Online Appendix to
"International Trade without CES: Estimating Translog Gravity"

by Dennis Novy, *Journal of International Economics*

This appendix contains additional material for the paper "International Trade without CES: Estimating Translog Gravity" by Dennis Novy. Appendix A derives the theoretical translog gravity equation. Appendix B contains the empirical appendix. Whenever I refer to equations, tables, figures or sections from the main text, they carry the same numbering as there.

Appendix A. Deriving the translog gravity equation

This appendix outlines the derivation of the translog gravity equation (7). Substituting the expenditures shares implied by (4) into the market-clearing condition (6) yields

$$y_i = \sum_{j=1}^{J} x_{ij} = \sum_{j=1}^{J} y_j \sum_{m=N_{i-1}+1}^{N_i} s_{mj} = \sum_{j=1}^{J} y_j \sum_{m=N_{i-1}+1}^{N_i} \left(\alpha_m + \sum_{k=1}^{N} \gamma_{km} \ln(p_{kj}) \right).$$

Use $p_{kj}=t_{kj}p_k$ and define world income as $y^W \equiv \sum_{j=1}^{J} y_j$ to obtain

$$y_i = \sum_{j=1}^{J} y_j \sum_{m=N_{i-1}+1}^{N_i} \left(\alpha_m + \sum_{k=1}^{N} \gamma_{km} \ln(t_{kj}) \right) + y^W \sum_{m=N_{i-1}+1}^{N_i} \left(\sum_{k=1}^{N} \gamma_{km} \ln(p_k) \right),$$

which can be rearranged as

$$\sum_{m=N_{i-1}+1}^{N_i} \left(\sum_{k=1}^{N} \gamma_{km} \ln(p_k) \right) = \frac{y_i}{y^W} - \sum_{s=1}^{J} \frac{y_s}{y^W} \sum_{m=N_{i-1}+1}^{N_i} \left(\alpha_m + \sum_{k=1}^{N} \gamma_{km} \ln(t_{ks}) \right),$$

where the first summation index on the right-hand side is changed from j to s. Then substitute the last equation back into the import share (5):

$$\begin{aligned}
\frac{x_{ij}}{y_j} &= \sum_{m=N_{i-1}+1}^{N_i} \left(\alpha_m + \sum_{k=1}^{N} \gamma_{km} \ln(t_{kj}) \right) + \frac{y_i}{y^W} - \sum_{s=1}^{J} \frac{y_s}{y^W} \sum_{m=N_{i-1}+1}^{N_i} \left(\alpha_m + \sum_{k=1}^{N} \gamma_{km} \ln(t_{ks}) \right) \\
&= \frac{y_i}{y^W} + \sum_{s=1}^{J} \frac{y_s}{y^W} \sum_{m=N_{i-1}+1}^{N_i} \left(\sum_{k=1}^{N} \gamma_{km} \ln(t_{kj}) - \sum_{k=1}^{N} \gamma_{km} \ln(t_{ks}) \right) \\
&= \frac{y_i}{y^W} + \sum_{s=1}^{J} \frac{y_s}{y^W} \sum_{m=N_{i-1}+1}^{N_i} \left(\sum_{k=1}^{N} \gamma_{km} \ln\left(\frac{t_{kj}}{t_{ks}} \right) \right).
\end{aligned}$$

Use (3) to arrive at

$$\frac{x_{ij}}{y_j} = \frac{y_i}{y^W} + \sum_{s=1}^{J} \frac{y_s}{y^W} \sum_{m=N_{i-1}+1}^{N_i} \left(\sum_{k=1,k\neq m}^{N} \frac{\gamma}{N} \ln\left(\frac{t_{kj}}{t_{ks}}\right) - \frac{\gamma}{N}(N-1)\ln\left(\frac{t_{mj}}{t_{ms}}\right)\right)$$

$$= \frac{y_i}{y^W} + \sum_{s=1}^{J} \frac{y_s}{y^W} \sum_{m=N_{i-1}+1}^{N_i} \left(\sum_{k=1}^{N} \frac{\gamma}{N} \ln\left(\frac{t_{kj}}{t_{ks}}\right) - \gamma \ln\left(\frac{t_{mj}}{t_{ms}}\right)\right).$$

