed on indicators for exporter distance from the US, other controls, and fixed effects for 10 digit HS codes. Estimation is by OLS, *t*-statistics are in italics. Y/L is aggregate real GDP per worker of the exporter, price level is the exporters aggregate price level, landlocked is an indicator for the exporter having no port, and tariff and transport costs are ad valorem percentages. *N* is sample size, "HS codes" is the number of 10 digit HS code fixed effects, and "R^2 within" is the R^2 after removing HS10 means from the data.

Table 6 - Regression of U.S. import value/kilo on distance and other controls, all available observations

	1990		1995		2000		2003	
more than 4000km	0.400		0.353		0.551		0.510	
	36.1		*30.2*		*50.2*		*45.8*	
1-4000km		0.296		0.069		0.160		0.165
		14.5		*3.4*		*8.0*		*8.3*
4000-7800km		0.571		0.508		0.657		0.656
		41.6		*36.5*		*51.4*		*50.8*
7800-14,000km		0.339		0.229		0.461		0.345
		23.3		*15.7*		*32.8*		*24.1*
more than 14,000km		0.443		0.334		0.598		0.494
		24.3		*19.8*		*35.1*		*28.9*
log Y/L	-0.057	-0.056	-0.096	-0.153	0.058	-0.016	0.151	0.042
	-4.9	*-4.5*	*-8.6*	*-12.6*	*5.8*	*-1.4*	*14.7*	*3.6*
log price level	0.718	0.643	0.746	0.721	0.523	0.550	0.488	0.511
	43.4	*36.7*	*53.3*	*49.1*	*39.2*	*37.8*	*35.7*	*34.9*
landlocked	0.362	0.310	0.327	0.249	0.374	0.320	0.478	0.384
	18.9	*16.2*	*18.0*	*13.7*	*22.2*	*18.7*	*27.2*	*21.6*
tariff	-0.013	-0.014	-0.009	-0.010	-0.002	-0.003	-0.001	-0.001
	-10.8	*-11.5*	*-7.7*	*-8.8*	*-2.4*	*-3.7*	*-1.5*	*-1.1*
transport cost	-0.014	-0.015	-0.004	-0.004	-0.015	-0.016	-0.014	-0.014
	-33.9	*-35.2*	*-21.8*	*-22.2*	*-41.9*	*-42.1*	*-39.2*	*-39.3*
R^2 within	0.219	0.230	0.212	0.225	0.197	0.202	0.213	0.225
N	52,028		66,366		74,271		78,910	
HS codes	7,422		8,518		8,910		9,139	

Notes to Table 6: Estimates of equation 21 in the text. For each year, log U.S. import value/kilo is regressed on indicators for exporter distance from the US, other controls, and fixed effects for 10 digit HS codes. Estimation is by OLS, *t*-statistics are in italics. Y/L is aggregate real GDP per worker of the exporter, price level is the exporters aggregate price level, landlocked is an indicator for the exporter having no port, and tariff and transport costs are ad valorem percentages. *N* is sample size, "HS codes" is the number of 10 digit HS code fixed effects, and "R^2 within" is the R^2 after removing HS10 means from the data.

Table 7- Regression of U.S. import unit value on distance and other controls, restricted sample

	1990		1995		2000		2003	
more than 4000km	0.197		0.117		0.328		0.310	
	13.4		*7.5*		*23.3*		*21.9*	
1-4000km		-0.130		-0.321		-0.029		-0.079
		-4.6		*-11.7*		*-1.1*		*-3.0*
4000-7800km		0.394		0.336		0.482		0.489
		22.7		*19.1*		*30.8*		*31.1*
7800-14,000km		-0.179		-0.300		0.031		-0.073
		-9.9		*-16.4*		*1.8*		*-4.2*
more than 14,000km		0.122		-0.043		0.269		0.167
		5.4		*-2.1*		*13.0*		*8.1*
log Y/L	-0.032	-0.129	-0.058	-0.249	0.288	0.068	0.343	0.86
	-2.1	*-8.1*	*-4.0*	*-16.3*	*21.8*	*4.5*	*25.7*	*5.7*
log price level	0.855	0.764	0.780	0.817	0.307	0.427	0.305	0.423
	39.8	*33.7*	*44.4*	*43.8*	*17.7*	*22.4*	*17.4*	*22.3*
landlocked	0.269	0.121	0.330	0.111	0.358	0.205	0.463	0.266
	11.1	*5.0*	*14.2*	*4.8*	*17.0*	*9.6*	*21.3*	*12.2*
tariff	-0.015	-0.015	-0.006	-0.008	0.005	0.004	-0.001	0.002
	-13.9	*-13.9*	*-3.8*	*-4.8*	*3.6*	*3.0*	*-0.7*	*1.2*
transport cost	-0.017	-0.017	-0.012	-0.012	-0.026	-0.024	-0.018	-0.017
	-20.6	*-20.4*	*-20.34*	*-19.5*	*-34.3*	*-32.0*	*-30.0*	*-26.1*
R^2 within	0.156	0.189	0.146	0.185	0.130	0.147	0.141	0.165
N	62,834		76,235		86,364		88,756	
HS codes	7,713		8,513		9,076		9,264	

