Econometrica Supplementary Material

SUPPLEMENT TO "NONHOMOTHETICITY AND BILATERAL TRADE: EVIDENCE AND A QUANTITATIVE EXPLANATION"

(*Econometrica*, Vol. 79, No. 4, July 2011, 1069–1101)

BY ANA CECÍLIA FIELER

Appendix A generalizes the utility function in the original paper, Appendix B describes the data, Appendix C presents robustness checks, and Appendix D presents Monte Carlo simulations.

APPENDIX A: AN ALTERNATIVE FORM FOR THE UTILITY FUNCTION

UTILITY FUNCTION (1) in the main paper captures the empirical finding that poor households spend most of their income on basic goods, while rich ones spend it on luxuries. The same parameter σ_τ, however, controls both the elasticity of substitution across goods and the income elasticity of demand—two objects that need not be linked in reality. A more general form for preferences relaxes this link:

$$\text{(S1)} \qquad \sum_{\tau=1}^{S}\left\{\alpha_\tau \frac{\sigma_\tau}{\gamma_\tau(\sigma_\tau-1)}\left[\int_0^1 q(j_\tau)^{\sigma_\tau-1/\sigma_\tau}\,dj_\tau\right]^{\gamma_\tau}\right\}.$$

Denote by λ the consumer's Lagrange multiplier and by P_τ the CES price index for goods of type $\tau=1,\dots,S$. I consider two (not exhaustive, but instructive) cases.

CASE 1—$\gamma_\tau=\sigma_\tau/(\sigma_\tau-1)$ for All τ: The first-order conditions imply that for all τ, $\lambda=\frac{\alpha_\tau}{P_\tau}$ if $q(j_\tau)>0$ for some j_τ. Since these equalities do not hold simultaneously for arbitrary prices, the consumer only demands goods of the type with the highest value for $\frac{\alpha_\tau}{P_\tau}$. Importantly, the Lagrange multiplier and hence consumer demand do not depend on income: preferences are homothetic.

CASE 2—$\gamma_\tau\neq\sigma_\tau/(\sigma_\tau-1)$ for All τ: Spending on any two types of goods, 1 and 2, satisfies

$$\frac{x_1}{x_2}=\lambda^{\varphi_1-\varphi_2}\left[\frac{(\alpha_1)^{\varphi_1}P_1^{\phi_1}}{(\alpha_2)^{\varphi_2}P_2^{\phi_2}}\right],$$

where $\varphi_\tau=-\sigma_\tau+\frac{\sigma_\tau(1-\sigma_\tau)(\gamma_\tau-1)}{\sigma_\tau+\gamma_\tau-\sigma_\tau\gamma_\tau}$ and $\phi_\tau=\frac{(1-\sigma_\tau)\gamma_\tau}{\sigma_\tau+\gamma_\tau-\sigma_\tau\gamma_\tau}$ for $\tau=1,2$. As in equation (2), the term in brackets determines the level of x_1/x_2, and $(\lambda^{\varphi_1-\varphi_2})$ determines how it changes with consumer income. Note, however, that the parameters added to equation (S1) relative to the original utility function are not

 DOI: 10.3982/ECTA8346

identified. Consider the case with two types, A and B. For any set of prices P_τ and parameters $\{\sigma_A, \sigma_B, \gamma_A, \gamma_B\}$, the parameter α_A can be judiciously chosen to match any level of the ratio x_A/x_B. The rate of change of x_A/x_B, in turn, is governed by the exponent of the Lagrange multiplier, $(\varphi_A - \varphi_B)$. Parameters σ_A, σ_B, γ_A, and γ_B, therefore, all play the same role, and only one of them is sufficient to pin down $(\varphi_A - \varphi_B)$; the rest can be normalized. In the paper, γ_A and γ_B are set to 1.

In sum, utility function in the paper assumes that the type of good with a high σ_τ has a high income elasticity of demand and a high elasticity of substitution across goods. Case 2 shows that this assumption is not necessary for any of the results. Without changing predicted trade flows, a different normalization of the parameters may imply that the type of good with a *high* income elasticity of demand has a *low* elasticity of substitution across goods.

