online appendix to

Airplanes and Comparative Advantage

James Harrigan*
June 2010

This appendix includes

1. A three-country version of the theory model.
2. Additional data description.
3. Additional tables of statistical results.
4. Econometric methodology discussion.

A three country version of the model

Setting up the three country model

Without putting more structure on the model, I can not say much about equilibrium wages and gains from trade, nor can I do comparative statics. In this section of the appendix, I show that useful insights can be obtained by computing the full equilibrium for a special 3-country case. Three countries are the minimum required for relative transport costs to matter to the equilibrium of the model, and the example serves to illustrate how distance affects comparative advantage and real wages. The simplifying assumptions are

1. Country 1 and 2 are adjacent, country 3 is distant (it may be useful to keep in mind examples such as the U.S., Mexico and China, or the U.S., Canada, and Europe).
2. All countries share the same aggregate technology parameter T and the same labor supply L.
3. The degree of timeliness preference is constant across goods, $f(z) = f$.
4. Air shipment costs increase more rapidly with weight than surface transport costs.

The first three simplifications are for tractability only, but the fourth is more substantive. As noted above, the fourth assumption is empirically validated in the empirical analysis that follows. Since what matters for comparative advantage is the ratio of air to surface transport costs, I economize on notation by assuming that surface transport costs are the same for all goods on a given trade route.

Because countries 1 and 2 are identical in every way, they will have the same wage in equilibrium. I take this wage as the numeraire, and w_3 denotes the nominal wage in country 3.

The geography of the three country model means that the structure of transport costs is very simple. Countries 1 and 2 can ensure timely delivery to each others' markets without using expensive air shipment, so all trade between 1 and 2 takes place by surface transit at an iceberg cost of s_1. Some long distance trade involving country 3 will be by air at an iceberg cost of $a(z)$, and some by ship at a cost of s_3 where $a(z) > s_3 > s_1 > 1$.

Order goods so that $a(z)$ is increasing in z, and let $\bar{z}$ be the cutoff for air shipment between 3 and 1 or 2. By equation (2), $\bar{z}$ is determined by

$$\frac{a(\bar{z})}{s_3 f} = 1 \tag{A1}$$

A three country version of the model

For goods shipped to or from 3, the transport mode will be air for $z \in [0, \bar{z})$ ("light" goods) and surface for all other ("heavy") goods.

Expressions for the three price parameter functions $\Phi_d(z)$ and the nine market share functions $\pi_{od}(z)$ are found using equations (4) and (5). What is of interest here are the results for trade patterns. I focus on the results of competition in market 1, where the two possible foreign suppliers are nearby country 2 and faraway country 3.[1]

Consider 3's market share in light goods in 1 relative to 2's market share. These are goods where both sellers provide timely delivery, but using different transport modes. From (9),

$$\frac{\pi_{31}(z)}{\pi_{21}(z)} = \left(\frac{s_1}{w_3 a(z)} \right)^{\theta}, \qquad z < \bar{z}.$$

Next, consider the same competition in heavy goods. These are goods where 2 provides timely delivery and 3 does not. From (8),

$$\frac{\pi_{31}}{\pi_{21}} = \left(\frac{s_1}{f\, w_3 s_3} \right)^{\theta}, \qquad \bar{z} \le z.$$

Both of these relative market shares are decreasing in 3's wage and in the relative cost of transport used by 3. Now define the ratio of ratios as in (11), which gives comparative advantage:

$$\frac{\pi_{31}(z)/\pi_{21}(z)}{\pi_{31}/\pi_{21}} = \left(\frac{f\, s_3}{a(z)} \right)^{\theta} \ge 1$$

As expected from Proposition 1, country 3 has a comparative advantage in air-shipped goods in market 1, and 2 has a comparative advantage in heavy goods. The lighter the good (that is, the lower is $a(z)$), the greater is 3's comparative advantage.

Equilibrium in the 3 country model

Because of symmetry and low dimensionality, there are only two endogenous variables to be solved for in the three country model, with the remainder found by substitution. The first is the cutoff $\bar{z}$, which is trivial to solve for by equation (A1). The second endogenous variable is the

---

[1] Country 1 may of course supply its own market in any good z, but this has no bearing on the relative chances of 2 or 3 winning the competition.

nominal wage in country 3, which can be found using the national income identity for any one of the countries.

To compute equilibrium requires a functional form for the air transport cost schedule $a\left(z\right)$. It is convenient to use

$$a(z) = \beta s_3 f^{1+z}$$

where the shift parameter β has a range of $\left[f^{-1},1\right]$. Recall that the condition for airfreight to be profitable for trade with country 3 in good z is $a(z) \le f s_3$. For low values of β air freight is always profitable for country 3 trade,

$$\beta = f^{-1} \rightarrow \qquad a(0) = s_3 \qquad a(1) = f s_3$$

while for high values it is never profitable:

$$\beta = 1 \rightarrow \qquad a(0) = f s_3 \qquad a(1) = f^2 s_3$$

Substituting into the cutoff condition (13) gives the solution for $\bar{z}$:

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

Given the functional form for the air transport cost schedule and numerical values for the parameters of the model, computation of equilibrium is straightforward.[2] Figures A1 and A2 show how the equilibrium changes as the cost of air transport moves from low (such that country 3 conducts all its trade by air) to prohibitive (such that country 3 conducts all its trade by surface).

