Duernecker, Georg; Meyer, Moritz; Vega-Redondo, Fernando

Working Paper

Trade Openness and Growth: A Network-Based Approach

CESifo Working Paper, No. 9319

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Duernecker, Georg; Meyer, Moritz; Vega-Redondo, Fernando (2021) : Trade Openness and Growth: A Network-Based Approach, CESifo Working Paper, No. 9319, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/245500

CESifo WORKING PAPERS

9319
2021

September 2021

Trade Openness and Growth: A Network-Based Approach

Georg Duernecker, Moritz Meyer, Fernando Vega-Redondo

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9319*

Trade Openness and Growth: A Network-Based Approach

Abstract

In this paper, we propose a novel approach to the study of international trade that leads to a measure of country openness that is quite different from the various alternatives proposed by the received literature. In contrast to these, our measure does not use indicators of aggregate trade intensity, trade policy, or trade restrictiveness but relies on a broad systemic viewpoint on the effects of trade. More specifically, it goes beyond direct trade connections and measures a country's level of integration in the world economy through the full architecture of its second, third, and all other higher-order connections in the world trade network. We apply our methodology to a sample of 204 countries spanning the period from 1962 to 2016 and perform a Bayesian analysis of model selection to identify the most important correlates of growth. The analysis finds that there is a sizable and significant positive relationship between our integration measure and a country's rate of growth, while that of the aforementioned traditional measures of outward orientation is only minor and statistically insignificant. We perform several sensitivity checks and conclude that our baseline findings are very robust to either different data sets or alternative variations of the integration measure. Overall, this suggests that a network-based approach to measuring country openness may provide a valuable perspective on economic growth.

Keywords: globalization, trade integration, economic growth, network analysis, dynamic panel model, Bayesian model averaging.

Georg Duernecker
Goethe University Frankfurt / Germany
duernecker@em.uni-frankfurt.de

Moritz Meyer
World Bank, Washington DC / USA
meyer.moritz@gmail.com

Fernando Vega-Redondo
Bocconi University, Milan / Italy
fernando.vega@unibocconi.it

July 1, 2021
An earlier version of this paper was circulated under the title "The Network Origins of Economic Growth".

1 Introduction

What are the main factors that underlie the sharply contrasting growth performance of different countries? This question has generated a heated controversy in the literature on economic growth over the last decades. Much of the analysis that has been undertaken – both theoretically and empirically – to study the issue has revolved around the identification of some appropriate measure of openness to international trade. In a nutshell, the main objective of this paper is to revisit the question by adopting a *dynamic* and *systemic* (network-based) viewpoint to assessing the extent to which a country is effectively open to (or integrated in) the world market. And by doing so, we shall argue, new and valuable insights arise.

Starting with the early work of Baumol (1986) and Barro (1991), the empirical growth literature has devoted countless efforts to the question of whether some suitable notion of country openness is one of the essential conditions for fast economic growth. However, despite a long and intense debate on the issue, no shared position has been reached. At some point in the late 1990s, the so-called *Washington Consensus* emerged, holding that greater country openness to international trade leads to faster growth and higher living standards. This view was based on the influential work by Dollar (1992), Sachs and Warner (1995), and Frankel and Romer (1999).[1] Yet, shortly thereafter, the thorough re-investigation of existing evidence undertaken by Rodríguez and Rodrik (2001) largely dissolved that consensus. In particular, these authors argued that, in the absence of a supporting theory, all of the received indicators used to measure a country's outward-orientation did not convincingly capture truly relevant dimensions of economic growth[2]

On the other hand, from a theoretical perspective, it is clear that a fruitful analysis of the relationship between openness and growth needs to move beyond the traditional (largely static) Ricardian framework. A first step in this direction was proposed by Melitz (2003), who applies and intrinsically dynamic viewpoint to study the economic implications of openness. Specifically, Melitz focuses on the role played by international competition in promoting the (market-based) selection of more efficient firms.[3] Subsequent papers, however, such as those Chaney (2014) and Alvarez, Buera, and Lucas (2017), have developed alternative theories that, more in line with ours, highlight the function of cross-border trade flows in transferring economically relevant information. Thus, while Chaney's model posits that international trade provides domestic firms with information that allows them to access to foreign markets, that of Alvarez *et al.* assumes that trade with other countries enhances the technology and operational know-how of domestic firms.

In contrast with the two aforementioned papers, our approach here stresses that a country's access to innovation opportunities depends on how well it is integrated in the *whole* world trade network, as captured by its proximity to other (important) countries in terms of the overall

[1] It is also the position supported by more recent studies such as Dollar and Kraay (2003), Alcala and Ciccone (2004), and Feyrer (2019)

[2] For a review, see Winters (2004), Rodríguez (2007), and Estevadeordal and Taylor (2013).

[3] See Melitz and Redding (2014) for a comprehensive survey and discussion of this literature.

(and evolving) network of trade flows. Our approach, therefore, is both dynamic *and* global in that not only direct trade connections but also *indirect* ones are taken to channel valuable information that generates growth. More specifically, we build on the theory of globalization developed by Duernecker and Vega-Redondo (2018) to formulate an operational model where inter-country connections (and hence knowledge flows) are supported by trade. Thus, under the assumption that foreign ideas are complementary to domestic ones in fueling innovation, a high rate of growth (which can only be sustained by a high rate of innovation) depends on having a rich pattern of direct and indirect trade-based connections to other countries.

To test the model, we use a dynamic-panel data set including 204 countries and spanning the period from 1962 to 2016. On the basis of it we construct the set of evolving matrices of inter-country trade flows that are central to the theory, which in turn allows us to compute the empirical counterpart of the variables that measure trade-based integration for each country. A preliminary study of the data already reveals a number of interesting patterns. One of them is that, across the time span being considered, the world as a whole has become not only more integrated but also more unequal. In particular, we find that while the group of most integrated countries shows a persistent tendency to increase their integration, the majority of less integrated countries stagnates or even displays an opposite trend towards lower integration. An additional interesting finding is that our measure of trade-based integration is essentially uncorrelated with the classical trade-share variable for openness used by the literature. This early observation already provides some support to the notion that a systemic/network perspective to understanding outward orientation is qualitatively distinct from the local one pursued the received literature and policy discussions.

The paper then proceeds with a systematic analysis of the problem by positing an econometric model where the GDP of each country (in log terms) is conceived as a linear projection of this same lagged variable and a number of covariates. The latter include our measure of trade-based integration as well as other 33 variables highlighted by the growth literature as potentially important factors correlated to country growth. These variables comprise the most commonly used openness indicators (such as the Trade Share and the Sachs-Warner Index) along with a wide set of prominent country characteristics (such as, for example, government share, geographical location, population growth, life expectancy, or political indicators).

The consideration of such a wide range of variables poses the important problem of model selection. In other words, it raises the question of what is the right model specification in terms of which the implied conditional correlations must be evaluated. To tackle the problem, we follow the standard methodology known as Bayesian Model Averaging (BMA). This is a procedure that, given the empirical evidence, associates to every possible model specification a posterior probability that it is the correct (or "true") one. Then, on the basis of those model-associated probabilities, we can also readily compute the posterior probabilities that each of the different variables under consideration belongs to the correct model, as well as the corresponding posterior mean estimates of their regression coefficients. In essence, our main conclusion will be that, according to both of the former two covariate-associated criteria, the trade-integration

measure derived from our theory strongly outperforms any other covariate (except for lagged GDP). This, we shall explain, suggests that the correlation between integration and growth is indeed a robust one, even when we account for possible model-selection bias. Such a conclusion is further reinforced through a number of sensitivity checks. For example, we confirm that our baseline findings are robust to the use of different data sets, or to the application of a number of variant formulations of our integration measure (e.g. one that restricts to trade in capital goods alone).

As already indicated, the theory developed in this paper is most closely related to that proposed by Alvarez, Buera, and Lucas (2017). For, as we do, they also assume that trade is the channel through which ideas diffuse. In particular, their model posits that "trade puts domestic producers in contact with (...) foreign and domestic producers from which they can learn and improve their technologies" (see also Buera and Oberfield (2020)). In the empirical realm, on the other hand, there has been quite a broad literature that has tested the implications of this idea in a variety of different contexts and from a diverse set of perspectives. We summarize it very succinctly, concentrating on just a small set of representative papers.

A seminal contribution to the empirical growth literature was provided by Coe and Helpman (1995), who showed – for a sample of the 21 OECD countries plus Israel, during the period 1971-90 – that "foreign R&D has beneficial effects on domestic productivity, and that these are stronger the more open an economy is to foreign trade." This view has been strengthened further by Coe *et al.* (1997) and Coe *et al.* (2009) by confirming that this finding still holds when trade is restricted to just machinery and equipment, and it is also robust to controlling for institutional factors and human capital.[4] Their analysis is naturally related to ours, as we also sustain the notion that trade flows spanning a network across countries transmit knowledge globally. Moreover, as in Coe et al. (1997, 2009), we also measure openness in terms of the (GDP-normalized) import volume.

Our approach, however, adds to the aforementioned papers and the received literature along several dimensions. The key one is that we go well beyond the first-order effects associated to *direct* trade flows. As advanced, this provides a much wider view of the problem, which highlights that the *overall pattern* of international trade is indeed prominently related to countries' performance. We also show that not only trade in capital goods tends to be an important vehicle for knowledge transmission, but that trade in commodities is *not*. In general, at an admittedly high level, our analysis provides a trade-based foundation for the major cross-country spillovers that are well-known to be a key component of technological change (see e.g. Eberhardt *et al.* (2013)). In fact, we find that these effects are not just strong – as found by Etur and Musolesi (2017) – but also highly heterogeneous across countries, as they are largely dependent on the position each country occupies in the overall trade network.[5]

[4]Keller (1998) challenged the validity of the results in Coe and Helpman (1995) by arguing that similar findings can be obtained in an analysis where the links of the observed trade network were replaced by randomly created trade. We address this point in our empirical analysis in Section F.3 were we perform a spurious analysis that considers random permutations of the trade network.

[5]We thank an anonymous referee for suggesting this general formulation of the phenomenon.

The remainder of this paper is structured as follows. Section 2 introduces the theory and explains the measure of country integration we shall use throughout the paper. Section 3 describes the data and presents summary statistics of our globalization index across countries and over time. Section 4 compares our integration measure to other openness measures considered in the literature. In Section 5, we present the econometric model, while Section 6 reports the main results. These results are then discussed in Section 7, with a special focus on understanding the role played by our integration measure. Section 8 discusses our robustness analysis. Section 9 concludes. All supplementary materials are included in Appendices A-I.