APPENDIX B: DATA

I use data on bilateral merchandise trade flows in 2000 from the U.N. Comtrade data base (United Nations (2008)). In compiling the data, I give precedence to trade flows reported by the importing country. If the importer's report is not available, I use the exporter's. I keep in the sample only countries with matching data on population and GDP from the World Bank (2008).[1] Countries also report total trade flows to the World Bank. These flows should be (weakly) larger than the sum of trade flows in the U.N. Comtrade data base, because the U.N. Comtrade data base does not contain all countries. Accordingly, I exclude countries whose total trade flows in the U.N. Comtrade data base are more than 20% larger than in the World Bank data. Seventeen countries and 0.3% of world trade are excluded using this criterion.[2]

The resulting data comprise 162 countries and 95% of world trade in 2000. Of these countries, 145 directly report trade to the United Nations. All trade flows to and from reporting countries are observed, and trade flows between the remaining 17 countries (marked with a dagger in Table S.I) are missing. Hence, of all possible importer–exporter pairs, 25,810 ($= 162^2 - 17^2 - 145$) are observed and 272 ($= 17^2 - 17$) are missing.

[1]Neither the United Nations nor the World Bank officially reports statistics for Taiwan, but in practice, the U.N. country classification "Other Asia, not elsewhere specified" (code 490) refers to Taiwan. Data on Taiwan's population and income, in turn, are taken directly from the Taiwanese government's website http://eng.stat.gov.tw.

[2]These countries are Brunei Darussalam, Comoros, Djibouti, Gabon, Georgia, Guinea-Bissau, Guatemala, Honduras, Kiribati, Moldova, Mauritania, Panama, Sierra Leone, Timor-Leste, St. Vincent and Grenadine, Vanatu, and Samoa.

TABLE S.I
LIST OF COUNTRIES[a]

Albania	Denmark*	Lebanon	Serbia and Montenegro
Algeria	Dominica	Lesotho	Seychelles†
Angola†	Dominican Republic	Libya†	Singapore
Antigua and Barbuda	Ecuador	Lithuania	Slovak Republic
Argentina	Egypt, Arab Rep.	Luxembourg	Slovenia
Armenia	El Salvador	Macao	Solomon Islands†
Australia*	Equatorial Guinea†	Macedonia	South Africa
Austria*	Eritrea	Madagascar	Spain*
Azerbaijan	Estonia	Malawi	Sri Lanka†
Bahamas, The	Ethiopia	Malaysia	St. Kitts and Nevis
Bahrain	Fiji	Maldives	St. Lucia
Bangladesh	Finland*	Mali	Sudan
Barbados	France*	Malta	Suriname
Belarus	French Polynesia	Mauritius	Swaziland
Belgium*	Gambia, The	Mexico	Sweden*
Belize	Germany*	Mongolia	Switzerland
Benin	Ghana	Morocco	Syrian Arab Republic
Bhutan†	Greece*	Mozambique	Taiwan†
Bolivia	Grenada	Namibia	Tajikistan
Bosnia and Herzegovina†	Guinea	Nepal	Tanzania
Botswana	Guyana	Netherlands*	Thailand
Brazil	Haiti†	New Caledonia	Togo
Bulgaria	Hong Kong	New Zealand*	Tonga
Burkina Faso	Hungary	Nicaragua	Trinidad and Tobago
Burundi	Iceland	Niger	Tunisia
Cambodia	India	Nigeria	Turkey
Cameroon	Indonesia	Norway*	Turkmenistan
Canada*	Iran, Islamic Rep.	Oman	Uganda
Cape Verde	Ireland	Pakistan	Ukraine
Central African Republic	Israel	Papua New Guinea	United Arab Emirates
Chad†	Italy*	Paraguay	United Kingdom*
Chile	Jamaica	Peru	United States*
China	Japan*	Philippines	Uruguay
Colombia	Jordan	Poland	Uzbekistan†
Congo, Dem. Rep.†	Kazakhstan	Portugal*	Venezuela
Congo, Rep.†	Kenya	Qatar	Vietnam
Costa Rica	Korea, Rep.	Romania	Yemen, Rep.†
Cote d'Ivoire	Kuwait	Russian Federation	Zambia
Croatia	Kyrgyz Republic	Rwanda†	Zimbabwe
Cyprus	Lao PDR†	Saudi Arabia	
Czech Republic	Latvia	Senegal	

[a]The dagger (†) denotes countries that do not report trade to the United Nations; the asterisk (*) denotes OECD countries.

These data used in the estimation contain total trade for 145 countries. To construct Figures 3(a), 4(a), and 5(a), I complement them with data from the World Bank (2008) on the total value of merchandise trade flows for 16 of the

17 nonreporting countries (Taiwan is missing). For the 145 reporting countries, the correlation between trade shares in the World Bank and in the U.N. Comtrade data is 0.95. This use of extraneous data is a small out-of-sample check on the model, and it does not change any of the empirical results.

