Appendix: The Effects of Political Institutions on the Extensive and Intensive Margins of Trade

In Song Kim* John Londregan† Marc Ratkovic‡

February 12, 2019

Abstract

This Web Appendix contains supplemental materials. Section A1 provides the details of the network model and formal proofs of Proposition 1. In Section A2, we describe the proposed two-stage Bayesian LASSO estimator. Section A3 formalizes the concept of first order stochastic dominance that we utilize to make model comparisons. Section A4 provides robustness check results accounting for the importance of oil rich countries. The rest of this Appendix provides further details for the *Polity IV* measure, the list of 131 countries used for our analysis along with a detailed description of the parameters used in our empirical analysis.

Key Words: extensive and intensive margins of trade, Polity, democracy, autocracy, international trade, variable selection, LASSO

*Associate Professor, Department of Political Science, Massachusetts Institute of Technology, Cambridge, MA, 02139. Email: insong@mit.EDU, URL: http://web.mit.edu/insong/www/

†Professor of Politics and International Affairs, Woodrow Wilson School, Princeton University, Princeton NJ 08544. Phone: 609-258-4854, Email: jbl@princeton.edu

‡Assistant Professor, Department of Politics, Princeton University, Princeton NJ 08544. Phone: 608-658-9665, Email: ratkovic@princeton.edu, URL: http://www.princeton.edu/∼ratkovic

A1 Formal Analysis of the Margins of Trade

We provide further details about the analysis in Section 2.

A1.1 Welfare Gains from International Trade

We begin by deriving the Marshallian demand and indirect utility functions given the CES utility given in expression (3). Denote the price of good F relative to the numeraire good D by p. The budget constraint is:

$$q_D + p \cdot q_F = I \tag{16}$$

where, for an individual with endowment (ι_D, ι_F), income is given by:

$$\iota_D + p \cdot \iota_F = I$$

Maximizing the utility given in expression (3) with respect to the budget constraint (16) gives rise to the standard Marshallian demand functions:

$$q_D(p, I) = \frac{1}{1 + p^{1-\sigma}} I \qquad \text{and} \qquad q_F(p, I) = \frac{p^{-\sigma}}{1 + p^{1-\sigma}} I$$

Substituting these demands into the utility function, see expression (3) in the text, gives the "indirect utility function" that prevails at an income level of I and a relative price for good F of p:

$$\psi(p, I) \equiv U\Big(q_D(p, I), q_F(p, I)\Big) = \big(1 + p^{1-\sigma}\big)^{\frac{1}{\sigma-1}} I. \tag{17}$$

For an individual endowed with one unit of good D and no units of good F, an increase in p causes welfare to fall:

$$\frac{d}{dp}\psi(p, 1)\frac{\partial}{\partial p}\psi(p, 1) = -\big(1 + p^{1-\sigma}\big)^{\frac{\sigma}{1-\sigma}} p^{-\sigma} < 0$$

That is, a decline in p from an opening to trade leads to an increase in wellbeing.

In contrast, for a person in the import competing sector, with an endowment of one unit of good F and no units of good D, welfare increases in p:

$$\frac{d}{dp}\psi(p, p) = \frac{\partial}{\partial p}\psi(p, p) + \frac{\partial}{\partial I}\psi(p, p) = \big(1 + p^{1-\sigma}\big)^{\frac{2-\sigma}{\sigma-1}} > 0$$

hence the drop in p induced by free trade reduces welfare. So absent transfers among the citizens, an opening to trade benefits the people endowed entirely with good D and hurts those whose endowment consists solely of good F.

Distribution under free trade. We now show that it is possible to spread the gains from trade in such a way that everyone benefits. This result is well known in the literature on international trade (e.g., Helpman and Krugman, 1985), we reproduce it here to make the exposition more self-contained. First, notice that after opening to trade the country can still afford its endowment, which is also the autarkic consumption bundle, so that one could design a transfer scheme that simply forced people to consume as they had without trade and leave them no worse off than they had been before. We now show that there is a feasible post trade allocation that actually leaves everyone strictly better off.

We start by introducing the "expenditure function," which gives the minimum cost of achieving a utility level U when relative prices are p:

$$e(p,U) = \left(1 + p^{1-\sigma}\right)^{\frac{1}{1-\sigma}} U \tag{18}$$

Notice that:

$$\frac{\partial}{\partial p} e(p,U) = q_F^H(p,U) = \left(1 + p^{1-\sigma}\right)^{\frac{\sigma}{1-\sigma}} U \cdot p^{-\sigma} > 0.$$

This is the Hicksian demand for good F–the quantity consumed as part of the cheapest bundle of goods that attains a utility level of U.

Not surprisingly, the Hicksian demands slope downwards:

$$\frac{\partial^2}{\partial p^2} e(p,U) = \frac{\partial}{\partial p} q_F^H(p,U) = -\sigma\left(1 + p^{1-\sigma}\right)^{\frac{2\sigma-1}{1-\sigma}} U \cdot p^{-(1+\sigma)} < 0. \tag{19}$$

What's more, we have the following useful result:

LEMMA 1 *$e(p,\psi(q,I))$ is linear in I for any prices p and q*

Proof of Lemma 1: From expression (18) we have: $e(p,\psi(q,I)) = \left(1 + p^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \psi(q,I)$. Substituting from (17) this becomes:

$$e(p,\psi(q,I)) = \left(1 + p^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \left(1 + q^{1-\sigma}\right)^{\frac{1}{\sigma-1}} I = \left(\frac{1 + p^{1-\sigma}}{1 + q^{1-\sigma}}\right)^{\frac{1}{1-\sigma}} I \ \square$$

With free the available resources are given by $I(p^{\mathsf{trade}})$, while the cost of providing a utility level of $\psi(p^{\mathsf{autarky}}, p^{\mathsf{autarky}})$ for the n_F individuals endowed with good F and a utility level of $\psi(p^{\mathsf{autarky}}, 1)$ for the n_D individuals with good D is, by Lemma 1:

$$n_F e\Big(p^{\mathsf{trade}}, \psi(p^{\mathsf{autarky}}, p^{\mathsf{autarky}})\Big) + n_D e\Big(p^{\mathsf{trade}}, \psi(p^{\mathsf{autarky}}, 1)\Big) = e\Big(p^{\mathsf{trade}}, \psi(p^{\mathsf{autarky}}, n_D + n_F p^{\mathsf{autarky}}\Big)$$

We can now state and prove the main result:

LEMMA 2 *There always sufficient resources available to leave everyone better off with trade than under autarky. Formally,* $I(p^{\mathsf{trade}}) > e\Big(p^{\mathsf{trade}}, \psi(p^{\mathsf{autarky}}, n_D + n_F p^{\mathsf{autarky}}\Big)$