To ease notation define the geometric mean of trade costs in country *j* as

$$T_j \equiv \left(\prod_{k=1}^{N} t_{kj} \right)^{1/N}$$

so that

$$\frac{x_{ij}}{y_j} = \frac{y_i}{y^W} + \sum_{s=1}^{J} \frac{y_s}{y^W} \sum_{m=N_{i-1}+1}^{N_i} \left(\gamma \ln\left(\frac{T_j}{T_s}\right) - \gamma \ln\left(\frac{t_{mj}}{t_{ms}}\right)\right).$$

Recall that $t_{mj}=t_{ij}$ if $m \in [N_{i-1}+1, N_i]$ so that the previous equation can be rewritten as

$$\frac{x_{ij}}{y_j} = \frac{y_i}{y^W} + \sum_{s=1}^{J} \frac{y_s}{y^W} n_i \left(\gamma \ln\left(\frac{T_j}{T_s}\right) - \gamma \ln\left(\frac{t_{ij}}{t_{is}}\right)\right)$$

$$= \frac{y_i}{y^W} - \gamma n_i \ln(t_{ij}) + \gamma n_i \ln(T_j) + \gamma n_i \sum_{s=1}^{J} \frac{y_s}{y^W} \ln\left(\frac{t_{is}}{T_s}\right),$$

where $n_i \equiv N_i - N_{i-1}$ denotes the number of goods of country *i*. Note that $\ln(T_j)$ can be rewritten as a weighted average of trade costs over the trading partners of country *j*:

$$\ln(T_j) = \frac{1}{N} \sum_{k=1}^{N} \ln(t_{kj}) = \sum_{s=1}^{J} \frac{n_s}{N} \ln(t_{sj}).$$

Appendix B. Empirical appendix

B.1. Robustness checks for the extensive margin measure

The right-hand side variables of translog gravity equation (16) contain a measure of the extensive margin, n_i. But the specific measure based on Hummels and Klenow (2005) that I use for n_i in the regressions in columns 1 and 2 of Table 1 might not be suitable. As a robustness check, I therefore re-estimate equation (16) non-parametrically in two ways.

First, I replace n_i by a set of exporter dummy variables. Thus, I interact the bilateral distance and adjacency variables with these dummy variables. But given that $\ln(T_j^{dist})$ and T_j^{adj} are importer-specific, it becomes impossible to estimate the two variables involving these terms because all degrees of freedom would be exhausted. They are therefore dropped. All 28 resulting distance coefficients are negative and almost all of them are significant at the one percent level. The average distance coefficient is 0.0062 in absolute value. The smallest distance coefficient in absolute value is 0.0001 for Iceland, and the largest distance coefficient in absolute value is 0.0310 for Australia followed by 0.0286 for Germany. As to the interacted adjacency variables, some drop out of the estimation because the corresponding exporting countries are not adjacent to any other countries in the sample. These countries are Australia, Greece, Iceland, Japan, Korea and New Zealand. But the estimated interacted adjacency coefficients have the expected positive sign and most of them are significant. Only one coefficient has a negative sign but it is insignificant. Overall, the R-squared stands at 77 percent.

The translog model in equation (16) suggests that the trade cost coefficients should be correlated with the extensive margin. I therefore compare the interacted distance and adjacency coefficients to empirical measures of the extensive margin. Specifically, I compute the correlation between the estimated coefficients and three different measures. First, I consider the extensive margin measure by Hummels and Klenow (2005). The correlation is 44 percent with the absolute values of the distance coefficients and 50 percent with the adjacency coefficients. Second, I consider the unweighted count of six-digit product categories as an extensive margin measure (see section 3.2). The correlations are 47 percent and 53 percent, respectively. Third, I use (the logarithm of) the exporting country's income (GDP), $\ln(y_i)$, as a simple proxy of the extensive margin. This proxy captures the idea that larger countries tend to export a larger range of goods. The correlations are 58 percent and 64 percent, respectively. Figure B1 plots the individual distance and adjacency coefficients against the three extensive margin measures. In general I find that the trade cost coefficients are related to the extensive margin proxies as suggested by the translog model.

As the second, related way of re-estimating equation (16), I stratify based on intervals of the extensive margin. I divide the sample into five intervals and allow the coefficients on distance and adjacency to vary across these intervals. This stratification is carried out for the three extensive margin measures mentioned above. As the translog model suggests, the intervals representing the smallest extensive margins, denoted by h=1, are expected to have the lowest coefficients in absolute value, and vice versa for the intervals representing the largest extensive margins, denoted by h=5.