2 Theoretical framework

2.1 The model

Our empirical analysis relies heavily on some of the ideas underlying the theory developed by Duernecker and Vega-Redondo (2018), henceforth designated by DV. Here, we provide only a concise description of their model, which is discussed in detail in the working-paper version, Duernecker *et al.* (2020). At the end of this section, we shall also explain in some detail how the present approach differs from that of DV, as well as its contrast with other related models in the literature, such as those proposed by Buera and Oberfield (2016) and Alvarez, Buera, and Lucas (2017).

The model views the world economy as a directed network defined over a fixed set of nodes, $N = \{1, 2, ..., n\}$, each of these conceived as an individual country. Every such country $i \in N$ is populated by a given number of firms that produce the goods exported by this country, with the links across countries representing inter-country trading relationships. More specifically, a link exists from country i to country j if i exports to j. Naturally, we are interested in the intensity of these trade flows, so those links are weighted accordingly. A natural way to represent the inter-country pattern of such weighted links is through an adjacency matrix $A = (a_{ij})_{i,j \in N}$ where each $a_{ij} \geq 0$ measures the export volume from i to j (if $i = j$, the transaction reflects domestic trade). Through appropriate normalization, it is useful to have these flows normalized to add up to one, so that A becomes a row-stochastic matrix.

The setup is dynamic, with time t modeled continuously in $[0, \infty]$, and the state of the system at any t given by the vector $\mathbf{y}(t) = (y_i(t))_{i=1}^n$ that specifies the current GDP $y_i(t)$ of each country i. The evolution of $\mathbf{y}(t)$ depends on the amount (measure) of growth-generating projects $z_i(t)$ that are active in each country $i \in N$. Specifically, we posit the following simple law of motion:

$$\dot{y}_i(t) = \xi \, z_i(t) \qquad\qquad (i = 1, 2, ..., n) \tag{1}$$

for some constant $\xi > 0$. That is, we assume that the growth rate of a country is proportional to the mass of its ongoing projects.

In view of (1), a key step in building the model must be to specify the mechanism by which

the stock of active projects changes over time. In line with DV, we assume that new projects arise through *innovation* while old ones dissipate due to *obsolescence*. These two opposing forces are formulated as follows.

Innovation: At any t, every firm in each country i receives an innovation opportunity at a fixed rate η (formally, with probability $\eta \mathrm{d}t > 0$ for a time interval of infinitesimal length $\mathrm{d}t$). This innovation actually materializes only if the firm is able to access some complementary information (or know-how) that lies somewhere in the world – specifically, information that originates in country j with probability proportional to the economic size of this country, i.e. $y_j(t)$. Then, the question arises of how that information is transferred from j to i. In line with existing literature (both theoretical and empirical)[6] we posit that such transfer is channeled through – or embodied by – trade. More precisely, it is assumed to flow downstream from j to i with a probability proportional to the volume of exports from the former country to the latter. This gives rise to a diffusion process that, mathematically, defines a random walk on the directed export network, with the transition probabilities at each stage being determined by the normalized (i.e. relative) volumes of trade that any exporting country sells to each of its customer countries. In the end, if the network is connected, every piece of information originating in each i arrives to every other j. The *value* of this information, however, is postulated to decrease with the time it takes to arrive, due to a number of possible complementary factors, e.g. obsolescence, noisy transmission, or delay. *Ex ante*, of course, the actual length between each pair of countries is uncertain (i.e. random), so our focus is on the expected time it takes, depending on their position in the overall production network.

Formally, if we denote by $\nu_{ij}(t)$ the rate of new projects actually initiated in country i at t that rely on information available in j, and $\dot{z}_i^+(t)$ stands for i's aggregate (gross) rate of project creation, we can write:

$$\dot{z}_i^+(t) = \sum_{j \neq i} \nu_{ij}(t) = \sum_{j \neq i} \eta \; y_i(t) \; \frac{y_j(t)}{\sum_{k \neq i} y_k(t)} \; f(\varphi_{ji}(A(t))), \tag{2}$$

where

- $\eta \, y_i(t)$ is the rate of innovation opportunities arising in country i at t,
- $\frac{y_j(t)}{\sum_{k \neq i} y_k(t)}$ stands for the probability that the complementary information required to materialize the aforementioned opportunities is available in country j,
- and $f(\varphi_{ji}(A(t)))$ is the decay associated to the expected length $\varphi_{ji}(A(t))$ of the diffusion path from j to i, with $f : \mathbb{R} \to [0, 1]$ being a decreasing function.

Obsolescence: As a countervailing force, we posit that ongoing projects become obsolete and hence are discontinued at a fixed rate λ. Thus, if $\dot{z}_i^-(t)$ denotes the aggregate (gross) rate at

[6] On the theoretical side, the two aforementioned papers, Buera and Oberfield (2016) and Alvarez, Buera, and Lucas (2017), provide good illustrations. Concerning empirical work, on the other hand, we refer to Caselli and Coleman (2001) or Acharya and Keller (2009) for the study of specific cases.

which standing projects at t are terminated in country i, we can write:

$$\dot{z}_i^-(t) = \lambda\, z_i(t) \qquad\qquad (i = 1, 2, ..., n). \tag{3}$$

Then, combining (2) and (3), the *net* rate of project creation in i at t, $\dot{z}_i(t)$, is given by:

$$\begin{aligned} \dot{z}_i(t) = \dot{z}_i^+(t) - \dot{z}_i^-(t) &= \sum_{j \neq i} \eta\; y_i(t)\; \frac{y_j(t)}{\sum_{k \neq i} y_k(t)}\; f(\varphi_{ji}(A(t))) - \lambda\, z_i(t) \\ &= \eta\; y_i(t) \sum_{j \neq i} \phi_{ji}(t) - \lambda\, z_i(t), \end{aligned} \tag{4}$$

where $[\phi_{ji}(t)]_{j=1}^n$ represents the vector of decay-discounted flows of information that arrive at country i from every other $j \neq i$. The sum of all such flows, $\Phi_i(t) \equiv \sum_{j \neq i} \phi_{ji}(t)$, captures how well integrated is country i with the rest of the world, so we refer to it as i's *Globalization Index* (GI).

Consider now a stationary growth path where, for each country $i = 1, 2, ..., n$, the number of projects z_i active in every country remains unchanged, so that

$$\dot{z}_i(t) = 0 \qquad \forall i = 1, 2, ..., n, \tag{5}$$

and such stationarity also applies to its growth rate $\rho_i \equiv \dot{y}_i / y_i$, and its pattern of information flows Φ_i. Then, combining (1), (4), and (5), we find that the following relationship holds at a stationary state:

$$\rho_i^* = \frac{\eta \xi}{\lambda}\; \Phi_i^* \qquad\qquad (i = 1, 2, ..., n). \tag{6}$$

Expression (6) highlights the prominent role played in our theory by the information flows channeled into each country through its trade pattern. Indeed, the central prediction following from that expression is stark: countries that are better integrated in the world economy (i.e. have a higher GI) grow faster. Formally, the induced relationship between globalization and growth can be simply stated as follows:

$$[\mathbf{GG}] \qquad \forall i, j \in N, \quad \rho_i^* \geq \rho_j^* \Leftrightarrow \Phi_i^* \geq \Phi_j^*.$$

Of course, to test this prediction we still need to articulate a useful operationalization of the theoretical framework. This is the task undertaken in the ensuing subsection.

2.2 Operationalization of the theory

To render the theory operational, we need to construct the matrix A that, as explained above, governs the information diffusion process and consequently determines the expected path lengths φ_{ji} that underlie the GIs Φ_i. The construction of these objects involves the following steps.

The ***first step*** of the procedure involves the construction of the matrix of trade flows $X \equiv (x_{ij})_{i,j=1}^n$ between every pair of countries, where x_{ij} stands for the exports from i to j.

(Naturally, along the main diagonal of X we have $x_{ii} = 0$ for all $i \in N$.) To measure the *relative* importance of the trading partners of each country i, we simply normalize i's export flows $(x_{ij})_{j \neq i}$ by their total exports so that the induced magnitudes $\tilde{x}_{ij} \equiv \frac{x_{ij}}{\sum_{j \neq i} x_{ij}}$ satisfy $\sum_{j \neq i} \tilde{x}_{ij} = 1$. This leads to the row-stochastic matrix $\tilde{X} \equiv (\tilde{x}_{ij})_{i,j=1}^{n}$, which describes the distribution of export *shares* across the different countries and is one of the key components in the construction of the matrix A.

The ***second step*** focuses on the computation of a suitable indicator of openness for each country. To do so we follow Arribas et al. (2009) and measure the *openness* of any given country i by

$$\theta_i \equiv \frac{\sum_{j \neq i} x_{ij}}{(1 - \beta_i) y_i} \tag{7}$$

where β_i stands for the weight of country i's GDP in the world economy, i.e. $\beta_i = y_i / Y$, where $Y \equiv \sum_{j \in N} y_j$. In contrast with the received measures of openness, the denominator of (7) subtracts from y_i the part of a country's demand that, in the absence of foreign-trade bias, would be satisfied domestically, i.e. $(y_i / Y) y_i$. By doing so, the case where $\theta_i = 1$ corresponds to a situation where country i is *fully open*, in that its trade is "blind" to international borders. To see this note that, in such a border-blind case, the share of i's final output that is exported – i.e. $(1/y_i) \sum_{j \neq i} x_{ij}$ – is exactly equal to the weight of the rest of the world in the overall economy: $(1/Y) \sum_{j \neq i} y_{ij}$.

The ***third step*** combines the previous two as follows. Denote by Θ the diagonal matrix with the vector $(\theta_i)_{i \in N}$ along its main diagonal and let I be the identity matrix. Then, we define the matrix A as follows:

$$A = (I - \Theta) + \Theta \tilde{X} \tag{8}$$

The resulting matrix $A = (a_{ij})_{i,j=1}^{n}$ is non-negative and row-stochastic ($\sum_{j=1}^{n} a_{ij} = 1$), as required. Along the main diagonal, the entries $a_{ii} = 1 - \theta_i$ capture the extent of ***closedness*** of each country i. In line with our former explanation of θ_i, we can interpret a_{ii} as the fraction of trade that in an "unbiased" trade pattern would be directed abroad but in the case under consideration is steered towards the domestic market (hence incapable of channeling useful information elsewhere). In contrast, off the main diagonal, the entries $a_{ij} = \theta_i \tilde{x}_{ij}$ $(i \neq j)$ capture how international trade – and therefore the information embodied by it – is diffused to other countries.

