LEARNING THE WEALTH OF NATIONS

FRANCISCO J. BUERA, ALEXANDER MONGE-NARANJO, AND GIORGIO E. PRIMICERI

Abstract. We study the evolution of trade policies over time and across countries. We consider a model in which own and neighbors' past experiences influence policy choices, through their effect on policymakers' beliefs. We estimate the model using a large panel of countries. We find that there is a strong geographical component to learning, which is crucial to explain the slow adoption of liberal policies during the post-war period. Our model also predicts that there would be a substantial reversal to protectionism if nowadays the world was hit by a shock of the size of the Great Depression.

1. Introduction

The search for principles that lead to the wealth of nations is the central goal of economic policy. Yet, despite seemingly pursuing a common objective, growth policies differ enormously across countries. In addition, these differences tend to persist over time and often reflect differences in the views and beliefs that are dominant in the countries. Trade policies, the focus of this paper, are the prototypical example of these facts.

In this paper we argue that the wide dispersion in trade policies largely reflects divergent beliefs about the benefits of free trade. Policy decisions, in turn, influence beliefs through their effects on economic development. Stanley Fisher suggests the rise of socialist ideas in the 20th century as a clear example of the interaction of ideas, policies and outcomes:

> "It is not hard to see why views on the role of the state changed between 1914 and 1945 [...] A clear-headed look at the evidence of the last few decades at that point should have led most people to view the market model with suspicion, and a large role for the state with approbation- and it did." [**?**, p. 102]

Date: September 24, 2006. We would like to thank Michael Clemens for providing some of the tariff data of Clemens and Williamson (2004) and Scott Kastner for sharing the Hiscox and Kastner (2007) openness measure.

In a similar vein, the dismantlement of these ideas was a consequence of the change in views that "resulted from a combination and interaction of research and experience with development and development policy." (**?**) The testimony of a key actor of this episode attests this fact:

> "I remember the foreign minister and the finance minister from another country saying to me: 'You're the first prime minister who is ever tried to roll back the frontiers of socialism. We want to know what's going to happen. Because if you succeed others will follow.'" [Margaret Thatcher in **?**]

Trade policies constitute, perhaps, the best example of policies that are fervently debated along ideological lines about the role of the state vs. markets. The effective influence of those ideological lines on actual policies is also driven by recent growth experiences. We aslo focus on trade policies because of two addtional reasons. First, despite important problems, there exist relatively good measures of these policies for many countries. Second, trade liberalization is often seen as the "sine qua non" part of a liberal package of policies that are is believed to foster economic development (**?**).

The crucial connection among beliefs, policies and economic development has received little formalization in the literature. In this paper we aim to help closing this gap considering a world in which the trade policies in each country are (partly) driven by beliefs about trade in the country and in which the evolution of beliefs is driven by past outcomes of the countries and their neighbors.

More specifically, we are motivated by two sets of questions. First, we want to understand whether and why beliefs about trade policies have changed over time: Can these changes be captured by a rational learning process? How localized is this learning? How fast does this process converge? How dispersed are today's beliefs about the benefits of free trade? Secondly, we want to predict how likely are massive policy reversals to protectionism due to possibly large shocks hitting the world economy, e.g. shocks of the size of the Great Depression.

To answer these and similar questions, we model the behavior of benevolent policymakers. Policymakers start with some prior beliefs about the effect of trade openness on economic growth and update these beliefs with the arrival of new information. They decide to open the country to international trade if they think that free trade fosters growth and if political costs entailed by free trade policies are not too large. The introduction of random political costs allows us to quantify the explanatory power of the simple learning mechanism. We estimate the parameters of the model (prior beliefs, geographical structure of learning, and the distribution of political costs) using a panel of 128 countries for the postwar period, 1950-1998.

Our estimation results indicate that there is a strong geographical component to learning. The weight assigned to the experience of other countries declines by approximately 40% for every 1000Km (621.37 miles) of physical distance between countries. To illustrate the importance of this estimate, observe that the median country in our dataset has only four countries within 1000Km of distance.[1] The geographical nature of learning is crucial to explain the dynamics of trade policies in the post-war period. Indeed, the model fits very well the slow adoption of liberal policies and this is attributed to the persistence of beliefs. The learning mechanism rationalizes most of the policy choices (92%) observed in the data. Moreover, it explains 60% of the numerous trade liberalizations towards the end of the century, because the generalized bad growth outcomes of late 1970s and early 1980s were even worse for closed economies. However, the model also indicates that there still exists a substantial dispersion of beliefs by the end of the sample.

This motivates our concern that large shocks might induce a sequence of policy reversals to protectionism. We address this concern by performing a series of counterfactual simulations. Our model predicts that about 10% of the countries would revert to protectionism within 5 years, following a world-wide shock of the size of the Great Depression. We conclude from our exercises that understanding the evolution of beliefs is a central ingredient for the dynamics of policies.

Our paper relates to a large literature that studies the determinants of trade policies. This literature mainly explores political economy dimensions, like redistributional issues, interest-group politics, the role of multilateral institutions (see, for example, **?**, **?**, **?** and **?**). Our paper complements this literature by studying the formation and evolution of beliefs about the benefits of different policies. In this respect, our work is more closely related to **?**, where policy choices are related to the behavior of rational agents learning from past experience. However, while the analysis of **?** is solely theoretical, we attempt to explore the quantitative role played by the evolution of beliefs for policy outcomes. Therefore, we purposely

[1] The relative development of different countries, as measured by the differences in their per-capita GDP, also determines the relevant information neighborhood.

abstract from the political economy aspects, concentrating on tractable models of beliefs formation.

Models of policymakers as rational learning agents have been successfully applied to explain the rise and fall of US inflation (see, for instance, **?**, **?**, **?** and **?**). Differently from this literature, policymakers in our paper do not face a complex trade-off between alternative policy objectives. Our focus is instead on a multicountry model and on the role of learning spillovers among countries. In this respect, our paper is related to the literature on social learning and information spillovers in technology adoption and diffusion (see, for instance, **?**, **?** or **?**).

Finally, this paper draws from the empirical literature studying the connection between trade and growth (see, for instance, **?**, **?**, **?** or **?**). However, while most of this literature investigates the impact of free trade on economic development, our focus is exactly on the converse, i.e. understanding the determinants of trade policies. In this respect, our objective is more similar to **?**, **?** and, to a lesser extent, **?**, although none of these papers stresses the importance of past growth performances for current policy choices.

The rest of the paper is organized as follows. Section 2 provides a long-term perspective on the evolution of trade policies and beliefs about the benefits of free trade. Section 3 examines the dynamics and geography of economic growth and trade policies during the postwar period. Section 4 and 5 present the theoretical model and the estimation methodology. The estimation results and counterfactual exercises are discussed in sections 6 and 7 respectively. Section 8 contains some robustness checks and section 9 concludes.

2. Trade Policies and Beliefs: A Long-Term View

In this section we examine the behavior of trade policies of leading countries during the 19th and 20th centuries. We also look at the evolution of the dominant views regarding free-trade and protectionism during these centuries. We draw two main conclusions. First, trade policies exhibit large and long lasting cycles. Important liberalization episodes are followed by large reversals to protectionism. Second, policy reversals and changes in mainstream views about the optimality of openness to trade followed large aggregate shocks. For instance, after an interval of relative freer trade, the world shifted to protectionism after both the “Long” (1870-1892) and the “Great” (1930s) depressions.

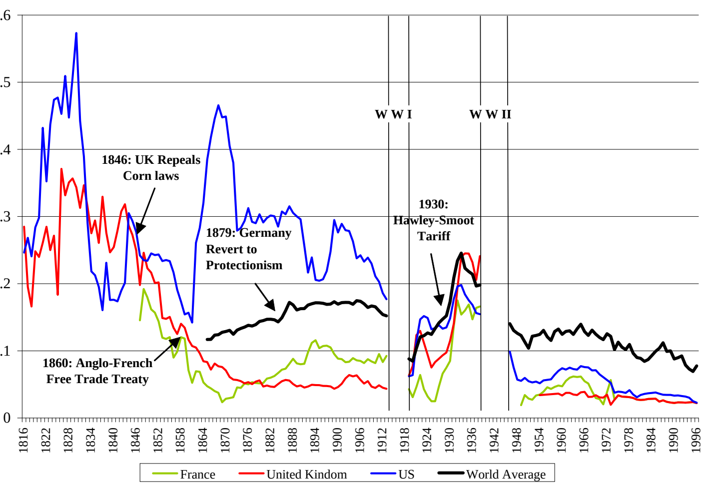

FIGURE 1. Evolution of average tariffs: world average and selected countries, 1816-1996

2.1. **Evolution of trade policies.** Figure 1 plots the behavior of average import tariffs (weighted by volume) in France, the United Kingdom, the United States. The figure also displays a world (unweighted) average.[2] Average tariffs are commonly used to capture openness to trade in different periods of history (e.g. **?**, **?**, **?**). Despite important limitations –discussed below– average tariffs remains a useful measure when interested in a long historical view.

Figure 1 identifies three large cycles of liberalization-reversion to protectionism during the 19th and 20th centuries. The first cycle starts after the Napoleonic Wars, with a slow transition towards economic liberalism in Great Britain that leads to the repealing of the Corn Laws in 1846. Such trend towards free trade reaches its peak with the Cobden-Chevalier free trade agreement between Great Britain and France in 1860. The Cobden-Chevalier is one of many agreements which by 1860s involved most of European countries. By the 1860s, the large fraction of the world economy represented by Europe went from a restricted regime policy where trade prohibitions were the rule (especially for manufactured goods) to moderately low tariffs. Trade policies in non-European countries follow somehow independent trends. For example, by the middle of the 19th century the US move towards more liberal policies, but this trend was reversed after the Civil War.

[2] The data is taken from **?** for the United Kindom, from **?** for the United States and from **?** for France. The world average is taken from **?**. For 1865-1939 this is the unweighted average of 35 countries and for 1950 onwards the unweighted average of 182 countries.

This period of liberalization translated into a large increase in trade. European exports went from an average growth rate of 3.5% per year during 1830 and 1846 to an average of 6% growth rate per year during 1846 and 1860. Other measures of effective openness such as the ratio of exports-to-GDP also increased substantially. For instance, that ratio in Great Britain went from 3% in 1820 to 12% percent in 1870.[3]

This cycle of liberalization is reverted in the late 1870s. This was at least partially a consequence of the "Long Depression," a long period of economic stagnation of most European economies that lasted from around 1870 until at least 1892.[4] Germany, with a major tariff reform in 1879, was the first large European country that reverted towards protectionism.

The second movement towards freer trade was mainly driven by the decline of transportation costs, rather than explicit trade liberalizations (**?** and **?**). The renewed trend towards openness can be seen in the aggregate dynamics of trade, and the record ratio of trade-to-GDP right before the first world war. European exports grew at 4% on average in the period 1890/2-1913 after a period of relatively slower growth, during the shift to protectionism, 1877/9-1890/2 (**?**). The ratio of Trade-to-GDP reached 18% in Great Britain by 1913 (**?**). This period of growing trade was interrupted by the First World War, and definitely terminated after the world reversal toward protectionism following the Great Depression in the 1930s.