The U.N. Comtrade data are available until 2005. I use the year 2000 for two reasons. First, there are more and more reliable data for 2000. Twenty-four countries that report trade for 2000 have not yet reported for 2005, and the inconsistencies between the U.N. Comtrade and the World Bank data are larger in 2005 than in 2000. Second, several countries trade close to or more than 100% of their GDP in 2005. The simple model of Section 2 does not account for trade flows beyond 100% of a country's GDP; only the model with intermediate inputs in Section 5.2 does. But as discussed there, intermediate inputs do not qualitatively change the analysis: they make the model computationally more demanding and they compromise the identification of parameters. Still, it is worth noting that all the stylized facts of the data exploited in Section 3 also hold in the 2005 data.

All main results are robust to the use of data from 1999 or 2001, instead of 2000, or to the use of data from Feenstra, Lipsey, Deng, Ma, and Mo (2005) or from the International Monetary Fund (2008), instead of the U.N. Comtrade data base.

APPENDIX C: ROBUSTNESS

C.1. *Weighting*

Santos Silva and Tenreyro (2006) discussed the weighting of observations in trade. Trade flows among large countries presumably have large variances because their values are large, and trade flows among small, typically poor, countries may have large measurement errors. To account for size in the empirical Section 3, I simply divide trade flows by the product of the importer and exporter total income, $z_{ni} = \frac{X_{ni}}{X_n X_i}$. Efficiency gains from estimating an optimal weighting matrix with NLLS are, to my knowledge, not established because numerical errors from the first stage enter the estimated matrix.[3] In addition, here, it would require ad hoc assumptions on the structure of the matrix, since error terms are heteroskedastic and clustered.

I experiment with several alternative weights. For all weights, the message of the empirical results remains: The new model is closer to the data than the EK model because it reconciles the large volume of trade among rich countries with the small volume among small, poor countries. Normalizing trade flows to $z_{ni} = \log(1 + X_{ni})$ does not change the results.[4] The results change only in

[3]Even in linear estimations, Altonji and Segal (1996) used Monte Carlo simulations to show that the gains from estimating an optimal weighting matrix are not clear in small samples.

[4]Gravity-type regressions typically involve taking the logarithm of trade flows. I take the logarithm of 1 plus trade flows because 7034 of the 25,810 observations are zeroes.

specifications that put more weight on trade flows among large countries. For example, if $z_{ni} = \frac{X_{ni}}{(X_n X_i)^{1/2}}$, the EK model overestimates trade of small countries, instead of underestimating trade of rich countries, as in Section 3. The average trade share of the 30 smallest countries is 92% in the EK model, 58% in the new model, and 38% in the data. If even more weight is put on large countries (e.g., $z_{ni} = X_{ni}$), then the optimization algorithm captures only trade flows among large rich countries; parameter estimates approach those of the OECD subsample and the difference between the two models narrows.

C.2. *Normalization of Parameters θ_A and σ_A*

In the paper, I estimate the model by fixing $\theta_A = 8.28$ and $\sigma_A = 5$ in specification 1 (equation (13) of the paper) and $\theta_A = 8.28$ in specification 2 (equation (15) of the paper). Table S.II shows the parameter estimates of both specifications when θ_A equals 3.60 and 12.86, the other estimates in Eaton and Kor-

TABLE S.II

ESTIMATES OF THE NEW MODEL WITH DIFFERENT VALUES FOR θ_A AND σ_A[a]

	Specification 1				Specification 2	
θ_A	8.28	8.28	**3.60**	**12.86**	**3.60**	**12.86**
σ_A	**2.00**	**8.00**	4.00	5.00		
γ_1	1.40	1.37	1.61	1.26	1.38	1.20
	(0.05)	(0.03)	(0.09)	(0.03)	(0.22)	(0.03)
γ_2	0.20	0.19	0.40	0.14	0.18	0.09
	(0.03)	(0.03)	(0.06)	(0.02)	(0.12)	(0.02)
γ_3	−0.01	−0.01	−0.01	−0.01	0.00	0.00
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Border	0.98	0.99	1.01	0.99	0.95	0.93
	(0.07)	(0.08)	(0.08)	(0.07)	(0.05)	(0.05)
Language	0.94	0.93	0.91	0.94	0.90	0.93
	(0.05)	(0.06)	(0.06)	(0.05)	(0.04)	(0.05)
Trade agreement	1.20	1.23	1.24	1.22	1.24	1.21
	(0.12)	(0.12)	(0.14)	(0.11)	(0.09)	(0.08)
θ_B	13.73	14.67	9.49	19.65	15.02	26.01
	(1.02)	(0.80)	(0.71)	(1.01)	(1.39)	(1.69)
$(\alpha_A)^{1/\sigma_A}$	0.68	0.76	0.77	0.82		
	(0.04)	(0.02)	(0.03)	(0.01)		
σ_B	1.00	1.91	1.22	1.25		
	(0.03)	(0.07)	(0.05)	(0.03)		
R^2	42%	42%	42%	42%	67%	67%