The regression results are reported in Table B1. In all three specifications the distance coefficients are lowest in absolute value for the first interval and highest for the fifth interval. In column 2 the distance coefficients are strictly increasing in absolute value, although the pattern is more varied in columns 1 and 3. Formal tests of whether the distance coefficients are equal are rejected for all three specifications (p-values=0.00). A broadly similar pattern arises for the adjacency coefficients. As expected, the highest coefficients are estimated for the fifth intervals, and they tend to become lower for intervals representing smaller extensive margins. Formal tests of whether the adjacency coefficients are equal can be rejected for all specifications (the p-values range between 0.00 and 0.02).

In summary, I therefore conclude that the non-parametric regressions of equation (16) yield results that are consistent with the translog model.

B.2. A multiplicative error term

The traditional gravity specification, for instance in equation (18), typically has the trade flow variable in logarithmic form as the dependent variable with an additive error term. That is, in levels this would correspond to a multiplicative lognormal error term. However, the translog gravity equations (16) and (17) have the trade flow in levels as the dependent variable plus an additive error term.

Here I introduce a multiplicative error term for the translog specification. Instead of the additive error term in equation (14), I assume a multiplicative error term of the following form:

$$\frac{x_{ij}}{y_j} = \left[-\gamma n_i \ln(t_{ij}) + \gamma n_i \ln(T_j) + S_i\right] e^{\varepsilon_{ij}} ,$$

where ε_{ij} is assumed normally distributed. That is, $e^{\varepsilon_{ij}}$ and x_{ij}/y_j are log-normally distributed.

I carry out the estimation with nonlinear least squares (NLS) running the regressions that correspond to columns 1-4 of Table 1. The results are reported in Table B2. As expected, distance is negatively and adjacency is positively related to import shares. The signs and significance of the coefficients are exactly the same as in Table 1. But the values of the individual coefficients are of course not the same because the error term is specified differently (multiplicative in Table B2 as opposed to additive in Table 1). For example, the distance coefficient in column 1 is -0.0133 compared to -0.0296 in the corresponding column of Table 1.

The R-squareds are substantially higher. They fall in the range of 0.91-0.93 compared to the range of 0.50-0.59 in Table 1, indicating that the multiplicative error term produces a better fit. How can this fit be compared to that of the traditional gravity equations in Table 2? Instead of the log-linear specification underlying Table 2, the traditional gravity equations can also be estimated in levels with nonlinear least squares, based on equation (9) with the usual multiplicative error term. This yields exactly the same coefficients and R-squareds as reported in Table 2. Therefore, as the dependent variables are the same, the R-squareds in the range of 0.91-0.93 in Table B2 are directly comparable to the value of 0.89 in Table 2 for the traditional gravity specification. Based on a multiplicative error term the goodness of fit is thus not worse in the translog model than in the traditional model.[A1]

B.3. An alternative stratification procedure

The issue of stratification in the context of equations (19) and (B1) is important (see Appendix B.4 for details on equation B1). Stratifying in terms of *observed* import shares would come down to selection on the endogenous variable and would thus lead to an estimation bias. In line with results from Monte Carlo simulations, in the main part of the paper I resort to a two-step procedure whereby the sample is stratified in terms of intervals of *predicted* import shares. Nevertheless, it might still be a concern that the stratification is in terms of the (predicted) dependent variable. As a robustness check, I therefore adopt an alternative stratification procedure in terms of the right-hand side gravity variables.

[A1] Note that a logarithmic version of the above equation (i.e., taking logarithms on both sides) could also be estimated with nonlinear least squares. The crucial difference would be that the fitted values for import shares would have to be positive as the conditional expectation of import shares would be constrained to be positive due to the logarithmic form. This constraint would lead to lower R-squareds in the range of 0.43-0.74.

Translog gravity theory implies that trade cost elasticities vary across import shares (see equation 12). To be consistent with the guidance from theory, the stratification should thus be based on an indicator of right-hand side variables that is closely related to the import shares. But a problem with the estimating equations is that most right-hand side variables are fixed effects. It therefore makes sense to focus on right-hand side variables that are observable. More specifically, the standard gravity equation (18) has three such variables. The first two are logarithmic distance and the adjacency dummy. The third is the logarithmic income of the exporting country, which is part of the exporter fixed effect $\tilde{S}_i$. The remaining variables embedded in the fixed effects would be the multilateral resistance terms but those are unobservable.