Proof of Lemma 2: At $p^{\mathsf{trade}} = p^{\mathsf{autarky}}$ we have:

$$I(p^{\mathsf{autarky}}) = e\Big(p^{\mathsf{autarky}}, \psi\big(p^{\mathsf{autarky}}, I(p^{\mathsf{autarky}})\big)\Big) \tag{20}$$

and

$$q_F^H\Big(p^{\mathsf{autarky}}, \psi\big(p^{\mathsf{autarky}}, I(p^{\mathsf{autarky}})\big)\Big) = n_F \tag{21}$$

Differentiating $I(p^{\mathsf{trade}})$ we have:

$$\frac{\partial I}{\partial p^{\mathsf{trade}}}(p^{\mathsf{trade}}) = n_F \tag{22}$$

and:

$$\frac{\partial^2 I}{\partial p^{\mathsf{trade}2}}(p^{\mathsf{trade}}) = 0 \tag{23}$$

while:

$$\frac{\partial e}{\partial p^{\mathsf{trade}}}\Big(p^{\mathsf{trade}}, \psi\big(p^{\mathsf{autarky}}, I(p^{\mathsf{autarky}})\big)\Big) = q_F^H\Big(p^{\mathsf{trade}}, \psi\big(p^{\mathsf{autarky}}, I(p^{\mathsf{autarky}})\big)\Big) \tag{24}$$

and, see expression (19):

$$\frac{\partial^2 e}{\partial p^{\mathsf{trade}2}}\Big(p^{\mathsf{trade}}, \psi\big(p^{\mathsf{autarky}}, I(p^{\mathsf{autarky}})\big)\Big) = \frac{\partial q_F^H}{\partial p^{\mathsf{trade}}}\Big(p^{\mathsf{trade}}, \psi\big(p^{\mathsf{autarky}}, I(p^{\mathsf{autarky}})\big)\Big) < 0 \tag{25}$$

Combining expressions (21) and (24) and comparing with expression (22) we see that there is a tangency between $I(p^{\mathsf{trade}})$ and $e\Big(p^{\mathsf{trade}}, \psi\big(p^{\mathsf{autarky}}, I(p^{\mathsf{autarky}})\big)\Big)$ at $p^{\mathsf{trade}} = p^{\mathsf{autarky}}$, while expressions

(22) and (25) guarantee that this tangency is unique, with $I(p^{\mathsf{trade}})$ above the expenditure function for any $p \neq p^{\mathsf{autarky}}$. In particular, for $p^{\mathsf{trade}} < p^{\mathsf{autarky}}$:

$$I(p^{\mathsf{trade}}) > e\Big(p^{\mathsf{trade}}, \psi\big(p^{\mathsf{autarky}}, I(p^{\mathsf{autarky}})\big)\Big) \quad \square$$

A1.2 Rebellion

Our attention now turns to the question of whether the dictator will opt to open to trade, and this in turn depends on the leader's efforts to survival in the face of potential rebellion. We model this process as a game. Figure 3 in the main text illustrates how that game unfolds after the trade decision has been made. We'll work through the solution by backward induction, starting with the decision whether to rebel by the citizen drawn at random by nature, call this citizen j. This corresponds to the decision node marked j. If the citizen acquiesces her payoff will be $U_{\mathsf{j}}^{acquiesce} = \psi(p, A_{\mathsf{j}})$, whereas if she rebels her payoff $\tilde{A}_{\mathsf{j}}$ is uncertain.

$$\mathbb{E}\big\{U_{\mathsf{j}}^{rebel}\big\} = \mathbb{E}\big\{U(\tilde{A}_{\mathsf{j}})\big\}$$

Recall from expression (17) that ψ is a linear function of income, so we may rewrite the expected utility of rebellion as:

$$\mathbb{E}\big\{U(\tilde{A}_{\mathsf{j}})\big\} = U\big(\mathbb{E}\{\tilde{A}_{\mathsf{j}}\}\big)$$

Now let's calculate the expected income from rebellion. If the rebellion prospers, individual j gets an equal share of the income that survives the fray, $\omega \bar{I}(p)$, if the rebellion goes badly, she gets an income of 0:

$$\mathbb{E}\{\tilde{A}_{\mathsf{j}}\} = \frac{\rho\theta}{1+\rho\theta} \times \omega \bar{I}(p) + \frac{1}{1+\rho\theta} \times 0 = \frac{\rho\theta}{1+\rho\theta}\omega \bar{I}(p)$$

Define A_{j}^* as the allocation that would set $U_{\mathsf{j}}^{acquiesce}$ equal to $\mathbb{E}\big\{U_{\mathsf{j}}^{rebel}\big\}$:

$$A_{\mathsf{j}}^* = \tfrac{\rho\theta}{1+\rho\theta}\omega \bar{I}(p)$$

Any value of $A_{\mathsf{j}} > A_{\mathsf{j}}^*$ will make acquiescence even more attractive to j while any value less than that would lead that same individual to rebel. We suppose that in case of indifference the potential rebel leader chooses to acquiesce. We note that each citizen will have the same threshold for indifference:

$$A_i = A_{\mathrm{j}}^* = A^* \equiv \frac{\rho\theta}{1+\rho\theta}\omega\bar{I}(p) \; \forall \; i \tag{26}$$

Working our way back up the top of the game tree, the dictator will anticipate the value of A^* when deciding what to offer each individual. Offering $A_i > A^*$ is dominated by $A_i = A^*$ which still prevents i from rebelling if she is chosen by nature. Likewise, any offer on the inerval $(0, A^*)$ still results in the individual rebelling if she is chosen, and so it is dominated by $A_i = 0$. Thus, the dictator will make an offer to individual i of $A_i \in \{0, A^*\}$.

It remains to show to what proportion q of people the dictator offers 0, and what fraction get an offer of A^*. Suppose the dictator offers 0 to a fraction q of the population, and A^* to the rest. If nature chooses one of the people the dictator favored with a transfer (which happens with a probability of $1-q$), his income will be $I(p) - (n_D + n_F)(1-q)A^*$, while if nature chooses one of the individuals given 0 by the dictator (an occurance that arrises with probability q), the dictator's expected payoff is $\frac{\omega I(p)}{1+\rho\theta}$. Putting all of this together we have:

$$\begin{aligned}
\mathbb{E}(U_d(q)) &= (1-q)\psi\Bigg(p, \Big(I(p) - (n_D+n_F)(1-q)A^*\Big)\Bigg) + q\psi\Bigg(p, \left(\frac{\omega I(p)}{1+\rho\theta}\right)\Bigg) \\
&= (1-q)\psi(p,1)\Big(I(p) - (n_D+n_F)(1-q)A^*\Big) + q\psi(p,1)\left(\frac{\omega I(p)}{1+\rho\theta}\right) \\
&\propto (1-q)\Big(I(p) - (n_D+n_F)(1-q)A^*\Big) + q\left(\frac{\omega I(p)}{1+\rho\theta}\right)
\end{aligned}$$