The third long cycle starts with the end of World War II and goes to the present day. The period is characterized by a very slow and fragmented movement towards free trade that culminates with the widespread liberalizations in less developed countries during the 1980s and 1990s. Before looking more closely to this period –the focus of our analysis– it is instructive to show that world aggregate shocks shake and revert perceptions about openness to international trade.

2.2. **Evolution of beliefs.** Most of the discussion of trade policy focuses on political economy forces such as vested interests, contributions and rent-seeking (e.g. **?** and **?**). It is less well recognized that the public and policymakers' perception about the desirability of free trade is also an important determinant of trade policies. In turn, it is quite natural to expect that this perception about the benefits of

[3] Similar increases took place in most European economies. Austria, France and Spain went from being virtually closed economies exporting only 1% percent of GDP in 1820 to exporting 5% percent in 1870

[4] Continental Europe was the most severly affected region. Per-capita output went from growing 1.1% per year in 1850-1870 to 0.2% in 1870-1890 (**?**).

trade is affected by past experience. An early example of how the performance of alternative trade regimes shapes the evolution of beliefs is given by the debate on free trade in continental Europe following the liberalization of trade by the United Kingdom, as discussed in **?**:

> "[...]the European supporters of free trade did not fail to draw attention to the British example itself. For example, the Association Belge pour la Reforme Douaniere, which developed out of the Societe Belge d'Economie Politique, published in 1855 a manifesto for tariff reform which started as follows, 'Inspired by the results of economic science and by the experience of real facts, especially that of England, where, since the introduction of Sir Robert Peel's reforms, agriculture, navigation and industry, far from declining, have flourished in force and energy in the most unexpected way." [Bairoch (1989, p. 29)]

However, and quite interestingly given the solid, long-held and widespread free trade tradition, past performances can also shift and revert the perceptions of leading economists about trade. A clear illustration is the impact of the Great Depression on the rapid change of opinion of John M. Keynes on the perceived desirability of liberal policies. In 1919, Keynes wrote the following colorful ode to the liberal state of affairs preceding the war:

> "What an extraordinary episode in the economic progress of man that age was which came to an end in August 1914!...The inhabitant of London could order by telephone, sipping his morning tea in bed, the various products of the whole earth, in such quantity as he might see fit, and reasonably expect their early delivery upon his doorstep." [Keynes (1919, p. 10), quoted in Sachs and Warner (1995)]**?**

During the 1930s, when most market economies were stagnant while the Soviet Union was growing rapidly, his views on trade were radically different. In his essay on "National Self-Sufficiency" he drastically changes his views:

> "I sympathize, therefore, with those who would minimize, rather than with those who would maximize, economic entanglement among nations. Ideas, knowledge, science, hospitality, travel–these are the things which should of their nature be international. But let goods

> be homespun whenever it is reasonably and conveniently possible, and, above all, let finance be primarily national." [Keynes (1933, p. 3), quoted in Sachs and Warner (1995)]**?**

There, Keynes also observes that beliefs are shaken by the Great depression and that countries are looking at each to learn about the consequences of alternative policies:

> "But today one country after another abandons these presumptions. Russia is still alone in her particular experiment, but no longer alone in her abandonment of the old presumptions. Italy, Ireland, Germany have cast their eyes, or are casting them, towards new modes of political economy. Many more countries after them, I predict, will seek, one by one, after new economic gods. Even countries such as Great Britain and the United States, which still conform par excellence to the old model, are striving, under the surface, after a new economic plan. We do not know what will be the outcome. We are–all of us, I expect–about to make many mistakes. No one can tell which of the new systems will prove itself best." [Keynes (1933, p. 3)]

In sum, the two quite opposite discourses can only be understood as a process in which beliefs form and change from policy regimes and outcomes, including those of other countries.

3. Postwar Dynamics and Geography of Openness

In the rest of the paper, we formally explore the relationship between the dynamics of trade policies and believes by using a large panel of 128 countries from 1950-1998. We use a comprehensive measure of openness proposed by **?** and extended by **?**. In this section, we describe Sachs and Warner's openness indicator and present some reduced-form evidence on the connection between policies and past outcomes that further motivates our formal model of beliefs formation.

3.1. **Measuring openness.** Sachs and Warner's (1995, hereafter SW) construct a comprehensive measure of openness. They argue that a country dispose of a variety mechanisms to close its economy, beyond import tariffs. For instance, a country that heavily taxes exports would eventually dry up its imports by reducing the privately perceived terms of trade. Equally, international trade would be blocked if the country taxes or prohibits imports or if it blocks or taxes the convertibility of its currency. Following this logic, SW classify a country as open in a given year if the country meets *all* of the following five criteria: (i) The average tariff rate on imports is below 40%; (ii) Non-tariff barriers cover less than 40% of imports; (iii) The country is not a socialist economy; (iv) The state does not hold a monopoly of the major exports; and (v) The black market premium is below 20%. The SW openness indicator is a dichotomic variable that assigns a value of one to a country for each year that is open and zero otherwise.

The data is an unbalanced panel, partly because some countries started or stopped existing during the sample period. The large coverage of countries and years is a major advantage of the SW indicator over other indexes based on direct policy assessments, such as the World Bank's outward orientation index and the Heritage Foundation index of trade policy.[5]

Observe that the SW indicator has been criticized because items (iii)-(v) capture non-trade aspects of policy (see, for instance, **?**). As pointed out in the introduction, this would certainly be a concern if we were interested in investigating the causal relationship between free trade and growth. However, our focus is rather on understanding how countries performances affect policy choices. Therefore, the fact the SW indicator includes a broader set of policies, represents an important advantage for our analysis.

Of course, one disadvantage of SW is the discrete nature of the openness measure, especially compared to alternative approaches that look at the actual impact of trade distortions. For example, an alternative method is to use empirical models in order to assess the difference between the observed volume of trade and a "free-trade counterfactual." To this end, the well-known gravity equation is a successful empirical tool in replicating observed bilateral trade flows. Moreover, it has gained conceptual support from the work of **?** and **?**, who derive the gravity equation as the equilibrium condition of multicountry general equilibrium models. Along these lines, **?** have recently used bilateral trade flows to obtain estimates of the impact of policy distortions over time. Appendix A demonstrates that the SW indicator contains very similar information to the Hiscox and Kastner's measure of openness, for the set of countries for which the latter is available (about 60% of the countries present in the SW dataset).

[5] As reported by **?**, these other indicators of openness are highly correlated with the SW indicator.

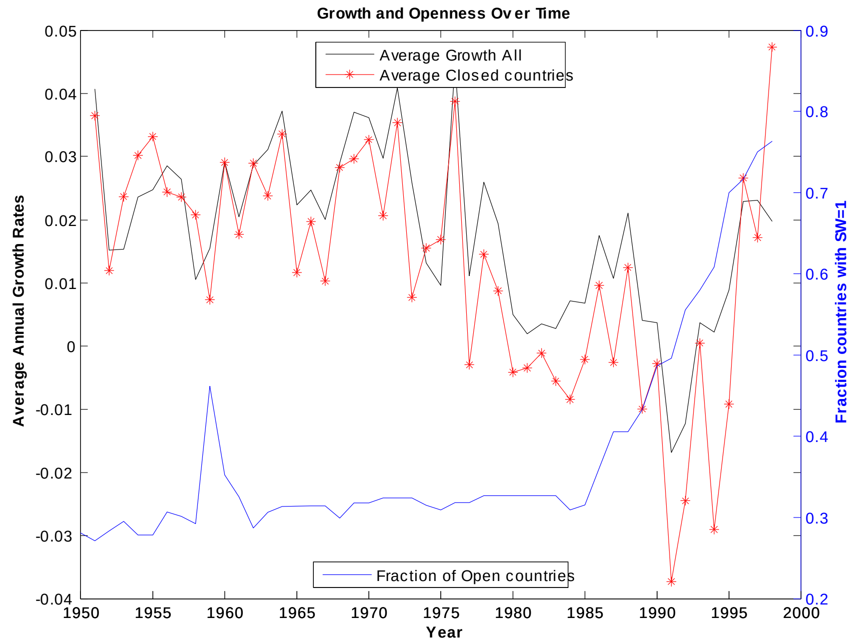

FIGURE 2. World average of openness and growth

3.2. **Reduced-form evidence.** Figure 2 shows the evolution of the fraction of open countries according to the SW openness indicator. While in 1950 less than 30% of the countries are open, by the end of the sample this fraction increases to almost 80%. The spike in the late 1950s reflects the introduction of the European Union and the entry of developing (and closed) countries into the sample. From 1963 to 1984, the fraction of open countries remains practically flat at around 32%. The year 1985 seems to be the commencement of the movement to global openness that carries out to the end of the sample.

The world-wide average hides interesting regional dynamics. Figure 3 displays the average SW indicator (with scale on the left-hand-side axis) for six regions in the world. Four of the regions open up only in the mid or late 1980s. Western Europe is an obvious exception since it is almost completely open since 1959 onwards. Asia and the Pacific is another exception since the trend towards openness starts in the 1960s and 1970s, much earlier than in the other regions. North and Central America start fairly open, but with the inception of the Central American common market and the strategy of import substitutions, only the U.S. and Canada remained open. South America starts implementing the strategy of import substitution earlier and

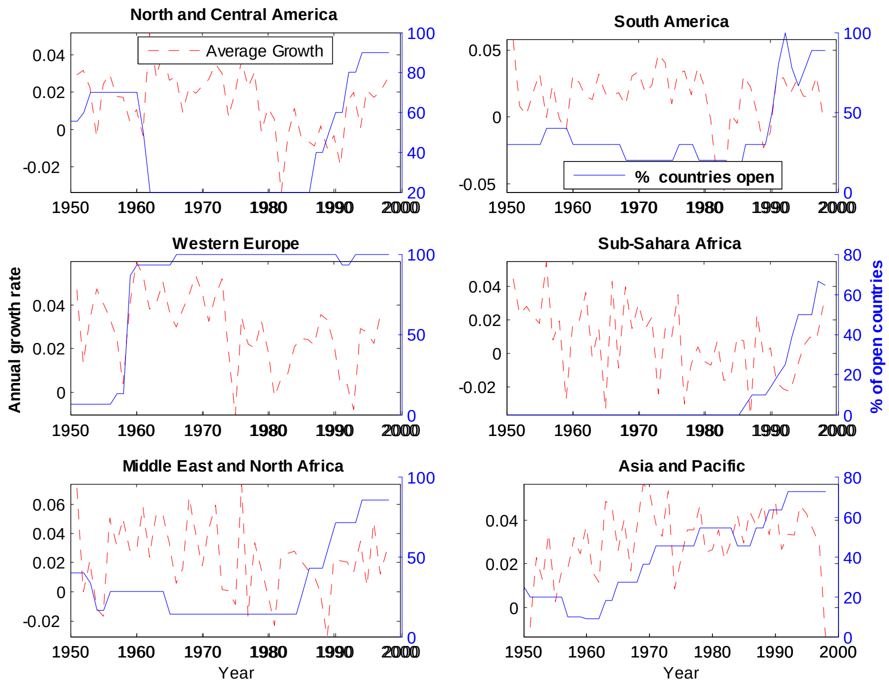

FIGURE 3. Average openness and growth across regions

exhibits a trend similar to the Middle East and North Africa: from the beginning to the mid-1980s, a small and rather constant fraction of countries remains open, and then it quickly shoots up in the later years. All of the African countries in the sample remain closed until also mid-1980s when this extreme closeness starts reverting.

