Politics, Institutions, and Trade*

In Song Kim†  John Londregan‡  Marc Ratkovic§

October 7, 2016

Abstract

A vast literature on international trade predicts that political institutions affect both the choice of trade partners (extensive margin) and trade volumes (intensive margin). Yet most analyses ignore the former margin, while their estimates of the latter neglect nontrading pairs and thereby suffer from selection bias. Also, the data on individual features of multifaceted political institutions are highly collinear, which impedes estimating their separate effects. We develop and implement a two-stage Bayesian LASSO estimator that lets us use detailed measures of institutional features, with highly disaggregated product-level trade flows encompassing over one hundred and thirty countries over a half century. We find that political institutions matter more for the extensive than the intensive margin, and that democratic political institutions are not always trade promoting; although trade appears to be fostered by political stability, other features of democracies – party competition and open executive recruitment – inhibit trade.

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

Word Count: 11905 (abstract: 147)

---

*We thank Benjamin E. Goldsmith, Joanne Gowa, Adeline Lo, Edward Mansfield and Walter Mebane for helpful comments. We also thank conference participants at the Asian Political Methodology and Midwest Political Science Association.

†Assistant Professor, Department of Political Science, Massachusetts Institute of Technology, Cambridge, MA, 02139. Email: insong@mit.EDU URL: [http://web.mit.edu/insong/www/](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](http://www.princeton.edu/~ratkovic)
1 Introduction

Scholars of international and comparative political economy have identified various mechanisms through which domestic political institutions affect international trade flows. On the extensive margin, domestic political institutions are important determinants of the choice of trading partners. Democracies are more likely to trade with one another (Bliss and Russett, 1998), and countries with stable domestic property rights and contractual institutions are more likely to trade certain products to begin with (Levchenko, 2007; Nunn and Trefler, 2014). On the intensive margin, some have argued that legislative constraints allow an executive make a credible commitment to liberalization and mutual reduction in trade barriers (Mansfield, Milner, and Rosendorff, 2000, 2002). Democratization also shifts the distribution of political power to the owners of abundant factors (the median voter), which induces more trade in line with countries’ comparative advantages (Milner and Kubota, 2005). In contrast, others argue that the large number of veto players and greater fragmentation of political authority in democracies makes them more sensitive than autocracies to protectionist demands (Frieden and Rogowski, 1996; Henisz, 2000; Isebelis, 2002; Henisz and Mansfield, 2006). That is, democracies might raise tariffs and erect non-tariff barriers to protect heterogeneous domestic interests (Kono, 2006, 2008; Tavares, 2008).

Despite the rich theoretical debates over institutional effects on trade, empirical research must surmount two significant hurdles. First, severe multicollinearity is endemic to trade data. Numerous factors that might affect bilateral trade flows are correlated with political institutions and with each other. A partial list includes economic size (Tinbergen et al., 1962); preferential trading blocks (Frankel, Stein, and Wei, 1997); membership in GATT/WTO (Rose, 2004; Gowa and Kim, 2005; Goldstein, Rivers, and Tomz, 2007; Subramanian and Wei, 2007); other domestic institutions (Nunn and Trefler, 2014); security alliances (Gowa, 1989); trade resistance (Anderson, 1979); and multilateral resistance (Anderson and Wincoo, 2003) to say nothing of importer, exporter, year, and even dyad-specific effects. Extreme multicollinearity among institutional features has spurred some scholars to seek refuge in aggregation, comparing portmanteau measures of institutional structure with total volumes of bilateral trade. Indeed, it is common for analyses to include a single measure of institutional structure, such as the Polity IV index (Gurr, Marshall, and Jaggers, 2010), or even more starkly, simply a dummy indicator for democracy (Przeworski et al., 2000). This refuge comes at a price, preventing researchers from distinguishing among the diverse political channels through
which political institutions might affect trade flows.

A second challenge to estimation is posed by pairs of countries that do not trade. These inactive dyads are often excluded from analyses (e.g., Mansfield, Milner, and Rosendorff 2000; Tomz, Goldstein, and Rivers 2007), a practice that engenders selection bias, while it also blurs important distinctions between the impact of institutions on the selection of trade partners (the extensive margin) and on the volume of trade amongst dyads that do engage in commerce (the intensive margin).

The main contribution of this paper is to disentangle the heterogeneous effects of political institutions on the extensive and intensive margins of trade. We adopt a variable selection methods developed in the statistical machine learning literature to investigate whether the relevance of particular institutional features arises as empirical regularities based on massive data set. First, we decompose the widely used Polity IV measure into its underlying components to capture the potentially disparate effects of institutional features such as executive selection and legislative constraint (Treier and Jackman 2008). We also include a brace of other variables that are often appended to the gravity model of trade, from WTO membership status to whether an importer and exporter share a former colonial relationship. Secondly, because we believe that international trade is governed by political and economic forces that vary by industry, (Krugman 1979; Rodrik 1995; Milner 1999), we work with directed dyads (including those with no trade) at the level of four digit SITC codes. We consider trade flows across 131 countries, for a total 17,030 ($131 \times 130$) directed dyads, across 449 Standard International Trade Classification (SITC) 4-digit products for 51 years, transcending the amounts of data that have been used in the literature.

We propose a two-stage Bayesian LASSO estimator to deal with the twin problems of multicollinearity and selection. First, the decomposition of the Polity IV measure, in addition to a drove of gravity variables as well as ancillary institutional variables, leaves us with a preponderance of highly collinear explanators. We cope with this by including LASSO (Least Absolute Shrinkage and Selection Operator) constraints in our estimation (Tibshirani 1996), which we embed in a Bayesian LASSO context (Park and Casella 2008) using the method of Ratkovic and Tingley (2015). Our estimation strategy shrinks coefficients and prunes superfluous explanatory variables, providing a principled means of selecting a parsimonious battery of explanatory variables from among the large set of highly correlated covariates that enter our analysis. Second, the disaggregation of bilateral
trade at the four digit product-level compels us to distinguish the extensive and intensive margins of trade because there are many countries that simply do not trade particular products with one another, an occurrence that becomes more pronounced at greater levels of disaggregation. The “two stage Tobit” framework (Heckman, 1979; Amemiya, 1984) encompasses both the intensive and extensive margins of trade. We first estimate the effects of political institutions on an indicator for whether the directed dyad in question trade at all (extensive margin). This estimate enables us to calculate a bias correction factor that we then incorporate in our stage estimator that examines the impact of institutions on trade volumes (intensive margin) amongst the directed dyads that actually trade.

We find that political institutions matter a great deal more on the extensive margin than they do on the intensive margin of trade. We find that the influence of political institutions on the selection of trading partners is on par with that of standard gravity variables such as the size of economy and “trade resistance.” In contrast, political institutions exercise but an exiguous impact on the variation in trade volumes for countries that actually trade. That is, the decision to trade at all is political, but given there is trade, the level of trade is largely economic. The result is stable and robust over time, industries, and dyads.

Consistent with Mansfield, Milner, and Rosendorff (2000), we find that constraints on the chief executive have positive effects on trade on both the extensive and intensive margins. However, our results also indicate that democratic institutions do not always promote trade. In particular, our estimates reveal that while the institutionalized transfer of executive power is associated with an increase in the extensive margin of imports, it diminishes exports. This may stem from irregular governments being captured by special interests that demand protection for imports and subsidies for exports. In addition, we find that countries whose de facto leadership succession is hereditary, corresponding to low values for the openness of executive recruitment measure, are predicted to be more likely to import on the extensive margin than countries with more open processes of executive recruitment. Our finding is consistent with Nunn (2007) who finds that countries with poor institutions are net importers of contract-intensive products such as intermediate products in manufacturing industries.