substituting for A^* from (26) this becomes:

$$\begin{aligned}
\mathbb{E}(U_d(q)) &\propto (1-q)\Big(I(p) - (n_D+n_F)(1-q)\frac{\rho\theta}{1+\rho\theta}\omega\bar{I}(p)\Big) + q\left(\frac{I(p)}{1+\rho\theta}\right)\omega \\
&= \bar{I}(p)\Bigg(\Big(n_D+n_F\Big)(1-q)\Big(1 - \frac{\rho\theta}{1+\rho\theta}\omega(1-q)\Big) + \frac{(n_D+n_F)q\omega}{1+\rho\theta}\Bigg) \\
&= \frac{I(p)}{1+\rho\theta}\Bigg((1-q)\Big(1+\rho\theta - \rho\theta\omega(1-q)\Big) + \omega q\Bigg) \\
&= \frac{I(p)}{1+\rho\theta}\Bigg(1+\rho\theta(1-\omega) - \big(1-\omega+\rho\theta(1-2\omega)\big)q - \rho\theta\omega q^2\Bigg)
\end{aligned}$$

Differentiating we find that whenever conflict is sufficiently destructive, with:

$$\omega < \frac{1+\rho\theta}{1+2\rho\theta} \tag{27}$$

we are guaranteed to have $V_d^{*\prime\prime}(q) < 0$ for all $q \in [0, 1]$, so the autocrat will find himself at a corner solution, with $q = 0$; the government will appease all of the potential rebels.

In Section 2.3 we confine ω to the interval $[0, \frac{1}{2}]$, which means that conflict is very destructive. Since $\frac{1}{2} < \frac{1+\rho\theta}{1+2\rho\theta}$, the restriction that $\omega \in [0, \frac{1}{2}]$ guarantees that condition (27) holds.

The autocrat's equilibrium income will be:

$$I(p)\frac{1+(1-\omega)\rho\theta}{1+\rho\theta}$$

While this expression is increasing in total income, $I(p) = n_D + n_F p$, it is also decreasing in ρ and θ. The welfare of the autocrat becomes:

$$U_d = \psi(p, 1)\Big((n_D + n_F p)\frac{1+(1-\omega)\rho\theta}{1+\rho\theta}\Big) \tag{28}$$

Notice that this is a decreasing function of both ρ and θ.

We have just proved:

LEMMA 3 *Given the assumptions of Section 2.3 the dictator will prefer to appease everyone rather than risk civil war, with each person receiving an appeasement level of A^* as given in expression* (26). *Any citizen who is chosen by nature as the leader of a potential rebellion will opt for insurrection if the transfer offered by the government falls below A^* from expression* (26), *while otherwise he acquiesces.*

Now we come to the question of whether the autocrat will opt for free trade. If he does the equilibrium outlined in Lemma 3 will unfold, with the leader's welfare given by expression (28), with a price for good F of $p = p^{\text{trade}}$ and spillovers of $\theta = \theta^{\text{trade}}$. If he instead chooses not to trade, the sequence of events again conforms to Lemma 3, but with $(p, \theta) = (p^{\text{autarky}}, \theta^{\text{autarky}})$. The leader will choose whichever option leads to the higher value of his welfare as given in expression (28):

Equilibrium: The autocrat will choose to open to trade provided $\theta^{\text{trade}} < \theta^*(p^{\text{trade}})$ where:

$$\psi(p^{\text{trade}}, 1)\big(n_D + n_F p^{\text{trade}}\big)\frac{1+\rho\theta^*(p^{\text{trade}})(1-\omega)}{1+\rho\theta^*(p^{\text{trade}})} = \psi(p^{\text{autarky}}, 1)\big(n_D + n_F p^{\text{autarky}}\big)\frac{1+\rho\theta^{\text{autarky}}(1-\omega)}{1+\rho\theta^{\text{autarky}}} \tag{29}$$

In the lefthand panel of Figure 4, which appears in Section 2.3 of the text, the bold curved line depicts $\theta^*(p^{\text{trade}})$. When the free trade price for good F is p^{trade} the blue dashed line corresponds to the set of $(p^{\text{trade}}, \theta^{\text{trade}})$ pairs for which the autocrat is willing to trade rather than content himself with the insular outcome at $(p^{\text{autarky}}, \theta^{\text{autarky}})$. If trade causes the level of network connectivity to rise above $\theta^*(p^{\text{trade}})$, corresponding with the $(p^{\text{trade}}, \theta^{\text{trade}})$ pairs depicted in the red line, the autocrat will forgo trade, preferring to remain poorer but more secure at $(p^{\text{autarky}}, \theta^{\text{autarky}})$.

A1.3 Proof of Proposition 1

PROPOSITION 1 (REGIME SECURITY AND THE EXTENSIVE MARGIN OF TRADE) *The maximum degree of network spillovers an insecure authoritarian regime will accept in order to trade, $\theta^*(p^{\textit{trade}})$, is a decreasing function of $p^{\textit{trade}}$.*

Proof: Let's define the function W as:

$$W(p, \rho, \theta) \equiv \psi(p, 1)I(p)\left(\frac{1 + \rho\theta(1 - \omega)}{1 + \rho\theta}\right)$$

Notice that this coincides with the righthand side of equation (29) when we evaluate W at $(p, \rho, \theta) = (p^{\text{autarky}}, \rho, \theta^{\text{autarky}})$.

$$\frac{\partial W}{\partial \theta} = \frac{-\rho\omega}{(1 + \rho\theta)^2}\psi(p, 1)I(p) < 0$$

while:

$$\frac{\partial W}{\partial p} = \left(\frac{1 + \rho\theta(1 - \omega)}{1 + \rho\theta}\right)\Big(\psi_p(p, 1)I(p) + n_F\psi(p, 1)\Big)$$

However, this expression is negative:

$$\begin{aligned}
\psi_p(p, 1)I(p) + n_F\psi(p, 1) &= \psi_p\big(p, I(p)\big) + \psi(p, n_F) \\
&= -p^{-\sigma}I(1 + p^{1-\sigma})^{\frac{2-\sigma}{\sigma-1}} + (1 + p^{1-\sigma})^{\frac{1}{\sigma-1}}n_F \\
&= (1 + p^{1-\sigma})^{\frac{2-\sigma}{\sigma-1}}\big(-p^{-\sigma}(n_D + pn_F) + n_F(1 + p^{1-\sigma})\big) \\
&= (1 + p^{1-\sigma})^{\frac{2-\sigma}{\sigma-1}}\big(-p^{-\sigma}n_D + n_F\big) \\
&= (1 + p^{1-\sigma})^{\frac{2-\sigma}{\sigma-1}}n_D\big(\frac{n_F}{n_D} - p^{-\sigma}\big) \\
&= (1 + p^{1-\sigma})^{\frac{2-\sigma}{\sigma-1}}n_D\big((p^{autarky})^{-\sigma} - p^{-\sigma}\big) \\
&< 0,
\end{aligned}$$

where the last inequality follows from equation (5) in the main text. Therefore,

$$\frac{\partial W}{\partial p} < 0.$$

So when $\theta = \theta^{\text{autarky}}$, it follows that welfare is decreasing in p, and so the autocrat will opt for trade whenever $p^{\text{trade}} < p^{\text{autarky}}$.