In sum, while most regions display a late and fast opening, they also exhibit well defined regional trends. This fact suggests, rather strongly, that regional factors play a key role in determining the policies that countries end up undertaking.

In order to relate policy choices to past growth performances, figure 2 also plots the (unweighted) average growth rate of per-capita GDP[6] for all countries and only the closed ones. Notice that it is only after the growth collapse of the 1970s and the early 1980s, driven mainly by the closed economies, that the trend towards openness starts. A similar pattern can be observed at the regional level (figure 3). Initially, most regions were closed and exhibited high average growth. However, South and Central America, Africa and the Middle East decided to switch to openness after

[6] GDP data are obtained from Penn World Table 6.2.

the bad performance of the 1970s and 1980s. In addition, observe that openness in the later years of the sample is associated with the resumption of growth.

In order to examine all of these issues more formally, we run a probit model using the SW indicator and GDP growth data for the 128 countries of our sample. The dependent variable of the probit regression is the probability of being open. This reduced-form analysis delivers two main results. First, trade policies are spatially correlated. Second, policy choices are highly correlated to the past performance of trade regimes: for each additional point of per-capita GDP growth of open countries in the neighborhood of country i, the probability that country i is open increases by approximately 10 percentage points. All the details of this exercise are presented in appendix B.

This reduced-form evidence leads us to model policies as being chosen by rational policy-markers that update their beliefs after observing a new vintage of data. The rest of the paper develops and estimates this model.

4. Model

In this section we consider a simple model in which the continuous arrival of information affects policymakers' beliefs and trade policy decisions in each of the N countries in the world economy. Whether a country opens or not to international trade is determined by the perceived consequences of openness on growth, as well as the political costs entailed by liberal policies.

4.1. The policy decision problem.

We simplified the choice of trade policies to the dichotomic case of countries that are either open or closed. Policies are chosen period by period and we let $\theta_{i,t}$ be an indicator variable that equals one if country i is open in period t and zero otherwise. Moreover, let $Y_{i,t}$ denote the level of GDP in country i at time t and $y_{i,t} \equiv \log Y_{i,t} - \log Y_{i,t-1}$ its growth rate. Policymakers choose the sequence $\{\theta_{i,s}\}_{s=t}^{\infty}$ to maximize:

$$\begin{aligned} & \max_{\{\theta_{i,s}\}_{s=t}^{\infty}} E_{i,t} \sum_{s=t}^{\infty} \delta^{s-t} \left[\log Y_{i,s} - \theta_{i,s} K_{i,s}\right] \\ \text{s.t.} \quad & y_{i,s} = \beta_i^c \left(1 - \theta_{i,s}\right) + \beta_i^c \theta_{i,s} + \varepsilon_{i,s} \\ & K_{i,t} \overset{i.i.d.}{\sim} N\left(0, \sigma_{k,i}^2\right) \\ & \varepsilon_{i,s} \overset{i.i.d.}{\sim} N\left(0, \sigma_i^2\right), \qquad \text{all } s > t. \end{aligned} \tag{4.1}$$

From the perspective of policymakers (but not necessarily from ours) openness and growth are linked by a linear, causal relationship. In particular, β_i^c and β_i^o

represent the average growth rates in country i under closeness and openness respectively. Policymakers do not know the values of β_i^c and β_i^o and use past and current data to learn about them.

The variable $K_{i,t}$ is an exogenous random variable that captures political and social costs of being open at time t. The constant $\sigma_{k,i}^2$ denotes the variance of these costs in country i. In our estimation procedure, we will treat $\left\{\sigma_{k,i}^2\right\}_{i=1}^N$ as a set of unknown parameters. A small estimate of $\sigma_{k,i}^2$ indicates that large variations in exogenous political costs are not needed to rationalize the observed dynamics of trade policies over time. In other words, we can interpret $\sigma_{k,i}^2$ as a metric for the fit of the model for country i.

The timing of events is as follows: at time $t-1$, policymakers in country i observe data on openness and growth of all N countries and update their beliefs about β_i^c and β_i^o. At the beginning of time t, they observe the realization of $K_{i,t}$ and decide whether to open or close.

Given our assumptions, policy decisions are independent over time. Optimal policy at time t is given by

$$\theta_{i,t} = 1\left\{E_{i,t-1}\left(\beta_i^o\right) - E_{i,t-1}\left(\beta_i^c\right) > K_{i,t}\right\},$$

where $1\{\cdot\}$ is the indicator function. Notice that the optimal policy decision only depends on the expected average growth rates and not the entire distribution of beliefs: policymakers choose a free trade regime if their expectation of the difference between average growth under openness and closeness is higher than the political cost of being open.

4.2. **Learning.** To render tractability in this multicountry learning problem, we impose additional restrictions. We assume that, in period $t = 0$, policymakers of country i start off with a Normal-Inverse Gamma prior density[7] on the coefficients of (4.1), i.e. $\beta_i \equiv [\beta_i^c, \beta_i^o]'$ and σ_i^2. More precisely,

$$\begin{aligned} \sigma_i^2 &\sim IG\left(s_0, d_0\right) \\ \beta_i|\sigma_i^2 &\sim N\left(\hat{\beta}_{i,0}; \sigma_i^2 \cdot P_{i,0}^{-1}\right), \end{aligned} \tag{4.2}$$

where IG and N denote the Inverse Gamma and the Normal distributions respectively. Here s_0 and d_0 are the scale and the degrees of freedom parameterizing

[7] Normal-Inverse Gamma priors are standard in linear Gaussian models because they are conjugate priors. For an introduction to Normal-Inverse Gamma distributions and the concept of conjugate prior, see **?**.

the Inverse Gamma density, while $\hat{\beta}_{i,0}$ and $\sigma_i^2 \cdot P_{i,0}^{-1}$ represent, respectively, the expected value and the variance of the conditional prior on β_i. We choose the following parameterization for the inverse of the precision matrix $P_{i,0}$:

$$P_{i,0}^{-1} = \nu^2 \cdot \begin{bmatrix} 1 & \rho_i \\ \rho_i & 1 \end{bmatrix}, \tag{4.3}$$

where ρ_i captures the correlation coefficients of initial beliefs. Notice that we are making the simplifying assumption that the diagonal elements of the covariance matrix are the same. This means that policymakers start with a similar degree of uncertainty about the effects of openness and closeness on economic growth. The coefficient ν^2 parameterizes this uncertainty.

Priors are recursively updated with every new vintage of data. In updating their beliefs, policymakers of country i might use data from other countries, depending on how useful such data are perceived to be to learn about β_i. For example, if Argentinian policymakers believe that the effect of free trade on Argentinian growth is fundamentally different from the rest of the world, they will update their beliefs using only Argentinian data. On the contrary, if they believe that the growth effect of trade openness is approximately the same in the whole world, the data for every country will carry the same weight as Argentina's own data to update their beliefs.

In order to capture this idea in a flexible way, we assume that policymakers of country i believe that the relationship between openness and growth in other countries is described by the following equations:

$$\begin{aligned}
y_{j,s} &= \beta_{j|i}^c (1-\theta_{j,s}) + \beta_{j|i}^o \theta_{j,s} + \varepsilon_{j|i,s} && (4.4)\\
\beta_{j|i}^c &= \beta_i^c + \sigma_j \sqrt{q_{ij}} \eta_j^c && (4.5)\\
\beta_{j|i}^o &= \beta_i^o + \sigma_j \sqrt{q_{ij}} \eta_j^o && (4.6)\\
\eta_j^c &\sim N(0,1) \\
\eta_j^o &\sim N(0,1), \qquad j = 1, ..., N
\end{aligned}$$

where the subscript $j|i$ denotes country i view about country j variables.

Under this formulation, policymakers of country i believe in the existence of a linear causal relationship between openness and growth also in other countries. However, the effect of openness on growth might differ across countries. The variable q_{ij} (which scales the "noise" variables η_j^c and η_j^o) determines how useful data of country j are for country i. In particular, if $q_{ij} = 0$, the effect of openness in country i and j is the same. Policymakers of country i would then use both sources of data symmetrically to update beliefs. As q_{ij} increases, data from country j become less and less informative about the growth effect of openness in country i.

We assume that q_{ij} is a parametric function of a vector of covariates z_{ij}:

$$q_{ij} = \exp\left[-2 \cdot z'_{ij}\gamma\right] - 1.$$

We borrow this formulation from the literature on geographically weighted regressions (see, for instance, **?**). The vector z_{ij} may include various measures of physical as well as cultural distance between country i and j. This specification captures the idea that policymakers attach more weight to countries that are closer geographically and culturally to the home country.

Under the additional assumption that the vector $\{\sigma_j\}_{j=1}^N$ is known to policymakers,[8] Bayes law induces simple updating rules for $\hat{\beta}_{i,t} \equiv E_{i,t}\left([\beta_i^c, \beta_i^o]'\right)$, i.e. the expectation of policymakers' beliefs in country i about the average effect of closeness and openness on growth. Appendix C derives these formulas and shows that this formulation of the problem is equivalent to a weighted least squares estimation problem, in which policymakers of country i assign a weight

$$w_{ij} = \frac{\sigma_i}{\sigma_j} \exp\left[z'_{ij}\gamma\right]. \tag{4.7}$$

to data coming from country j.

5. Inference

Like the agents of our model, we (the econometricians) are also Bayesian and wish to construct the posterior distribution for the unknown parameters of the model. These unknown coefficients are:

$\left\{\hat{\beta}_{j,0}^c\right\}_{j=0}^N$	:	expectations of initial beliefs about the effect of closeness
$\left\{\hat{\beta}_{j,0}^o\right\}_{j=0}^N$	:	expectations of initial beliefs about the effect of openness
$\left\{\rho_j\right\}_{j=0}^N$	:	correlation of initial beliefs about the effect of closeness and openness
$\{\sigma_{j,k}\}_{j=1}^N$	:	standard deviation of the political cost
γ	:	coefficients of the weighting function

[8] The purpose of this assumption is simplifying the Bayesian learning mechanism. To assign values to $\{\sigma_j\}_{j=1}^N$ we run a panel regression of growth on openness, time and fixed effects, using the entire postwar sample. If anything, this assumption works against us because assigns more knowledge than they actually have to policymakers.

If we collect the set of unknown coefficients in the vector α and denote by D the entire set of available data on openness and growth, standard application of Bayes rule delivers:

$$p(\alpha|D) \propto \mathcal{L}(D|\alpha) \cdot \pi(\alpha),$$

where $p(\cdot)$, $\mathcal{L}(\cdot)$ and $\pi(\cdot)$ denote the posterior, sampling and prior densities respectively. We now turn to the description of the priors and the construction of the likelihood function.

5.1. **Priors.** Our model is quite heavily parameterized. The use of somewhat informative priors helps preventing overfitting problems. For instance, we would like to avoid cases in which we fit the data well, but only due to estimates of policy-makers' initial beliefs which are clearly implausible. As an example, consider the literature on macroeconomic forecasting: highly parameterized models do well in-sample, but perform poorly out-of-sample. The use of priors considerably improves the forecasting performance of these models (see, for instance, **?** or, more recently, **?**). The role of priors is similar in our context, as our ultimate goal is using the model to conduct a set of counterfactual experiments.