The ***fourth step*** computes the expected lengths φ_{ji} required for the information generated in a country j to arrive, directly or indirectly, to any other country i, when such information is channeled through trade as reflected by the matrix A in (8). The precise derivation of the expected lengths $\left(\varphi_{ji}\right)_{i,j(i \neq j)}$ can be found in Appendix A. There we show that such expected path lengths can be computed as follows:

$$\begin{aligned} \Big(\varphi_{ij}\Big)_{\substack{i=1,2,\ldots,n \\ i \neq j}} &= \left(I - A_{-j}\right)^{-2} \left(I - A_{-j}\right) e \\ &= \left(I - A_{-j}\right)^{-1} e, \end{aligned} \tag{9}$$

where A_{-j} stands for the $(n-1)\times(n-1)$-matrix derived from A by deleting the jth row and column, and e is the column $(n-1)$-vector whose components are all equal to 1.

Recall that, in our theory, such path lengths determine the informational decay $f(\varphi_{ij}) \in [0,1]$ induced by any indirect connection from any country i to some other country j, where $f(\cdot)$ is a decreasing function. For concreteness, in our empirical analysis we rely on a variation of the canonical exponential form typically considered by the network and international-trade literature: the so-called "iceberg costs."[7] That is, we suppose that a constant fraction of informational value is lost for every additional order of magnitude traveled, so that that $f(s) = \delta^{\log(s)}$ for some $0 < \delta < 1$. The results reported in the paper are obtained for the specific factor $\delta = 0.93$, but the gist of our analysis is essentially unaffected by the specific value being considered.

Finally, in the ***fifth step***, we compute the Globalization Index of country i, $\Phi_i(t)$, as the weighted sum of all decay-discounted flows of information that arrive at country i from every other $j \neq i$:

$$\Phi_i \equiv \sum_{j \neq i} \beta_j f(\varphi_{ji}). \tag{10}$$

3 Data

Our data on bilateral trade flows is taken from the United Nations Commodity Trade Statistics Database (UN Comtrade) and it covers 204 countries on an annual basis for the period from 1962–2016. See Table 11 in the Appendix for the countries included in the sample. For each year and for every pair of countries, we use the information on the total value of exports, measured in current USD. The export flows for the countries in our sample cover, on average, 98% of the total yearly world export flows over the period from 1962–2016. Likewise, the GDP coverage ratio in our sample is high and very stable over time, with an average of 99% of world GDP. The high and stable coverage ratios, both in terms of trade flows and GDP, are reassuring because they suggest that our dataset allows for an accurate description of the world trade network[8].

Figure 1 provides a schematic visualization of the world trade flows in the year 2015 through a discretely represented network where, for the sake of clarity, links are binary (i.e. ignore the trade-based weight). Thus, in this network each link represents the existence of *some* bilateral trade flow between two countries, while the size of a country's label is taken to be proportional to the country's aggregate GDP. A number of observations arise. Most interestingly, the figure shows that the world trade network is far from complete. That is, many countries are connected to just a relatively small fraction of other countries and the variation in this respect is not necessarily related to country size. Even though the figure accounts for no direction in the trade links, this information could also be provided, with the direction of the links indicating the origin and destination of the flows. Naturally, this would lead to the distinction between

[7] This assumption is widely used in the theories of international trade and economic geography. It was first proposed by Samuelson (1954) and then adopted in the well-known paper by Krugman (1991).

[8] To cope with missing values, we use the observed import flows from country j to country i to impute the missing export flow from i to j. On average, 5.8% of the annual trade flows are imputed.

in-degree and out-degree of any given country, where the former reflects its imports and the latter its exports. Since both perspectives yield an equivalent representation of the network,[9] we choose the import-based one and define simply the *degree* of a country as the number of other countries from which it receives its imports.

Figure 1: World-trade discrete network in 2015

Table 1 provides more information on the properties of the trade network for the years 1965 and 2015. The first column – labeled "Avg" – shows the average degree of the discrete trade network (expressed as the fraction of the total number of countries in the network). For example, in 1965, countries imported on average from 41.3% of all countries. This value implies that the global trade network was far from complete. We also have that the connectedness of countries varied substantially. To show this, the next three columns show the 25^{th}, 50^{th} and the 75^{th} percentile of the distribution of countries according to their degree. The numbers in the first row indicate that, at the 25^{th} percentile, the countries imported in 1965 from only 26% of all countries, whereas at the 75^{th} percentile they imported from more than half of all countries.

		Avg	25^{th}	50^{th}	75^{th}
1965	All	41.3	26.0	35.8	52.4
	Poor	38.8	27.6	36.8	48.6
	Rich	74.3	70.5	81.2	91.7
2015	All	63.3	49.7	62.8	76.8
	Poor	60.3	49.6	60.4	71.9
	Rich	79.0	69.4	82.3	94.8

Table 1: Summary statistics on the degree distribution of the discrete world trade network, expressed as a percentage of the total number of countries in the sample.

[9]Note that the aggregate out-degree is equal to the aggregate in-degree, even if the overall distribution in each case may of course be quite different.

Table 1 also shows that there is a marked contrast between rich and poor countries. For concreteness, we classify countries as rich (poor) if, in the year 2015, their GDP per capita was above (below) 50% of the U.S. level. Then, the second row of the table shows that in 1965 poor countries imported, on average, from only 38.8% of all countries, while rich countries did so from an average of 74.3% of countries. Even the most connected poor countries at the 75^{th} percentile imported from less than 50% of countries whereas at that same percentile the rich countries imported from almost every country in the world. Overall, the pattern described for the year 1965 is essentially maintained for the year 2015 although the extent of average connectivity grows, with the increase being especially significant for the poorer countries.

Next, we take a complementary perspective on the description of the data that focuses on the weight of the links, as given by the normalized row-stochastic matrix A described in Section 2.2. Recall that, in this matrix, each entry a_{ij} represents the fraction of the exports of country i that are imported by country j. We are interested, in particular, on assessing how polarized are the imports of each country j towards a relatively small subset of other countries, in contrast with having a more diversified set of import providers. To this end, we consider the following statistics. First, denoting by m_j the median value of the distribution of weights $\left(a_{ij}\right)_{i\neq j}$, we define by

$$\lambda_j = \sum_{\{i:a_{ij}\leq a_{m_j j}\}} a_{ij} \tag{11}$$

the aggregate import weight of country j for countries i lying no higher than that of the median m_j. On the other hand, we denote by ν_j^u the total import weight of country j associated to its top u importers, where we consider the specific values of $u = 1, 3, 10$. Finally, we average those magnitudes and obtain $\bar{\lambda}$ and $\bar{\nu}^u$, where the averages are taken either at the whole world level or are separately computed for rich or poor countries, as defined before. The results are displayed in Table 2, with the magnitudes expressed in percentage terms over the total import weight attained by each country.

		$\bar{\lambda}$	$\bar{\nu}^{10}$	$\bar{\nu}^{3}$	$\bar{\nu}^{1}$
1965	All	3.3	79.6	55.9	34.0
	Poor	2.9	81.3	57.6	34.2
	Rich	4.4	62.3	36.5	20.0
2015	All	1.3	77.3	54.0	31.2
	Poor	1.1	80.9	57.3	33.4
	Rich	2.1	64.3	41.7	23.3

Table 2: Summary statistics on the distribution of import weights (expressed in percentage terms) for the world trade network.

The results show a strong concentration of countries' imports on only a few links for both the years 1965 and 2015. For example, according to the values in the first row, the weakest 50% of the import connections in 1965 account, on average (for the population as a whole), for only about 3% of the total import weight a country. At the same time, the strongest single

connection of a country accounts, on average, for 34% of the total weight in that same year. Again, we observe a quite different pattern for rich and poor countries. The import connections for poor countries are more highly concentrated than for rich ones, with the patterns being quite stable across both of the years considered.

4 Trade integration and alternative openness measures

As explained in Section 2, we interpret the *Globalization Index* (GI) derived from our theory as a measure of *trade-based integration*, similar in spirit (although, as we shall see, not in the details) to other measures of country openness that have been considered in the literature. The objective of the present section is to rely on the operationalization of this index explained in Subsection 2.2, to compute the GI, Φ_{it}, for every country i in our sample and every year $t = 1962, ..., 2016$, then contrasting it with two of the leading openness measures proposed in the literature: Trade Share (TS) and the Sachs-Warner Index (SWI). This exercise should clarify the nature of our proposed measure of trade integration, and the extent to which it incorporates features that are quite distinct from those displayed by such alternative measures.

Table 3 shows the value of the GI for a representative set of countries and for the years 1965, 1990 and 2015. The countries in the table are ranked in a descending order according to their 2015-value of the GI. Quite interestingly, we find that the most integrated countries have become more integrated over time, and that the ranking among these countries has remained quite stable, with the important exception of China. In contrast, several of the least integrated countries have become even less integrated over time, while among countries lying in the middle range the pattern is a diverse one with some countries becoming more integrated while others becoming less so. An additional interesting observation transpiring from Table 3 is that our measure of trade integration appears quite unrelated to either TS or the SWI. We elaborate on this feature below.

Further insights on our measure of trade integration are provided by Figure 2, where the four panels display information on its world distribution, its evolution over time, and its correlation with both economic performance and the trade share. The main observations derived from each panel can be summarized as follows.

- Panel (a) shows that the world is, and has been, very unequal in terms of the level of integration, as measured by the GI. The world integration distribution in 1965, 2005 and 2015 is very dispersed but relatively stable over time. If anything, between 1965 and 2015, there has been a general shift towards more integration at the world level.

- Panel (b) indicates that, with few exceptions, the ranking of countries in terms of integration has remained rather stable and, generally speaking, the richest countries show higher integration than poorer ones. (Each circle represents a country and the size of a circle is proportional to country's per capital GDP relative to U.S. GDP per capita.)