How does θ^* change as we change p? When $p < p^{\mathsf{autarky}}$, we can invoke the implicit function theorem to conclude:

$$\theta^{*\prime}(p) = -\frac{\frac{\partial W}{\partial p}}{\frac{\partial W}{\partial \theta}} < 0$$

The lower the world price of good F compared with the autarkic price, the greater the degree of network spillover the autocrat is prepared to tolerate and still embrace international trade. □

Notice that:

$$\theta^{*\prime}(p) \propto \frac{(1+\rho\theta)(1+\rho\theta(1-\omega))}{\rho\omega}$$

When an authoritarian regime has a vice grip on power, so that ρ is near 0, the slope of $\theta^*(p)$ becomes arbitrarily steep:

$$\theta^{*\prime}(p) \to \infty \text{ as } \rho \to 0$$

So only extreme network externalities will restrain the highly secure regime, with a very small value for ρ, from trading. The right panel of Figure 4 from Section 2.3 of the paper illustrates the consequences of a steepened slope for $\theta^*(p)$. It shows that the set of (θ, p) pairs for which the autocrat refrains from trade shrink to encompass only the gray region.

Contrasting the regime depicted on the left side of the graph, with a shallow slope for $\theta^*(p)$, with the regime on the right with a steeply pitched $\theta^*(p)$ we see that if trade reduced the price to p^{trade} then the regime on the left would only open to trade for values of θ^{trade} below $\theta^*(p)$, depicted in blue, while the regime on the right will be receptive to the same opportunity to trade for all values of θ. Thus it will be the insecure authoritarian regimes, with relatively high values of ρ, that will discriminate across products, eschewing trade in goods fraught with high network externalities while still embracing commerce in commodities with low externalities.

A2 Bayesian LASSO Selection Model

We give an overview of our Bayesian LASSO regression and then our exact selection model.

A2.1 Bayesian LASSO Regression

Before describing the two-stage model, we give the full hierarchical representation of our Bayesian regression model. We give the model for a normal regression, noting that extension to the probit is straightforward. This is the model that we adapt to the selection model immediately following.

For generic outcome Y_{ijtk} and covariate p-vector $\mathbf{x}_{ijtk}$, associated mean parameters β_k, arbitrary element β_{qk}, and subscripts following what is in the text, we model the outcome as

$$Y_{ijtk}|\beta_k, f_{ik}, g_{jk}, h_{tk} \sim \mathcal{N}\left(\mathbf{x}_{ijtk}^{\top}\gamma_k + f_{ik} + g_{jk} + h_{tk}, \sigma_k^2\right) \tag{30}$$

$$\beta_{qk}|\lambda_k, \sigma_k \overset{\text{i.i.d.}}{\sim} Dexp\left(\frac{\lambda_k}{\sigma_k}\right) \tag{31}$$

$$\lambda_k^2 \overset{\text{i.i.d.}}{\sim} \Gamma\left(n\log(p) + p, 1\right) \tag{32}$$

$$\sigma_k^2 \propto 1/\sigma_k^2 \tag{33}$$

$$f_{ik} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{fk}^2);\ g_{jk} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{gk}^2);\ h_{tk} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{hk}^2) \tag{34}$$

$$\sigma_{fk}^2 \propto 1/\sigma_{fk}^2;\ \ \sigma_{gk}^2 \propto 1/\sigma_{gk}^2;\ \ \sigma_{hk}^2 \propto 1/\sigma_{hk}^2 \tag{35}$$

which is the Bayesian representation of a LASSO model (Tibshirani, 1996) with p the number of covariates.

The normal likelihood is not conjugate with the double exponential prior. To restore conjugacy, we augment the data again representing the double exponential as a scale mixture of normals with an exponential mixing density (Park and Casella, 2008). We parameterize the conditional prior on λ_k^2 in terms of the sample size n in order to guarantee that the posterior mode takes on a value zero with positive probability in the limit. This prior is also constructed such that it is proper and the posterior mean of λ_k grows as $\sqrt{n \log p}$, the minmax optimal rate. The parameters b_{ik}, c_{jk}, and d_{tk} are random effects, each with a Jeffreys' prior over the scale parameter. See Ratkovic and Tingley (2017) for details.

A2.2 Our Selection Model Implementation

We give precise details next, but our estimator uses a latent normal representation of the probit regression in the first stage. As described in text, we construct a draw of the inverse Mills ratio from a draw of the probit model. This covariate is then entered as a second stage covariate, but changes at each scan. In this way, we give a more honest accounting of the impact of the inverse mills ratio, incorporating first stage (extensive margin) uncertainty in our second stage (intensive margin) estimates.

For our selection model, we consider two outcomes, one for each margin. On the extensive margin, the binary variable δ_{ijtk} takes on a value of 1 if country i imports good k from country j at time t, else it equals 0. For the intensive margin $\widetilde{Y}_{ijtk}$ is the flow of imports of k to i from j at time

t. We will work with the more tractable log-transformed outcome $Y_{ijtk} = \log(1 + \widetilde{Y}_{ijtk})$.[1].

We let $\mathbf{x}_{ijtk}$ denote the observed covariates, which we decompose into a vector of gravity variables $\mathbf{x}_{ijtk,G}$, augmented with an intercept, and another consisting of POLITY variables $\mathbf{x}_{ijtk,P}$: $\mathbf{x}_{ijtk} = [\mathbf{x}_{ijtk,G}^\top : \mathbf{x}_{ijtk,P}^\top]^\top$.

Estimating the Extensive Margin We use a probit specification at the extensive margin for each industry k, comparing two versions: $\mathcal{E}_k^P$, which includes gravity and political variables, and $\mathcal{E}_k^G$, which is indentical with the first, save that it excludes the political variables.