Our results also show that political institutions exert heterogeneous effects across different types

---

1 An example of a four digit product includes SITC 6532: “fabrics, woven, of synthetic staple fibers, containing 85% or more by weight of such fibers (other than chenille fabrics).”
of industries. For example, the positive effects of democratic institutions on the selection of trading partners are particularly pronounced in manufacturing industries while they have relatively small effects on the exchange of crude materials. It is of little surprise that the variation of effects across industries is not limited to political variables. We find that most variables included in our model tend to exhibit heterogeneous effects. Our findings cast serious doubt on the implicit assumption that trade volumes of different products are equally affected by the standard gravity variables that undergird much of the empirical literature on trade.

We are not the only authors to encompass the extensive margin in our analysis. Helpman, Melitz, and Rubinstein (2008) define the extensive and intensive margins at the firm-level and then control for the fraction of exporting firms while maintaining the assumption that this fraction is constant across industries. Our approach differs from theirs in that we define both margins at the product-level and then estimate them separately across industries. We also include a large number of variables for measuring political institutions. Our method is also related to Silva and Tenreyro (2006), who propose the Poisson pseudo-maximum likelihood (PPML) estimator instead of the log-linearized gravity equations for estimating intensive margins. Their primary concern is the bias in the presence of heteroscedasticity and non-normality of errors. We focus instead on the selection bias in the extensive margins. Despite the differences, we share the concerns of these authors who warn of the danger of ignoring country pairs with zero trade.

The rest of the paper is organized as follows. The next section shows the prevalence of country pairs with no trading relations. We then introduce the product-level trade data used for our analysis and describe the decomposition of the Polity IV data into its components. Section 3 describes our two-stage Bayesian LASSO model that deals with the selection and multicollinearity problems simultaneously. Empirical results will be presented in Section 4. The last section concludes.

2 More and Better Data

Much of the existing scholarship uses aggregated economic data and, when the institutional environment is not entirely ignored, employs portmanteau measures of political institutions. These two varieties of agglomeration present a blurred image of the political economy of trade. Here we push

\footnote{Note that fraction of exporting firms is unobservable. Helpman, Melitz, and Rubinstein (2008) estimate this quantity based on a measure derived directly from their model. See equation (14) in Helpman, Melitz, and Rubinstein (2008, page 456).}
on both frontiers, making use of trade flow data at a more disaggregated level than is customary in
the literature, while simultaneously unpacking specific institutional variables, which has been used
frequently to summarize the nature of a country’s political institutions.

This greater clarity and precision come at a cost: the greater the level of disaggregation, the more
commonly one encounters product-specific country dyads with zero trade, while the proliferation of
explanatory variables saddles the analyst with severe collinearity issues. We respond to the first of
these challenges by explicitly incorporating the extensive margin of trade into our analysis. And
we surmount the second hurdle by employing the LASSO, an algorithm that provides a principled
means of sparsifying densely interwoven specifications.

2.1 Trade Data
It is well-known that the politics of international trade is driven by heterogeneous and often conflict-
ing interests across different sectors (Krugman 1979; Grossman and Helpman 1994; Rodrik 1995;
Milner 1999). In fact, countries are endowed differently with factors of production, and hence sec-
tors face different distributional consequences of international trade (Rogowski 1987; Hiscox 2002).
Moreover, differences in production technology also imply that even the typical economic factors
in the gravity models of trade, such as distance and size of the economy, might have differential
effects on trade flows across dissimilar products on both margins. For instance, distance might
matter more for industries with high transportation costs, whereas the importance of the size of
the importing country’s market will be magnified for industries with greater increasing returns to
scale.

Despite the indisputable heterogeneity across industries, remarkably few studies have examined
differential effects of gravity variables on trade flows, let alone covariates for political institutions
(see Dutt, Mihov, and Van Zandt 2013 for a notable exception). Instead, most studies typically
regress the total volume of bilateral trade, aggregating across a diverse set of industries, on a set
of country-level and dyad-level covariates, a procedure that assumes homogeneous effects of each
covariate (e.g., Rose 2004; Goldstein, Rivers, and Tomz 2007). The paucity of analysis at a more
disaggregated level is due in large part to the computational difficulties associated with collecting
and analyzing such a massive body of data, consisting of millions of data points and encompassing
hundreds of disaggregated industries for each country pairing in each year. Even when such data is
collected, one then needs carefully to match product categories across years because new products
enter the dataset while others disappear.

We collect all SITC 4-digit product-level trade data for all country pairs from 1962 to 2012, which is the most finely disaggregated trade data available that spans the entire time period. We organized the data by first identifying countries that have existed as sovereign states since 1961 (a number of countries have been created or disbanded over time, such as the USSR). This results in a total of 131 countries whose trade flows account for more than 90% of the total trade. The list of countries is presented in Web Appendix A4. Next, we created 449 unique SITC 4-digit product categories that are comparable across years to deal with the appearance and disappearance of product categories over time. We used the concordance table available in the United Nations Statistics Division for this task. We then computed the volume of trade for each product so that a dataset of 449 products for each country pair is available across each year. Since we consider all directed dyads even when there is no trade, we have a total of 389,969,970 (131 × 130 × 449 × 51) observations for our analysis. Both the original and the aggregated SITC 4-digit datasets are made available through our Web Appendix webpage. The open-source software concordance: Product Concordance for International Trade, for creating the concordances across different product categories, is made available through the Comprehensive R Archive Network (http://cran.r-project.org/package=concordance).

We include a set of “gravity” variables widely used in the literature. To control for market size, we use population for both the importer and the exporter (imp_POP, exp_POP), and logged income for both the importer and the exporter (imp_tcgdp, exp_tcgdp), from Penn World Tables 7.0. We also include the dyad-level covariates: the logged distance between the members of the dyad (ldist), the log of the product of the land area of the two countries in the dyad (lareap), the number of landlocked countries in the dyad (landl ∈ {0, 1, 2}), the number of island nations in the dyad (island ∈ {0, 1, 2}), indicators for a contiguous land border (border), an indicator for whether both members of the dyad were territories of the same nation during the entire sample period (comctry), e.g. France and Guadeloupe, a common language (comlang), a past colonial relationship (colony), and a common colonizer (comcol). In order to control for the effects of institutional membership in

---

3 We used the data extraction API available from the UN Comtrade Database to handle the large automatic download volume. Typically, the size of data amounts to more than 150MB for each year.

4 Notice that some of these variables, such as comcol, might reasonably be considered to proxy political institutions, although they are included among the canonical controls commonly applied to trade data.
GATT/WTO, we include indicators for formal membership as well as non-member participants for both importers and exporters (e.g., *imp_fmember*, *exp_participant*). Furthermore, we control for whether both elements of a dyad are formal members of the WTO (*bothf*) or participants (*bothp*), and for whether only one of the countries in a dyad is a formal WTO member (*onef*), or a WTO participant (*onep*). These “gravity” variables are from Rose (2004) and Goldstein, Rivers, and Tomz (2007). Finally, we include an indicator for the presence of an alliance relationship (*defense*) in order to control for security relations among trading partners (Gowa 1989). The alliance data is from the Correlates of War Project (Gibler 2013).