We assume the following prior densities:

$$\begin{aligned}
\pi(\hat{\beta}^c_{i,0}) &= N\left(\bar{\beta}^c_0, \omega^2\right), \quad i = 1, ..., N \\
\pi(\hat{\beta}^o_{i,0}) &= N\left(\bar{\beta}^o_0, \omega^2\right), \quad i = 1, ..., N \\
\pi\left(\frac{\rho_i + 1}{2}\right) &= \text{beta}(a, b), \quad i = 1, ..., N \\
\pi(\sigma_{i,k}) &= \text{IG}(s, d), \quad i = 1, ..., N \\
\pi(\gamma) &= \text{Uniform}.
\end{aligned}$$

We now turn to the description of the parameterization of the priors:

- We set $\bar{\beta}^c_0 = 0.0275$ and $\bar{\beta}^o_0 = 0.0125$. We have chosen these numbers using the Maddison data (**?**). First of all, we have computed the average annual growth rate of GDP using all countries present in the Maddison dataset in the period 1901-1950 (excluding the years corresponding to the two wars). This number corresponds approximately to 2% (the exact number is 2.16%). Then we have noticed that the average growth rate (between 1946 and 1950) of those countries that, according to the Sachs and Warner indicator, result closed in 1950 is approximately 1.5% higher than the average growth rate of the countries that result open in 1950. Notice that starting with a prior that most countries believed that closeness fosters growth is consistent with the findings of **?** and the fact that only about one forth of the countries are open in 1950. The value of ω is set to 0.025, which implies a quite agnostic view about the mean of initial beliefs.

- We choose a and b such that ρ_i has an a-priori mean equal to 0 and a standard deviation equal to 0.4. Overall, this prior is very diffuse and does not play much of a role for the results of the paper.[9]
- We select s and d so that $\sigma_{i,k}$ has an a-priori mean and standard deviation equal to 0.01. The idea here is trying to discourage the model from fitting the data using very large variances of the exogenous political cost K_{it}. This prior distribution implies that, if policymakers believe that growth under openness is 1% higher than under closeness, they will open the country to free trade with probability 87% on average (standard deviation 10%).
- As the coefficients γ are common to all countries, we use a flat prior for γ.

An important parameter for the evolution of beliefs is ν^2, i.e. policymakers' initial uncertainty about the effects of openness and closeness on economic growth. This parameter is important because it affects the speed of learning, especially for those countries for which fewer data are available. In our baseline estimation we set this parameter to a fixed value. This is because weak identification makes it very hard to estimate simultaneously ν^2 and ρ in (4.3). Since the SW indicator is not available before 1950, it is much easier to come up with a reasonable value for ν^2 rather than for ρ. In calibrating ν^2 we first observe that $\sigma_i^2 \cdot \nu^2$ should be approximately equal to $var(\bar{y}_i)$,[10] i.e. the variance of the average growth rate of GDP. We obtain an estimate of $var(\bar{y}_i) = 0.0175^2$ as the variance of the average growth rates of the countries present in the Maddison dataset between 1901 and 1950 (excluding the wars).[11] To obtain an estimate of σ_i^2 based on pre-sample observations, we use again the Maddison data and run a regression of GDP growth on time and fixed effects. We then compute the variance of the residuals for each country and calculate the mean of these variances (which equals 0.0044). Therefore, we set $\nu^2 = \frac{(0.0175)^2}{0.0044} = 0.0696$. Given the potential importance of this parameter,

[9] We use a beta prior on $\frac{\rho_i+1}{2}$ as opposed to ρ_i directly because the support of the beta distribution is the $[0, 1]$ interval.

[10] This can be seen by combining (4.2) and (4.3) and noticing that we cannot distinguish between open and closed countries in the pre 1950 data.

[11] There is a huge outlier in the distribution of the average growth rates across countries. Therefore, this variance is estimated with a robust method (squared average distance from the median of the 16 and 84th percentiles).

in section 8 we show the robustness of our results to this choices. In particular, we show that a procedure based on estimating ν^2 and fixing ρ delivers similar results.

5.2. **The likelihood function.** In order to derive the posterior distribution of the unknown coefficients, we update these priors with the likelihood information. If we denote by D^s the available data up to a generic time s, the likelihood function can be written as the following product of conditional densities

$$\mathcal{L}(D^T|\alpha) = C \cdot \prod_{i=2}^{N} \left[\mathcal{L}(\theta_{i,1}|\alpha) \cdot \prod_{t=2}^{T} \mathcal{L}(\theta_{i,t}|D^{t-1}, \alpha) \right],$$

where C is a constant which does not depend on α. In turn, the conditional density $\mathcal{L}(\theta_{i,t}|D^{t-1}, \alpha)$ can be written as

$$\mathcal{L}(\theta_{i,t}|D^{t-1}, \alpha) = \Phi\left(\frac{\hat{\beta}^o_{i,t-1} - \hat{\beta}^c_{i,t-1}}{\sigma_{i,k}}\right)^{1(\theta_{i,t}=1)} \cdot \left(1 - \Phi\left(\frac{\hat{\beta}^o_{i,t-1} - \hat{\beta}^c_{i,t-1}}{\sigma_{i,k}}\right)\right)^{1-1(\theta_{i,t}=1)},$$

where $\Phi(\cdot)$ denotes the cdf of a standard Gaussian density. These results are derived in appendix D.

6. Results

This section presents the estimation results and various measures of fit for our baseline specification of the model. Extensions of the baseline model and robustness checks are presented in section 8.

6.1. **Estimation results: the weighting function.** In our baseline specification, the weight that country i assigns to the data of country j (w_{ij}) is a function of a constant and two additional variables: d_{ij}, physical distance (in thousands of Km) between the capital of country j and country i, and ℓ_{ij}, a dummy variable equal one if countries i and j have the same official language. In other words, $z_{ij} = [1, d_{ij}, \ell_{ij}]$ in expression (4.7). Table 1 reports the estimates of the corresponding coefficients γ:

	Posterior mode (Posterior std)
constant	-0.5182 (0.0382)
d_{ij}	-0.4725 (0.0245)
ℓ_{ij}	1.0418 (0.0719)
log-likelihood	-817.0118

TABLE 1: *Estimates of the coefficients of the weighting function in the baseline model.*

Everything else equal, countries put more weight on data from countries nearby. Figure 4 plots the weight that country i puts on country j (w_{ij}) as a function of distance (d_{ij}). For instance, the weight that a country would put on data from another country distant 2000Km is approximately one third of the weight on own data. The other conclusion that can be drawn from table 1 is that language matters: speaking the same language increases the weight by a factor larger than two.

With this information in hand we first ask: *Is learning globalized across the countries of the world?* The answer is *no*. Learning appears to be quite localized instead. Figure 5 plots the histogram of $\frac{\{\sum_{j\neq i} w_{ij}\}_{i=1}^{N}}{w_{ii}}$, i.e. the total weight assigned by each country to the rest of the world, relative to the weight assigned to their own data. There are essentially no countries that weight equally the data from all the other countries. However, learning is not isolated either, as the weight put on data from other countries is substantial: there are very few cases (only three cases) in which own data receive more weight than the rest of the world.

The weighting function plays an important role in our analysis. In the next sections we will analyze its contribution to the evolution of policymakers' beliefs

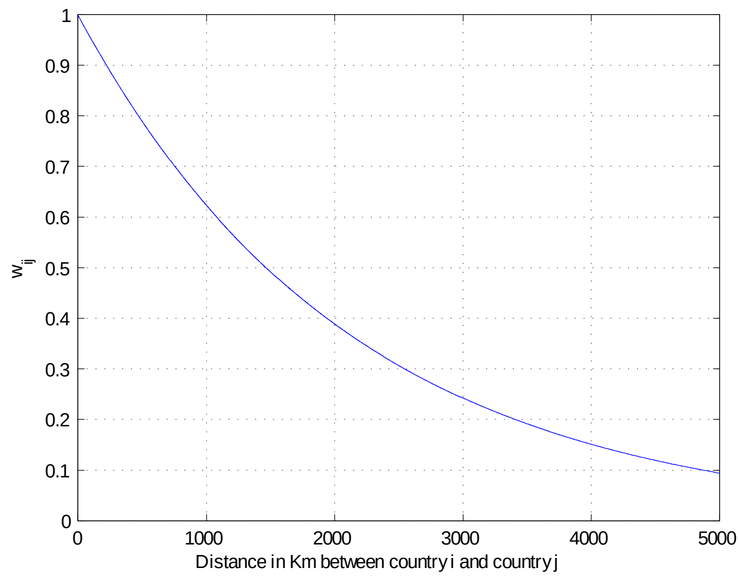

FIGURE 4. Weights on data from other countries (relative to own country) as a function of distance in Km.

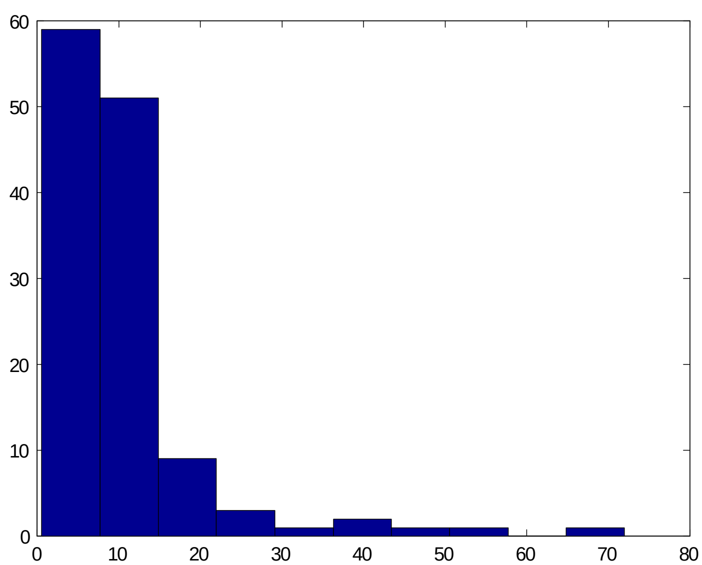

FIGURE 5. Histogram of total weight put on other countries relative to own country.

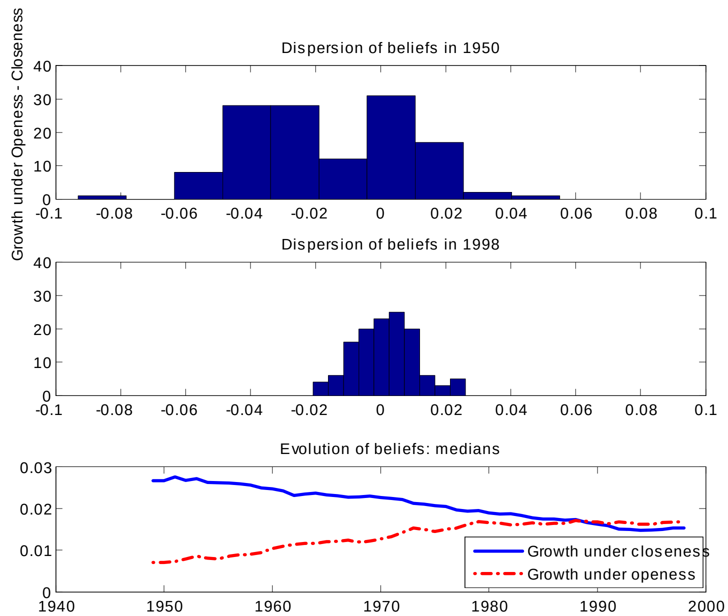

FIGURE 6. Estimates of policymakers' beliefs: (a) histogram of the posterior mode of the difference of expected growth under openness and closeness in 1950; (b) histogram of the posterior mode of the difference of expected growth under openness and closeness in 1998; (c) evolution of the posterior mode of expected growth under closeness and openness for the median country.

about the effect of trade openness on growth. Moreover, in section 8, we will analyze the robustness of our results to alternative specifications of this weighting function.