Figure A1 shows real and nominal wages. As air transport becomes more expensive, all countries suffer real wage declines, but country 3 suffers the most. The welfare effect of higher air transport costs on country 3 works both through lower nominal wages (the global demand for country 3 labor declines) and a higher consumption price index (greater timeliness-adjusted transport costs).

The fact that the distant country has lower nominal wages and a higher price level is a result familiar from economic geography models (see Redding and Venables (2004) for theory and supportive empirical evidence), and lower real income as a function of distance from the center is also a feature of the equilibrium in Limao and Venables (2002). An interesting

---

[2] Calculations were done in *Mathematica*, and the programs are available on request.

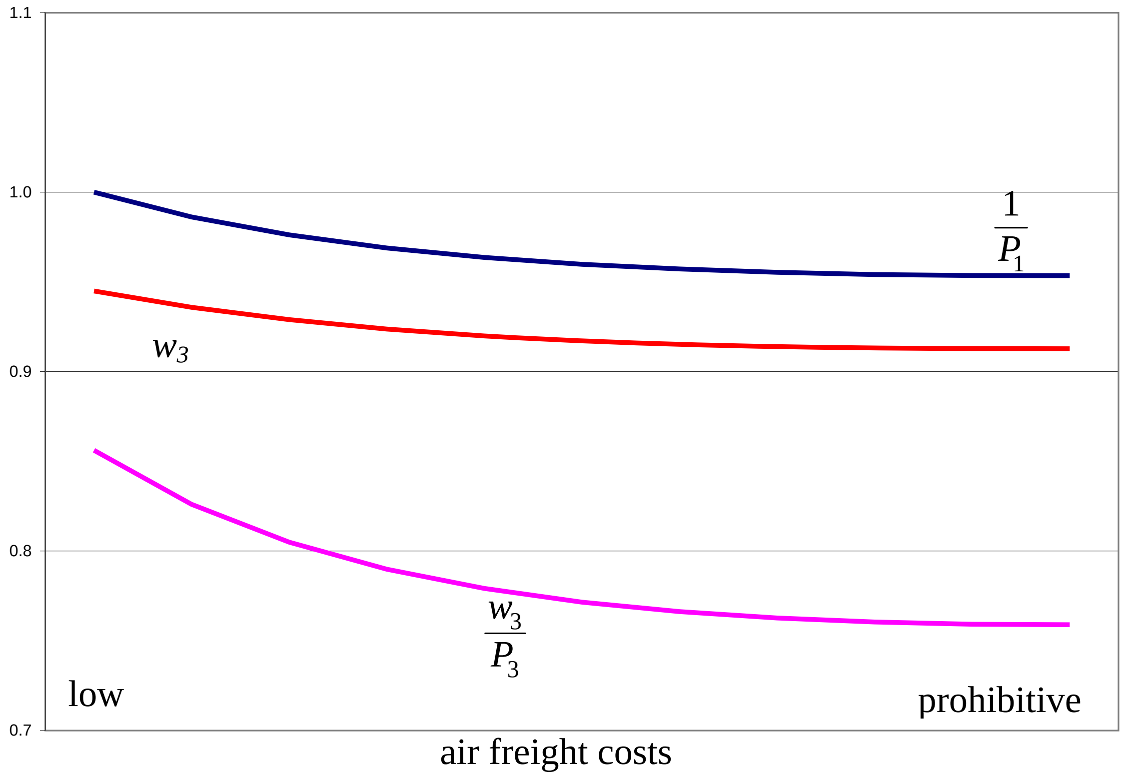
difference from Limao-Venables is the result here that all countries gain from a drop in air transport costs, though the gain is disproportionately larger for the distant country. In Limao-Venables there is a terms-of-trade effect which can hurt countries that are near the center when more distant countries enter the world trading system, as the nearby countries now face greater competition in their export market. This terms-of-trade effect is absent in my model because the nature of competition is different: in response to greater competitiveness from remote country suppliers, the centrally located economies become more specialized, tending to move out of products where the remote country has become more competitive.

Figure A2 shows the corresponding effects on aggregate trade patterns. The overall openness of the two nearby countries doesn't fall by much when air transport becomes more expensive: the nearby countries trade more with each other (higher π_{12}), as well as consuming more of their won production (higher π_{11}). By contrast, the elimination of air transport as a viable option causes country 3 to reduce its total trade substantially.

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 A3. 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. A similar result is derived by Limao and Venables (2002), who find that countries closer to the center will specialize more strongly in the transport-intensive good when overall transport costs decline.