In specification $\mathcal{E}_k^P$ a dyad trades, so that $\delta_{ijtk} = 1$, whenever the latent propensity to trade, T_{ijtk}, is sufficiently high:

$$\delta_{ijtk} = \begin{cases} 1; & T_{ijtk} > 0 \\ 0; & T_{ijtk} \leq 0 \end{cases} \tag{36}$$

Adding structure to the propensity, we get have the following structure:

$$\mathcal{E}_k^P : T_{ijtk} = \mathbf{x}_{ijtk,G}^\top \boldsymbol{\beta}_{k,G} + \mathbf{x}_{ijtk,P}^\top \boldsymbol{\beta}_{k,P} + b_{ik} + c_{jk} + d_{tk} - \tilde{u}_{ijtk} \ , \ \tilde{u}_{ijtk} \sim N(0,1) \tag{37}$$

where $\{b_{ik}, c_{jk}, d_{tk}\}$ are importer, exporter, and time random effects, respectively,[2] while $\tilde{u}_{ijtk}$ is a standard normal random variable. The structure for $\mathcal{E}_k^G$ is identical save that $\boldsymbol{\beta}_{k,P} = \mathbf{0}$.

This expression connects with the observed data *via* the probit probability:

$$\mathcal{E}_k^P : \Pr(\delta_{ijtk} = 1|\cdot) = \Phi(\mathbf{x}_{ijtk,G}^\top \boldsymbol{\beta}_{k,G} + \mathbf{x}_{ijtk,P}^\top \boldsymbol{\beta}_{k,P} + b_{ik} + c_{jk} + d_{tk}) \tag{38}$$

where $\Phi(z)$ denotes the distribution function of a standard normal random variable.

Estimating the Intensive Margin Here we also consider two specifications for each product k: $\mathcal{I}_k^{*P}$, which contains both the gravity and political variables, and $\mathcal{I}_k^{*G}$, with the same structure, save that it excludes the polity variables.

[1] Liu (2009) advocate Poisson regression over the Tobit, however the trade flow outcome is virtually continuous and our theory predicts heterogeneous zero-inflation, both of which would confound a Poisson specification. Another alternative, the Poisson pseudo-maximum likelihood estimator (Silva and Tenreyro, 2006), does not differentiate between the two margins.

[2] The random effects account "multilateral resistance" (Anderson and Wincoop, 2003) induced by the panel structure of the data.

$$\mathcal{I}_k^{*P}: Y_{ijtk} = \mathbf{x}_{ijtk,G}^{\top}\gamma_{k,G} + \mathbf{x}_{ijtk,P}^{\top}\gamma_{k,P} + f_{ik} + g_{jk} + h_{tk} + \epsilon_{ijtk} \ , \ \epsilon_{ijtk} \sim N(0, \sigma_k^2) \tag{39}$$

$$\mathcal{I}_k^{*G}: Y_{ijtk} = \mathbf{x}_{ijtk,G}^{\top}\gamma_{k,G} + f_{ik} + g_{jk} + h_{tk} + \epsilon_{ijtk} \ , \ \epsilon_{ijtk} \sim N(0, \sigma_k^2) \tag{40}$$

where $\{f_{ik}, g_{jk}, h_{tk}\}$ are importer, year, and time random effects, respectively.

We only estimate this specification for dyads for which $T_{ijtk} > 0$, so our estimator for $\mathcal{I}_k^{*P}$ must be adapted to contend with selection bias. Specifically, we must contend with possible correlation between the intensive margin residuals, ϵ_{ijtk} and the residuals on the extensive margin, $\tilde{u}_{ijtk}$.[3] Whereas the unconditional expected value for the extensive margin shock equals zero, if we condition on trade actually taking place the expected value for u_{ijtk} will be negative[4]:

$$E\{\tilde{u}_{ijtk}|\delta_{ijtk} = 1, \mathcal{E}_k^P\} = -\frac{\phi(\mathbf{x}_{ijtk,G}^{\top}\boldsymbol{\beta}_{k,G} + \mathbf{x}_{ijtk,P}^{\top}\boldsymbol{\beta}_{k,P} + f_{ik} + g_{jk} + h_{tk})}{\Phi(\mathbf{x}_{ijtk,G}^{\top}\boldsymbol{\beta}_{k,G} + \mathbf{x}_{ijtk,P}^{\top}\boldsymbol{\beta}_{k,P} + f_{ik} + g_{jk} + h_{tk})} = m_{ijtk}^P \tag{41}$$

If we add m_{ijtk}^P as an additional covariate[5], as we do in the following intensive margin specification, our estmates of the remaining intensive margin coefficients will be unbiased (Heckman, 1979; Olsen, 1980):

$$\mathcal{I}_k^{P}: Y_{ijtk} = \mathbf{x}_{ijtk,G}^{\top}\gamma_{k,G} + \mathbf{x}_{ijtk,P}^{\top}\gamma_{k,P} + f_{ik} + g_{jk} + h_{tk} + \theta_k m_{ijtk}^P + \epsilon_{ijtk}; \ \ \epsilon_{ijtk} \sim \mathcal{N}(0, \sigma_k^2) \tag{42}$$

$$\mathcal{I}_k^{G}: Y_{ijtk} = \mathbf{x}_{ijtk,G}^{\top}\gamma_{k,G} + f_{ik} + g_{jk} + h_{tk} + \theta_k m_{ijtk}^P + \epsilon_{ijtk}; \ \ \epsilon_{ijtk} \sim \mathcal{N}(0, \sigma_k^2) \tag{43}$$

Again, recall we only estimate $\mathcal{I}_k^P$, or $\mathcal{I}_k^G$, for

A3 First Order Stochastic Dominance

Similarly to Figure 13, Figure A1 compares the distribution of Bayes factors associated with the inclusion of our *Polity* variables, and it suggests the evidence on the extensive margin tends to be stronger for differentiated products, while there does not seem to be much effect on the intensive margin. However, Figure A1 uses box and whisker plots, and provides less information than does Figure 13 which applies the powerful notion of **first order stochastic dominance**.

[3]We expect that random shocks that dispose countries to trade at all are associated with higher trade volumes, as we shall see turns out to be the case for for all but a handful of our dyads.

[4]For the sake of illustration We calculate this expectation using the full specification $\mathcal{E}_k^P$.

[5]We use the full specification $\mathcal{E}_k^P$ rather than restricted version $\mathcal{E}_k^G$ in order to avoid omitted variable bias induced by misspecification at the extensive margin. Our theory supports this decision as do the data, see the results presented below.