2.2 Unpacking Polity IV

Many authors consider some aspect of regime type to be important for trade, and there are myriad measures for such regime characteristics as democracy and autocracy. Some of these measures provide a scale (Coppedge and Reinicke 1991; Bollen, 1993; House, 2014), whilst others work with dichotomous measures (Przeworski et al. 2000), and a handful, such as Olmsted, Signorino, and Xiang (2014), work with less aggregated measures. Yet, with varying degrees of transparency, almost all of the measures combine multiple characteristics of political institutions into a single portmanteau measure. For example, the *Polity IV* (Gurr, Marshall, and Jaggers 2010) measure of government type, *POLITY*, is calculated as the difference between *DEMOCRACY*, a cardinal scale that calibrates how democratic a country’s institutions are, and *AUTOCRACY*, which measures the degree to which institutions are autocratic. We focus on Polity IV because it is relatively transparent and because it is available for a wide range of countries over a large number of years.

The underlying *DEMOCRACY* and *AUTOCRACY* scales are themselves constructed from five component categorical variables: *XRCOMP* (competitiveness of executive recruitment), *XROPEN* (openness of executive recruitment), *XCONST* (constraint on chief executive), *PARCOMP* (competitiveness of political participation), and *PARREG* (regulation of participation). Thus we can represent the *POLITY* score as an amalgam of these five subscales. (See Web Appendix A3 for a more detailed description of the contribution of each to *POLITY*.)

Four of the scales that contribute to the *POLITY* measure posses natural orders of their own. The *XRCOMP* scale ranges from designation of the executive by hereditary selection or rigged elections, corresponding to a value of 1 to selection of the executive in competitive elections, which receives a score of *XRCOMP* = 3. The *XROPEN* measure looks at who can be chosen as
executive, ranging from a 1 for hereditary monarchies, to a score of 4 for polities in which anyone can be elected or designated as head of state. The PARCOMP measure of party competition ranges from 1, for polities where competition with the rulers is prohibited not only in word but also in deed, to PARCOMP = 5 at the other end of the spectrum, with stable parties, a lack of coercion, and regular transfers of political power among groups. All three of the XRCOMP, XROPEN and PARCOMP scales are set equal to zero when the executive is chosen by irregular means, and during transitional periods. The XCONST variable measures the degree to which the executive is constrained, ranging from a value of 1 for dictators and absolute monarchs to 7 for executives whose authority is eclipsed by other political actors, such as we observe when most legislation is initiated by a legislature, a party or a council of nobles.

In contrast with the linearity of the preceding four subscales, the PARREG variable, a five-point scale that purports to measure the degree to which political participation is “regulated” in a particular country, does not resolve into a one-dimensional continuum, with only a fraction of the PARREG scale contributing to the POLITY measure\(^5\) (see Vreeland (2008) for a more extensive discussion of the idiosyncratic nature of the PARREG variable). We set PARREG aside for the purposes of this analysis, but we incorporates two other measures from the Polity IV data set, XRREG and DURABLE. The first of these is a three-point scale measuring the “regularity” of the current leader’s accession to power, with XRREG = 1 corresponding to the seizure of power, whereas a value of 3 indicates executive succession by regular, if not necessarily democratic, means. The DURABLE variable measures the length of time the current regime has been in power.

Some analysts work with an abbreviated version of POLITY. For example, Mansfield, Milner, and Rosendorff (2000) dichotomize Polity IV, setting values above 6 to equal 1 (democracy) and assigning a value of zero otherwise (autocracy). Olmsted, Signorino, and Xiang (2014) dispense with any attempt to amalgamate the subscales and treat each Polity component as a distinct explanator. Here we adapt their approach to our context, interacting our Polity IV variables with indicators for the exporter and the importer. In Section \[\footnote{PARREG \in \{1, 2, 5\} all leave POLITY unaffected, while PARREG \in \{3, 4\} lead to lower POLITY scores.}\] we include as separate explanatory variables our six Polity subscales: XRCOMP, XROPEN, XCONST, PARCOMP, XRREG and DURABLE measures. When we do, we observe considerable heterogeneity among the Polity coefficients for a particular product and direction of trade, thereby confirming our suspicions that disentangling the
components of POLITY is important to understanding the impact of institutions on trade.

2.3 Extensive and Intensive Margins of Trade

The overarching empirical regularity in international trade is that for most products and most dyads, trade flows are zero. According to Helpman, Melitz, and Rubinstein (2008), more than half of country pairs do not engage in any trade at all! Figure 1 displays the absence of positive trade across the country dyads in our analysis. The graph reveals an interesting trend: more dyads are engaged in some level of trade by the end of the sample period. Yet even during the years with the most extensive diffusion of trade, we see that fewer than twenty percent of product-specific dyads actively trade. Amongst those goods that are traded, a majority of country-years exhibit unidirectional trade.

Despite the fact that the absence of trade is the dominant pattern, the majority of empirical studies focus exclusively on samples of countries that exhibit positive trade flows, i.e., discarding observations with zero trade entirely. Moreover, the variables commonly used in gravity equations also affect the probability that two countries trade with each other (Helpman, Melitz, and Rubinstein 2008), extinguishing any reasonable hope that restricting one’s sample to the nonzero trade volumes will produce unbiased estimates (Heckman 1979). Indeed, one of the few studies that
takes the zero trade outcome seriously (Dutt, Mihov, and Van Zandt, 2013) finds that the impact of the WTO comes almost exclusively by way of expanding the set of goods that are traded. We will adopt the useful dichotomy of trade activity into the extensive margin of trade, which examines whether goods are traded at all, and the intensive margin, which pertains to fluctuations in trade volumes among goods that are positively traded. The model we will estimate is an extension of the gravity model of trade, expanded to incorporate institutional variables. To deal with the extreme sample selection at work in our data we employ Heckman’s “two stage Tobit” estimator (Heckman 1979; Amemiya 1984). This estimator nests the widely used Tobit model (Tobin 1958) as a special case.6

3 The Two-stage Bayesian LASSO Model

To achieve our objective—a more accurate picture of the relationship between political institutions and trade—it is not enough to work with more exhaustively disaggregated product data, to incorporate more fully articulated measures of political institutions, and to distinguish between the intensive and extensive margins of trade. We must also bring to bear an estimation technique that can cope with estimating the extensive and intensive trade margins at the same time, even as it copes with the severe multicollinearity the results from working with multiple measured features of political institutions, along with a roster of features such as distance and trade treaties that are standard elements of the gravity model for trade. To accomplish this we combine Heckman’s type II Tobit model with the Bayesian LASSO. This section provides an overview of our estimation strategy; further details appear in the Appendix A.

As foreshadowed in subsection 2.3, to contend with the vast number of non-trading dyads, we adopt the strategy outlined by Heckman (1979) and dubbed by Amemiya (pp.31–36, 1984) as “the type two Tobit model.” Heckman’s procedure takes place in two stages.

3.1 Stage 1: The Extensive Margin

The first stage of our procedure involves estimating a probit model of whether dyads trade. Let $\delta_{ijtk} = 1$ if there are positive imports by country $i$ from exporting country $j$ in year $t$ for SITC 4-digit.