	Trade integration			Δ	Rank			Rank_{TS}	SWI
	1965	1990	2015	65-15	1965	1990	2015	2015	1990
United States	0.72	0.76	0.77	0.05	1	1	1	121	1
China	0.59	0.64	0.74	0.15	24	18	2	112	0
Germany	0.71	0.74	0.72	0.01	2	2	3	41	1
United Kingdom	0.70	0.71	0.70	0.00	3	4	4	91	1
France	0.68	0.72	0.69	0.01	4	3	5	84	1
Mexico	0.60	0.64	0.68	0.08	21	15	9	60	1
Hong Kong	0.59	0.65	0.68	0.09	20	12	10	1	1
South Korea	0.54	0.64	0.67	0.12	52	13	12	44	1
India	0.63	0.61	0.66	0.03	14	29	14	109	0
Brazil	0.58	0.60	0.64	0.06	33	31	22	123	0
Argentina	0.58	0.54	0.59	0.01	27	50	40	124	0
Nigeria	0.57	0.55	0.57	0.00	37	48	51	125	0
Guatemala	0.52	0.52	0.55	0.03	66	61	56	96	1
Ghana	0.55	0.49	0.53	-0.02	47	80	68	29	1
Yemen	0.39	0.47	0.52	0.13	123	98	77	119	1
Congo	0.53	0.49	0.51	-0.02	59	79	89	64	0
Liberia	0.56	0.53	0.51	-0.05	44	53	90	15	-
Uganda	0.47	0.47	0.50	0.03	97	92	94	101	1
Gambia	0.43	0.43	0.45	0.02	119	114	119	55	1
Central Afr. Rep.	0.44	0.42	0.41	-0.03	115	117	124	103	0

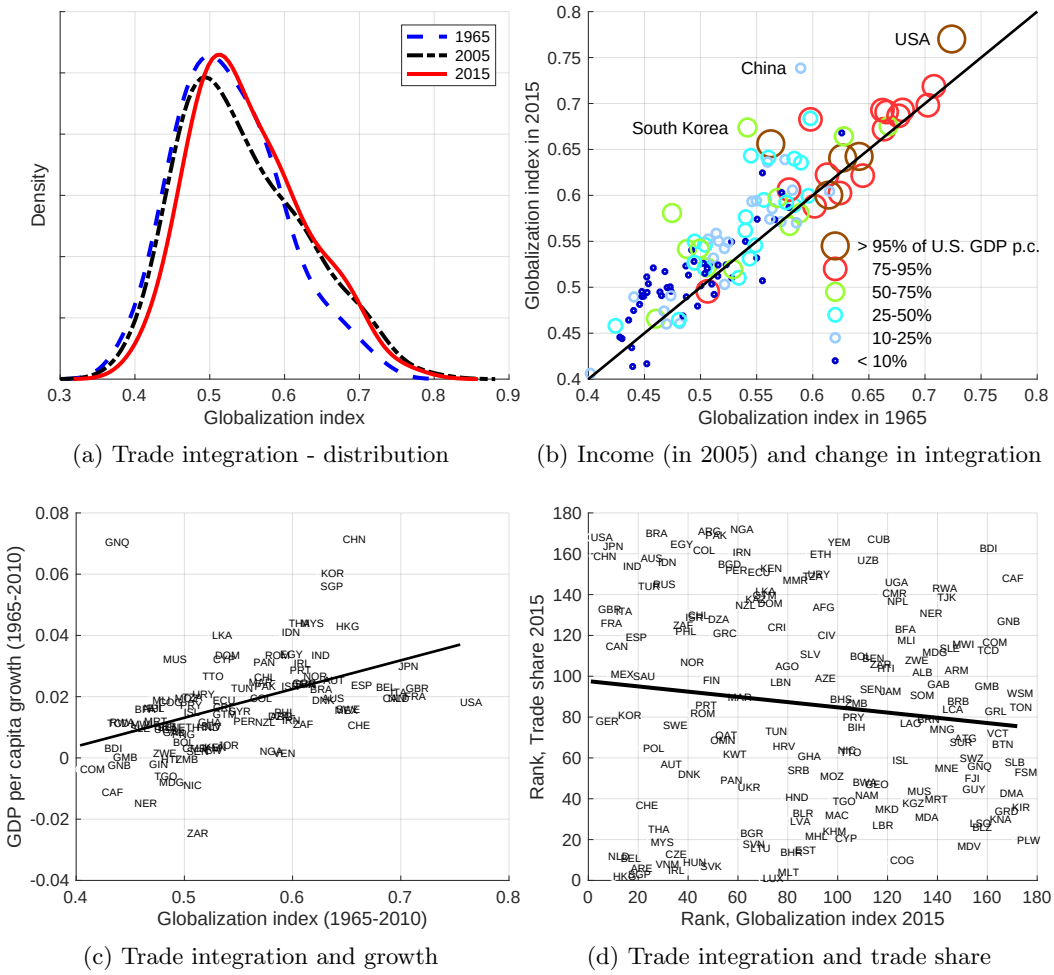
Rank: Ranking of each country in terms of the GI in a given year, $\textbf{Rank}_{TS}$: Ranking of each country in terms of the trade share (country with largest trade share is no. 1). Δ is the absolute change in the value of the GI between 2015 and 1965. **SW**: Sachs-Warner dummy variable - is 1 (0) if country is open (closed) to trade. Rankings are based on the sample of 125 countries for which data are available in 1965, 1990 and 2015.

Table 3: The Globalization Index – summary statistics and comparison with other openness indices: Trade Share and the Sachs-Warner Index.

- Relatedly, Panel (c), displays a strong relationship between a country's 1965-2010 average of the GI (x-axis) and the annual GDP growth rate (y-axis). That is, countries which are better integrated into the world trade network also grow faster. Section 5 explores this relationship more systematically and in greater detail.

- Finally, Panel (d) bears on a very interesting and somewhat striking fact. It shows that trade integration, as measured by the GI, is essentially uncorrelated with the TS, the classical measure of openness. Each 3-letter acronym in the figure represents a country and the location of a given country is determined by its position in the ranking of countries in the year 2015 based on the GI (x-coordinate) and the TS (y-coordinate). Countries that rank highly according to each measure are considered as open in terms of that measure. The rank correlation between our measure of trade integration and the trade share is very low, and equal to -0.06. Furthermore, some of most integrated economies in the world, such as the United States, France and the United Kingdom are classified as relatively closed according to the TS measure. Instead, at the opposite end, many of the countries that display a low GI (thus are not well trade-integrated) rank highly in terms of their TS

and therefore should be considered open according to it.



(a) Trade integration - distribution

(b) Income (in 2005) and change in integration

(c) Trade integration and growth

(d) Trade integration and trade share

Figure 2: Trade integration - across the world and over time.

In line with the absence of correlation between the GI and TS highlighted in the point above, we may go back to Table 3 to find that, for the selected set of countries, a similar state of affairs applies for the SWI. There, we can observe that several of the most (least) integrated countries according to our indicator are classified as closed (open) – i.e. have an index of 0 or 1, respectively – according to the SWI. Such a disconnect between the two indicators also holds more broadly for a larger sample of 109 countries for which we have the data on both measures. Specifically, we find that among the top-50 % of the most integrated countries according to our measure, one third of the countries are classified as closed according to the SWI. And, reciprocally, one third of the bottom 50 % of countries are classified as open.

Such a lack of correlation between our Globalization Index and the traditional openness indicators is a remarkable and somewhat surprising observation. At this point, therefore, we find it necessary to explore this puzzling issue in greater depth. We start by investigating the relationship between the GI and the SWI. In their highly influential study, Sachs and Warner

(1995) construct their dummy variable for openness by classifying a given country as open if none of the following five criteria hold: (i) the country has average tariff rates above 40 percent; (ii) non-tariff barriers cover more than 40 percent of its imports; (iii) the country operates under a socialist economic system; (iv) there is a state monopoly of the country's major exports; and (v) the black-market premium on its official exchange rate exceeds 20 percent. In view of criteria (i)-(v), a useful basis to understand the weak relationship between the SWI and the Gl is provided by the work of Rodriguez and Rodrik (2001), Harrison (1996) and Harrison and Hanson (1999). For, as these authors show, most of the explanatory power of the SWI comes from the two non-trade components: the existence of a state monopoly of the country's major exports, and the black-market premium on its official exchange rate. In view of this, Rodriguez and Rodrik (2001) argue that the SWI acts, in essence, as a dummy variable for Sub-Saharan countries and therefore should not be regarded as a suitable measure of a country's outward orientation.

Somewhat more subtle is the relationship between the GI and TS (defined as the sum of exports and imports of a country as a fraction of its GDP). To explore the low correlation between the two measures, we first present a stylized numerical example of a hypothetical world composed of just three countries which are linked by exports and imports. This example will be useful to demonstrate what features of trade determine a country's level of integration and its trade share, and how a low trade share can coexist with a high integration level and vice versa. For expositional convenience, we support our discussion through Table 4, where in three separate columns we specify the key magnitudes involved for the following three cases:

- In column (A) we specify the general expressions that are used in the calculation of the GI for a world with an arbitrary number n of countries.
- In column (B), we particularize the general expressions to our three-country example when the three countries are fully symmetric: each country $i = 1, 2, 3$ has a GDP $y_i = 1$ and its exports to the other two countries $j \neq i$ are $x_{ij} = 1/3$.
- In column (C), we consider again the three-country example but suppose that while countries 2 and 3 are as before, country 1 exports less, i.e. $x_{1j} = 1/10$ for $j = 2, 3$.

On the other hand, the different expressions listed vertically in Table 4 can be succinctly described as follows:

- in row (1), we have the matrix of bilateral trade flows;
- in row (2), the vector of GDP's for each country;
- in (3), their trade shares;
- in (4), their individual openness;
- in (5), the diffusion matrix;
- in (6), the expect path lengths involved in joining every pair of countries;
- in row (7) the GI of every country.