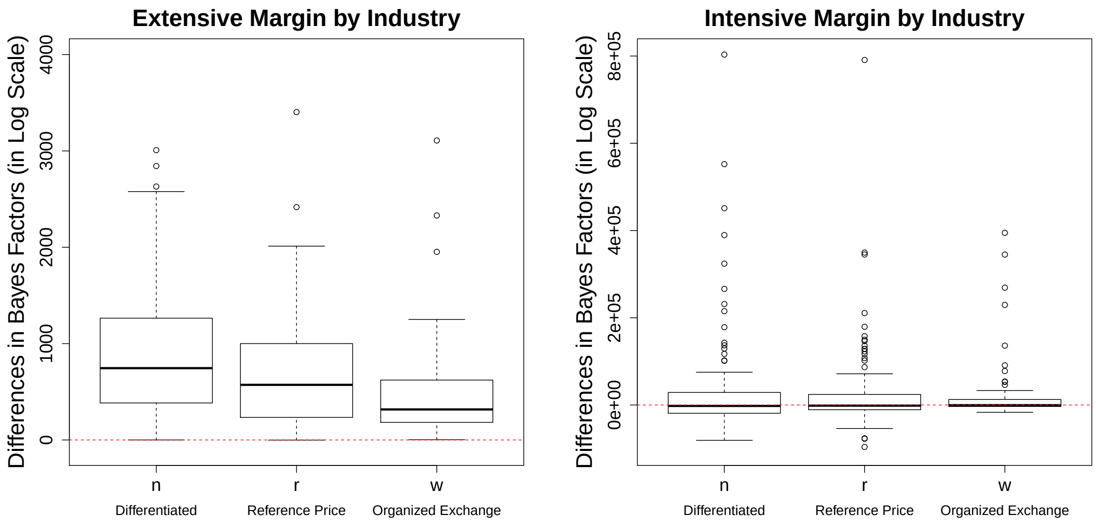

Figure A1: **Distribution of Bayes Factors by Rauch Categories**

We now go into greater detail about this important concept. that we applied in Section 4.3. Let $F^c_{org}(B)$ denote the fraction of products traded on organized exchanges for which $\mathsf{BF}^{\mathcal{E}}_{k\in org} > B$. This is called the *complementary cummulative density function*, and it's often referred to as the cpomplementary cdf. We define $F^c_{ref}(B)$ and $F^c_{diff}(B)$ analogously for products with reference price and differentiated goods, respectively.

We expect that for *any* value of B, we will find that the fraction of differentiated products with Bayes factors in excess of B is at least as great as the fraction of referenced priced products with Bayes factors above B, and that there will be some values of B for which the inequality is strict:

$$F^c_{diff}(B) \geq F^c_{ref}(B) \ \forall \ B \quad \text{and} \quad F^c_{diff}(B') > F^c_{ref}(B') \text{ for some } B' \tag{44}$$

When the cumulative distribution functions F^c_{diff} and F^c_{ref} satisfy condition (44), which was set forth by Hanoch and Levy (p.37 1969), we say that F^c_{diff} *first order stochastically dominates* F^c_{ref}. Making reference to the underlying populations $\{\mathsf{BF}^{\mathcal{E}}_k\}_{k\in diff}$ for F^c_{diff} and $\{\mathsf{BF}^{\mathcal{E}}_k\}_{k\in ref}$ for F^c_{ref} we can express this more succinctly as:

$$\{\mathsf{BF}^{\mathcal{E}}_k\}_{k\in diff} \succ \{\mathsf{BF}^{\mathcal{E}}_k\}_{k\in ref} \tag{45}$$

While the notion of stochastic dominance is widely applied in finance, to guage the attractiveness of different assets, we use it here to allow us to compare the distribution of Bayes factors across the population of products in each Rauch category. Condition (45) tells us that the evidence for including our regime type variables is systematically stronger for the population of differentiated products than it is for the reference priced goods.

A4 Robustness Checks Controlling for Oil Wealth

In this section, we provides a series of robustness check results accounting for the uniqueness of oil reach countries. Our analysis shows that natural resource wealth does not affect the main findings.

Controlling for Oil Wealth. We include `% of oil rent per GDP` as additional covariates of the analysis in order to directly account for oil wealth (Lange, Wodon, and Carey, 2018). Specifically, we include the variable for both exporter and importer. Instead of list-wise deleting observations with missing values, we created indicators to account for the missingness.

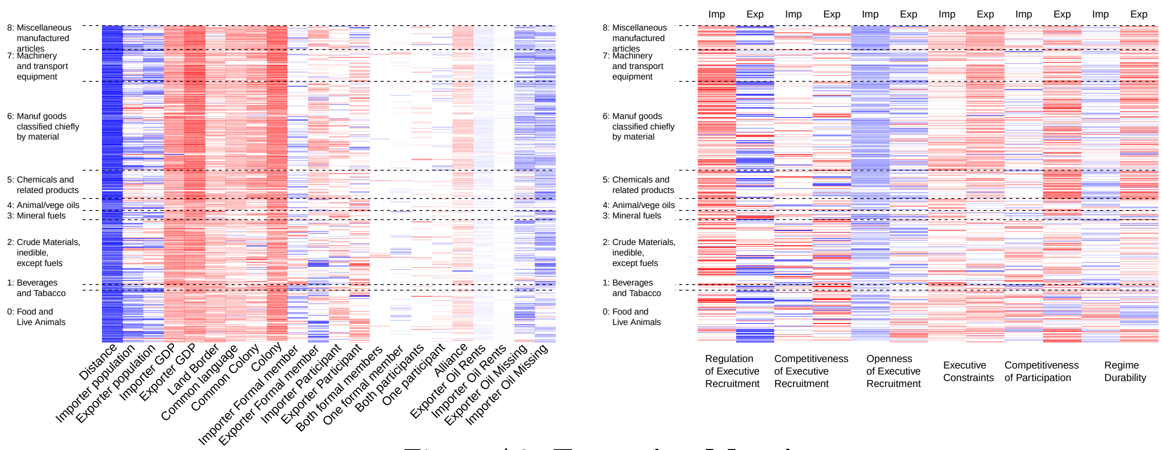

Figure A2: **Extensive Margin**

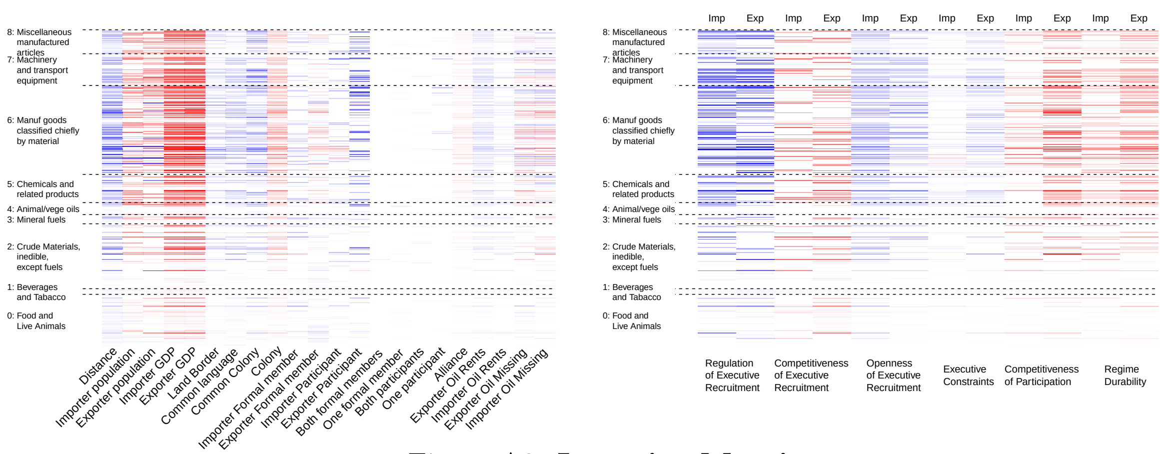

Figure A3: **Intensive Margin**

Note that the missingness indicators are found to be negative (the two blue columns in the extensive margin). This shows that missing oil wealth variable is not random, justifying our choice of not list-wise deleting the variable. Comparing these results with Figures 7–10, we find that including these variables do not change the results.