---

6In his review, Liu (2009) advocated using Poisson regression over the Tobit, however, given the virtually continuous nature of reported trade flows, the Poisson is not an attractive option, and in any case, it does not allow for the differences we find between the determinants of selection and the process governing positive trade flows.
product \( k \), while otherwise \( \delta_{ijtk} = 0 \). We index our \( P \times 1 \) vector of explanatory variables \( X_{ijtk} \) in a similar fashion. Given the panel structure of the data, our model incorporates random effects: let \( d_{tk} \) denote a random shock for year \( t \) that affects product \( k \). In order to control for the “multilateral resistance” to international trade identified by Anderson and Wincoop (2003), we include both importer and exporter specific random shocks denoted by \( b_{ik} \) and \( c_{jk} \) respectively. The random effects are assumed to follow a standard normal distribution. Finally, \( \tilde{u}_{ijtk} \) is an observation-specific shock. We summarize the variables used in our model in Table A3 in Web Appendix.

Our first stage probit model of the extensive margin then corresponds to:

\[
Z_{ijtk}^* = X'_{ijtk} \beta_k + b_{ik} + c_{jk} + d_{tk} - \tilde{u}_{ijtk} \tag{1}
\]

where, \( Z_{ijtk}^* \) is a latent variable: we observe \( \delta_{ijtk} = 1 \) when \( Z_{ijtk}^* > 0 \) and \( \delta_{ijtk} = 0 \) otherwise, while \( \beta_k \) is the \( P \times 1 \) vector of coefficients.

We estimate this first stage model by Bayesian methods for observations with at least one year of positive trade and at least one year without trade during the 51-year time frame our data covers—we dub these the “sometimes traders.” Notice that countries that always or never trade provide little useful information about the parameters of equation (1).

### 3.2 Stage 2: The Intensive Margin

Now let’s turn to the intensive margin of trade. Let \( Y_{ijtk}^0 \) denote the volume of trade between importer \( i \) and exporter \( j \) in year \( t \) for SITC 4-digit product \( k \). We will work with the following transformation of trade flows: \( Y_{ijtk} = \log(1 + Y_{ijtk}^0) \). By construction, when \( Y_{ijtk}^0 = 0 \) we will also have \( Y_{ijtk} = 0 \).

In the model of Heckman (1979) we only observe \( Y_{ijtk} > 0 \) when \( \delta_{ijtk} = 1 \), while otherwise there are zero year \( t \) product \( k \) imports to country \( i \) from country \( j \). Following Heckman, we allow for correlation between \( \tilde{u}_{ijtk} \) and the residual from the \( Y_{ijtk} \) equation. Accordingly we include \( m_{ijtk} \), the expected value of \( \tilde{u}_{ijtk} \) given \( Y_{ijtk} > 0 \), among the explanatory variables in our estimating equation for the intensive margin:

\[
\mathbb{E}(\tilde{e}_{ijtk}|\beta_k, X_{ijtk}, b_{ik}, c_{jk}, d_{tk}) = \frac{\phi(X'_{ijtk} \beta_k + b_{ik} + c_{jk} + d_{tk})}{1 - \Phi(X'_{ijtk} \beta_k + b_{ik} + c_{jk} + d_{tk})} \equiv m_{ijtk} \tag{2}
\]

\(^7\)The reader will recognize this as the inverse Mills ratio.
This brings us to our model for the intensive margin, which we fit only to observations for which
\(Y_{ijkt} > 0\):
\[
Y_{ijtk} = X'_{ijtk}\gamma_k + m_{ijtk}\theta_k + f_{ik} + g_{jk} + h_{tk} + \bar{\epsilon}_{ijtk}
\]  
(3)

where \(\bar{\epsilon}_{ijtk}\) is an observation-specific shock that is independent of \(\bar{u}_{ijtk}\).

Notice that, in contrast with the standard Tobit model (Tobin, 1958), Heckman’s type II model allows the explanators for the intensive margin, see expression (3), to differ from those used for the extensive margin, corresponding to expression (1).

Across the different products and model specifications, we estimate over 90,000 coefficients that correspond with political or economic covariates, plus parameters associated with the three-way random effects.

### 3.3 The Bayesian LASSO

For each product we are still left with an extensive and highly collinear roster of explanatory variables, including both gravity variables and different measured features of each country’s political institutions. To contend with this extreme collinearity, we turn to the LASSO, a variable selection strategy pioneered by Tibshirani (1996) to take models with hundreds or thousands of covariates and return a small subset that are useful for explaining the outcome. We implement a Bayesian version that allows us to fit a probit model and include random effects with greater ease. The Bayesian LASSO, first introduced by Park and Casella (2008), also allows researchers to draw statistical inference based on posterior distributions of estimated parameters. We implement the estimator of Ratkovic and Tingley (2015), which adjusts the prior structure of Park and Casella (2008) to guarantee better variable selection properties, and which, unlike standard implementations (Friedman, Hastie, and Tibshirani 2010) incorporates random effects into variable selection. For further details, the reader is encouraged to consult Appendix A and Ratkovic and Tingley (2015).

### 4 Empirical Results

This section presents estimates from our models of international trade. Our basic model (Model 1) includes a set of 22 standard gravity variables described in Section 2.1. Incorporating the six Polity components, interacted with exporter and importer dummy variables, adds the 12 more variables we discussed in Section 2.2 to the specification. In Model 2, we include interactions between the Polity variables for importers and exporters to examine whether bilateral trade is affected by
complementaries between participants’ political institutions. Model 3 explores time-varying effects by interacting each gravity and Polity variable with indicators for specific time periods such as the Cold War and the WTO era. We find that Polity variables matter more on extensive margins (choice of trading partners) than on intensive margins (trade volumes conditional on positive trade). We also find highly heterogeneous effects: the effects of democratic institutions vary across industries and over time, and they are not always trade promoting.

4.1 Heterogeneous Effects of Component Variables: Model 1

We first present the result from Model 1. We estimate the extensive and intensive margins of trade separately for each of the 449 SITC 4-digit industries. Consequently, we have a large number of coefficient values and corresponding credible intervals \((13,470 = 449 \times (18 + 12))\) to report for each margin. We use a “heat map” to present our findings. Coefficients for which 95% of the posterior mean falls on one side of zero are represented by narrow blue bands; the greater the fraction of the density below zero, the darker the share of blue. Industries for which the posterior allots greater probability to positive coefficient values are colored in shades of red, with darker hues representing posteriors that accord higher probabilities to positive coefficient values. To understand how our results are presented in heat maps, we first present Figure 2 and Figure 3 that zoom in on a column and a row, respectively.

Figure 2 portrays our estimated coefficients for the log of distance, the key element of the gravity model, in our model of the extensive margin. The industries are sorted by their SITC 4-digit codes, and we label the one-digit industry groups on the y-axis. The graphical display on the lefthand side of the panel portrays the coefficient estimates for each product; the dot represents the median of the posterior density, while the lines present 95% credible intervals—in many cases these are so narrow with the large number of observations that they appear to coincide with the dots representing the medians. On the righthand side of the figure we color code each coefficient (the heat map representation).

Figure 2 shows that the effects of distance vary by product, although it generally reduces the likelihood of trade. Consistent with the gravity model of trade, we do not find positive effects of distance on the extensive margin in any product, which can be easily seen by the lack of red stripes. However, our model predicts that distance does not matter for some industries. Consider the coefficient for “Railway locomotives with steam tenders” that appears among the results for the
Figure 2: Effects of Distance on the Extensive Margin of Trade: This figure presents the heterogeneous effects of distance (logged) on the extensive margin. We label the four industries with the smallest effects of distance on trade. The left side of the panel shows the posterior medians (black circles) with 95% credible intervals. The righthand side is the heatmap representation, which corresponds to the first column in Figure 4.
Figure 3: **Effects of Gravity and Polity Variables on Product-Level Trade:** This figure shows the effects of each explanator on the extensive margin of trade for a product in the textile industry (SITC 6554). The heatmap representation at the top corresponds to a row in Figure 4.