Notes to Table 7: Estimates of equation (20) in the text. For each year, log U.S. import unit value is regressed on indicators for exporter distance from the US, other controls, and fixed effects for 10 digit HS codes. Estimation is by OLS, *t*-statistics are in italics. The sample is restricted to SITC 6, 7, and 8, and small observations (imports of one unit and/or with value less than $10,000) are also dropped. Y/L is aggregate real GDP per worker of the exporter, price level is the exporters aggregate price level, landlocked is an indicator for the exporter having no port, and tariff and transport costs are ad valorem percentages. *N* is sample size, “HS codes” is the number of 10 digit HS code fixed effects, and “R^2 within” is the R^2 after removing HS10 means from the data.

Table 8- Regression of U.S. import value/kilo on distance and other controls, restricted sample

	1990		1995		2000		2003	
more than 4000km	0.366		0.318		0.531		0.485	
	26.1		*23.2*		*41.7*		*37.4*	
1-4000km		0.147		-0.070		0.016		0.001
		6.4		*-3.23*		*0.7*		*0.0*
4000-7800km		0.560		0.491		0.647		0.647
		35.3		*30.3*		*43.5*		*43.5*
7800-14,000km		0.220		0.116		0.367		0.199
		13.3		*7.0*		*23.0*		*12.2*
more than 14,000km		0.352		0.27		0.509		0.371
		17.2		*14.3*		*26.7*		*19.6*
log Y/L	-0.041	-0.050	-0.109	-0.199	0.109	-0.008	0.164	-0.001
	-3.3	*-3.8*	*-8.6*	*-14.6*	*9.5*	*-0.6*	*13.9*	*-0.1*
log price level	0.736	0.622	0.779	0.760	0.466	0.510	0.496	0.533
	41.0	*32.9*	*48.4*	*44.97*	*30.5*	*30.7*	*31.3*	*32.0*
landlocked	0.394	0.317	0.337	0.224	0.440	0.354	0.515	0.370
	17.1	*13.9*	*15.8*	*10.6*	*22.5*	*17.9*	*24.6*	*17.7*
tariff	-0.010	-0.013	-0.004	-0.006	0.002	0.001	0.001	0.002
	-6.4	*-8.5*	*-3.0*	*-4.5*	*2.35*	*0.9*	*0.59*	*1.86*
transport cost	-0.020	-0.020	-0.015	-0.015	-0.028	-0.028	-0.021	-0.020
	-27.1	*-27.5*	*-28.6*	*-29.1*	*-42.5*	*-41.8*	*-35.6*	*-34.7*
R^2 within	0.287	0.310	0.281	0.308	0.259	0.271	0.273	0.301
N	31,194		40,039		45,367		47,223	
HS codes	4,046		8,618		4,933		5,080	

Notes to Table 8: Estimates of equation (20) in the text. For each year, log U.S. import value/kilo is regressed on indicators for exporter distance from the US, other controls, and fixed effects for 10 digit HS codes. Estimation is by OLS, *t*-statistics are in italics. The sample is restricted to SITC 6, 7, and 8, and small observations (imports of one unit and/or with value less than $10,000) are also dropped. Y/L is aggregate real GDP per worker of the exporter, price level is the exporters aggregate price level, landlocked is an indicator for the exporter having no port, and tariff and transport costs are ad valorem percentages. *N* is sample size, "HS codes" is the number of 10 digit HS code fixed effects, and "R^2 within" is the R^2 after removing HS10 means from the data.