Naturally, for the symmetric three-country world considered in column (B) all values – vectors and matrices – listed in rows (1)-(7) are symmetric across the three countries. It may

	(A) General formula	(B) Symmetric case	(C) Country 1 exports less
(1)	$X = \begin{pmatrix} 0 & x_{1,2} & ... & x_{1,n} \\ x_{2,1} & 0 & ... & x_{2,n} \\ ... & ... & ... & ... \\ x_{n,1} & x_{n,2} & ... & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & \frac{1}{10} & \frac{1}{10} \\ \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}$
(2)	$y_i \quad (i \in N)$	$[1, 1, 1]$	$[1, 1, 1]$
(3)	$TS_i = \frac{\sum_j (x_{ij} + x_{ji})}{y_i} \quad (i \in N)$	$[1.33, 1.33, 1.33]$	$[0.87, 1.1, 1.1]$
(4)	$\theta_i = \frac{\sum_j x_{ij}}{y_i} \frac{1}{1-\beta_i} \quad (i \in N)$	$[1, 1, 1]$	$[0.3, 1, 1]$
(5)	$A = \begin{pmatrix} 1-\theta_1 & \theta_1 \tilde{x}_{1,2} & ... & \theta_1 \tilde{x}_{1,n} \\ \theta_2 \tilde{x}_{2,1} & 1-\theta_2 & ... & \theta_2 \tilde{x}_{2,n} \\ ... & ... & ... & ... \\ \theta_N \tilde{x}_{n,1} & \theta_n \tilde{x}_{n,2} & ... & 1-\theta_n \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{pmatrix}$	$\begin{pmatrix} 0.7 & 0.15 & 0.15 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{pmatrix}$
(6)	$\varphi_{ji} = (I - A_{-i})^{-1} e \quad (i, j \in N,\ i \neq j)$	$\begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 5.1 & 5.1 \\ 2 & 0 & 3.6 \\ 2 & 3.6 & 0 \end{pmatrix}$
(7)	$\Phi_i = \sum_{j \neq i} \beta_j f(\varphi_{ji}) \quad (i \in N)$	$[0.95, 0.95, 0.95]$	$[0.95, 0.9, 0.9]$

Table 4: A simple example for a world with three countries.

also be worth noting that their individual openness $\theta_i = 1$ for all of them since their domestic trade is proportional to their individual weight in the world as a whole.

In contrast, column (C) considers the more subtle case where country 1 exports less to both other countries, while everything else remains equal. Consequently, the trade share is lower than before for all countries, but more so for country 1. Due to lower exports, country 1 is no longer fully open as reflected by the value of $\theta_1 < 1$ and the corresponding change in the first row of the transition matrix A. Importantly, the pattern of imports of country 1 is as before and therefore the expected lengths of the diffusion paths from countries 2 and 3 to country 1 are unchanged and so is the value of Φ_i, country 1's GI. However, as country 1 is less open than before, the *direct* (one-link) diffusion paths flowing from country 1 to 2 and 3 are longer than before. And, moreover, so are the *indirect* (multiple-link) paths flowing *into* countries 2 or 3 since they involve the indirect connection via country 1. As a result, while the GI is unchanged for country 1, it is reduced for countries 2 and 3.

The simple numerical example described in Table 4 is useful to shed light on the contrast between TS and the GI. Specifically, it shows that, while for the determination of a country's TS what matters is the volume of its exports and imports, for the GI it is its pattern of import links (direct and indirect) that plays the key role. This means that, in general, a country can have both a low TS and a high GI – such as country 1 in the example described in column (C) – or a high TS and a low GI – as for countries 2 and 3 in that same example. To explore at a broader level whether, in effect, the full pattern of combinations in the relation between the TS

and the GI is indeed possible in a truly large and complex setup such as the real world, we turn again to the whole set of countries in our world sample.

It will be useful to divide the sample into four groups that reflect different combinations of the TS (high/low) and the GI (high/low). Table 5 shows the results for a selected set of countries belonging to these four groups for the year 2015. Panel (a) [(d)] shows countries with a high [low] TS and with a high [low] value of the GI. These countries are considered open [closed] according to both measures. Instead, for the countries in Panels (b) and (c) the two measures disagree about the countries' level of openness. One of the important observations that arises from the table is the following: the countries with a high GI are characterized by many import links. Thus, in the column labelled "#" that reports the share of all countries that a given country receives imports from, the most globalized countries import from almost all other countries. In contrast, we note that the least globalized countries import from substantially less and in many cases less than half of all countries.

		Rank								
		GI	TS	$\frac{IMP}{GDP}$	#	$\bar{a}_i$	$\bar{a}_i^2$	$\bar{a}_i^5$	$\bar{a}_i^{25}$	$\bar{a}_i^{50}$
(a) High GI / high TS	Netherlands	8	12	0.91	98	1.03	1.08	1.02	0.96	0.95
	Hong Kong	10	2	2.41	93	0.56	0.74	0.92	1.04	1.07
	Belgium	13	11	0.71	95	1.34	1.79	1.95	1.76	1.74
	Singapore	16	3	2.17	95	0.62	0.58	0.55	0.59	0.60
(b) High GI / low TS	USA	1	168	0.20	98	0.42	0.79	1.73	4.26	4.76
	China	2	159	0.19	99	5.35	8.14	11.41	12.30	12.55
	France	5	126	0.40	97	1.75	3.07	5.39	6.46	6.19
	Japan	6	164	0.22	96	1.73	2.88	4.87	7.79	8.18
	Canada	7	115	0.41	96	4.78	8.87	18.51	42.14	46.74
(c) Low GI / high TS	Liberia	114	27	0.80	50	0.02	0.03	0.03	0.02	0.02
	Congo	121	10	1.12	52	0.02	0.03	0.03	0.04	0.04
	Lesotho	153	28	0.80	26	0.00	0.01	0.01	0.01	0.01
(d) Low GI / low TS	Burkina Faso	123	123	0.54	59	0.01	0.01	0.03	0.07	0.07
	Niger	133	131	0.51	57	0.00	0.01	0.02	0.02	0.02
	Rwanda	138	143	0.42	58	0.01	0.01	0.05	0.24	0.35
	Burundi	157	163	0.33	46	0.03	0.06	0.13	0.16	0.12

Year: 2015, #: Number of import links, $\bar{a}_{ji} \times 100$: mean of import links, $\bar{a}_i^s$: s-order weight of import links

Table 5: Trade Share, the Globalization Index, and higher-order trading connections.

As a complement of the previous point, it is also important to emphasize that it is *not* the volume of imports that determines per se whether a country is considered open according to our measure. To see this, consider the column labeled "$\frac{IMP}{GDP}$" that shows the import share for each country. Many of the countries in Panels (c) and (d) display import shares that are higher than those of the most globalized countries in our dataset, such as the U.S., China, or France. Nevertheless, the number of import links of these countries is substantially lower, which is the reason why these countries rank low according to our GI.

But why is the number – not just the weight – of import links important? The reason is that the number of direct (first-order) links in turn determines the number of higher-order

links that a country has and therefore the total weight associated to more distant sources of information. To explore empirically this idea, in the last five columns of Table 5 we list the total diffusion weight reached by each country through direct import links as well as import links of order 2, 5, 25, and 50. While the total direct (or first-order) weight of a country i is given by $\bar{a}_i = \sum_{j \neq i} a_{ji}$, for any other $s \in \{2, 5, 25, 50\}$ we define the corresponding s-order weight by $\bar{a}_i^s = \sum_{j \neq i} a_{ji}^s$, where the elements a_{ji}^s are the ij entries of the diffusion matrix A multiplied by itself $s-1$ times, i.e. A^s. We observe that it is not only the value of $\bar{a}_i$ for the least globalized countries in Panels (c,d) that is a small fraction of the value of the most globalized countries in Panels (a,b). In addition, the countries in the two groups differ even more markedly in terms of their higher order links, as represented by the total diffusion weight. The value of $\bar{a}_i^s$ is uniformly much higher for the most globalized countries indicating that the indirect links contribute significantly to those countries' overall (direct and indirect) connectivity. Instead, the very low value of $\bar{a}_i^s$ for the least globalized countries shows that their trading partners are not generally well connected themselves.

Having established that the GI embodies a fundamentally different perspective of country openness than the traditional measures, we now turn to the quantitative analysis to revisit the empirical debate on the openness-growth nexus. The objective of the analysis is to assess the relative strength of the relationship between our proposed measure of country openness and economic growth. To this end, we compare it with a large set of alternative variables (33 of them) that have been highlighted by the empirical growth literature (see below for details). At this point, it is important to emphasize that we do not claim to establish a causal connection between economic growth and the GI – or, for that matter, concerning any of the additional variables considered. For, as is well known, growth empirics is generally plagued with a number of serious problems, above all the endogeneity concern. Our primary goal, therefore, is to explore the correlation patterns between country integration and economic performance, embedding the analysis into a carefully specified dynamic model that accounts for the whole structure of possible conditional dependencies on any subset of the alternative variables.

5 Econometric model

The empirical analysis is based on the following econometric model

$$y_{it} = \alpha y_{it-1} + \boldsymbol{\beta} \mathbf{x}_{it} + \boldsymbol{\delta} \mathbf{z}_i + \eta_i + \zeta_t + v_{it}. \tag{12}$$

where y_{it}, which denotes the log of GDP per capita of country i in period t, is modeled as a linear projection on its own lag and a set of covariates. The vectors $\mathbf{x}_{it}$ and $\mathbf{z}_i$ are vectors of dimensions $k \times 1$ and $m \times 1$, respectively. with the first being time-varying and the second one constant over time. On the other hand, η_i is a country fixed effect, ζ_t is a time effect that is common across all countries, and v_{it} is the random disturbance term which is assumed to satisfy $E[v_{i,s} \cdot v_{j,t}] = 0$ for all i, j, s, t.

Following Moral-Benito (2013, 2016) we assume that only the time-invariant variables in $\mathbf{z}$ are strictly exogenous and we treat all variables in $\mathbf{x}$ as potentially predetermined. Hence, to complete the model we augment it by an unrestricted feedback process which relates the predetermined variables in period t, $\mathbf{x}_t$, to all lags of the explained variable y, all lags of the predetermined variables, and the time-invariant variables $\mathbf{z}$. As we explain in detail in Appendix B, the estimation of parameters of the model pursues a limited-information maximum likelihood (LIML) approach.

Clearly, a key step in estimating the model in (12) concerns the choice of variables to be included in $\mathbf{x}$ and $\mathbf{z}$. This issue has proven to be a difficult challenge in the empirical growth literature. Partly, this problem derives from the fact that the theoretical literature lacks guidance about what factors are ultimately related to growth. As a consequence, researchers have often specified the empirical model in a more or less *ad-hoc* fashion. Over the years, this practice has led to the proposal of a large number of variables as possible growth correlates. For example, Durlauf *et al.* (2005) conducted a survey of the empirical growth literature and identified a total of 145 regressors that were found to be statistically significant in at least one study.