I proceed by constructing an indicator based on these three variables. Adjacency and the exporter's income are supposed to be positively related to import shares, while distance is expected to be negatively related. The difficulty is how to combine these three variables. I first standardize them to remove differences in units of measurement. I then construct a simple unweighted indicator for import shares by adding the standardized variables for adjacency and exporter's income and subtracting the standardized variable for distance. The resulting indicator has a correlation of 77 percent with the observed logarithmic import shares. For example, the three biggest values of the indicator are for the US share of Canadian imports, the Austrian share of Slovakian imports and the French share of Belgian imports, all of which seem sensible. Most importantly, the indicator is based on a combination of arguably exogenous right-hand side variables. It is not based on a first-stage prediction of import shares.

As the final step, I construct five import share intervals based on the indicator and run regression (19). The results are reported in columns 1 and 2 of Table B3. They correspond to columns 1 and 2 of Table 3. The smallest import shares according to the indicator are in interval h=1, and the largest import shares are in interval h=5. As before, hypothesis A puts forward the equality of distance coefficients, i.e., $\lambda_1 = \lambda_2 = \ldots = \lambda_5$. The alternative, consistent with the translog gravity model, is a declining pattern in the absolute distance coefficients, i.e., $\lambda_1 > \lambda_2 > \ldots > \lambda_5$. In column 1 of Table B3 the distance coefficients clearly decline in absolute value except for the last interval. But once adjacency is added as a control in column 2, the last coefficient shrinks in magnitude and becomes the smallest. The hypotheses that the distance coefficients are equal are rejected (the p-values are 0.02 and 0.00 in columns 1 and 2, respectively). As in section 3.3.1, the results therefore seem inconsistent with the standard gravity model.

I carry out a similar procedure for translog gravity and hypothesis B. Specifically, I focus on observable right-hand side variables in equation (17). These are distance, adjacency and the exporter's income and, through the importer fixed effect $\hat{S}_j$, also the two multilateral resistance terms $\ln(T_j^{dist})$ and T_j^{adj}. I standardize them and construct an indicator by adding the variables for adjacency, exporter's income and the multilateral resistance term for distance and by subtracting the variables for distance and the multilateral resistance term for adjacency. I then run regression (B1). The results are reported in columns 3 and 4 of Table B3. They correspond to columns 1 and 2 of Table B4a. Hypothesis B states that the distance coefficients should be equal, i.e., $\kappa_1 = \kappa_2 = \ldots = \kappa_5$ (see Appendix B.4 for details on Table B4a and hypothesis B). In column 3 the distance coefficients show an increasing pattern in absolute value although their equality marginally cannot be rejected (p-value=0.14). Once adjacency is added as a control in column 4, the monotonically increasing pattern disappears and the equality of the coefficients clearly cannot be rejected (p-value=0.35).

Overall, I conclude that the alternative stratification procedure in terms of right-hand side variables leads to similar results as those discussed in section 3.3.1.

B.4. Hypothesis B (heterogeneous distance coefficients for the translog model)

Hypothesis B is based on the translog gravity estimating equation (17). Its premise is that the translog specification is correct and that trade cost coefficients in that estimation should not vary systematically across import shares. I adopt the same strategy as above in that I allow the trade cost coefficients to vary across intervals h=1,...,H of import shares per good, also adding interval fixed effects. For simplicity, I again drop the adjacency variable from the notation so that the estimating equation becomes

$$(B1) \quad \frac{x_{ij} / y_j}{n_i} = -\kappa_h \ln(dist_{ij}) + \hat{S}_i + \hat{S}_j + \hat{S}_h + \upsilon_{ij},$$

where κ_h denotes the trade cost coefficients, $\hat{S}_h$ denotes the interval fixed effect and υ_{ij} is an error term. Hypothesis B states – as predicted by the translog gravity model – that the κ_h distance coefficients should not vary across intervals of import shares per good, i.e., $\kappa_1 = \kappa_2 = \ldots = \kappa_H$. The alternative is – consistent with the standard gravity model – that the magnitude of the κ_h distance coefficients should increase in the import share per good.[A2]

As with hypothesis A, one needs to be careful in constructing the intervals. If they were chosen based on *observed* values of import shares per good, one would incur an upward endogeneity bias in the coefficients of interest, κ_h. But this bias can be avoided if one first estimates equation (17) to obtain common trade cost coefficients, predicts the corresponding import shares and then divides the sample into H intervals of *predicted* import shares per good. I verified the validity of this estimation strategy with Monte Carlo simulations.[A3]