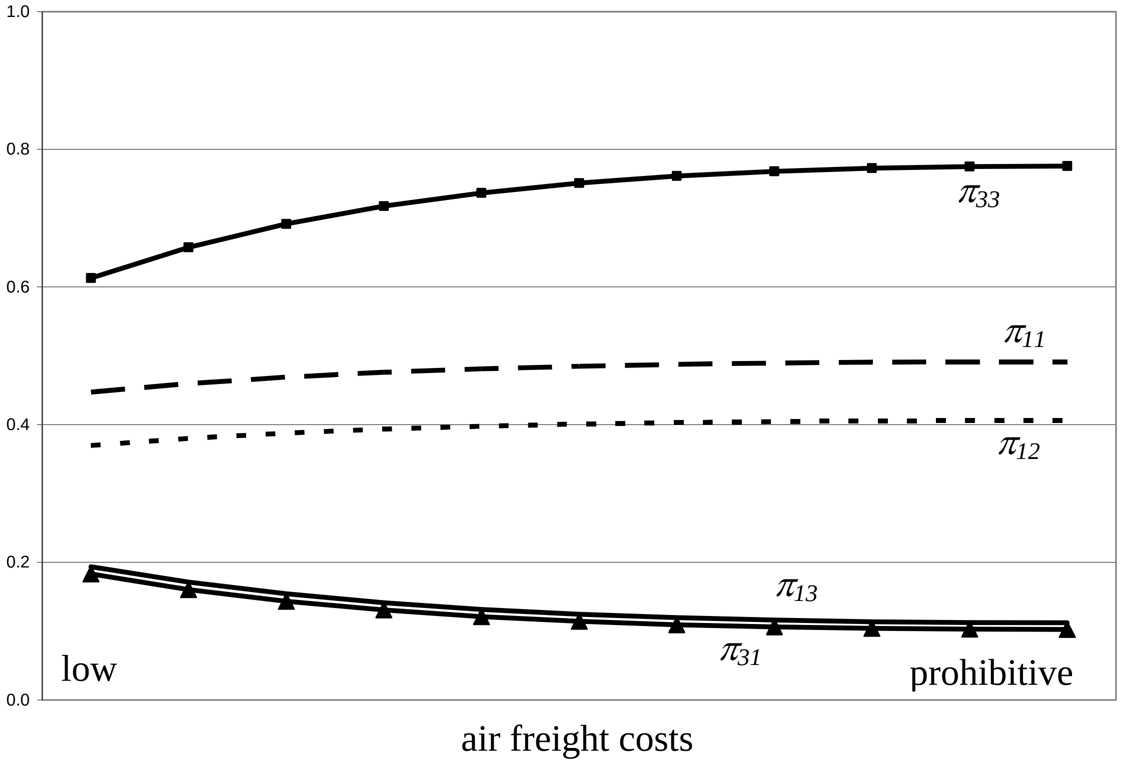
A three country version of the model

Figure A1 - Real and nominal wages as a function of air transport costs



<u>Notes to Figure A1</u>**:** Illustrates equilibrium wages as a function of air freight costs for a numerical example, with air freight costs varying from prohibitive (right axis) to low enough so that country 3 always uses air freight (left axis). The nominal wage in country 1 is the numeraire and is set equal to one, and country 1 real wages are normalized to one when air freight costs are low. Equilibria are computed as the shift parameter β falls from 1 to f^{-1}. Parameter values are $f = \theta = 2$, $\tau_3 = 1.2$, $\tau_1 = 1.1$, $L_1 = L_2 = L_3 = 1$.

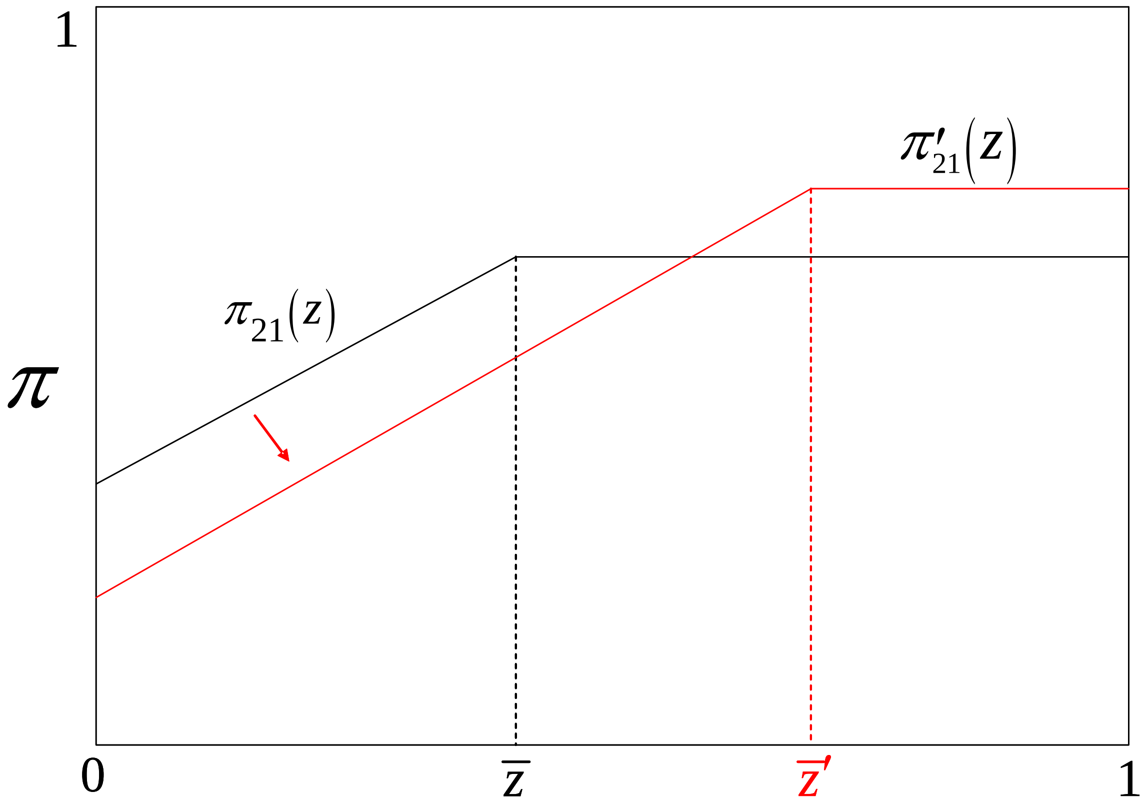
Figure A2 - Aggregate Market shares as a function of air transport costs