Removing the OPEC Members. We further investigate the robustness by removing the members of the OPEC (Organization of the Petroleum Exporting Countries) entirely from the analysis. In particular, we construct a time-varying OPEC membership dataset, and them remove all dyad-year observations if either the importer or the exporter or both are members of the organization. We then replicate Figure 11 on the non-OPEC observations as Figure A4:

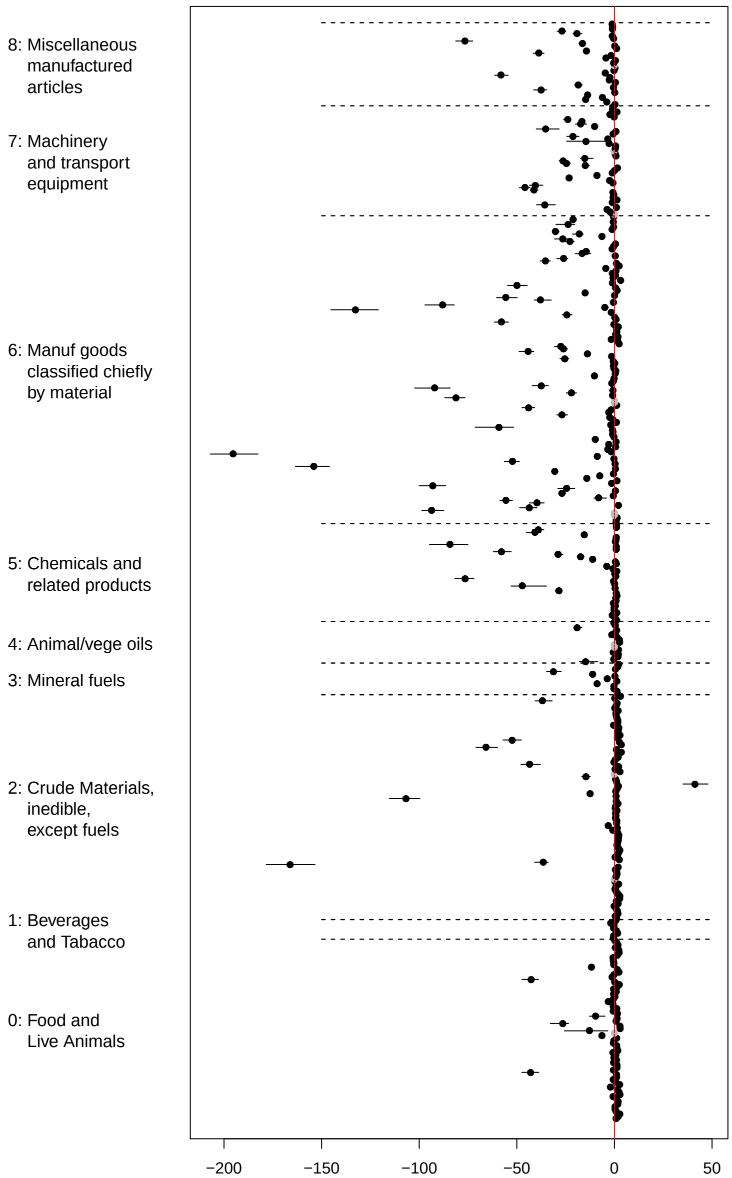

Figure A4: **The Impact of Inverse Mills Ratio:** This figure shows that there exists selection in the extensive margin even after removing OPEC members.

A5 Connecting **POLITY** with the *Polity IV* Component Variables

The *Polity IV* codebook uses terse acronyms for the *Polity IV* subscales, whereas we use more descriptive variable names. In the interest of transparency, here is the concordance:

Our label	*Polity IV* aronym
Competitiveness of Executive Recruitment	XRCOMP
Openness of Executive Recruitment	XROPEN
Constraint on Chief Executive	XCONST
Competitiveness of Political Participation	PARCOMP
Regulation of Executive Recruitment	XRREG
Regime Durability	DURABLE

The *Polity IV* variable that we doen't use is PARREG, which captures the extent to which political participation is "regulated." This is because the variable is on a monotonic scale that hinders interpretation. Other component variables take on values according to the following scheme:

Components	Value	Democracy	Autocracy	POLITY
Competitiveness of Executive Recruitment	Selection (1)	0	+2	-2
	Dual/Transitional (2)	+1	0	+1
	Election (3)	+2	0	+2
Openness of Executive Recruitment	Closed (1)	0	+1	-1
	Dual Executive-Designation (2)	0	+1	-1
	Dual Executive-Election (3)	+1	0	+1
	Open (4)	+1	0	+1
Constraint on Chief Executive	Unlimited Authority (1)	0	+3	-3
	Intermediate Category (2)	0	+2	-2
	Slight to Moderate Limitation (3)	0	+1	-1
	Intermediate Category (4)	+1	0	+1
	Substantial Limitations (5)	+2	0	+2
	Intermediate Category (6)	+3	0	+3
	Executive Parity or Subordination (7)	+4	0	+4
Competitiveness of Political Participation	Not Applicable (0)	0	0	0
	Repressed (1)	0	+2	-2
	Suppressed (2)	0	+1	-1
	Factional (3)	+1	0	+1
	Transitional (4)	+2	0	+2
	Competitive (5)	+3	0	+3
regulation of political participation	Unregulated (1)	0	0	0
	Multiple Identity (2)	0	0	0
	Sectarian (3)	0	+1	-1
	Restricted (4)	0	+2	-2
	Regulated (5)	0	0	0

Table A1: **Weights on *Polity IV* Component Variables**: This table presents the original weights used to construct *Polity IV* score. Each component variable contributes either to Democracy or Autocracy, which will then be used to produce a 21 point cardinal scale.