6.2. **Estimation results: evolution of beliefs.** Figure 6 presents a summary of the estimated evolution of beliefs over time and across countries. Figure 6a plots an histogram of $\left\{\hat{\beta}^{o*}_{i,0} - \hat{\beta}^{c*}_{i,0}\right\}_{i=1}^{N}$, i.e. the difference between the posterior mode of the mean of policymakers' prior beliefs about the effect of openness and closeness on growth. Notice that initial beliefs were quite negative on openness and characterized by considerable dispersion across countries. According to our estimates, in 1950 about the 80% of the countries believed that trade openness had a negative effect on economic growth. This is consistent with the findings in **?**.

Figure 6b shows that, by 1998, beliefs have shifted quite considerably. By then, the histogram of implied $\left\{\hat{\beta}^{o*}_{i,T} - \hat{\beta}^{c*}_{i,T}\right\}_{i=1}^{N}$ has moved to the right and a perceptible majority of countries believe openness to be growth enhancing. However, quite interestingly, the dispersion of beliefs across countries has declined but certainly not disappeared.

Figure 6*c* provides a time series perspective on the evolution of beliefs, by plotting expected growth under openness and closeness in the median country. As anticipated, the median country started off with beliefs biased toward closeness and slowly has shifted towards favoring free trade regimes. By 1998, the median beliefs are slightly in favor of openness.

We now ask the important question: *Has the world reached a point in which the vast majority of policymakers are convinced that free trade is beneficial for economic growth?* Figure 6*b* and *c* tell us that this is certainly not the case. There still exists a considerable amount of negative views about openness, and many of those with a favorable view would not believe that the gains of openness to be quantitatively large.

6.3. **The model's fit.** Before reporting our counterfactual experiments, it is important to assess how well our model fits the data. Using different criteria, we will argue that the model explains quite well the dynamics of trade policies over time and across countries.

Figure 7 presents an histogram of $\left\{\sigma^*_{i,k}\right\}_{i=1}^{N}$, i.e. the posterior mode of the standard deviation of the political cost of staying open for the 128 countries of our postwar sample. Notice that the vast majority of countries are characterized by a value of $\sigma^*_{i,k}$ in the order of 0.5%, which is very low. To be fair, the estimates of $\sigma_{i,k}$ are within the range of values that are plausible according to our prior. While this indicates that identification might be weak in some instances, it also suggests that large political costs are essentially never needed to fit the data on the evolution of policies.

Figure 8 presents an histogram of $\left\{\frac{\sigma^*_{i,k}}{\sqrt{Var(y_{i,t})}}\right\}_{i=1}^{N}$, i.e. the ratio between the estimated $\sigma_{i,k}$ and the standard deviation of the growth rate of GDP in each country. Figure 8 makes the point that in essentially every case the variability of the exogenous political cost necessary to explain the dynamics of policies is substantially lower (on average 10 times lower) than the typical variability of GDP growth in the same country.

Another measure of model fit can be obtained by the following counterfactual experiment. Suppose that the political cost variable is zero at each point in time

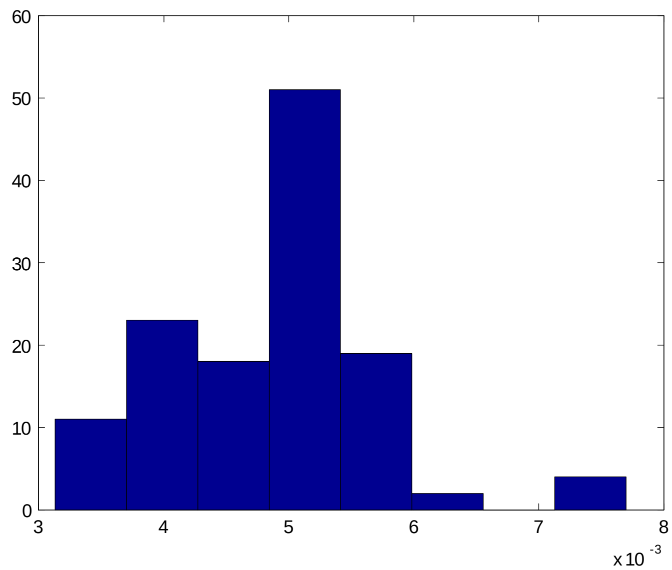

FIGURE 7. Histogram of $\left\{\sigma^*_{i,k}\right\}_{i=1}^{N}$, i.e. the posterior modes of the standard deviations of polititical cost in each country.

and for every country. Given the estimates of the model's parameters, how many times would the model predict the policy actually observed in the data? The answer to this question is quite striking: the model predicts correctly 92% of the policies adopted by the countries. Figure 9 plots an histogram of the share of wrong model predictions across countries. For most countries, the model predicts the right policy all (or almost all) the time. While for a few countries the number of wrong predictions is larger than 10%, in very few cases this number is larger than 40%. In our sample this is true for only 2 countries: $Jamaica$ and $Morocco$.

We conclude this section by comparing the fit of our baseline model to two alternative specifications. First, we assume that countries only learn from their own past data ($\mathbf{M}_{own}$). Second, we assume that countries learn globally, putting the same weight to the data of every country in the world as its own ($\mathbf{M}_{all}$). Table 2 reports the value of the log-likelihood of our baseline specification and of these two alternative models. There are two things to notice: first, our model dominates

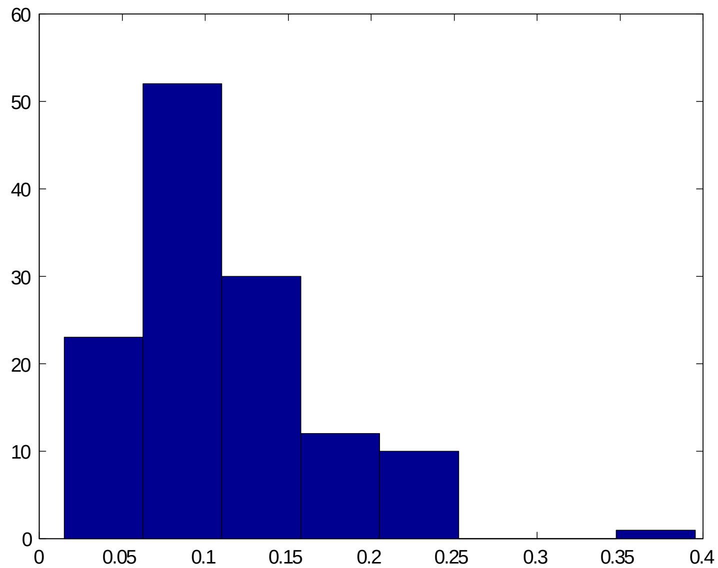

FIGURE 8. Histogram of $\left\{ \frac{\sigma^*_{i,k}}{\sqrt{Var(y_{i,t})}} \right\}_{i=1}^{N}$, i.e. the posterior modes of the standard deviations of polititical cost relative to the standard deviation of GDP growth in each country.

the alternatives;[12] second, among the alternatives, the fit of the autarkic model is substantially better to the opposite extreme.

	Baseline model	$\mathbf{M}_{own}$	$\mathbf{M}_{all}$
log-likelihood	-817.01	$-1,049.13$	$-1,806.2$
share of correct predictions	91.6%	88%	62.2%

TABLE 2: *Measures of fit for the baseline model, a model in which countries learn only from their own past data* ($\mathbf{M}_{own}$) *and a model in which countries put equal weight on every other country of the world* ($\mathbf{M}_{all}$).

[12] Compared to the alternatives, our model has 3 additional parameters. Nevertheless, both a likelihood ratio test or the Bayesian Information Criterion (BIC) would easily reject the restricted models. Of course the formal way of performing a model comparison exercise in a Bayesian framework would be based on the comparison of the marginal data densities. However, the computation of the marginal data density is computationally quite expensive for these models. We leave this extension for future work.

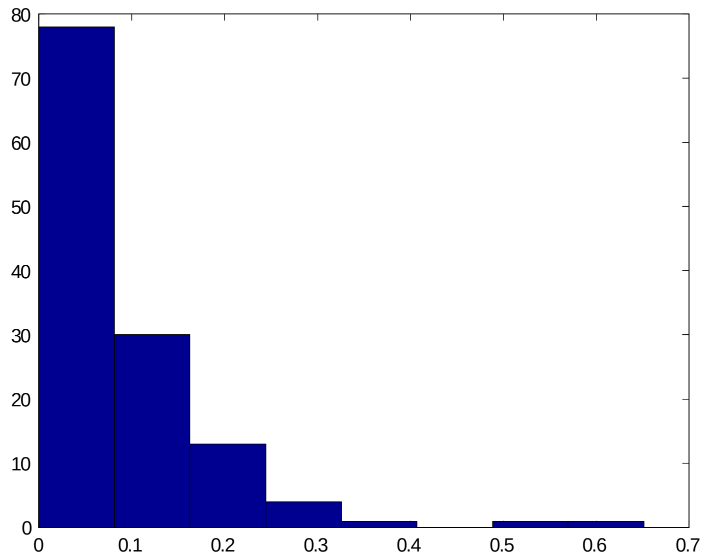

FIGURE 9. Histogram of the shares of wrong model prediction for each country.

7. COUNTERFACTUALS

In this section we answer two questions:

(1) *Do spillovers of information matter for the diffusion of free trade?*
(2) *Would the world revert to protectionism if it was hit by another Great Depression?*

We will argue that the answer to both of these questions is *yes*.

We answer these questions using counterfactual simulations of the model. Suppose that time τ is the starting point of the counterfactuals. These simulations are constructed as follows: based on time τ beliefs and the realization of the exogenous political cost at time $\tau+1$, policymakers choose the value of the policy variable; this policy choice contributes to the realization of the value of GDP growth in period $\tau+1$; a new vintage of data is now available and policymakers form time $\tau+1$ beliefs by updating their priors with the new information; and so on.