To address such model uncertainty, we apply the approach known as Bayesian model averaging (BMA).[10] In a nutshell, its objective is to develop a systematic way of assessing the probability that a given model specification be the "true" one. More specifically, suppose there are K candidate regressor variables. Hence, in total, there are 2^K possible combinations of regressors, where each combination gives rise to a different model. Let M_j $(j = 1, .., 2^K)$ denote any given such model that relates the outcome of interest y to a particular set of regressor variables. Then, given a prior $P(\cdot)$ over the space of those models and any collection of observed data $\mathbf{y}$, we may apply the logic of Bayesian inference to derive a posterior probability over any specific model M_j. That is, using Bayes Rule, such a posterior probability $P(M_j|\mathbf{y})$ is computed as follows:

$$P(M_j|\mathbf{y}) = \frac{p(\mathbf{y}|M_j)P(M_j)}{P(\mathbf{y})} \tag{13}$$

where $P(\mathbf{y})$ is the likelihood of the data and $p(\mathbf{y}|M_j)$ is their corresponding marginal (or integrated) likelihood.

Ultimately, we are interested in assessing the importance of each of the K candidate variables in explaining the data variable y. Thus, identifying such measure of "importance" of a variable v with the posterior probability that this variable belongs in the "true" growth model we compute:

$$P(v \in \mathbf{M}|\mathbf{y}) = \sum_{k \in M_j} P(M_j|\mathbf{y}). \tag{14}$$

[10] Bayesian model averaging is based on work by Raftery (1995) and was first applied by Sala-i-Martin et al. (2004) to determine which regressors should be included in linear cross-country growth regressions. An alternative to BMA is *Weighted Average Least Squares* (WALS) which was recently introduced by Magnus et al. (2010). WALS is superior to BMA in several respects; most importantly it outperforms BMA in terms of computational burden. Hence, it seems to be the preferable tool for model selection. For our purpose, however, it is not an adequate choice because it cannot (yet) deal with multivariate systems as the one in Expression (18). Moreover, unlike the BMA, it does not provide a metric that is useful to gauge the overall importance of a variable for explaining the data. See Moral-Benito (2015) for a survey on model-averaging techniques.

This probability is known as the (*posterior*) *inclusion probability* of variable v. Those variables with a high inclusion probability may then be considered as robustly related to economic growth.

In practice, as a choice for the model priors $P(M_j)$, we follow Ley and Steel (2009) and use the so-called Binomial-Beta prior structure (named after the implied model-size prior distribution) which has been shown to limit the effects of weak priors. Then, on the basis of it, the implementation of the BMA requires the estimation of all possible models associated to any given combination of co-variates. Clearly, this is computationally unfeasible when the number K of regressors is large – for instance, in our case we consider 34 potential regressors, which give rise to 2^{34} different models to assess. Thus we resort to the approach developed by Madigan and York (1995) known as Markov-Chain Monte-Carlo Model-Composition (MC^3). The MC^3 approximates the posterior probability distribution through an ergodic stochastic process that evolves according to a transition kernel that compares the posterior probabilities of neighboring models. And when this Markov chain is simulated for a sufficiently long time – so that the model-to-model transition probabilities become stationary – it can be be taken to have converged to the desired posterior distribution. In Appendix E, we explain this procedure more precisely and also report on how it performs in our particular case.

The set of regressors considered in our BMA analysis includes variables covering institutional, geographical, economic and demographic factors. Table 10 shows the complete list of variables. Naturally, among the candidate regressor variables contemplated, we have always included our measure of trade integration introduced in Section 2.2. And, of course, all the models under study consider the same dependent variable: the logarithm of real GDP per capita. To reduce the problem of serial correlation, we group the data into time intervals. That is, for a given time period, the dependent variable is the end-of-period value of per-capita GDP, whereas for the the regressor variables we take their within-period average values. In the benchmark case we use 10-year intervals, but as a robustness check we also use 5-year intervals.[11]

Finally, let us mention that we use data from 82 countries (covering all regions of the world and, as mentioned, 99% of its overall GDP) for the period 1960-2000.[12] See Table 11 for the list of countries in the sample. We have yearly observations for the dependent variable and all the candidate regressor variables. Using 10-year intervals, gives us a balanced panel with T=4 observations for every country. In Table 10 we report the data source and some descriptive statistics.

[11] We follow Caselli et al. (1996) and measure the flow variables (such as population growth) as 10-year averages while for the stock variables (such as life expectancy), we use the value of the variable in the first year of each 10-year period. To fix ideas consider, as an example, the period from 1960-1969. In this case, the dependent variable is the value of real per-capita GDP of a given country in the year 1970 and the lagged dependent variable is the 1960-value of real per-capita GDP. Moreover, the value of the variable representing a country's "population growth" is the 1960-1969 average of the country's population growth rate and the value of the variable representing a country's "life expectancy" is the value of the life expectancy in the year 1960.

[12] Ideally, we would preferred to consider a longer time horizon. However, due to data limitations there is a trade-off between the length of the time period considered and the number of variables included in the sample. Extending the time horizon would have considerably reduced the number of observations. For example, the data for the SWI is not available after 1992.

6 Main results

The main results of our analysis are reported in Table 6. The rows of this table correspond to each of the 34 regressor variables considered in our econometric exercise. They are ordered according to their posterior inclusion probability (PIP) and the row corresponding to the GI measure is highlighted for the sake of clarity. On the other hand, concerning the seven columns in the table, for the moment we focus only on the first four of them that correspond to our benchmark globalization measure, while the last three columns will be discussed in Subsection 7.4 when we consider an alternative globalization measure relying just on higher-order trade flows. The four columns under consideration specify, for each of the variables contemplated, the following values:

- The posterior mean, $E(\theta_v|\mathbf{y})$, of the coefficient θ_v estimated for the variable v. This mean is computed as $E(\theta_v|\mathbf{y}) = \sum_{v \in M_j} P(M_j|\mathbf{y})\hat{\theta}_v^j(M_j)$, where $\hat{\theta}_k^j(M_j)$ denotes the value estimated under model M_j.
- The posterior inclusion probability, PIP, as computed by (14) for each variable v.
- The fraction of models, $\%_{sig}$, where coefficient estimates are significant at the 5% level.
- The standardized coefficients, "beta", obtained by using standarized data in the analysis.[13]

In addition note that, on the first column, we also indicate the statistical significance of each estimated coefficient by computing the posterior variance, relying on the usual convention: 10% (*), 5% (**), 1% (***).[14]

A number of interesting observations emerge from the results in Table 6 for the benchmark case.

1. The posterior mean estimate for the coefficient of the GI regressor is not only positive but significant at the 1% level. To reinforce the latter point, we also note that the estimate coefficient is significant in 99% of all the models that include that variable.
2. A sizable PIP of more than 50% is attained by only eight variables. This value is line with the estimated posterior model size of 8.7. But more importantly for our purposes, our GI scores a very high inclusion probability of 85%.

[13]The standardization is achieved by de-meaning and normalizing the original data so that each variable has mean zero and a unit standard deviation. Doing so, therefore, the value of the coefficient specifies by how many standard deviations the dependent variable changes when the associated independent variable changes by one standard deviation.

[14]Following Leamer (1978), the posterior variance is computed as $V(\theta_k|\mathbf{y}) = \sum_{k \in M_j} P(M_j|\mathbf{y})V(\theta_k|\mathbf{y}, M_j) + \sum_{k \in M_j} P(M_j|\mathbf{y}) \left[E(\theta_k|\mathbf{y}, M_j) - E(\theta_k|\mathbf{y})\right]^2$. Sala-i-Martin et al. (2004) note that, having a ratio of posterior mean to standard deviation of around two (in absolute value) indicates an approximate 95-percent Bayesian coverage region that excludes zero. Using this "pseudo-t" statistic, we associate the levels of significance of 10%, 5%, 1% to the ratios of posterior mean to standard deviation of 1.645, 1.960 and 2.576, respectively.

Description of variable	Benchmark				Higher-order trade		
	$E(\theta_k\|\mathbf{y})$	PIP	$\%_{sig}$	beta	$E(\theta_k\|\mathbf{y})$	PIP	$\%_{sig}$
Lagged logarithm of real GDP per capita	0.8351***	1.00	100	0.7924	0.8497***	1.00	99
Investment share of real GDP	0.5903	0.92	17	0.0196	0.6553	0.81	32
1/0 dummy for Sub-Saharan country	-0.0789*	0.88	47	-0.0285	-0.0864**	0.85	58
Globalization Index	6.2887***	0.85	99	0.4361	5.7298**	0.79	95
1/0 dummy for armed conflict	-0.0681	0.75	6	-0.0192	-0.0821	0.53	11
Population share in the geographic tropics	-0.0538	0.72	25	-0.0195	-0.0635	0.58	42
Land area within 100km of navigable water	13.8364	0.68	97	0.0406	13.8946	0.60	92
Total population	0.4379	0.58	1	0.0197	0.3088	0.42	2
1/0 dummy for Latin-American country	-0.0237	0.34	16	-0.0066	-0.0446	0.34	21
Life expectancy at birth	1.3664**	0.23	80	0.1005	1.3966**	0.52	78
Sachs & Warner Index	0.1801***	0.16	98	0.0661	0.1753***	0.30	93
1/0 dummy for East Asian country	0.0627	0.15	39	0.0030	0.0881**	0.32	55
Government share of real GDP	-1.5067***	0.11	91	-0.0745	-1.6364***	0.11	88
Land share in the geographic tropics	-0.0327	0.11	15	-0.0116	-0.0538	0.15	27
Land share in Koeppen-Geiger tropics	0.0368	0.11	1	0.0130	0.0287	0.09	2
Labor force participation rate	1.1940*	0.08	54	0.0973	1.1159	0.12	48
Democracy index	-0.0869	0.07	4	-0.0204	-0.0983	0.08	7
1/0 dummy for former Spanish colony	-0.0609*	0.06	45	-0.0125	-0.0697**	0.11	58
Population share aged 0-14 years	-0.6378	0.05	21	-0.0898	-0.5311	0.05	27
Average years of secondary schooling	-0.0566	0.05	18	0.0318	0.0206	0.05	25
Land area in km^2	-0.0801	0.05	7	-0.0163	-0.1153	0.06	34
Exports plus imports as a share of GDP	-0.0668	0.04	19	-0.0311	0.0895	0.04	15
1/0 dummy for Western European country	0.0488	0.04	9	0.0118	0.0780*	0.05	44
Population density	-0.0474	0.04	1	-0.0287	-0.0181	0.04	1
Annual growth rate of population	-2.1608	0.03	69	-0.0486	-1.2932	0.06	69
Population share aged 65 years and above	2.1308	0.03	38	0.1009	3.2982	0.05	69
Consumption share of real GDP	-0.3434	0.03	12	-0.0249	-0.3048	0.02	8
Average years of primary schooling	-1.3586*	0.02	71	-0.1976	-1.4052*	0.05	79
Urban population	-0.2327	0.02	54	0.0095	-0.5013	0.02	64
Air distance to NYC, Rotterdam, Tokyo	-0.0160	0.02	9	0.0058	-0.0506	0.05	12
1/0 dummy for landlocked country	-0.0301	0.02	6	-0.0066	-0.0440	0.03	9
1/0 dummy for socialist rule in 1950-95	-0.0103	0.02	0	0.0001	-0.0175	0.02	3
Price level of investment	0.0305	0.01	6	0.0934	0.0271	0.02	1
Timing of national independence	-0.0053	0.01	3	0.0030	-0.0101	0.01	3

Table 6: Results of the Bayesian model averaging analysis.