Table 9- Explaining exporter fixed effects

	sample	1990		1995		2000		2003	
distance > 4000km	1	0.274		0.243		0.424		0.431	
		2.99		*2.4*		*4.01*		*3.7*	
	2	0.339		0.261		0.485		0.473	
		2.89		*2.03*		*3.57*		*3.3*	
	3	0.246		0.273		0.524		0.499	
		3.44		3.54		*6.01*		*4.93*	
	4	0.289		0.314		0.594		0.574	
		3.57		*3.69*		*6.13*		*5.19*	
log distance	1		0.004		-0.053		0.072		0.040
			0.1		-0.8		*1.0*		*0.5*
	2		0.009		-0.081		0.068		0.001
			0.12		*-1.01*		*0.7*		*0.0*
	3		0.074		0.075		0.232		0.160
			1.57		1.47		*3.57*		*2.16*
	4		0.088		0.079		0.251		0.157
			1.65		*1.39*		*3.99*		*1.88*
log aggregate price level	1	0.741	0.773	0.684	0.710	0.586	0.609	0.638	0.654
		15.7	*16.1*	*13.6*	*14.7*	*10.9*	*10.1*	*10.9*	*10.1*
	2	0.813	0.850	0.736	0.761	0.617	0.641	0.677	0.686
		13.7	*13.9*	*11.7*	*12.3*	*9.0*	*8.4*	*9.6*	*8.8*
	3	0.649	0.648	0.645	0.684	0.562	0.622	0.639	0.683
		17.1	*17.6*	*17.2*	*12.3*	*12.3*	*11.9*	*12.2*	*11.5*
	4	0.700	0.742	0.672	0.719	0.572	0.641	0.673	0.718
		16.2	*16.5*	*15.3*	*15.8*	*11.0*	*10.5*	*11.7*	*10.6*

Notes to Table 9: Estimates of equations (21) and (22) in the text. In the first stage, log import unit value or value/weight is regressed on exporter and HS10 fixed effects, as well as tariffs and transport costs. In the second stage, the estimated country fixed effects are regressed on exporter distance from the U.S. and the exporter's aggregate price level. The results in the table are computed by weighted least squares, where the weights are the inverse of the standard errors of the estimated fixed effects from the first stage.

Results from four different samples are reported, and for each sample distance of the exporter from the U.S. is measured in two ways (as an indicator for greater than 4000km, or as log kilometers). The samples are

1. dependent variable is log unit value, all available observations.
2. dependent variable is log value/weight, all available observations.
3. dependent variable is log unit value, non-manufacturing and small observations excluded.
4. dependent variable is log value/weight, non-manufacturing and small observations excluded.

Manufacturing includes SITC 6, 7 and 8. "small" observations are imports of one physical unit and/or with value less than $10,000. *t*-statistics are in italics.

Table 10 – Conditional fixed effects logit regression of shipment choice on distance and other controls, all available observations

	1990		1995		2000		2003	
more than 4000km	1.36		1.41		1.86		1.90	
	36.3		*37.4*		*53.7*		*56.3*	
1-4000km		2.86		2.28		2.11		2.25
		38.1		*35.5*		*34.6*		*39.0*
4000-7800km		2.81		2.72		2.82		2.80
		53.5		*57.4*		*66.0*		*66.9*
7800-14,000km		1.41		1.31		1.69		1.72
		26.9		*27.4*		*38.8*		*39.9*
more than 14,000km		2.16		1.77		2.05		2.12
		34.8		*33.9*		*42.2*		*44.5*
log Y/L	0.19	0.19	0.18	0.11	0.40	0.12	0.32	0.06
	5.4	*5.6*	*6.2*	*3.7*	*15.2*	*4.2*	*12.8*	*2.4*
log price level	0.67	0.29	0.67	0.37	0.47	0.46	0.61	0.59
	13.7	*6.2*	*19.2*	*9.9*	*14.2*	*13.1*	*18.2*	*16.7*
landlocked	1.16	0.91	1.12	0.82	1.04	0.75	1.02	0.73
	27.5	*20.3*	*27.7*	*19.1*	*28.2*	*19.2*	*28.4*	*19.3*
tariff	-0.012	-0.016	-0.013	-0.008	0.013	0.011	0.001	0.012
	-1.8	*-1.9*	*-3.7*	*-2.2*	*4.4*	*3.4*	*0.2*	*4.0*
N	80,869		102,110		115,485		120,118	
HS codes	6,674		7,918		8,628		8,614	
%HS used	50.2		54.6		56.7		55.5	

Effect of distance on probability of air shipment, implied odds ratios

	1990	1995	2000	2003
more/less than 4,000km	4.9	5.7	7.6	7.3
1-4,000km /NAFTA	11.9	7.6	6.3	7.0
4,000-7,800/ NAFTA	20.2	19.8	21.2	21.0
7,800-1,4000km/ NAFTA	3.5	3.6	4.9	4.9
more than 14,000km/ NAFTA	7.3	5.3	5.9	6.2

Notes to Table 10: Estimation of equation (23) in the text. Robust t-statistics in *italics*. Dependent variable is $a_{ic} = 1$ if product i is shipped by air from exporter c. Odds ratios are average ratios of predicted probabilities. See text for details.