Contrary to the rest of the paper, for our counterfactuals we need to postulate a true data generating process for GDP growth. To keep things simple, we simulate the realization of GDP growth in every period using the following stochastic process:

$$\begin{aligned} y_{i,s} &= \beta^c (1-\theta_{i,s}) + \beta^c \theta_{i,s} + f_{i,s} + e_{i,s}, \qquad s = \tau+1, ..., \tau+H \\ e_{i,s} &\sim N\left(0, \varpi_i^2\right), \end{aligned} \tag{7.1}$$

where H denotes the length of the simulation and $f_{i,s}$ is an exogenous forcing variable that allows us, for instance, to simulate the effects of the Great Depression. We obtain specific values for β^c, β^o and $\left\{\varpi_i^2\right\}_{i=1}^N$ by running a simple panel regression of GDP growth on the Sachs and Warner openness indicator. This procedures delivers $\beta^c = 0.0113$, $\beta^o = 0.0269$ and $\left\{\varpi_i^2\right\}_{i=1}^N$ equal to the variances of the residuals. Unless otherwise noticed, the other model parameters (initial beliefs, volatility of the political costs and coefficients of the weighting function) are set to the estimated posterior mode.

Before turning to the description of the results of these experiments, we want to stress that the results of this section are conditional on (7.1), i.e. the particular data generating process that we have chosen for GDP growth. The process in (7.1) has two main advantages. First, it is transparent and easy to cast into our model. Second, it closely resembles the growth regressions that we assume our policymakers estimate to update their beliefs. Therefore, if the forcing variable is set to $f_{i,s} = 0$, then our learning model eventually converges to a self-confirming equilibrium in which everybody knows the truth. On the other hand, in using (7.1), we are making the strong assumption that the residuals of the GDP growth equation ($e_{i,s}$) are uncorrelated with the policy variable ($\theta_{i,s}$). While this is the case in our model,[13] it is possible to imagine situations in which this condition fails. The most obvious of these situations is, for instance, a case in which the shocks to GDP growth ($e_{i,s}$) are not independent from the political cost ($K_{i,s}$).

7.1. **Autarkic and global learning.** Our baseline model allows for a rather flexible specification of the weighting function (4.7). In this subsection, we analyze two restricted versions of the model: one in which countries learn only from their own past experience ($\mathbf{M}_{own}$) and another one in which countries put equal weight on their own past data as well as on the data from all other countries in the world ($\mathbf{M}_{all}$). In particular, we perform counterfactual simulations to understand how beliefs would have evolved under these two alternative scenarios.

[13] Policy decisions are predetermined in our model.

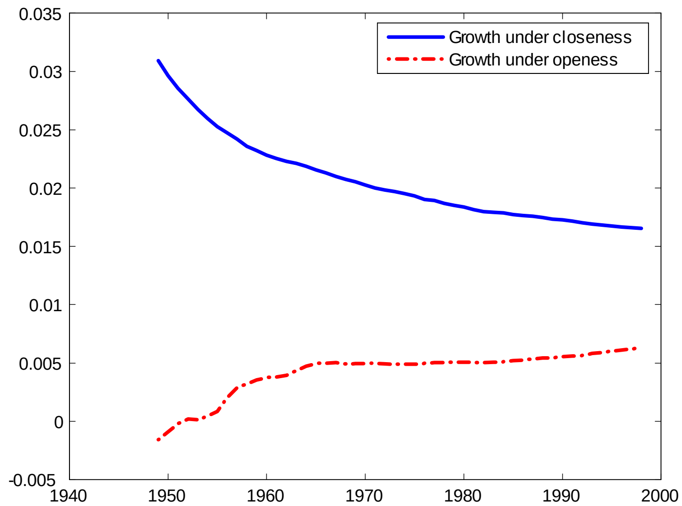

FIGURE 10. Evolution of beliefs (expected growth under openness and closeness in the median country) under the counterfactual scenario in which countries learn only from their own past data.

Figure 10 plots the evolution of median beliefs about the effect of openness and closeness in the $\mathbf{M}_{own}$ model. Compared to the evolution of beliefs estimated in section 6 (see figure 6*c*), learning here is extremely slow, due to the fact that countries do not gain any knowledge from the experience of other countries. It is easy to infer that this learning scheme would have slowed down the diffusion of free trade in the world even more.

On the contrary, if countries had weighted foreign data as much as their own, learning would have been very fast. This point is nicely illustrated in figure 11, which plots the evolution of median beliefs about the effect of openness and closeness under the $\mathbf{M}_{all}$ model. Free trade regimes would have prevailed in most countries already since the mid 1950s.

In sum, we conclude that the localization of learning might have severely slowed down the global diffusion of openness to international trade.

7.2. **Policy reversals: another Great Depression.** Motivated by the historical evidence in section 2.1, we now consider the impact of negative global shocks on the openness of countries. To accomplish this task we also conduct a counterfactual simulation exercise. Contrary to the previous subsection, we keep the estimated

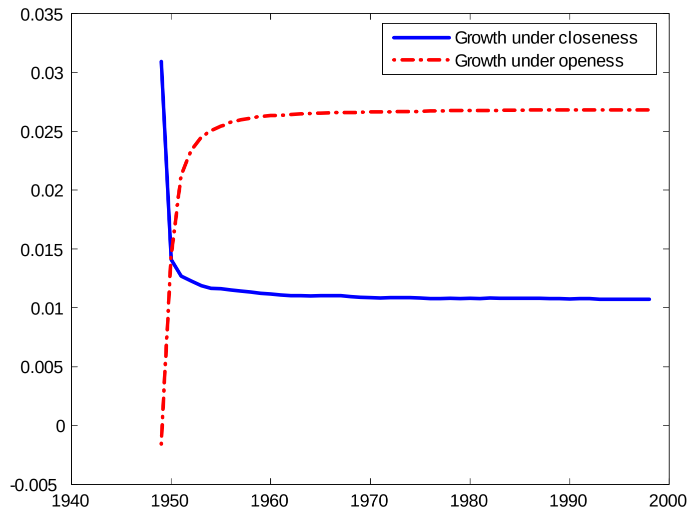

FIGURE 11. Evolution of beliefs (expected growth under openness and closeness in the median country) under the counterfactual scenario in which countries learn putting equal weight on their own past data as well as the data from all other countries in the world.

weighting scheme, but introduce a forcing variable $\{f_{i,s}\}_{s=1}^{H}$ in the growth data generating process (7.1) to induce a global depression. We calibrate $\{f_{i,s}\}_{s=1}^{H}$ to match the size of the Great Depression in the 1930s. Using the Maddison dataset we compute

$$\{f_{i,s}\}_{s=1}^{H} = [-0.0047, -0.0517, -0.0867, -0.0687, 0.0023, 0, ..., 0], \qquad i = 1, ..., N.$$

which corresponds to the average deviation across countries of the growth rate in 1929-1933 relative to the average growth between 1919 and 1928.

Figure 12 plots the impulse response of the proportion of open countries to the global shock. Each line corresponds to a different starting point for the experiment (different τ). For instance, had such global depression hit the world in 1951, when the world was mostly closed, it would have spawn the diffusion of free trade. A deep recession would have cast serious doubts on the growth perspectives of closed economies. The opposite would happen with a global recession in 1998, when most countries are open. The recession would persuade almost 10% of the countries to revert to protectionism and it would take almost forty years for this effect to wash

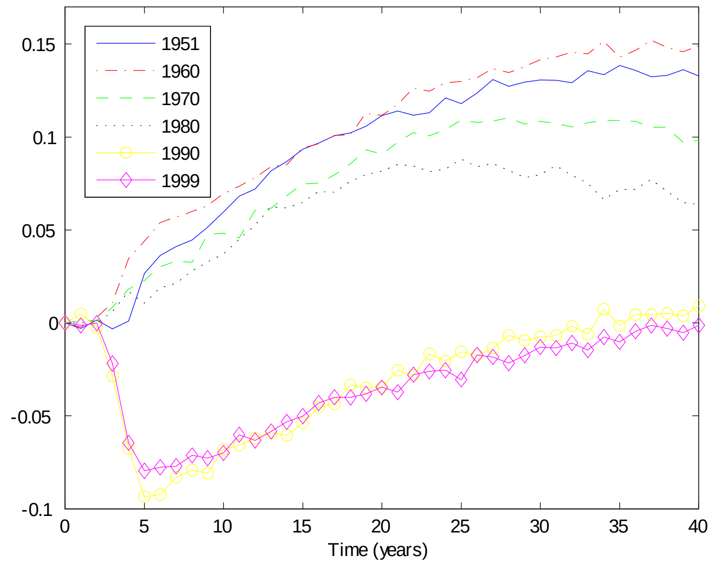

FIGURE 12. Proportion of countries opening to international trade as a conseguence of a counterfactual severe recession hitting at different possible points in time.

out. Of course things would be even worse if for some reason the recession was more severe in open countries relative to closed ones.

Summing up, *would we observe policy reversals towards protectionism if the world was hit by a severe recession?* The answer of our model is *yes*.

8. Robustness Checks

8.1. **Alternative weighting schemes.** In this subsection we examine the robustness of our results to changes in the empirical specification of the countries' weighting function (4.7). In particular, variables related to country size might be important determinants of the weight that a country puts on the experience of others. For instance, larger countries usually receive more attention in the news. To take this kind of considerations into account, table 3 (columns 2 and 3) reports the estimation results of the weighting function, when this is augmented with two new variables: s_{ij}, logarithm of total GDP of country j as a percentage of world GDP, and $a\iota_{ij}$, absolute value of the difference in log-GDP per-capita between country i and country j.

	baseline ($\mathbf{M}_1$)	$\mathbf{M}_2$	$\mathbf{M}_3$	$\mathbf{M}_4$	$\mathbf{M}_5$
constant	-0.5182	-0.5583	-1.2121	-1.0208	
	(0.0382)	(0.0478)	(0.0612)	(0.0681)	
d_{ij}	-0.4725	-0.4635	-0.4854	-0.4225	-0.3749
	(0.0245)	(0.0233)	(0.0228)	(0.0225)	(0.0227)
ℓ_{ij}	1.0418	1.0357	0.2155	0.324	1.4571
	(0.0719)	(0.0682)	(0.0569)	(0.0729)	(0.1556)
s_{ij}		-0.0261	-0.1427		
(total GDP)		(0.0155)	(0.016)		
$a\iota_{ij}$			0.2137		
(GDP per-capita)			(0.0393)		
c_{ij}				-5.082	
				(31.344)	
log-likelihood	-817.0118	-814.2882	-756.76	-792.4437	-833.63

TABLE 3: *Estimates of the coefficients of the weighting function in the baseline model and various alternative models (described in sections 8.1 and 8.2).*

Observe that the inclusion of log-total GDP (s_{ij}) alone leaves the results essentially unchanged. However, the fit improves substantially when the weighting function is augmented to include income differences between countries ($a\iota_{ij}$). In particular, notice that the coefficient on $a\iota_{ij}$ is positive, indicating that policymakers find more informative the experience of countries with a different level of development. This is probably due to the desire to imitate more developed countries and stay away from the negative experiences of less developed countries. Finally, observe that the coefficient on s_{ij} is negative, perhaps contrary to intuition. In any case, we will see below that the quantitative effect is rather small.