“Machinery and transport equipment” industry. The plot of the posterior indicates that the 95% credible interval for this coefficient barely excludes zero, and the corresponding band of the heatmap takes on a pallid shade. This is likewise the case for “Finished structural parts of zinc” and for “Ores and concentrates of uranium and thorium.” The median of the “Ammoniacal gas liquors produced in gas purification” coefficient actually coincides with 0, and the corresponding color band is white.

Figure 3 zooms in on a row of the subsequent heatmap representations. It presents extensive margin coefficients for “Coated or impregnated textile fabrics” (SITC 6554). The 95% credible interval for the exporter GDP coefficient corresponding to this product lies far above zero, and it is represented by a dark red band of color, while the posterior for the distance coefficient is concentrated well below zero, thus earning a dark blue streak in the corresponding line of the heatmap. The posterior density for the XRREG coefficient is relatively small and positive, thus it is represented on the heatmap as a light red stripe.

Figures 4 and 5 portray the coefficients for the extensive and intensive margins of trade, respectively, for all covariates and all industries. The industries are arrayed by SITC 4-digit code, with industry groupings labeled on the vertical axis. Explanators are listed along the horizontal axis of
each panel in the same order as in Figure 3. We depict the coefficients for the standard gravity variables in the lefthand panel and the Polity coefficients in the righthand panel.

Figure 4: Model 1 Extensive Margin: This figure presents our coefficient estimates and corresponding credible intervals for all variables included in Model 1 on the extensive margin. The left panel contains the results for conventional gravity variables, and the right panel depicts the coefficients for Polity component variables. Red (blue) color represents positive (negative) correlation between a given explanatory variable on the x-axis and the occurrence of trade of a particular SITC 4-digit product. Darker shade represents a greater degree of correlation. White color is used when there is no statistically significant relationship. We will use heatmap representations in the subsequent figures similarly in order to make the comparison across industries (along vertical direction) and explanatory variables (along horizontal direction) easier.

Figure 5: Model 1 Intensive Margin

A quick perusal of the heat map in the lefthand panel of Figure 4 confirms that, across industries, the key variables of the standard gravity model have their customary effects. As we saw in Figure 2, distance impedes the extensive margin of trade, while increased income (GDP) in either
the exporting or the importing country is associated with a greater likelihood that two countries trade in a given product. We find that higher GDP of exporting country is associated with even higher likelihood of trade in chemical and manufactured industries indicated by the darker lines in the top part of the fifth column. We also find, consistent with the literature on trade, that the existence of a former colonial relationship increases the likelihood that two countries will engage in trade. On the intensive margin, displayed in the lefthand panel of Figure 5, income for both the exporter and the importer is strongly trade promoting, while distance impedes trade, albeit less dramatically than on the extensive margin, and the effect of former colonies fades relative to the extensive margin. The sporadically negative coefficients earned by our population measures on the extensive margin, represented by the blue bands, indicate that most traded goods have positive elasticities with respect to per capita income.

Notice that even for the log of distance variable in our extensive margin estimates (lefthand side panel of Figure 4), the staccato hatching indicates that there are some industries in which distance is not even selected by our estimator as an explanator. The darker blue bars, indicating a large negative coefficient, are more prevalent among the manufacturing industries than they are among primary products.

Importantly, we find that political institutions matter even after controlling for a large number of gravity variables used in the literature. If the Polity variables provide no additional explanatory power, the righthand panel of Figures 4 and 5 should consist predominantly of white space, with a smattering of mostly pastel-colored coefficient bands. A cursory glance reveals that the figures contrast vividly with the panorama that would accompany a null result.

Heterogeneity is even more conspicuous among our Polity variables than among our gravity variables. We find that not all democratic institutions are trade promoting. The righthand panel of Figure 4 reveals that the most vivid colors of the heat map pertain to the XRREG variable: regular governments import more and export less than irregular governments that came to power by extraordinary means, typically via a coup d’état. Another consistent pattern is the effect XROPEN exerts to reduce trade. This variable is essentially an inverse nepotism measure, gauging the degree to which the executive is not chosen on a hereditary basis.

The XRREG effects, which push in different directions for exports and for imports, are possibly a byproduct of protectionism: autocratic leaders, for whom the “hurdle factor” of
Myerson (1999) is low, may be succumbing to protectionist pressure from anti-trade special interest groups (Acemoglu and Robinson 2000). This pattern is also consistent with the successful export-oriented industrialization process in countries with autocratic leaders, such as contemporary China and, during the 1970s, South Korea.

Turning to the XROPEN variable for importing country, we interpret the preponderance of negative coefficients, which are especially prevalent amongst manufacturing imports, as stemming from the reliance of the most kleptocratic regimes on the international economy. When the protections afforded by the rule of law are nearly absent, producers of tradeable goods that could be produced domestically will prefer to flee, relocating behind the protective shield of an international border and exporting to the autocratically governed country. This relocation limits producers’ risk to the potential confiscation of a few shipments rather than of an entire factory. This is also consistent with Nunn (2007) who finds that countries with unstable property rights institutions are net importers of contract-intensive products such as intermediate products in manufacturing industries.

In general, however, we find it is democratic institutions that help promote trade on extensive margins. For example, we observe that exporting countries tend to have have meaningful constraints on executives (XCONST) and competitive political participation (PARCOMP). However, the size of the positive effects vary across industries. This is especially true for our exporter coefficients corresponding to the elected chief executive, (XRCOMP), for which the predominantly red bands representing positive coefficient values are punctuated by streaks of blue–see for example product 5151, encompassing radioactive chemicals and isotopes, or product 5123, consisting of ethers, epoxides, and acetals, exports of which are attenuated, ceteris paribus when the chief executive is elected.

Our remaining Polity coefficients on the extensive margin in Figure 4 appear either as very pallid bands, such as the XCONST variable for importers, corresponding to insignificant coefficient estimates, or as intermittently striped bars, such as our XRCOMP variable for importers, indicating effects whose direction differs across industries, or both, as in the case of the importer’s values for DURABLE and for PARCOMP.

On the other hand, the effects of Polity variables on the intensive margin are substantially attenuated. The paler color scheme and greater prevalence of white space represents the relatively reduced significance of our measures of political institutions for the intensive margin of trade.
Notice also that even for the columns in which significant coefficients are more prevalent, such as the \( XRREG \) and \( XRCOMP \) columns, the signs of the coefficients vary, complicating generalization across industries.

We note that two of our Polity exporter variables, \( XRCOMP \), which is higher for elected leaders, and \( PARCOMP \), which measures the degree to which opposition is tolerated, tend to promote the extensive margin for trade but impede the intensive margin. While these effects are not consistent across industries, and their impact on the extensive margin of trade is less substantial than we observe for \( XRREG \) and \( XROPEN \), this is a regularity worth noting.

We suspect the complementary opposite signs for the \( PARCOMP \) coefficients are a byproduct of the tendency of more democratic governments to be willing and ready to commit to negotiate trade agreements, while at the same time they are more susceptible to moderately protectionist pressures from special constituencies [Dixit and Londregan 1995], so that in democracies constituents in import competing industries are more successful at extracting protections as a condition of allowing trade.