<u>Notes to Figure A2</u>**:** Illustrates equilibrium aggregate market shares (that is, the share of country *o* production in country *d* expenditure) as a function of air freight costs for a numerical example, with air freight costs varying from prohibitive (right axis) to low enough so that country 3 always uses air freight (left axis). Parameter values are the same as in Figure 2.

A three country version of the model

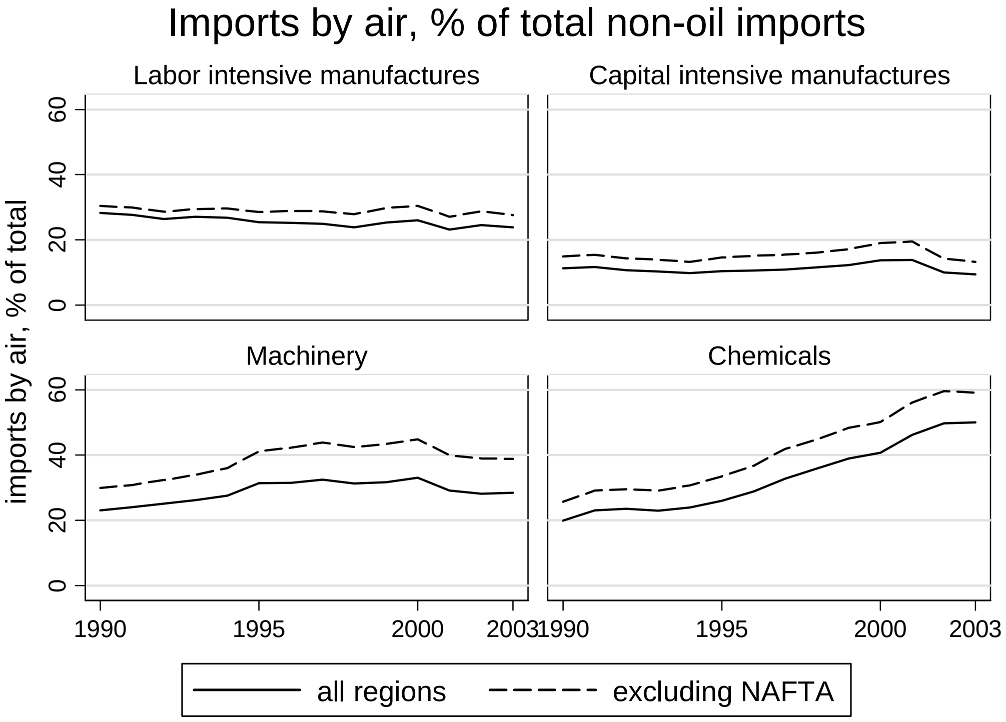
Figure A3 - Change in market shares when air freight costs fall



Notes to Figure A3**:** Illustrates nearby country 2's import market share in country 1. The shift in the market share schedule from $\pi_{21}(z)$ to $\pi'_{21}(z)$ is a consequence of a fall in the cost of air shipment, which causes a shift in the cutoff for air shipment by faraway country 3 from $\bar{z}$ to $\bar{z}'$. The result is lower nominal wages for country 2 relative to country 3, and a greater specialization in heavy goods by country 2.