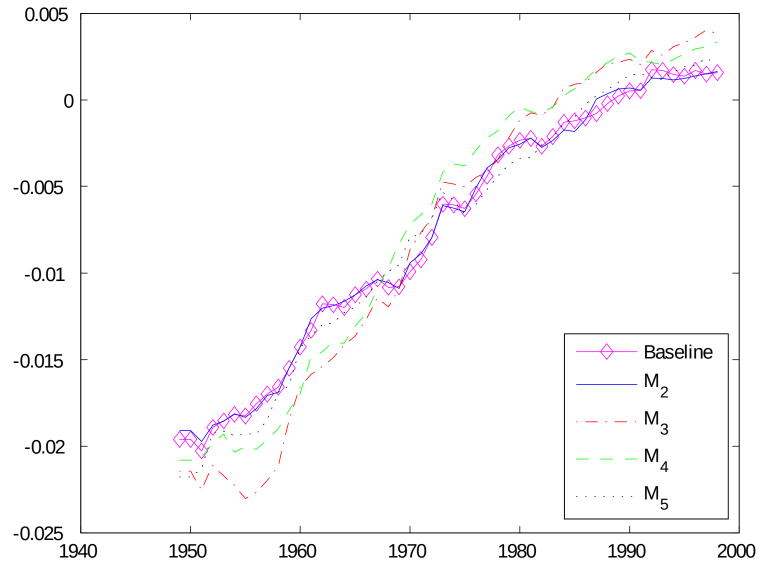

FIGURE 13. Evolution of beliefs (difference between expected growth under openness and closeness in the median country) implied by different models.

Another concern might be that colonizers receive a disproportionately large or small weight from previous colonies. Column 4 in Table 3 reports the estimation results using the new variable c_{ij}, a dummy variable equal one if country j has been a colonizer of country i. Notice that there is an improvement in fit with respect to the baseline specification and that the coefficient on distance remains basically unchanged. However, given the correlation between language and colonial links, the coefficient on language becomes smaller. Finally, notice that the weight depends negatively on having been a colonizer, although this coefficient is far from being statistically significant.

To understand the economic differences of alternative models, we now examine the evolution of beliefs implied by the different models. In particular, Figure 13 plots the evolution of the difference between median expected growth under openness and closeness. All the models tell essentially the same story about the evolution of beliefs.

Finally, figure 14 analyzes the probability of policy reversals after a severe recession. As in section 7.2, we simulate the effect of another Great Depression

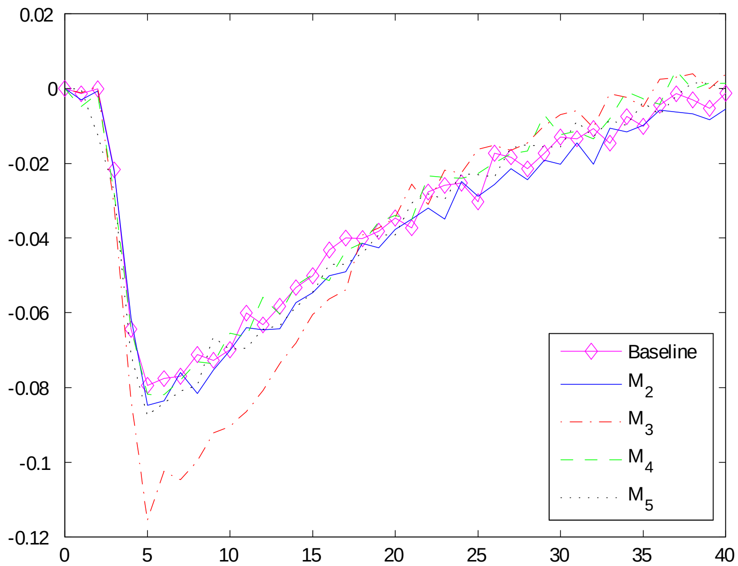

FIGURE 14. Proportion of countries opening to international trade predicted by different models, as a conseguence of a counterfactual severe recession hitting in 1998.

happening in 1998. Models $\mathbf{M}_2$, $\mathbf{M}_4$ and $\mathbf{M}_5$ (to be explained below) practically coincide with the baseline model. If anything, model $\mathbf{M}_3$ (the best fitting model) implies an even more pronounced movement to closeness.

8.2. **Estimating the degree of initial uncertainty.** In our baseline estimation, we have taken the uncertainty about initial beliefs (ν^2 in (4.3)) as given and estimated instead the correlation coefficient of initial beliefs in each country (ρ_i). This choice was motivated by weak identification and by the fact that ν^2 is easier to calibrate than ρ. To verify that the results are robust we take the opposite route: we estimate a different ν_i for each country and fix ρ to zero. In estimating ν_i we use an IG prior density with mean and standard deviation equal to $\sqrt{0.0696}$, which is the value we used to calibrate ν in the baseline exercise. Finally, observe that we eliminate the constant from the vector of covariates z_{ij} because it is not separately identified from ν_i.

The coefficient estimates of the weighting function and the model's fit are reported as $\mathbf{M}_5$ in Table 3. The value of the log-likelihood indicates that the baseline model is preferred. Notice also that the coefficients show some changes with respect to the baseline specification, but these changes have very moderate economic consequences. In fact, figure 13 and 14 show the similar implications of the model for the behavior of policymakers' beliefs and the consequences of a counterfactual severe recession.

We conclude that our results are robust.

9. Concluding Remarks

[To be written]

Appendix A. A Comparison Between Sachs and Warner (1995) and Hiscox and Kastner (2007) Openness Indicators

Hiscox and Kastner (2007, hereafter HK) have recently used bilateral trade flows to obtain estimates of the impact of policy distortions of in each country over time. They posit a basic gravity equation of the form

$$\log\left(\frac{M_{ijt}}{Y_{it}}\right) = \alpha_{it} + \beta \log Y_{jt} - \delta d_{ij} + u_{ijt},$$

where M_{ijt} are the imports of country i from country j in year t; Y_{it} and Y_{jt} denote GDP of the importing and exporting country, respectively; d_{ij} is the "distance" between importing country i and exporting country j and u_{ijt} is white noise error. Here, α_{it} is an intercept that varies across importing countries and across time but not across exporters to that country. HK use these estimated country-year dummy variables as an indicator of the openness to trade of countries. A large value of α_{it} indicates that the country imports more than predicted gravity and hence is more open to trade than the average of the countries-year. HK normalize the values of α_{it} as a fraction of the highest value in the sample, which for is Belgium in 1980. Such normalization is clearly arbitrary but helps indicate how far are the countries from a free-trade benchmark.[14]

The HK indicator (hereafter I_{HK}) is available for 82 countries for the entire sample of years between 1960 and 2000. As HK show, their indicator moves closely in line with well-documented liberalizations and provides a reasonable ranking across countries.

[14] They actually normalize the series by looking at positive percentual difference with the maximum, i.e. $100 \times \frac{\alpha_{\max}-\alpha_{it}}{\alpha_{\max}}$ to obtain a measure of barriers to trade to compare with actual collection of import duties. We use insted $\frac{\alpha_{it}}{\alpha_{\max}}$ to obtain a measure of openness easier to compare with SW.

Interestingly, I_{HK} and the SW indicator (hereafter I_{SW}) are highly correlated. Combining these datasets, we have information for both indicators for 73 countries for most of the years between 1960 and 1998.[15] For this sample, the simple correlation between these two indicators is 0.68, a very high number since it captures the correlation between a dummy and a continuous variable.

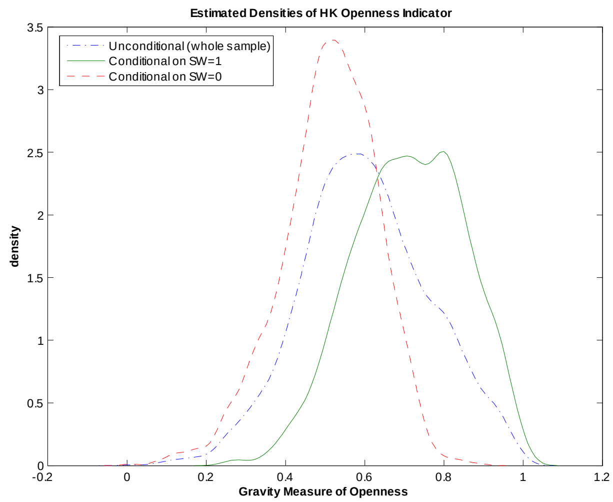

FIGURE 15. Densities of the HK gravity-based openness measure. Whole sample and conditional on the value of Sachs-Warner openness value.

I_{SW} captures much of the variation in I_{HK}. The sample average of I_{HK} is 60.2%, but is 71.2% for the country-years for which $I_{SW} = 1$, while it is only 51.1% when $I_{SW} = 0$. Under the interpretation of HK, openness in the sense of SW on average increases the ratio of imports-to-GDP by more than 20%, which is a value also higher than the standard deviation of 16.2% of I_{HK}. Figure 15 displays the density of values of I_{HK} for the entire sample and the conditional densities for $I_{SW} = 1$ and $I_{SW} = 0$, i.e. for the country-year values for which I_{SW} assumes either value. An overwhelming fraction of the lower tail of I_{HK} takes place when $I_{SW} = 0$ while an overwhelming fraction of the high realizations takes place when $I_{SW} = 1$. Indeed,

[15] The countries, by region, are: **North and Central America(10 countries):** USA, CAN, HTI, DOM, MEX, GTM, HND, SLV, NIC, CRI. **South America(10 countries):** COL, VEN, ECU, PER, BRA, BOL, PRY, CHL, ARG, URY. **Western Europe (15 countries):** GBR, IRL, NLD, BEL, FRA, CHE, ESP, PRT, AUT, ITA, GRC, FIN, SWE, NOR, DNK. **Sub-Sahara Africa(20 countries):** MLI, SEN, BEN, MRT, NER, CIV, GIN, BFA, GHA, TGO, CMR, NGA, GAB, CAF, TCD, COG, ZAR, ETH, ZAF, MDG. **Middle East and North Africa(7 countries):** MAR, TUN, IRN, TUR, EGY, JOD, ISR. **Asia and Pacific(11 countries):** CHN, KOR, JPN, IND, PAK, LKA, THA, PHL, IDN, AUS, NZL.

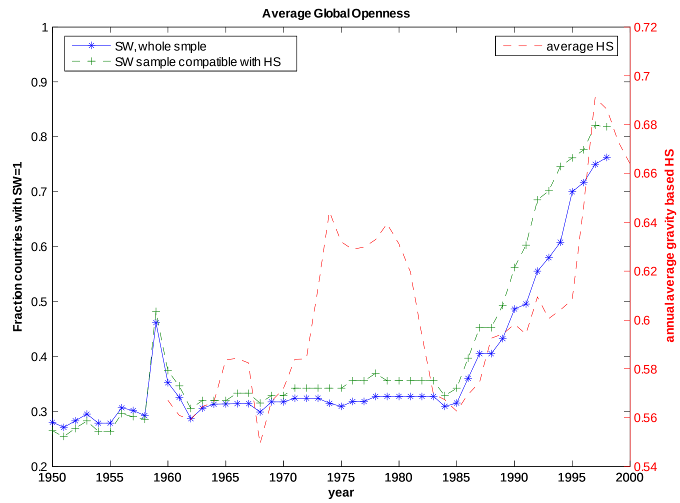

FIGURE 16. Fraction of open countries in the world and annual averages of the HS (gravity based) openness indicator.

if we look at the country-years with the lowest 33% values of I_{HK}, then 87.1% of those country-years have $I_{SW} = 0$, while if we look at the highest 33%, 84.4% of them have a $I_{SW} = 1$.