### 4.2 Alternative: The Extensive and Intensive Margins Contrasted

To formalize our finding that political institutions matter more for extensive margins than for intensive margins, we compare the Bayesian posterior density for each product \( k \) from a model that incorporates both gravity and Polity variables (\( M^P_k \)) against a model that includes only gravity variables (\( M^G_k \)). Applying the widely used procedure of Kass and Raftery (1995), we calculate the Bayes factor (BF) which tells us about how strongly the data favor the model incorporating the Polity variables compared with the model using the gravity variables by themselves:

\[
BF_{PG}^k = \frac{Pr(D_k|M^P_k)}{Pr(D_k|M^G_k)}
\]

where \( D_k = \{Y_k, X_k\} \) denotes the observed data for product \( k \). Appendix B provides more extensive computational details. Working with logged values of the Bayes factor, a positive difference implies that the model with Polity variable is preferred, while \( \log(BF_{PG}^k) \) greater than 2 is sometimes taken to imply “decisive” evidence in favor of \( M^P_k \) (Jeffreys 1998; Kass and Raftery 1995).

The first panel of Figure 6 shows the distribution of Bayes factors for Model 1 against a model that only includes economic gravity variables across each product \( k \). The left-hand box and whisker plot of the first panel confirms the importance of Polity for predicting the extensive margin of trade:
including the Polity variables in addition to conventional gravity variable leads to an improved Bayes factor for every single product in our dataset. Furthermore, the median Bayes factor is around 500 in log-scale, which provides strong evidence in favor of our finding. The right-hand results in the first panel, which relate to the intensive margin, are less overwhelming—in some industries, the improved fit to the data is not sufficient to justify the addition of so many parameters, though for the majority of industries it does. Our finding is robust to considering a model with only the POLITY variable. In the second panel, we compare the model with decomposed policy against the model with POLITY variable (and not its components). The last panel compares the model with POLITY against the model with only gravity variables. These results increase our confidence that political institutions matter for trade, especially on the extensive margins.

4.3 Interactions among the Polity Variables: Model 2

Scholars of international political economy have emphasized that political institutions of country pairs mutually affect international cooperation (e.g., Dixon [1994], de Mesquita et al. [1999], Mansfield, Milner, and Rosendorff [2000], [2002]). In particular, Mansfield, Milner, and Rosendorff [2000] theorize that exporters’ and importers’ values for XCONST interact, with countries that impose greater constraints on their executives tending to trade more with each other. To explore this
possibility, as well as the possibility that other political variables interact in a meaningful way, we expand our basic model by including interactions between the importer’s and the exporter’s Polity variables for each product category. In order to maximize the probability of finding interaction effects with the LASSO estimation, if they are present, we use the transformation of the interaction terms recommended by [Ratkovic and Tingley](2015). Specifically, we work with interaction terms that are constructed to be uncorrelated with their main effects, allowing us effectively to distinguish main effects from interactive effects. Doing so maximizes the probability that if there are interaction effects among our variables we will find them.

**Figure 7: Model 2 Extensive Margin**

**Figure 8: Model 2 Intensive Margin**
We find a surprising lack of evidence for interaction effects (see Figures 7 and 8). This can be seen by the prevalence of white columns that correspond to the interaction terms. Of the twenty interaction effects we estimate for the extensive margin, only three show the least hint of a stable effect across product classes; careful examination of the righthand panel of Figure 7 suggests a possible trade promoting effect on the extensive margins for manufacturing goods if levels of importers’ $XRCOMP$ and exporters’ $PARCOMP$ are high, but even this effect is a feeble one. Even more pallid are the sporadically trade promoting effects of the interaction between the importer’s value of $XRCOMP$ and the level of $XCONST$ for the exporter, and the interaction of importer $XRREG$ and exporter $XROPEN$, leaving us with hints of an effect for but two of the twenty interactions we estimate on the extensive margin of trade. The striking regularity of our results is that while the direct effects of the explanators from our basic model persist, when we add the interactions, the interaction effects themselves add almost nothing.

When we check the robustness of the Bayes factors for the Polity variables, in the left panel of Figure 9 we continue to observe that for the extensive margin the Polity variables improve the fit of our model for every product in our dataset, whereas we encounter languid results on the intensive margin, with negative Bayes factors for many, albeit a minority, of industries. As expected by the lack of interaction effects, the right panel shows that including the full interaction terms often do not improve the fit.

![Figure 9](image-url)  
**Figure 9: Institutional Variables Matter for Extensive Margins:** The left figure compares the importance of including the Polity component variables in explaining extensive and intensive margins of trade in Model 2. The baseline model is the model with only gravity variables. It confirms the importance of political institutions on the extensive margin. The right figure compares Model 2 against Model 1. It shows that adding additional interaction terms does little to improve the fit.
4.4 Stability of our Results Across Time: Model 3

Next we consider the possibility that the mechanisms that promote trade have shifted over time. We divide our sample period into four subperiods, which we label as “pre-OPEC,” encompassing the dozen years between 1962 through 1973, “Cold War,” which covers 1974 through 1989, “post–Cold War,” encompassing 1990 through 2001, and “WTO,” the eleven years between 2002–2012 that followed China’s December 11, 2001 WTO entry. The four periods are chosen to reflect important changes in international trading system as well as the heterogeneity across different “waves” of democratization (Huntington 1993). Indicator variables for each subperiod are then fully interacted with the covariates at each time. We use “Cold War” period as our baseline for ease of interpretation.

![Figure 10: Model 3 Extensive Margin](image)

The main effects in this model continue to resemble the parameter estimates from our basic model both for the standard gravity variables and for the polity effects. Our estimates for the main effects of the gravity variables appear in the first columns of the left-hand panel of Figure 10. The estimated main effects largely echo the coefficients in our basic model, see Figure 4, across product classes the estimated main effect for distance $ldist$ is consistently negative, while main effects for the $comlang$ (common language) and $concol$ (common colonizer) variables earn positive coefficients, especially for chemicals, manufactured goods, and transportation equipment (SITC

---

8We provide an enlarged version of the results from Model 3 in Web Appendix A1 where axis labels are presented with bigger font size.
codes 5 and higher).

Now turn to the righthand panel of Figure 10. As with the gravity variables, the estimated main effects for the polity variables also largely coincide with the estimates for our basic model, as reported in Figure 4: the EXP.XRREG earns consistently positive coefficients across industries, whereas the estimated effects of IMP.XRREG are negative. Likewise, our estimated main effects for EXP.XCONST and for EXP.DURABLE are positive, as are both the importer and exporter main effects for PRCOMP, while the estimated main effect for IMP.XROPEN remains negative when we include the subperiod interaction terms.

Estimated Subperiod Interactions

In contrast with our analysis in the preceding subsection that revealed a marked paucity of significant interactions between importers’ and exporters’ political characteristics, here do find various interactions between our subperiod indicators and the other variables in our model.

Amongst the standard gravity variables, our estimates indicate that during the early pre-OPEC part of our sample, the trade promoting impact of having shared a common colonizer, comcol, was more pronounced, whereas the same effect is substantially attenuated during the WTO subperiod, as is the trade enhancing impact of colony, the variable coding for whether one county in the dyad was a colony of the other. That is, the effects of colonial relationship on bilateral trade have faded out over time.