Table B4a presents regression results for equation (B1) under the assumption of H=5, i.e., with five import share intervals. In column 1 where distance is the only trade cost regressor, the distance coefficients appear to generally rise in magnitude across import shares and the hypothesis that they are equal can be rejected (p-value=0.00). However, this rejection is driven by the coefficient for the first interval (equal to -0.0449), which deviates most from the other coefficients. Indeed, the hypothesis that the coefficients for intervals 2-5 are equal cannot be rejected (p-value=0.44). Neither can the hypothesis of equality between all distance coefficients be rejected when I rerun regression (B1) with more intervals.[A4]

In column 2, I add adjacency. Since all adjacent country pairs in the sample fall into the fifth interval, the adjacency variables for the other intervals drop out. With adjacency included, I no longer obtain a monotonic pattern of distance coefficients. In fact, the point estimates for intervals 4 and 5 are smaller in magnitude than for interval 3, and they are not statistically

---

[A2] To see this, divide the constant elasticity gravity equation (9) by y_j and take the derivative with respect to $\ln(t_{ij})$. The result is $d(x_{ij}/y_j)/d\ln(t_{ij})=-(\sigma-1)x_{ij}/y_j$, implying that the absolute value of this derivative is increasing in x_{ij}/y_j. In the translog gravity equation (7), this derivative is given by $d(x_{ij}/y_j)/d\ln(t_{ij})=-\gamma n_i$. If constant elasticity gravity were the true specification, then γn_i should also be increasing in x_{ij}/y_j, or equivalently γ should be increasing in $(x_{ij}/y_j)/n_i$. Thus, in equation (B1) the κ_h distance coefficients should be increasing in $(x_{ij}/y_j)/n_i$.

[A3] I simulated import shares under the assumption that the translog gravity equation (7) is the true model, using distance as the trade cost proxy based on the trade cost function (13) and assuming various arbitrary values for the distance elasticity ρ and the translog parameter γ. The variance of the error term was chosen to match the R-squared of around 55 percent as in Table 1. I divided the sample into intervals based on either the simulated import shares or predicted import shares from a first-stage regression of equation (17). I then ran regression (B1) with both types of intervals, replicating this procedure 1000 times. Forming intervals based on the simulated import shares leads to a severe upward bias in the κ_h coefficients.

[A4] For example, with H=10 the test of coefficient equality cannot be rejected (p-value=0.24).

different from each other (p-value=0.69). This evidence is inconsistent with the pattern of distance coefficients that one would expect under the constant elasticity gravity model.[A5]

In Table B4b I present corresponding results based on equation (16) with x_{ij}/y_j as the dependent variable. Multilateral resistance terms now appear as regressors. As in Table B4a, in columns 1 and 2 intervals are chosen based on predicted import shares per good. As a robustness check, the intervals in columns 3 and 4 are chosen based on predicted import shares only. Distance is the only trade cost regressor in columns 1 and 3. Adjacency is added in columns 2 and 4.

As noted above, if gravity with a constant elasticity were the true underlying model, one should observe a monotonic increase in the absolute distance coefficients across the intervals. However, such a pattern is generally not supported by the estimations. For example, in column 1 the distance coefficient for the first interval (equal to -0.0535) is larger in absolute size than those for intervals 2 and 3 but smaller than those for intervals 4 and 5. In column 2 the smallest distance coefficient is associated with the second interval (equal to -0.0351); in column 3 the smallest coefficient is for the fourth interval (equal to -0.0332); in column 4 the smallest coefficient is for the second interval (equal to -0.0327). Nevertheless, formal tests of coefficient equality across intervals (i.e., hypothesis B) can still be rejected because the coefficients are tightly estimated.

Overall, the tests of heterogeneous distance coefficients in Tables 3, B4a and B4b appear inconsistent with coefficient patterns one should expect under the constant elasticity gravity model combined with a standard log-linear trade cost function. At least qualitatively, they seem more consistent with the predictions of the translog gravity model under the maintained assumption of a log-linear trade cost function.

[A5] I also estimated equation (B1) with nonlinear least squares and a multiplicative error term instead of the additive error term (see Appendix B.2 for details of this estimation procedure). The results are qualitatively similar to those in Table B4a.