Appendix - Notes on computation of the regression estimates

The data used in this paper have three dimensions: time, product, and exporting country. All regressions are run separately for each year, so the key statistical issue is how to pool the cross-product and cross-country variation within each year. Let $i = 1,...,N$ index products and $c = 1,...,C$ index exporting countries.

I can write one of my regression specifications generically as

$$y_{ic} = \mathbf{x}_c'\boldsymbol{\beta}_1 + \mathbf{z}_{ic}'\boldsymbol{\beta}_2 + \mu_i + u_{ic} \qquad i = 1,\ldots,N \quad c = 1,\ldots,C$$

The μ_i are the product fixed effects, $\mathbf{x}_c$ is a vector of exporting country characteristics (such as distance from the US, log aggregate price level, etc) and $\mathbf{z}_{ic}$ is a vector of variables that vary over both products and exporters (such as tariffs and transport charges). For each country, we can stack N_c observations as follows:

$$\mathbf{y}_c = \mathbf{X}_c\boldsymbol{\beta}_1 + \mathbf{z}_c'\boldsymbol{\beta}_2 + \boldsymbol{\mu}_c + \mathbf{u}_c \qquad c = 1,\ldots,C$$

where $N_c \leq N$ is the number of products exported by country c. The $N_c \times 1$ vector $\boldsymbol{\mu}_c$ is composed of the product fixed effects for products exported by c.

The standard OLS assumption is that $E\left(\mathbf{u}_c\mathbf{u}_c'\right) = \sigma^2 I_c, c = 1,\ldots,C$, but this is *a priori* implausible in this case: we would expect some correlation across products for each country. A standard approach to this statistical issue is to use a robust covariance matrix which allows for arbitrary cross-commodity correlation within each country, a solution which might be called "clustering by country". The asymptotic theory behind robust covariance matrices with clustering relies on holding the number of observations per cluster fixed while increasing the number of clusters; in my notation, holding N fixed and letting C go to infinity (see, for example Wooldridge (2002), page 328-331). As a consequence, relying on this asymptotic theory in my application is not appropriate, since the number of countries (about 100) is very small relative to the number of products (as many as 14,000).

An alternative approach is to specify a two-way error components model. The generic regression model then becomes

$$y_{ic} = \mathbf{x}_c'\boldsymbol{\beta}_1 + \mathbf{z}_{ic}'\boldsymbol{\beta}_2 + \mu_i + \delta_c + u_{ic} \qquad i = 1,\ldots,N \quad c = 1,\ldots,C$$

Because $\mathbf{x}_c$ has no cross-commodity variation and $\boldsymbol{\beta}_1$ is the main parameter of interest, to identify $\boldsymbol{\beta}_1$ I need to make the random effects assumption that δ_c is a random variable which is orthogonal to $\mathbf{x}_c$ and $\mathbf{z}_{ic}$.

Computation of this mixed fixed effects-random effects model is straightforward: first remove the product means from all variables, and then run generalized least squares on the transformed data. The GLS-RE estimator runs OLS on transformed data. The transformation subtracts θ times the country-specific means from the raw data, where

$$\theta_c = 1 - \sqrt{\frac{\sigma_u^2}{N_c \sigma_\delta^2 + \sigma_u^2}}$$

and σ_δ^2 and σ_u^2 are the variance components. Notice that if σ_δ^2 is very small, then θ is close to zero, and the GLS transformation will leave the data almost unchanged. The estimator that I use for σ_δ^2 is

$$\hat{\sigma}_\delta^2 = \max\left\{0, \frac{1}{\bar{N}_c}\left(\frac{SSR_{between}}{\sum_{c=1}^{C} N_c - K} - \hat{\sigma}_u^2\right)\right\}$$

where *SSR* stands for "sum of squared residuals". Somewhat surprisingly, in my application the result of this estimator is invariably $\hat{\sigma}_\delta^2 = 0$. The intuitive reason is that $SSR_{between}$ is small relative to SSR_{within}. Despite the fact that $\hat{\sigma}_\delta^2 = 0$, the null hypothesis $\sigma_\delta^2 = 0$ can nonetheless be tested using a chi-square statistic. This null hypothesis is invariably *rejected*. Thus, the data analysis gives an odd message: there are random country effects, but they are too small to adjust for.

To summarize, as a consequence of $\hat{\sigma}_\delta^2 = 0$, the results reported in Tables 5 to 8 are computed by OLS with product fixed effects.

Appendix reference

Wooldridge, Jeffrey M., 2002, *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press.