Our analysis also sheds light on the controversy about whether GATT/WTO membership increases international trade. Unlike the original works that consider dyad as unit of analysis [Rose 2004; Goldstein, Rivers, and Tomz 2007], we estimate the effects of membership for importer and exporter separately. We find heterogeneous effects over industries and time. First, in all three models, we find that formal membership of exporters (exp_fmember) increases the trade of manufactured products (industries six, seven, and eight) both on extensive and intensive margins. Second, non-member participants trade more heavily in crude materials and food, but not of other products. This latter effect results from two sets of coefficients at work in our data: firstly, the exp_participant variable that encompasses both members and to non-member participants, tends to earn a positive coefficient for most crude products and food. But these coefficients are offset by negative coefficients on exp_fmember, so that it is only the non-member participants who exhibit the enhanced trading effect. Finally, we find increasing time-varying effects of par-
participant status on the extensive margins of trade as shown by a red vertical bar associated with \textit{imp\_participant\_WTO} (concentrated particularly on manufactured goods). These findings suggest that the effects of the WTO are not ubiquitous, but rather that they are driven by formal members’ exports of manufacturing goods (i.e., imports by developing non-member participants), and non-member participants’ exports of raw materials. This trade pattern in which the metropolis exports manufactures while the countries of the periphery ship primary products is reminiscent of the pattern of globalization during the century preceding the first world war and documented by 
\textbf{Williamson} (2011) in which a core of countries industrialized, even as the periphery “deindustrialized”, specializing in the export of raw materials, albeit that as of the second half of the Twentieth Century the list of core countries has expanded somewhat from the early industrializers. Plus ça change.

Turning to the Polity variables, we find that \textit{XCONST}, a variable measuring executive constraints, had a less pronounced trade promoting impact, especially on manufactured goods during the pre-OPEC period, whereas its trade enhancing impact has strengthened during the WTO subperiod. In contrast, party competition, \textit{PARCOMP}, was most positively related to trade during the pre-OPEC period, whereas its effect on trade have faded during the WTO subperiod. The export promoting impact of the executive having come to power by regular means, as measured by the coefficient for \textit{EXP\_XRREG}, is higher during the WTO period, whereas the import inhibiting effect of a regular executive, calibrated by the \textit{IMP\_XRREG} parameter, was more pronounced during the pre-OPEC period, but it has remained negative\footnote{Our analysis also identifies more ephemeral changes in the exporters’ coefficients for \textit{XRCOMP} and \textit{XROPEN}.}

The changing portrait of trade that emerges reveals that party competition has faded as a predictor of whether two countries trade, while executive constraint, a variable that was identified by \textbf{Mansfield, Milner, and Rosendorff} (2000) as important for trade, has become more important over time. The surprising dichotomous effect of \textit{XRREG}, which encourages exports and impedes imports, has remained over time, with the export promoting effects rising during the WTO subperiod, while the import discouraging effects have subsided somewhat relative to the pre-OPEC years, but still remain strong.

Figure\textsuperscript{11} depicts the intensive margin. Turning first to the lefthand panel, notice that over time the trade enhancing importance of sharing a language, and a former colonizer, have become less
important during the WTO subperiod, whereas the trade inhibiting impact of distance has waxed over time. The righthand panel of Figure 11 shows the trade promoting effects of XRREG, and the trade impeding impact of PARCOMP, have both attenuated during the WTO subperiod, so that over time the dichotomy between the importance of political institutions for the extensive margin, and their relative lack of impact on the intensive margin, has become even more pronounced.

Finally, the lefthandside panel of Figure 12 reveals that including temporal interactions leaves intact our basic finding that the Polity variables are more important for the extensive margin than they are for the volume of trade flows among countries that actually trade. The righthandside of Figure 12 provides evidence that including the temporal interactions is important for our political variables, whereas the constellation of gravity variables exhibit greater stability over time.

5 Concluding Remarks

We estimate the relationship between political institutions and both the extensive margin of trade, that is, whether a dyad trades at all, and the intensive margin, which measures the magnitude of the trade flows given they exist. To do this, and to cope with the vastitude of collinear data, we employ a two-stage Bayesian LASSO estimator that encompasses the selection issues raised by the large fraction of directed dyads with no trade whatsoever and with the substantial collinearity that emerges among the plethora of variables in our data set.

While our primary focus is on our substantive findings, we note that these results were revealed not only by having access to fine grained measures of trade and of regime characteristics, but
Figure 12: **Institutional Variables Matter for Extensive Margins:** This left panel compares the importance of including the Polity component variables in explaining extensive and intensive margins of trade from Model 3. The right panel shows that including time interaction improves the fit particularly on the extensive margin.

also by being able to deploy the relevant methodology—notably the bias correction developed by [Heckman (1979)](ref) allowed us separately to analyze the extensive and intensive margins. Likewise, the LASSO techniques ([Tibshirani, 1996](ref) [Ratkovic and Tingley, 2015](ref)) we brought to bear allowed us to analyze a much more extensive list of collinear explanators while still retaining the ability to draw statistically meaningful inferences. Because cross national data analysis often involves similarly lengthy lists of explanators, we believe that researchers interested in international comparisons will find useful the LASSO methods we apply here to the study of trade.

Controlling for the standard gravity model covariates, we find that political institutions have a robust and significant impact on the extensive margin trade, while there is weaker evidence for such a link with the intensive margin.

Working with data that span a half a century, hundreds of thousands of annual observations of directed dyads, and hundreds of SITC 4-digit product categories, we observe considerable heterogeneity in the effects of political institutions and economic variables across industries. Importantly, institutional features associated with democratic government are not uniformly trade promoting; rather, we find that the components of the widely used *Polity IV* measure of democracy vary in their influence. However, a few regularities transcend product categories: on the extensive margin, having a national leader who came to power by legal means is associated with increased imports but with attenuated exports, while a leader who was selected via other than hereditary means is linked with reduced imports. Moreover, our findings challenge the widely held belief that democracies are
especially predisposed to trade with one another. We estimate complementaries among the political institutions of trade partners, but we find that these are largely unrelated to either the extensive or the intensive margins.

We do, however, identify significant changes over time—during the half century we study, the importance of former colonial relationships for trade has faded, while after 2001 the importance of WTO participation for imports has waxed. Nevertheless, the pattern of trade under the WTO, which favors exports of manufactures for formal WTO members and primary product exports for non-member participants, bears a remarkable resemblance to the blueprint for commerce prior to the first World War, when the metropolitan powers imported raw materials and exported manufactures back to the countries on the economic outskirts.

We also find that since 2001, countries with long-lived political institutions are associated with less trade on the extensive margin. Similarly, we see that the mixed and insignificant coefficient estimates for party competition mask a dynamic effect in which, for the bulk of our sample period, party competition was related to the increased extensive margin of exports, and less powerfully, for imports as well, but since 2001 the trade promoting effect of this emblem of democracy has evaporated. In contrast, non-hereditary leadership selection, which discourages imports on the extensive margin, after 2001 appears to promote exports.

The political profile of countries with high levels of engagement with the international economy consists of regular, non-hereditary leadership succession, while the institutional durability and competitive politics that we associate with democracy have faded in importance. Our results suggest the emergence of a brave new world of international trade in which trade is fostered by political stability, whether it comes in the form of the stable authoritarian regimes of China and the Asian Tigers or instead emulates the longstanding deliberative democracies of the US and the EU.
References


Tinbergen, Jan et al. 1962. *Shaping the world economy; suggestions for an international economic policy*. Twentieth Century Fund, New York.