3. The standarized estimates reported in the fourth column also yield a high coefficient for the GI, several orders of magnitude larger than any other, with the exception of the lagged value of the dependent variable. This suggests that the estimated effect of the GI is not only statistically significant but also economically so.

Jointly considered, the above three points provide substantial support to the existence of a strong positive positive relation between trade integration and income per capita. In fact, note that this applies not only to the level but also to the growth rate of per-capita GDP, since our empirical model controls for the initial log-level of that variable in every period.

In contrast with the strong support obtained for the GI trade-integration measure, the results in Table 6 indicate that the relationship between economic growth and the conventional openness indicators, such as the TS and the SWI, is quite weak. In particular, both of these variables display quite low PIPs of 0.16 and 0.04, respectively. As explained in Section 4,

this is largely in line with the claim put forward by Rodriguez and Rodrik's (2001) that the traditional indicators of outward-orientation do not truly embody a notion of openness that is closely related to economic performance. In the robustness analysis conducted in Section F.2, we address the concern that the weak support enjoyed by those openness indicators might be driven by a potential dependence obtained between our GI and the traditional measures. A priori, such dependence is unlikely, since Section 4 already showed that GI is practically uncorrelated with TS and SWI. Nevertheless, we find it worthwhile to analyze this issue in the broader context of the BMA. Concretely, we include into the BMA different combinations of the different openness measures to check whether the exclusion of some variables significantly alters the results for the others. As shown in Table 15, by and large the analysis does not uncover any notable dependencies between the different measures.

Finally, an additional finding of some relevance to growth empiricists, is the discrepancy observed for some regressor variables in terms of the relevance attributed to them by their values of the posterior inclusion probability and the $\%_{sig}$-statistic. This applies, for example, to the variables representing the *Government share*, the *Average years of primary schooling*, the *Sachs-Warner* Index, or the *Annual population growth rate*. These are variables characterized by low values of the PIP – indicating that the models which include these variables receive only little support from the data – and high values of the $\%_{sig}$-statistic – indicating that the estimates of the variables' coefficients are significant in a large ("conditional") fraction of the models where the variable is included. For example, the *Government share* has a PIP of only 11% but the estimated coefficient is significant in 91% of the models that contain it. In Appendix G we explore this discordance in some detail and provide an explanation for it. We then argue that the striking disconnect observed for some of the variables considered illustrates and underlines the superiority of the model-averaging approach over the traditional approach: whereas the latter identifies robust estimates with those that are statistical significant within the models that include the corresponding variables, the former takes into account as well the support/likelihood that those models receive from the data in the first place (in comparison with the models that do not include the variables in question).

7 The key features of the Globalization Index

There are two essential features that underlie our theory and also make our measure of trade integration – the Globalization Index (GI) – stand apart from received measures of openness: (a) its view of trade as a channel of information flows; (b) its focus on the overall architecture of the trading network and hence the role of indirect links. In this section we provide some empirical support for the prominent role played by these two features as correlates of growth.

Concerning (a), we consider two different (complementary) routes. First, in Subsection 7.1 we show that while trading flows in capital goods are positively related to growth, trade in raw commodities and economic growth are largely unrelated. Since, conceivably, the first kind of trade embodies much more valuable information and know-how than the second, the indicated

evidence provides an intuitive basis for feature (a). A more direct support for it is then provided in Subsection 7.2. There we find that if we approximate the flow of ideas across two countries with the volume of patents in one that are cited in the other, such a variable is closely related to the corresponding trade-based distance from the former to the latter and therefore.

Pertaining to the feature of GI described by (b), on the other hand, we again explore two different avenues to assessing it. First, in Subsection 7.3, we measure the network distance from either *only* direct links or *only* indirect paths. Then we find that the latter capture the bulk of the relationship between trade-integration and growth while the former is relegated to a very subsidiary role. Finally, in Subsection 7.4, we arrive at a different, but similarly motivated, point by constructing a pseudo-measure of "globalization" that replaces the direct links that reflect trade with links that weigh geographic proximity alone. We then arrive at the conclusion that the induced measure of "trade-integration" that abstracts from direct-trade links enjoys a support from the data that is very close to that provided by our benchmark measure.

7.1 Good-specific trade flows

As advanced, here we investigate to what extent the relationship between trade integration and growth is associated to varying intensities to trade in different types of goods. Recall that our measure of country integration, the GI index, has been computed by using the ***total*** bilateral trade flows. Consequently, it treats, say, Brazilian coffee exports to Japan and Japanese computer equipment shipped to Brazil equivalently (conditional on having the same dollar value). Arguably, however, not all kinds of trade are equally meaningful and should have the same relation to a country's economic performance. In other words, if trade involves sophisticated goods (for example capital goods) it can be expected to embody valuable information and know-how and therefore have a stronger connection to long-run growth than trade in low-tech goods (say, raw materials). As a first step towards exploring this question, we consider here trade flows at the one-digit product level and separate it into the following two broad categories:

- **Capital goods**: Chemicals, Manufactured goods, Machinery and transport equipment, Miscellaneous manufactured articles
- **Commodities and processed raw materials**: Food and live animals, Beverages and tobacco, Crude materials (except fuels), Mineral fuels, lubricants and related materials, Animal and vegetable oils, fats and waxes.[15]

In analogy to how we compute our baseline GI measure, we use data on bilateral trade flows for each product type to obtain the corresponding type-specific GI. Such an index measures the connectedness of each country to global trade for that product type. Then, we include each of these type-specific GI measures in a separate BMA to explore how trade in the different product

[15]The classification scheme and the data on bilateral goods-specific trade are taken from the UN Comtrade database.

types is related to growth. Table 7 reports the posterior inclusion probability and the posterior mean of the GI coefficient for each product type.

	$E(\theta_k\|\mathbf{y})$	PIP	$\%_{sig}$
Benchmark	**6.2887*****	**0.85**	**99**
Capital goods	6.3352***	0.82	98
Machinery equipment	8.4218***	0.92	93
Manufactured goods	6.7548***	0.85	91
Chemicals	5.6109**	0.73	89
Other manufct goods	6.6491**	0.67	98
Raw material goods	1.8306	0.21	42
Beverages, tobacco	0.8214**	0.43	87
Mineral fuels	1.4935*	0.38	52
Oils and fats	0.7790	0.29	28
Crude materials	2.8673**	0.18	82
Food, live animals	1.4357*	0.15	19

Table 7: Statistics derived by BMA analysis for the various product-based globalization indices.

A number of observations are worth highlighting. First, the GI computed for the broad category of capital goods has a very high inclusion probability of 0.82 and a posterior mean for the regression coefficient that is somewhat higher than in the benchmark case. At the one-digit level, we find that *Machinery equipment* and *Manufactured goods* have by far the highest posterior mean and inclusion probability. The situation is drastically different for commodities and processed raw materials. For most of the product types in this category, the posterior inclusion probability is considerably lower than that for capital goods. Also the posterior mean is mostly insignificant.[16] Thus, taken together, these empirical findings suggest, in an admittedly indirect manner, that high growth is mostly associated to trade in goods that are expected to embody a larger amount of information and hence also diffuse more of that information. A complementary analysis of the phenomenon that focuses explicitly on information diffusion itself is discussed next.

[16]Our results on product-specific trade are complementary to those in Hausmann et al. (2007). In particular, they show that the composition of a country's production portfolio plays an important role for growth. That is, countries which specialize in the type of goods that rich countries typically export grow faster than countries that specialize in other goods. In order to put our result into the same perspective, we follow their approach and compute (for 2005) the weighted average of per-capita GDPs of the countries exporting a one-digit product type, where the weights are the revealed comparative advantage of each country in that product. According to this measure, a product type that is produced primarily by rich countries is associated with a higher income level than a product that is produced by poor countries. Interestingly, we find that there is a strong positive relation between the income level associated with a product and the posterior inclusion probability associated to the corresponding product-based GI. The results are available upon request. Elaborating upon Hausmann et al. (2007), a possible interpretation is that a country's growth is not only favored by having a production portfolio of goods that are similar to those of rich countries but also by being well-connected to world trade in terms of those goods as well.

7.2 Trade and the flow of ideas

In this section we aim at testing directly the postulated theoretical relationship between a country's trade integration and its global access to ideas. The challenge in this pursuit is how to operationalize the concept of the "global flow of ideas". For a long time economists have advocated the view that the global flow of ideas is inherently hard to track. For example, Krugman, in his *Geography and Trade*, stated that *"knowledge flows [...] are invisible; they leave no paper trail by which they may be measured and tracked"*. However, Jaffe et al. (1993) reacted to the previous statement by suggesting that

> *"[...] knowledge flows do sometimes leave a paper trail, in the form of citations in patents. Because patents contain detailed geographical information about their inventors, we can examine where these trails actually lead."*[17]

Here, we espouse the view of Jaffe et al. and use patent citations as a proxy for the flow of ideas. More specifically, we utilize the NBER's patent database which contains detailed information on all U.S. patents granted between 1963-1999 (roughly three million patents) and all citations made to these patents between 1975-1999 (over 16 million citations).[18] Furthermore, the data set also includes, for each patent (either if it is created by a single inventor or by a whole team of inventors), the identity and the address (country, city, zip code and street) of every inventor that was involved in it. Based on the aforementioned information, we construct two variables, Avg_{ij} and $Prob_{ij}^{inv}$ (see Appendix I for the details). On the one hand, the variable Avg_{ij} measures how many patents of country j are cited, on average, by patents of country i. On the other hand, the variable $Prob_{ij}^{inv}$ specifies the fraction of cross-country co-patenting (bilateral) relationships of inventors from country i to involve a co-inventor from country j.