In sum, despite their very different nature and construction, I_{SW} and I_{HK} appear to largely convey the same information with respect to openness. For instance, it is widely held that, after a number or reversals, the second half of the 20th century is a period in which the world advanced towards greater openness (e.g. **?** and **?**). Figure 16 shows, for each year, the fraction of open countries, i.e. $I_{SW} = 1$. This is done using the entire sample of 128 countries for which we have I_{SW} data, and also for the reduced sample of 73 countries for which we can combine it with I_{HK} data. The figure shows that the two I_{HK} series display essentially the same trends. The higher number in the restricted sample is the result that countries for which data were missing are typically transition or less developed countries, which tend to be initially closed. Obviously, the nature of these series is very different and the magnitudes of change are not comparable.

According to SW, initially, in 1950, less than 30% of the countries were open but by the end, in 1998, almost 80% were open. The spike in late 1950s reflects the introduction of the European Union that led those countries to open up (at least among themselves and to manufacturing goods). The early 1960s is a period in which some developing countries, specially Latin American countries, close down as part of the import substitution program. From 1963 to 1984, the fraction of open countries remains practically flat at around 32%. The year 1985 seems to be the commencement of the movement to global openness that carries out to the end of the sample. Interestingly the gravity-based I_{HK} (scaled in the right-side axis) tracks these episodes with surprising precision whenever there is data. The I_{HK} also captures the global macro instability of the 1970s and early 1980s, which is not captured by I_{SW} (most volatile countries in that period were closed).[16]

Appendix B. Probit Estimates

This appendix describes in detail the reduced-form probit model we refer to at the end of section 3.2.

We use the SW indicator and the per-capita GDP growth data for the entire sample of 128 countries to examine how past growth affects policy choices. Specifically, we model the trade policy of country i in period t as a function of its past policy choice, $\theta_{i,t-1}$, the distance weighted number of open countries in the previous period, $\bar{\theta}_{i,t-1}$, and measures of past performance of the two regimes, $\hat{E}_{i,t-1}(y|\theta=1)$ and $\hat{E}_{i,t-1}(y|\theta=0)$. The past performance of a regime (as perceived by country i in period t) is measured by the distance weighted average growth rate over the previous 3 years of countries that adopted this regime.

We consider the following reduced-form Probit model describing the evolution of trade policies

(B.1)

$$\Pr(\theta_{i,t}=1) = \Phi\left[\phi_1\theta_{i,t-1} + \phi_2\bar{\theta}_{i,t-1} + \phi_3\hat{E}_{i,t-1}(y|\theta=1) + \phi_4\hat{E}_{i,t-1}(y|\theta=0)\right]$$

where $\Phi(\cdot)$ denotes the CDF of a standard Gaussian density. Formally, we define

$$\bar{\theta}_{i,t-1} = \sum_{j\neq i} e^{-\frac{d_{ij}}{\delta}}\theta_{j,t-1},$$

$$\hat{E}_{i,t-1}(y|\theta=1) = \sum_{s=1}^{3}\sum_{j:\theta_{j,t-s}=1} e^{-\frac{d_{ij}}{\delta}} y_{j,t-s} \Big/ \left(\sum_{s=1}^{3}\sum_{j:\theta_{j,t-s}=1} e^{-\frac{d_{ij}}{\delta}}\right),$$

and

$$\hat{E}_{i,t-1}(y|\theta=0) = \sum_{s=1}^{3}\sum_{j:\theta_{j,t-s}=0} e^{-\frac{d_{ij}}{\delta}} y_{j,t-s} \Big/ \left(\sum_{s=1}^{3}\sum_{j:\theta_{j,t-s}=0} e^{-\frac{d_{ij}}{\delta}}\right).$$

[16] For this period the I_{HK} indicator could be slightly misleading. For instance, an sudden increase in imports induced by soaring oil prices would erroneously led to conclude that an important fraction of countries to have openned up.

The estimation results of various specifications of equation (B.1) are presented in table 4. We report the average marginal effects for close countries in the first 15 years of the sample. All specifications include fixed country effects and 7 time dummies.

Three features of the data come up in table 4. First, policies are very persistent (first row). The probability that a country who was opened in period $t-1$ is opened in period t is from 67 to 75 percentage points larger than that of countries that were closed in period $t-1$. Second, trade policies are spatially correlated (second row). For each additional (distance weighted) country that is open in the neighborhood of country i, the probability that country i is opened increases (in the long run) by approximately 3 percentage points.[17] Finally, past performance of trade regimes is highly associated with choices of trade policies (third and forth rows). For each additional point of per-capita GDP growth of open (close) countries in the neighborhood of country i, the probability that country i is open increases (decreases) by approximately 9 (11) percentage points in the long run. The effect of past performance of regimes remain large and significant even after controlling for the number of open countries, which measures reasons for conformity other than learning (column 4).

[17] At the beggining of the sample, the average country had 4.8 (with a standard deviation of 1.6) opened countries in its neighbourhood. By the end of the sample this number rose to 26.2 (with a standard deviation of 7.8).

	1	2	3	4
$\theta_{i,t-1}$	0.75 (0.02)	0.67 (0.04)	0.75 (0.03)	0.67 (0.05)
$\bar{\theta}_{i,t-1}$	...	0.01 (0.005)	...	0.01 (0.003)
$\hat{E}_{i,t-1}[y\|\theta=1]$	...	...	2.21 (0.83)	1.88 (0.61)
$\hat{E}_{i,t-1}[y\|\theta=0]$	...	...	-2.75 (1.11)	-1.42 (0.53)
δ	...	4596 (72.19)	4622 (73.80)	4375 (73.36)

TABLE 4: *Estimation results of reduced-form probit model*

Appendix C. Updating formulas

This appendix shows how to use the assumptions in section 4.2 to derive optimal (in the sense of Bayes) updating formulas for $\hat{\beta}_{i,t} \equiv E_{i,t}\left([\beta_i^c, \beta_i^o]'\right)$. The calculations are conducted from the perspective of policymakers of country i.

First, define the vector of regressors $x_{j,s} \equiv [1-\theta_{j,s}, \theta_{j,s}]'$ and slope coefficients $\beta_{j|i} \equiv \left[\beta_{j|i}^c, \beta_{j|i}^o\right]'$, and rewrite equation (4.4) as

$$y_{j,s} = x_{j,s}'\beta_{j|i} + \varepsilon_{j|i,s}, \qquad j = 1, ..., N. \tag{C.1}$$

We can now substitute equations (4.5) and (4.6) into (C.1). We obtain

$$y_{j,s} = x_{j,s}'\beta_i + \tilde{\varepsilon}_{j|i,s}, \qquad j = 1, ..., N, \tag{C.2}$$

where

$$\text{var}\left(\tilde{\varepsilon}_{j|i,s}\right) = \sigma_j^2\left(1 + q_{ij}\right), \qquad j = 1, ..., N.$$

Finally, rewrite (C.2) as

$$y_{j,s} = x'_{j,s}\beta_i + \frac{1}{w_{ij}}\varepsilon^*_{j|i,s}, \qquad j = 1, ..., N \tag{C.3}$$

where $\varepsilon^*_{j|i,s} \equiv w_{ij}\tilde{\varepsilon}_{j|i,s}$ and $w_{ij} \equiv \frac{\sigma_i}{\sigma_j}\frac{1}{\sqrt{1+q_{ij}}}$. As $w_{ii} = 1$, equation (C.3) holds for any j. Moreover, notice that $\text{var}\left(\varepsilon^*_{j|i,s}\right) = \sigma_i^2$. The estimation of equation (C.3) corresponds to a weighted least square estimation problem. If the weights $\{w_{ij}\}_{j=1}^N$ are known,[18] it is easy to show that the optimal updating formulas for the expectation of policymakers' beliefs in country i are:

$$\begin{aligned} P_{i,t} &= P_{i,t-1} + X'_t W^2 X_t \\ \hat{\beta}_{i,t} &= P_{i,t}^{-1}\left(P_{i,t-1}\hat{\beta}_{i,t-1} + X'_t W^2 y_t\right), \end{aligned}$$

where $y_t \equiv [y_{1,t}, ..., y_{N,t}]'$, $X_t = [x_{1,t}, ..., x_{N,t}]'$ and $W = \text{diag}([w_{i1}, ..., w_{iN}])$. The recursion is initialized at $\hat{\beta}_{i,0}$ and $P_{i,0}$ which denote the prior mean and precision matrix respectively.

Appendix D. The likelihood function

The likelihood function can be written as a product of conditional densities:

$$\mathcal{L}(D^T|\alpha) = \mathcal{L}(D_1|\alpha)\prod_{t=2}^T \mathcal{L}(D_t|D^{t-1},\alpha). \tag{D.1}$$

Under the assumption that the distribution of the vector $y_t \equiv [y_{1,t}, ..., y_{N,t}]'$ depends on α only through the vector $\theta_t \equiv [\theta_{1,t}, ..., \theta_{N,t}]'$, it follows that

$$\mathcal{L}(D_t|D^{t-1},\alpha) = \mathcal{L}(y_t|\theta_t)\mathcal{L}(\theta_t|D^{t-1},\alpha) \tag{D.2}$$

$$= \mathcal{L}(y_t|\theta_t)\prod_{i=1}^N \mathcal{L}(\theta_{i,t}|D^{t-1},\alpha) \tag{D.3}$$

Combining (D.1) and (D.3), we obtain the following result:

$$\mathcal{L}(D^T|\alpha) = C\cdot\prod_{i=2}^N\left[\mathcal{L}(\theta_{i,1}|\alpha)\cdot\prod_{t=2}^T\mathcal{L}(\theta_{i,t}|D^{t-1},\alpha)\right],$$

where C is a constant which does not depend on α.

Since the policy decision is given by

$$\theta_{i,t} = 1\left(E_i\left(\beta_i^o|D^{t-1}\right) - E_i\left(\beta_i^c|D^{t-1}\right) > K_{i,t}\right), \quad i = 1, ...N,$$

[18] This is not a trivial assumption. In practice we need to assume that policimakers of country i either know $\{\sigma_j\}_{j=1}^N$ or think that $\sigma_j = \sigma_i$ for every j.

it follows that

$$\begin{aligned}\Pr\left(\theta_{i,t}=1|D^{t-1},\alpha\right) &= \Pr\left(K_{i,t}<E_i\left(\beta_i^o|D^{t-1}\right)-E_i\left(\beta_i^c|D^{t-1}\right)\right)= \\ &= \Phi\left(\frac{\hat{\beta}_{i,t-1}^o-\hat{\beta}_{i,t-1}^c}{\sigma_{i,k}}\right).\end{aligned}$$

This implies

$$\mathcal{L}(\theta_{i,t}|D^{t-1},\alpha)=\Phi\left(\frac{\hat{\beta}_{i,t-1}^o-\hat{\beta}_{i,t-1}^c}{\sigma_{i,k}}\right)^{1(\theta_{i,t}=1)}\cdot\left(1-\Phi\left(\frac{\hat{\beta}_{i,t-1}^o-\hat{\beta}_{i,t-1}^c}{\sigma_{i,k}}\right)\right)^{1-1(\theta_{i,t}=1)},$$

where $\Phi(\cdot)$ denotes the cdf of a standard Gaussian density.

Northwestern University

Northwestern University

Northwestern University, NBER and CEPR