Table B1: Non-parametric estimation of equation (16)

	Intervals based on different extensive margin measures n_i		
	n_i=HK (2005)	n_i=unweighted count	n_i=ln(y_i)
Dependent variable	x_{ij}/y_j	x_{ij}/y_j	x_{ij}/y_j
	(1)	(2)	(3)
n_i ln($dist_{ij}$), h=1	-0.0009***	-0.0011***	-0.0011***
	(0.0003)	(0.0003)	(0.0003)
n_i ln($dist_{ij}$), h=2	-0.0150***	-0.0040***	-0.0061***
	(0.0042)	(0.0011)	(0.0015)
n_i ln($dist_{ij}$), h=3	-0.0085***	-0.0041***	-0.0144***
	(0.0033)	(0.0011)	(0.0048)
n_i ln($dist_{ij}$), h=4	-0.0074***	-0.0106***	-0.0066*
	(0.0021)	(0.0022)	(0.0035)
n_i ln($dist_{ij}$), h=5	-0.0201***	-0.0338***	-0.0268***
	(0.0047)	(0.0075)	(0.0056)
n_i adj_{ij}, h=1	0.0019	0.0046**	0.0046**
	(0.0011)	(0.0019)	(0.0019)
n_i adj_{ij}, h=2	0.0187*	0.0049*	0.0017
	(0.0110)	(0.0027)	(0.0028)
n_i adj_{ij}, h=3	0.0137	0.0182***	0.0245**
	(0.0097)	(0.0060)	(0.0098)
n_i adj_{ij}, h=4	0.0174***	0.0205*	0.0227
	(0.0063)	(0.0115)	(0.0159)
n_i adj_{ij}, h=5	0.0568***	0.0392***	0.0451***
	(0.0181)	(0.0136)	(0.0147)
R-squared	0.62	0.68	0.66
Observations	749	749	749

Notes: The index h denotes intervals in order of ascending extensive margin measures. The intervals in column 1 are based on the measure by Hummels and Klenow (2005). The intervals in column 2 are based on the unweighted count of six-digit product categories. The intervals in column 3 are based on the logarithmic income of the exporter. The n_i T_j^{dist} and n_i T_j^{adj} regressors are included but not reported here. Robust standard errors clustered around country pairs (378 clusters) reported in parentheses, OLS estimation. Exporter fixed effects not reported. * significant at 10% level. ** significant at 5% level. *** significant at 1% level.

Table B2: Translog gravity, nonlinear least squares estimation

Dependent variable	x_{ij}/y_j	x_{ij}/y_j	$(x_{ij}/y_j)/n_i$	$(x_{ij}/y_j)/n_i$
	(1)	(2)	(3)	(4)
$n_i \ln(dist_{ij})$	-0.0133***	-0.0119***		
	(0.0008)	(0.0010)		
$n_i \ln(T_j^{dist})$	0.0027***	0.0029***		
	(0.0004)	(0.0010)		
$n_i\, adj_{ij}$		0.0226***		
		(0.0050)		
$n_i\, T_j^{adj}$		-0.0071**		
		(0.0033)		
$\ln(dist_{ij})$			-0.0164***	-0.0133***
			(0.0011)	(0.0008)
adj_{ij}				0.0174***
				(0.0042)
R-squared	0.91	0.93	0.92	0.92
Observations	749	749	749	749

Notes: Robust standard errors clustered around country pairs reported in parentheses, NLS estimation. Columns 1 and 2: exporter fixed effects not reported. Columns 3-4: exporter and importer fixed effects not reported. ** significant at 5% level. *** significant at 1% level.

Table B3: Stratification in terms of right-hand side variables

	Constant elasticity gravity (Hypothesis A)		Translog gravity (Hypothesis B)	
Dependent variable	$\ln(x_{ij}/y_j)$	$\ln(x_{ij}/y_j)$	$(x_{ij}/y_j)/n_i$	$(x_{ij}/y_j)/n_i$
	(1)	(2)	(3)	(4)
$\ln(dist_{ij})$, h=1	-1.4742***	-1.4788***	-0.0122***	-0.0137***
	(0.1613)	(0.1570)	(0.0022)	(0.0021)
$\ln(dist_{ij})$, h=2	-1.3132***	-1.3366***	-0.0135***	-0.0164***
	(0.1323)	(0.1189)	(0.0041)	(0.0038)
$\ln(dist_{ij})$, h=3	-1.0595***	-1.1110***	-0.0170***	-0.0214***
	(0.1181)	(0.1115)	(0.0047)	(0.0044)
$\ln(dist_{ij})$, h=4	-0.8761***	-0.8920***	-0.0177***	-0.0207***
	(0.1366)	(0.1229)	(0.0049)	(0.0044)
$\ln(dist_{ij})$, h=5	-1.0044***	-0.7824***	-0.0285***	-0.0219***
	(0.1114)	(0.1187)	(0.0074)	(0.0067)
adj_{ij}, h=4				0.0211***
				(0.0079)
adj_{ij}, h=5		0.8003***		0.0501***
		(0.1745)		(0.0104)
R-squared	0.89	0.89	0.56	0.59
Observations	749	749	749	749

Notes: The index h denotes intervals in order of ascending stratified import shares. See Appendix B.3 for details of the stratification. The adj_{ij} regressors for intervals h=1-4 in column 2 drop out since no adjacent country pair falls into these intervals (intervals h=1-3 in column 4). Robust standard errors clustered around country pairs (378 clusters) reported in parentheses, OLS estimation. Exporter and importer fixed effects and interval fixed effects not reported. *** significant at 1% level.