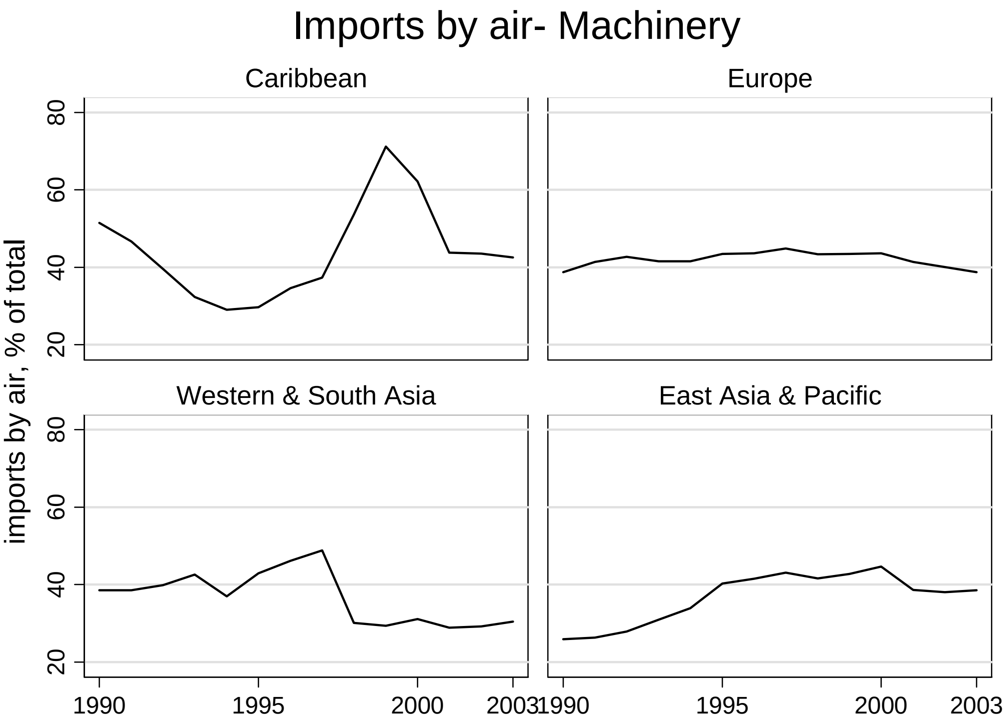
Additional data description

Chart A1



*petroleum removed

Chart A2



*petroleum removed

Additional data description