Appendix

A Modeling Details

We fit a type II Tobit model as two interconnected models. The first of these, for the extensive margin, is a probit model. The probit model can be augmented to be fit as a least squares model (Jackman 2009). The second model is for the intensive margin, but only observations with non-zero trade. An inverse-mills ratio from the probit model is used as a residual correction for the truncation at zero. Both models can be written as normal regression models, and the importer-, exporter-, and year- random effects can be fit within each model’s hierarchy.

Denote as \( Y_{ijtk} \) the logged level at which country \( i \) imports from country \( j \) product \( k \) in year \( t \). \( X_{ijtk} \) is a vector of covariates, combining the gravity and economic variables. Denote as \( S_{ijk} \) an indicator variable for whether country \( i \) imports any of product \( k \) from country \( j \) at any time \( t \), i.e., \( S_{ijk} = 1 \) (max \( t \{ Y_{ijtk} > 0 \} \) ), with \( N_{1k} = \sum_{i,j,t} S_{ijk} \).

For the first stage, in which we only include directed dyads that trade during some, but not all, of the years in our sample. We follow the standard latent representation of a probit model, assuming \( \Pr (Y_{ijtk} > 0) = \Phi (Y_{ijtk}^*) \). Associated with covariates \( X_{ijtk} \) is a \( P \times 1 \) vector of parameters, \( \beta_k \). Denote \( \beta_{kp} \) the \( p \)th element of \( \beta_k \) and \( p \)th element of \( X_{ijtk} \) as \( X_{ijtp} \). We model the latent outcome \( Y_{ijtk}^* \) as:

\[
Y_{ijtk}^* | \beta_k, b_{ik}, c_{jk}, d_{tk}, S_{ijk} = 1 \sim \mathcal{N} (X_{ijtk}^T \beta_k + b_{ik} + c_{jk} + d_{tk}, 1) \tag{5}
\]

\[
\beta_{kp} | \lambda_{1k} \overset{\text{i.i.d.}}{\sim} \text{Dexp} (\lambda_{1k}) \tag{6}
\]

\[
\lambda_{1k}^2 \overset{\text{i.i.d.}}{\sim} \Gamma (\sqrt{N_{1k}}, 1) \tag{7}
\]

\[
b_{ik} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{bk}^2); \quad c_{jk} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{ck}^2); \quad d_{tk} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{dk}^2) \tag{8}
\]

\[
\sigma_{bk}^2 \propto 1/\sigma_{bk}^2; \quad \sigma_{ck}^2 \propto 1/\sigma_{ck}^2; \quad \sigma_{dk}^2 \propto 1/\sigma_{dk}^2 \tag{9}
\]

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. See Park and Casella (2008) for details. We parameterize the conditional prior on \( \lambda_{1k}^2 \) 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. See Ratkovic and Tingley (2015) for details. The parameters \( b_{ik}, c_{jk}, \) and \( d_{tk} \) are random effects, each with a Jeffreys’ prior over the scale parameter.
Following the logic given above, we calculate the sparse estimate as

$$
\beta^{\text{sparse}}_{kp} \cdot = \beta_{kp} \mathbf{1} \left( \left| \sum_{i,j,t} \left( Y_{ijtk}^* - \sum_{p'} X_{ijtp'} \beta_{kp'} \right) X_{ijtk} \right| > \lambda_{1k} (N_{1k} - 1)^{1/4} \right) \tag{10}
$$

For the outcome model, we use all observed positive trade flows, including those for directed dyads that trade in all of the years, and so were omitted from the first stage of our estimation. We construct the selection correction variable in terms of our first-stage parameters, which allow us to calculate $Y_{ijtk}^*$ even for observations pertaining to directed dyads that were not used in the estimation of the first stage model. Denote the inverse-mills ratio $m_{ijtk} = \phi(Y_{ijtk}^*) / (1 - \Phi(Y_{ijtk}^*))$, with $\phi()$ and $\Phi()$ the density and distribution functions of a standard normal variable, respectively.

Denote as $N_{2k} = \sum_{i,j,t} \mathbf{1}(Y_{ijtk} > 0)$ the number of observations with non-zero trade. We model the level of trade for these observations as a linear regression in terms of parameters $\gamma_k$ with element $p$ denoted $\gamma_{kp}$. To account for the selection bias, we include an additional covariate $m_{ijtk}$ and associated parameter $\theta_k$.

$$
Y_{ijtk} | \gamma_k, f_{ik}, g_{jk}, h_{tk}, Y_{ijtk} > 0, m_{ijtk} \sim \mathcal{N} \left( X_{ijtk}^T \gamma_k + f_{ik} + g_{jk} + h_{tk} + m_{ijtk} \theta_k, \sigma_k^2 \right) \tag{11}
$$

$$
\Pr(\theta_k) \propto 1 \tag{12}
$$

$$
\gamma_{kp} | \lambda_{2k}, \sigma_k \overset{\text{i.i.d.}}{\sim} \text{Dexp} \left( \frac{\lambda_{2k}}{\sigma_k} \right) \tag{13}
$$

$$
\lambda_{2k}^2 \overset{\text{i.i.d.}}{\sim} \Gamma \left( \sqrt{N_{2k}}, 1 \right) \tag{14}
$$

$$
\sigma_k^2 \propto 1/\sigma_k^2 \tag{15}
$$

$$
f_{ik} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{f_k}^2); \ g_{jk} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{g_k}^2); \ h_{tk} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{hk}^2) \tag{16}
$$

$$
\sigma_{f_k}^2 \propto 1/\sigma_{f_k}^2; \ \sigma_{g_k}^2 \propto 1/\sigma_{g_k}^2; \ \sigma_{hk}^2 \propto 1/\sigma_{hk}^2 \tag{17}
$$

The model is nearly identical to the one above, except we now must model an unknown variance, $\sigma_k^2$. We place a Jeffreys’ prior over $\theta_k$, the parameter associated with the correction term $m_{ijtk}$. We again generate a sparse estimate as:

$$
\gamma_{kp}^{\text{sparse}} \cdot = \gamma_{kp} \mathbf{1} \left( \left| \sum_{i,j,t} \left( Y_{ijtk} - \sum_{p'} X_{ijtp'} \gamma_{kp'} \right) X_{ijtk} \right| > \lambda_{2k} \sigma_k^{\text{sparse}} \right) \tag{18}
$$

The Gibbs sampler was run for 8,000 iterations. The first 4,000 were discarded, and every 20th sample from the remaining 4,000 were saved. We analyze these 200 saved posterior draws.
B Bayes Factor Calculation

We calculate the Bayes factor (BF) as follows:

\[
BF_k^{PG} = \frac{\Pr(D_k|M_k^P)}{\Pr(D_k|M_k^G)} = \frac{\int \Pr(D_k|\theta_P, M_k^P) f(\theta_P|M_k^P) d\theta_P}{\int \Pr(D_k|\theta_G, M_k^G) f(\theta_G|M_k^G) d\theta_G}
\]  

(19)

where \(D_k = \{Y_k, X_k\}\) denotes the observed data for product \(k\), and \(\theta_\ell\) is the set of parameters in model \(M_\ell\) where \(\ell \in \{G, P\}\). Following Kass and Raftery (1995), we compute \(\Pr(D_k|M_\ell^k)\) for Model \(M_\ell\) based on the harmonic mean of the likelihood values given by a \(m^{th}\) sample from a posterior density as:

\[
Pr(D_k|M_\ell^k) = \left\{ \frac{1}{M} \sum_{m=1}^{M} \Pr(D_k|\theta_\ell^{(m)})^{-1} \right\}^{-1}.
\]  

(20)