Table B4a: Testing translog gravity (Hypothesis B)

	Intervals based on $(x_{ij}/y_j)/n_i$	
Dependent variable	$(x_{ij}/y_j)/n_i$	$(x_{ij}/y_j)/n_i$
	(1)	(2)
$\ln(dist_{ij})$, h=1	-0.0449***	-0.0347***
	(0.0068)	(0.0039)
$\ln(dist_{ij})$, h=2	-0.0518***	-0.0383***
	(0.0077)	(0.0042)
$\ln(dist_{ij})$, h=3	-0.0516***	-0.0412***
	(0.0078)	(0.0046)
$\ln(dist_{ij})$, h=4	-0.0543***	-0.0411***
	(0.0079)	(0.0045)
$\ln(dist_{ij})$, h=5	-0.0567***	-0.0380***
	(0.0084)	(0.0057)
adj_{ij}, h=5		0.0608***
		(0.0103)
R-squared	0.64	0.71
Observations	749	749

Notes: The index h denotes intervals in order of ascending predicted import shares. The intervals are based on predicted import shares divided by n_i. The adj_{ij} regressors for intervals h=1-4 drop out in column 2 since no adjacent country pair falls into these intervals. Robust standard errors clustered around country pairs (378 clusters) reported in parentheses, OLS estimation. Exporter and importer fixed effects and interval fixed effects not reported. *** significant at 1% level.

Table B4b: Testing translog gravity (Hypothesis B)

	Intervals based on $(x_{ij}/y_j)/n_i$		Intervals based on (x_{ij}/y_j)	
Dependent variable	x_{ij}/y_j	x_{ij}/y_j	x_{ij}/y_j	x_{ij}/y_j
	(1)	(2)	(3)	(4)
$n_i \ln(dist_{ij})$, h=1	-0.0535***	-0.0406***	-0.0403***	-0.0369***
	(0.0090)	(0.0064)	(0.0085)	(0.0061)
$n_i \ln(dist_{ij})$, h=2	-0.0446***	-0.0351***	-0.0338***	-0.0327***
	(0.0081)	(0.0052)	(0.0075)	(0.0054)
$n_i \ln(dist_{ij})$, h=3	-0.0507***	-0.0376***	-0.0334***	-0.0337***
	(0.0085)	(0.0054)	(0.0069)	(0.0053)
$n_i \ln(dist_{ij})$, h=4	-0.0585***	-0.0406***	-0.0332***	-0.0343***
	(0.0095)	(0.0062)	(0.0061)	(0.0055)
$n_i \ln(dist_{ij})$, h=5	-0.0627***	-0.0476***	-0.0601***	-0.0439***
	(0.0087)	(0.0077)	(0.0079)	(0.0084)
$n_i \ln(T_j^{dist})$, h=1	0.0430***	0.0286***	0.0291***	0.0258***
	(0.0076)	(0.0049)	(0.0065)	(0.0047)
$n_i \ln(T_j^{dist})$, h=2	0.0300***	0.0189***	0.0201***	0.0183***
	(0.0067)	(0.0037)	(0.0057)	(0.0039)
$n_i \ln(T_j^{dist})$, h=3	0.0343***	0.0199***	0.0195***	0.0189***
	(0.0072)	(0.0043)	(0.0058)	(0.0040)
$n_i \ln(T_j^{dist})$, h=4	0.0391***	0.0207***	0.0184***	0.0163***
	(0.0085)	(0.0055)	(0.0055)	(0.0044)
$n_i \ln(T_j^{dist})$, h=5	0.0417***	0.0256***	0.0413***	0.0242***
	(0.0084)	(0.0067)	(0.0083)	(0.0079)
$n_i\, adj_{ij}$, h=5		0.0536***		0.0529***
		(0.0161)		(0.0161)
$n_i\, T_j^{adj}$, h=5		-0.1309**		-0.0933*
		(0.0647)		(0.0501)
R-squared	0.64	0.69	0.64	0.68
Observations	749	749	749	749