Chart A3

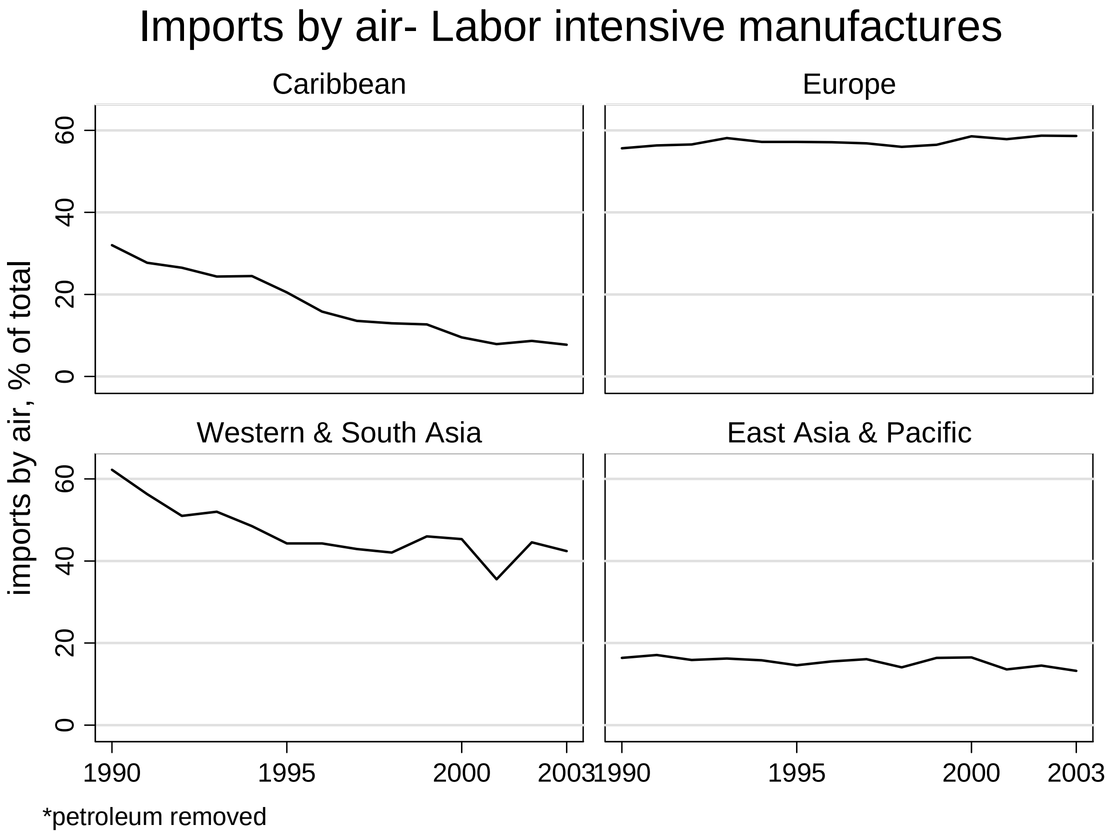
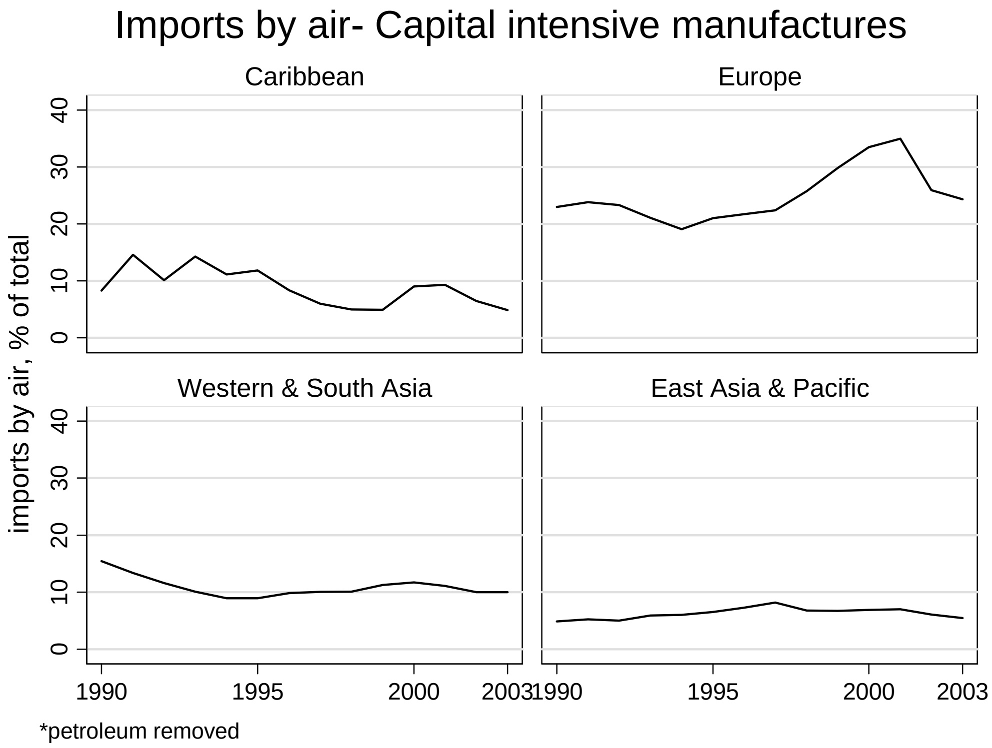
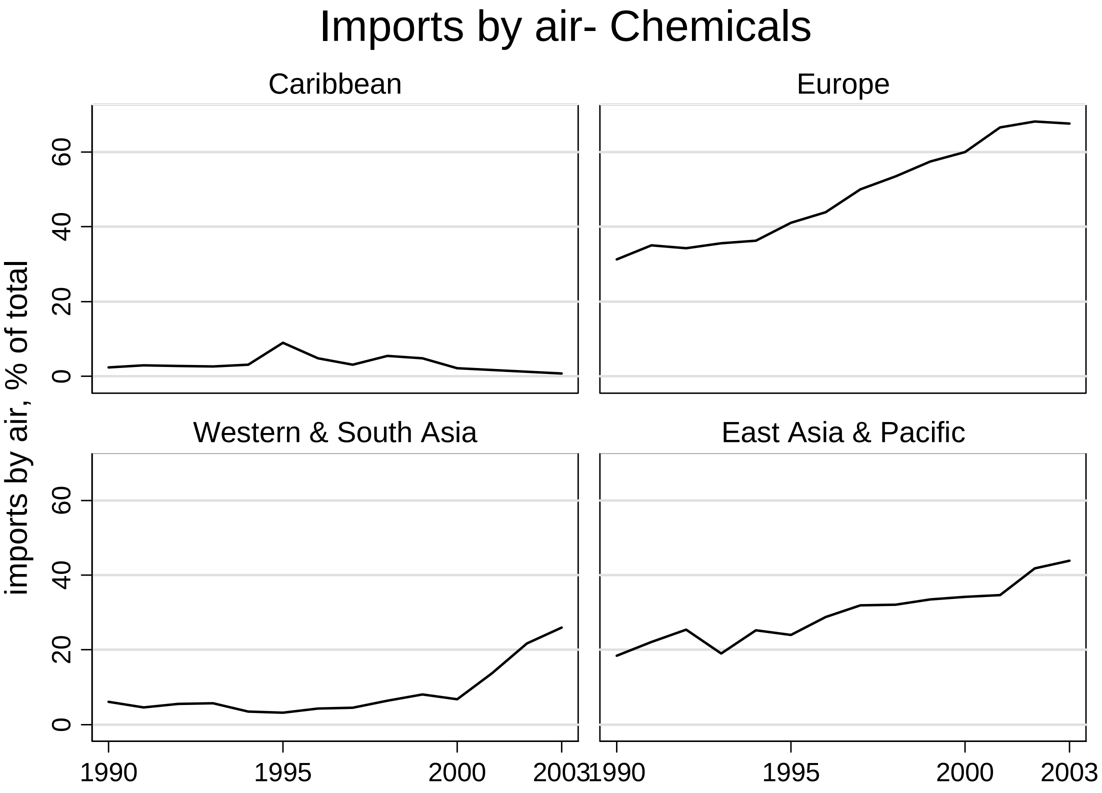