[a]Standard errors given in parentheses are clustered by importer and exporter. The number of observations is 25,810 in all cases.

tum (2002). Consistent with Section 3.1.2, an increase in θ_A decreases parameters γ_1 and γ_2 and increases θ_B. Estimates of Υ and θ_B are again similar in both specifications, suggesting that the model is correctly specified. The table also presents results from specification 1 with $\sigma_A = 2.0$ and 8.0, arbitrary values satisfying $1 < \sigma_A < (\theta_A + 1)$. An increase in σ_A increases σ_B. In all cases of Table S.II, predicted flows do not change, so neither the R^2 nor the stylized facts of Section 3 change.

Direct measures of trade costs can help pin down parameter θ_A. For example, Anderson and Van Wincoop (2004) estimated that trade costs across OECD countries are equivalent to an ad valorem tax of 74%. In the model, trade costs across OECD countries average 73% when $\theta_A = 12.86$. This result, however, is limited because trade costs have large measurement errors and vary a lot across goods and countries.

APPENDIX D: MONTE CARLO SIMULATIONS

To check for identification in the first empirical specification (equation (13) in the paper), I conduct Monte Carlo simulations. I make a random draw for each of the estimated parameters, $\{\Upsilon, \alpha_A, \sigma_B, \theta_B\}$, simulate data with the deterministic model of Section 2, and then run the optimization algorithm on simulated data.[5] I repeat this procedure 50 times and compare the parameter estimates with the original parameter draws. The parameters that distinguish the new model from EK, α_A, σ_B, and θ_B, are identified with a high degree of precision: in 98% of the simulations, the estimated parameter is within a 5% distance from its original draw. Identification of the iceberg cost parameters Υ is weaker because they are correlated. Still, in 84% of simulations the estimated parameters are within a 5% distance from the original draw.

REFERENCES

ALTONJI, J., AND L. M. SEGAL (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business & Economic Statistics*, 14, 353–366. [4]

ANDERSON, J., AND E. VAN WINCOOP (2004): "Trade Costs," *Journal of Economic Literature*, 42, 691–751. [6]

EATON, J., AND S. KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779. [6]

FEENSTRA, R., R. E. LIPSEY, H. DENG, A. C. MA, AND H. MO (2005): "World Trade Flows: 1962–2000," Working Paper 11040, NBER. [4]

INTERNATIONAL MONETARY FUND (2008): *Direction of Trade and Statistics CD ROM*. Washington, DC: Publication Services International Monetary System. [4]

SANTOS SILVA, J. M. C., and TENREYRO, S. (2006): "The Log of Gravity," *The Review of Economics and Statistics*, 88, 641–658. [4]

[5]For each parameter, I randomize over a uniform distribution with support $\gamma_1 \in [0.5, 2]$, $\gamma_2 \in [0.05, 0.5]$, $\gamma_3 \in [-0.1, 0]$, $\{\gamma_{\text{border}}, \gamma_{\text{language}}, \gamma_{\text{trade agreement}}\} \in [0.5, 1.2]$, $(\alpha_A)^{1/\sigma_A} \in [0.5, 0.9]$, $\sigma_B \in [1.0, 8.]$, and $\theta_B \in [5, 20]$.

UNITED NATIONS (2008): "UN Comtrade," UN Commodity Trade Statistics Database, Statistics Division, United Nations. Available at http://comtrade.un.org/. [2]

WORLD BANK (2008): "World Development Indicators 2008," Data Set, World Bank, Washington, DC. Available at http://web.worldbank.org/data. [2,3]

Dept. of Economics, University of Pennsylvania, 3718 Locust Walk, 434 McNeil, Philadelphia, PA 19104-6297, U.S.A.; afieler@econ.upenn.edu.

Manuscript received January, 2009; final revision received June, 2010.