Notes: The index h denotes intervals in order of ascending predicted import shares. The intervals in columns 1 and 2 are based on predicted import shares divided by n_i. The intervals in columns 3 and 4 are based on predicted import shares only. The $n_i\, adj_{ij}$ regressors for intervals h=1-4 drop out in column 2 since no adjacent country pairs fall into these intervals (intervals h=1, 2 and 4 in column 4). The $n_i\, T_j^{adj}$ regressors for intervals h=1-4 in columns 2 and 4 are included but not reported here. Robust standard errors clustered around country pairs (378 clusters) reported in parentheses, OLS estimation. Exporter fixed effects and interval fixed effects not reported. * significant at 10% level. ** significant at 5% level. *** significant at 1% level.

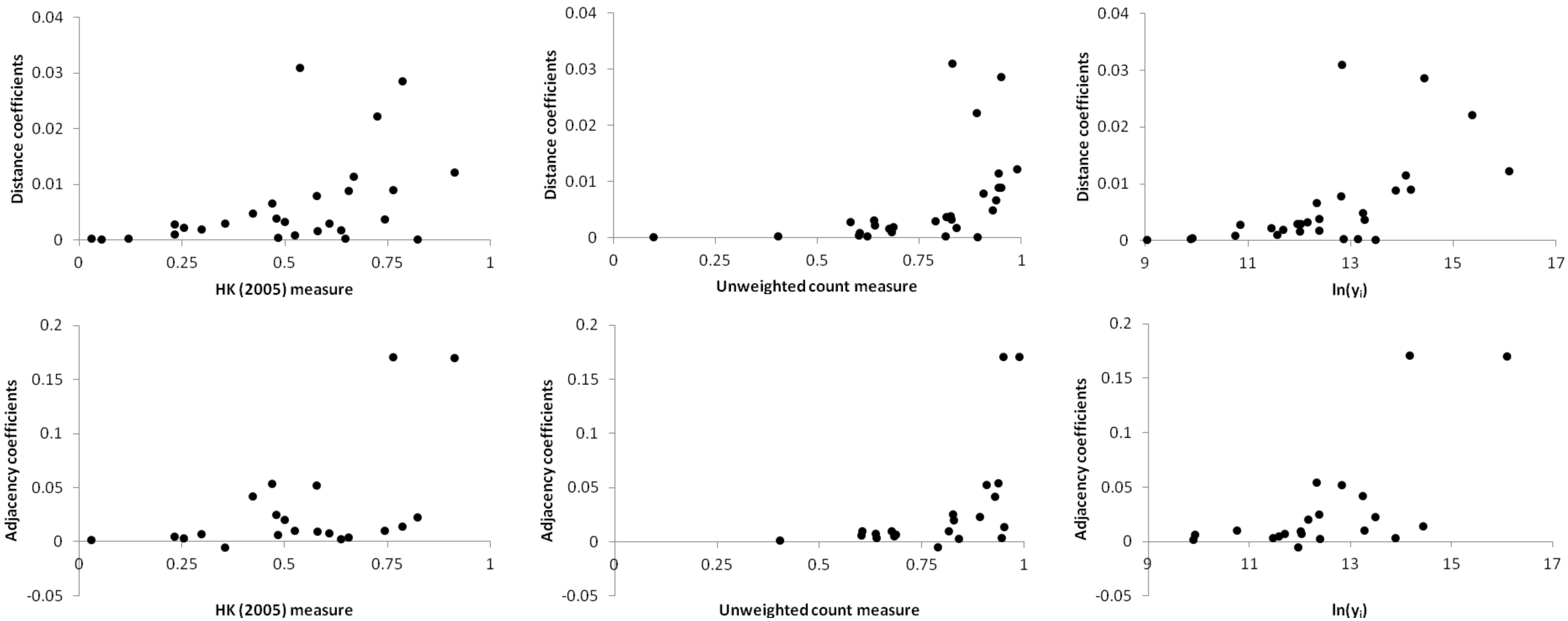

Figure B1: Non-parametric estimates of distance and adjacency coefficients. In the top panels the absolute values of distance coefficients are plotted against three extensive margin measures, and in the bottom panels the values of adjacency coefficients are plotted against the three measures. The extensive margin measures are by Hummels and Klenow (2005) in the left-hand side panels, an unweighted count of six-digit product categories in the middle panels and the logarithmic income of the exporter in the right-hand side panels.