Chart A4



Additional data description

Chart A5



*petroleum removed

Table A1 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		

Additional data description

Table A1 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

Additional data description

Table A2- 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 A3- 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

Additional statistical results

Table A4 Linear market share regressions

	all exporters				high income exporters				middle income exporters			
	1990	1995	2000	2003	1990	1995	2000	2003	1990	1995	2000	2003
all products, continuous air share												
air share	0.066	0.057	0.054	**0.058**	0.072	0.085	0.093	**0.111**	-0.034	-0.042	-0.067	**-0.050**
	9.6	*8.0*	*7.7*	***8.2***	*10.8*	*12.2*	*13.9*	***16.6***	*-2.9*	*-4.0*	*-6.8*	***-5.4***
all products, binary air share indicator												
air share = 0	-0.048	-0.046	-0.021	-0.038	-0.060	-0.063	-0.051	-0.068	0.034	0.044	0.069	0.053
	-9.3	*-8.4*	*-4.0*	*-7.3*	*-11.1*	*-10.4*	*-8.1*	*-10.6*	*4.7*	*7.1*	*12.1*	*10.5*
air share = 1	0.012	-0.010	0.027	0.022	0.003	-0.002	0.031	0.035	0.030	0.031	0.051	0.050
	1.1	*-0.9*	*2.6*	*2.0*	*0.3*	*-0.2*	*3.4*	*4.0*	*2.0*	*2.3*	*3.6*	*3.9*
(air share = 1) - (air share = 0)	0.060	0.036	0.049	**0.060**	0.063	0.061	0.081	**0.103**	-0.004	-0.013	-0.018	**-0.003**
	5.4	*3.1*	*4.3*	***5.2***	*5.9*	*5.6*	*7.9*	***10.0***	*-0.3*	*-0.9*	*-1.3*	***-0.3***
sample size	12,537	13,659	14,469	14,783	12,005	12,918	13,622	13,701	7,093	8,829	10,572	11,614
manufacturing products, continuous air share												
air share	0.078	0.071	0.064	**0.063**	0.080	0.090	0.095	**0.114**	-0.029	-0.040	-0.061	**-0.046**
	11.7	*9.5*	*8.5*	***8.4***	*12.6*	*12.2*	*13.1*	***15.9***	*-2.1*	*-3.3*	*-5.3*	***-4.4***
manufacturing products, binary air share indicator												
air share = 0	-0.043	-0.046	-0.012	-0.032	-0.056	-0.066	-0.044	-0.060	0.022	0.035	0.064	0.050
	-7.1	*-6.7*	*-1.8*	*-4.9*	*-8.9*	*-8.6*	*-5.6*	*-7.3*	*2.7*	*4.9*	*9.5*	*8.4*
air share = 1	0.057	0.033	0.050	0.048	0.042	0.024	0.040	0.055	0.020	0.025	0.047	0.057
	6.6	*3.0*	*4.5*	*4.2*	*5.3*	*2.4*	*4.2*	*6.1*	*1.2*	*1.5*	*2.8*	*3.8*
(air share = 1) - (air share = 0)	0.100	0.079	0.062	**0.080**	0.098	0.090	0.084	**0.114**	-0.003	-0.010	-0.017	**0.007**
	10.2	*6.4*	*5.0*	***6.4***	*10.3*	*7.6*	*7.2*	***10.0***	*-0.1*	*-0.6*	*-1.0*	***0.5***
sample size	9,042	9,783	10,453	10,610	8,838	9,495	10,104	10,120	5,335	6,557	7,852	8,526
non-manufacturing products												
air share	-0.050	-0.036	-0.019	**0.003**	-0.038	0.005	0.026	**0.045**	-0.029	-0.027	-0.048	**-0.044**
	-2.4	*-1.9*	*-1.0*	***0.2***	*-1.8*	*0.3*	*1.5*	***2.5***	*-1.2*	*-1.3*	*-2.4*	***-2.3***
non-manufacturing products, binary air indicator												
air share = 0	-0.013	-0.018	-0.011	-0.020	-0.017	-0.026	-0.022	-0.044	0.067	0.061	0.042	0.050
	-1.2	*-1.7*	*-1.0*	*-2.0*	*-1.5*	*-2.3*	*-1.9*	*-3.8*	*3.4*	*3.9*	*3.4*	*4.3*
air share = 1	-0.102	-0.113	-0.030	-0.027	-0.094	-0.081	0.008	-0.011	0.077	0.050	0.035	0.025
	-3.4	*-4.2*	*-1.2*	*-1.1*	*-3.2*	*-3.0*	*0.4*	*-0.5*	*2.5*	*1.8*	*1.3*	*1.0*
(air share = 1) - (air share = 0)	-0.089	-0.095	-0.020	**-0.006**	-0.076	-0.055	0.030	**0.033**	0.011	-0.011	-0.007	**-0.025**
	-3.0	*-3.5*	*-0.8*	***-0.3***	*-2.6*	*-2.0*	*1.3*	***1.4***	*0.4*	*-0.5*	*-0.3*	***-1.0***
sample size	3,495	3,876	4,016	4,173	3,167	3,423	3,518	3,581	1,758	2,272	2,720	3,088