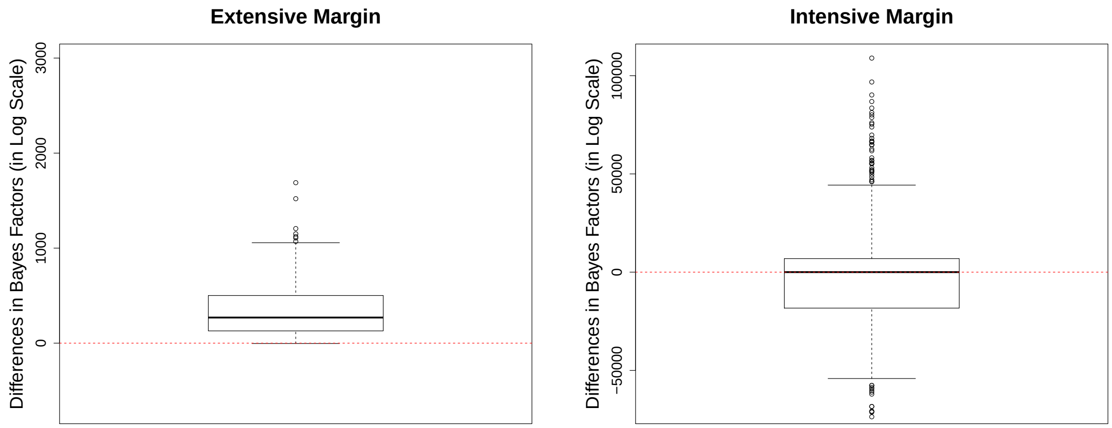

Figure A5: **Polity Component Variables vs. *Polity IV***: This figure compares the Bayes factors between the model with politcy component variables and the model with only the *Polity IV* variable. It shows that the former fits the extensive margin much better across all products than the latter. This provides additional support for our use of the polity component variables in the main analysis.

A6 List of Countries

Afghanistan	China, Hong Kong SAR	Guatemala	Mauritius	Senegal
Albania	China, Macao SAR	Guinea	Mexico	Sierra Leone
Algeria	Colombia	Guinea-Bissau	Mongolia	Singapore
Angola	Comoros	Guyana	Morocco	Somalia
Antigua and Barbuda	Congo	Haiti	Mozambique	Spain
Argentina	Costa Rica	Honduras	Myanmar	Sri Lanka
Australia	Cote d'Ivoire	Hungary	Nepal	Suriname
Austria	Cuba	Iceland	Netherlands	Sweden
Bahamas	Cyprus	Indonesia	New Caledonia	Switzerland
Bahrain	Dem. People's Rep. of Korea	Iran	New Zealand	Syria
Barbados	Dem. Rep. of the Congo	Iraq	Nicaragua	Thailand
Belize	Denmark	Ireland	Niger	Timor-Leste
Benin	Djibouti	Israel	Nigeria	Togo
Bermuda	Dominica	Italy	Norway	Trinidad and Tobago
Bolivia	Dominican Rep.	Jamaica	Oman	Tunisia
Brazil	Ecuador	Japan	Papua New Guinea	Turkey
Brunei Darussalam	Egypt	Jordan	Paraguay	Uganda
Bulgaria	El Salvador	Kenya	Peru	United Arab Emirates
Burkina Faso	Equatorial Guinea	Kuwait	Philippines	United Kingdom
Burundi	Fiji	Lao People's Dem. Rep.	Poland	United States
Cambodia	Finland	Lebanon	Portugal	Uruguay
Cameroon	France	Liberia	Qatar	Vanuatu
Canada	Gabon	Libya	Rep. of Korea	Venezuela
Central African Rep.	Gambia	Madagascar	Romania	
Chad	Ghana	Mali	Saint Pierre and Miquelon	
Chile	Gibraltar	Malta	Samoa	
China	Greece	Mauritania	Saudi Arabia	

Table A2: **List of 131 Countries**

A7 List of Variables

Variable	Description
δ_{ijtk}	indicator for positive imports
Z^*_{ijtk}	a latent import variable
Y^0_{ijtk}	the volume of positive imports
Y_{ijtk}	$log(1 + Y^0_{ijtk})$
X_{ijtk}	vector of explanatory variables
β_k	a vector of coefficients for the product k extensive margin
γ_k	a vector of coefficients for the product k intensive margin
d_{tk}	random shock for year t that affects product $k \sim N(0, \sigma^2_{dk})$
h_{tk}	random shock for year t that affects product $k \sim N(0, \sigma^2_{hk})$
b_{ik}	random shock for importer i that affects product $k \sim N(0, \sigma^2_{bk})$
f_{ik}	random shock for importer i that affects product $k \sim N(0, \sigma^2_{fk})$
c_{jk}	random shock for exporter j that affects product $k \sim N(0, \sigma^2_{ck})$
g_{jk}	random shock for exporter j that affects product $k \sim N(0, \sigma^2_{gk})$
$\tilde{u}_{ijtk}$	an observation-specific shock $\sim N(0, \sigma^2_u)$
$\tilde{\epsilon}_{ijtk}$	an observation-specific shock $\sim N(0, \sigma^2_\epsilon)$

Table A3: **Variable Descriptions**

In the quadruple $ijtk$

i	is the importing country
j	is the exporting country
t	is the year
k	is the SITC 4-digit product

Supplementary Appendix: References

Anderson, James E., and Eric van Wincoop. 2003. “Gravity with Gravitas: A Solution to the Border Puzzle.” *The American Economic Review* 93 (1): 170–92.

Hanoch, Giora, and Haim Levy. 1969. “The Efficiency Analysis of Choices Involving Risk.” *The Review of Economic Studies* 36 (3): 335–46.

Heckman, James J. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica* 47 (1): 153–61.

Helpman, Elhanan, and Paul Krugman. 1985. *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition and the International Economy.* Cambridge, MA: MIT Press.

Lange, Glenn-Marie, Quentin Wodon, and Kevin Carey. 2018. *The Changing Wealth of Nations 2018: Building a Sustainable Future.* Washington, D.C.: The World Bank.

Liu, Xuepeng. 2009. “GATT/WTO promotes trade strongly: Sample Selection and Model Specification.” *Review of international Economics* 17 (3): 428–46.

Olsen, Randall J. 1980. “A Least Square Correction for Selectivity Bias.” *Econometrica* 48 (7): 1815–20.

Park, Trevor, and George Casella. 2008. “The Bayesian Lasso.” *Journal of the American Statistical Association* 103 (482): 681–86.

Ratkovic, Marc, and Dustin Tingley. 2017. “Sparse Estimation and Uncertainty with Application to Subgroup Analysis.” *Political Analysis* 25 (1): 1–40.

Silva, J. M. C. Santos, and Silvana Tenreyro. 2006. “The Log of Gravity.” *The Review of Economics and Statistics* 88 (4): 641–58.

Tibshirani, Robert. 1996. “Regression Shrinkage and Selection via the Lasso.” *Journal of the Royal Statistical Society. Series B (Methodological)* 58 (1): pp. 267-88.