Additional statistical results

Table A5- Regression of U.S. import unit value on distance and other controls, restricted sample

	1990		1995		2000		2003	
more than 4000km	0.237		0.279		0.227		0.249	
	17.3		*18.4*		*15.8*		*17.8*	
1-4000km		-0.156		-0.062		0.091		0.023
		-6.6		*-2.77*		*3.8*		*0.9*
4000-7800km		0.499		0.582		0.420		0.462
		29.3		*34.1*		*26.8*		*29.3*
7800-14,000km		-0.123		-0.061		-0.008		-0.072
		-6.8		*-3.32*		*-0.5*		*-4.2*
more than 14,000km		0.006		0.034		0.079		0.029
		0.25		*1.53*		*3.7*		*1.4*
log Y/L	0.548	0.368	0.542	0.362	0.543	0.408	0.582	0.409
	74.0	*53.9*	*81.1*	*56.0*	*84.0*	*58.3*	*93.8*	*62.2*
landlocked	0.372	0.181	0.578	0.345	0.413	0.274	0.515	0.333
	14.2	*6.8*	*21.7*	*12.9*	*17.7*	*11.6*	*20.9*	*13.3*
tariff	-0.014	-0.014	-0.006	-0.007	0.003	0.002	-0.004	-0.002
	-6.0	*-6.4*	*-3.9*	*-4.6*	*2.4*	*1.96*	*-4.2*	*-2.0*
R^2 within	0.123	0.164	0.114	0.157	0.112	0.129	0.128	0.152
N	62,834		76,235		86,364		88,756	
HS codes	7,713		8,513		9,076		9,264	

Notes to Table A5: Estimates of equation 14 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 clustered by HS code 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, landlocked is an indicator for the exporter having no port, tariff is ad valorem percentage. 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 A6- Regression of U.S. import value/kilo on distance and other controls, restricted sample

	1990		1995		2000		2003	
more than 4000km	0.340		0.434		0.401		0.412	
	27.6		*30.4*		*28.7*		*32.6*	
1-4000km		0.113		0.156		0.151		0.112
		5.4		*8.3*		*8.1*		*5.7*
4000-7800km		0.597		0.693		0.568		0.621
		37.4		*42.3*		*34.7*		*41.7*
7800-14,000km		0.208		0.300		0.303		0.187
		12.7		*17.5*		*17.3*		*11.4*
more than 14,000km		0.193		0.314		0.291		0.214
		9.1		*14.7*		*13.8*		*10.7*
log Y/L	0.459	0.357	0.475	0.372	0.468	0.385	0.528	0.391
	64.2	*51.1*	*74.4*	*56.7*	*80.1*	*58.4*	*93.9*	*61.8*
landlocked	0.470	0.357	0.557	0.417	0.514	0.426	0.598	0.451
	16.9	*12.3*	*20.7*	*15.6*	*21.6*	*17.7*	*24.1*	*17.9*
tariff	-0.008	-0.012	-0.005	-0.006	0.0003	-0.001	-0.005	-0.004
	-4.5	*-16.7*	*-3.8*	*-4.7*	*0.25*	*-0.9*	*-5.6*	*-4.1*
R^2 within	0.220	0.260	0.215	0.250	0.206	0.220	0.233	0.263
N	31,194		40,039		45,367		47,223	
HS codes	4,046		4,618		4,933		5,080	

Notes to Table A6: Estimates of equation 14 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 clustered by HS code 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, landlocked is an indicator for the exporter having no port, tariff is ad valorem percentage. 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.

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, and a key statistical issue is how to pool the cross-product and cross-country variation within each year when estimating versions of equation (14). Let $i = 1,...,N$ index products and $c = 1,...,C$ index exporting countries.

I can write one of my regression specifications of equation (14) 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 GDP per capita, etc) and $\mathbf{z}_{ic}$ is a vector of variables that vary over both products and exporters (such as tariffs). 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 $N_\mathrm{c} \leq N$ is the number of products exported by country c. The $N_\mathrm{c} \times 1$ vector $\boldsymbol{\mu}_\mathrm{c}$ 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, as well as general heteroskedasticity. 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 4, 5, A4, and A5 are computed by OLS with product fixed effects and heteroskedasticity-robust standard errors.

Turning to equation (15), the same statistical issue arises: it is necessary to allow for cross-product correlation within a country, yet clustering by country has weak statistical justification. However, in this case I face an additional complication that an estimator for random effects probit with endogeneity is not available. Since endogeneity is a more pressing concern in this context, and since this model has no product fixed effects to soak up any of the cross-country variation, I report standard errors clustered by country, but these standard errors may be biased.

Appendix reference

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