If a country's trade integration is positively related to the global flow of ideas, then we would expect that the closeness (in the trading-network sense) of two countries should be associated with an intensified exchange of knowledge and more joint innovation activities. To test this hypothesis, we estimate by OLS the following model:

$$y_{ij} = \alpha + \beta f(\varphi_{ji}) + \gamma X_{ij} + \epsilon_{ij} \tag{15}$$

where $y_{ij} \in \left\{Avg_{ij}, Prob_{ij}^{inv}\right\}$, α is a constant term and $f(\varphi_{ji})$ is the measure of (network) closeness between countries j and i defined in Section 2.2. Our choice of covariates X_{ij} controls for the intercountry characteristics highlighted by Gravity Theory, the workhorse of much empirical work in international trade. According to the gravity equation the bilateral economic interaction between two countries is proportional to the size of the countries and inversely proportional to the distance between the countries. Thus, we include in X_{ij} the relative size of countries as measures by their relative GDP and the geographical distance between countries expressed in kilometers.

[17] Jaffe et al. (1993), p. 578

[18] See Hall et al. (2001) for a comprehensive description of the dataset.

The results for the baseline specification are in the columns labeled "All" in Table 8. Most importantly, we find that the estimate of β is highly significant and positive in both cases, suggesting that countries which are closer together in a network sense (as reflected by a higher value of $f(\varphi_{ji})$) are more likely to engage in joint innovation efforts ($Prob_{ij}^{inv}$) and are more likely to cite each others patents (Avg_{ij}). The (relative) size of the foreign country is also strongly and positively related, which is in line with the logic of the gravity equation. Interestingly, however, the coefficient estimate on km_{ji} is insignificant suggesting that the knowledge flow between countries is unrelated to geographical distance.

	Model 1: Avg_{ij}			**Model 2:** $Prob_{ij}^{inv}$		
	All	Cap	Raw	All	Cap	Raw
α	-0.813**	-0.238	0.775***	-0.398***	-0.488***	-0.086
	(0.317)	(0.295)	(0.239)	(0.078)	(0.092)	(0.073)
$f(\varphi_{ji})$	2.836***	1.845***	-0.321	0.618***	0.814***	-0.026
	(0.615)	(0.616)	(0.549)	(0.154)	(0.193)	(0.168)
y_j/y_i	0.092***	0.135***	0.201***	0.023***	0.023***	0.049***
	(0.023)	(0.020)	(0.019)	(0.006)	(0.006)	(0.006)
km_{ji}	0.481	0.267	-0.014	0.053	0.091	0.019
	(0.325)	(0.321)	(0.331)	(0.085)	(0.093)	(0.099)
N	3041	3041	3041	1812	1320	1344
R^2	0.18	0.19	0.18	0.25	0.26	0.25

Dependent variable in Model 1: Avg_{ij} average number of citations that a patent from country i makes to patents from country j; Dependent variable in Model 2: $Prob_{ij}^{inv}$ probability that inventor from country i has a joint patent with inventor from country j. Independent variables: α: constant, $f(\varphi_{ji})$: , y_j/y_i:, km_{ji}: distance in 100,000 km between countries i and j. All: Total trade, Cap: Trade in capital goods, Raw: Trade in raw materials. All variables are expressed as the 1975-1999 average.

Table 8: Country distance and patent citations.

In the spirit of our analysis of Subsection 7.1, we also explore matters further and check whether the bilateral knowledge flow between countries is related to the countries' involvement in trade in different types of goods. To this end, first we separately calculate the measure of bilateral network distance $f(\varphi_{ji})$ using data on the combined trade in all capital goods and in all raw materials (applying the same classification of goods listed in as in Table 7). Then we include these goods-specific closeness measures into the empirical model given by (15) and reestimate it. The results are in the columns labeled with "Cap" (for capital goods) and "Raw" (for raw materials) in Table 8.

The following interesting observations arise. While, the coefficient estimates for country size (y_j/y_i) and geographical distance (km_{ji}) are largely similar across product types – in terms of sign and significance – they differ fundamentally for countries' closeness measure. In particular, we find that for capital goods there is a robust and positive relation between countries' closeness and the bilateral knowledge exchange. In contrast, for raw materials, the coefficient estimates

suggest no significant relation.

Overall, our findings in this section provide some empirical support to the idea that, at least in part, trade integration is correlated with growth due to the knowledge flows embodied in the cross-country trade. More concretely, they show that close proximity between countries is significantly related to the bilateral exchange of ideas. There is, however, the important qualification that, as suggested also by the analysis of Subsection 7.1, such a phenomenon arises only when the goods traded are of the type we have generically labeled as capital goods, i.e. when they are likely to embody valuable information and know-how.

7.3 The role of direct and indirect trade links

As explained, a distinctive feature of our GI is that it measures a country's level of integration not only by its set of direct trade connections but also through the full architecture of its higher-order connections in the world trade network. An important question in this context is to what extent the countries' direct trade links as opposed to the indirect ones matter for the positive relationship between trade integration and growth.

In this section, we compare the results of two experiments that allows us to shed light on the relative role of direct and indirect links. In the first experiment, we calculate the network distance φ_{ji} between any pair (j, i) from a modified adjacency matrix A where we remove the direct connection between j and i by setting $a_{ji} = 0$. In that way, we compute the expected number of steps that it takes to get from j to i - without going through the direct connection between j and i. Then, we calculate the GI, Φ_i, as described above and include it (instead of the baseline index) in the Bayesian analysis. The second row in Table 9 shows that the results for this modified measure are almost the same as those for the baseline measure. In other words, the correlation of our baseline GI with growth does not seem to depend on the direct trade connections between countries.

	$E(\theta_k\|\mathbf{y})$	PIP	$\%_{sig}$
Baseline	6.289***	85	99
Only indirect links	6.136***	82	97
Only direct links	0.874	22	39

Table 9: Direct versus indirect links.

In contrast, in the second experiment we use only the direct trade connections of countries to compute the GI. More concretely, we compute the expected number of steps that it takes to get from any $j \neq i$ to i via the direct link as follows:

$$\Big(\tilde{\varphi}_{ji}\Big)_{j\neq i} = diag\Big(I - A_{-i}\Big)^{-2}(\cdot\times)\Big(a_{ji}\Big)_{j\neq i}$$

where A_{-i} is the adjacency matrix A from which we have deleted the ith row and the ith column, $diag(\cdot)$ denotes the vector of elements on the main diagonal of the matrix, and $(\cdot\times)$ is the element-by-element multiplication of two vectors.[19] Then, we compute the GI as described in Section 2.2 and include it into Bayesian model averaging analysis.The results are reported in the third row of Table 9.

A comparison of the results obtained when only indirect, or only direct, links are allowed suggests that the positive relation between trade integration and growth is largely driven by the countries' higher-order trade connections. That is, the direct links matter significantly less.

7.4 Globalization Index on higher-order trades

To approach the issue studied in the preceding section from a different perspective, here we study the implications of a variation of the baseline Globalization Index that uses only the higher-order trade connections of a country and replaces the first-order connections by a link that reflects purely exogenous (geographical) considerations.[20] More specifically, we first define $\varphi_{jm,-i}$ as the expected number of steps required to reach country $m \neq i$ from country j through trade-weighted links, conditional on not using any of the (direct) such links that involve country i. In place of those direct connections, we use the geography-based links whose weights ω_{mi} (appropriately normalized so at to add up to unity for each i across all $m \neq i$) are inversely proportional to the distance geo_{mi}. Thus, formally, we have:

$$\omega_{mi} = \frac{1/geo_{mi}}{\sum_{m' \neq i} 1/geo_{mm'}} \tag{16}$$

Finally, we compute the expected number of steps from country j to country i as the weighted average over $\varphi_{j,m,-i}$, where we use $\omega_{i,m}$ as weights, i.e.

$$\tilde{\varphi}_{ji} = \sum_{m \neq i} \omega_{mi}\ \varphi_{jm,-i}.$$

This is a directed-distance measure that computes the expected length of all trade-weighted paths arriving to country i through the countries m that export to it, directly and indirectly, then assessing the connection of j to those countries by exogenous (geographic, hence not trade-based) considerations.

On the basis of those magnitudes for every pair of countries, i and j, we have computed a Modified Globalization Index as before, then including it into the BMA analysis to investigate

[19] Notice that this formulation considers the connections from j to any other $k \neq j$, $k \neq i$ and back from k to j. Also note that for $a_{ji} = 0$ we obtain $\tilde{\varphi}_{ji} = \infty$. This would render the computation of Φ_i infeasible. To resolve this issue, we replace $\tilde{\varphi}_{ji}$ by largest finite distance that prevails for country i in the given year. In addition to such largest finite distance, we also experimented with other imputation methods including, for example, the maximum distance across countries and years. None of these had any significant quantitative effect on the results.

[20] In Appendix H.2, we further advance on the approach of using geographical distance as an exogenous proxy for bilateral trade flows. More concretely, following the approach by Frankel and Romer (1999) we construct a modified GI, which relies on the geographical distance between each pair of countries i and j, and we use this measure to instrument for the baseline GI. We thank an anonymous referee for suggesting this exercise.

its relation to growth. The results, which are reported in the column labeled *Higher-order trade* in Table 6, are very similar to those of the baseline benchmark measure. This provides further support (complementary to that of Subsection 7.3) to our suggestion that higher-order links play a prominent role in capturing the essential component of the relationship between trade integration and growth.

8 Robustness analysis

We have conducted a broad robustness analysis of our baseline results and explored various extensions. In the interest of space, we relegate the details and results of this analysis to Appendix F, providing here just a succinct advance of it. First, we test the robustness to different data inputs by considering alternative data sources and utilizing different waves of datasets. Second, we apply alternative ways of measuring the degree of trade integration of a country. More concretely, while our baseline measure of integration reflects a notion of network centrality that is known as closeness centrality, in our robustness analysis, we have considered other prominent notions of centrality such as PageRank centrality. Third, we have modified the set of regressor variables and experimented with the model priors. Overall, the analysis confirms our main finding that trade integration is strongly positively correlated with economic growth, and that this result is not sensitive to different data sources, data treatment, alternative measures of network centrality, or assumptions about priors and the set of covariates included in the empirical model.

9 Conclusion

In this paper, we propose a new approach to evaluating a country's outward orientation, and then investigate the relationship of the induced measure to its growth performance. Previous work has mostly used indicators involving aggregate trade intensity, trade policy, or trade restrictiveness of the country in question. Instead, we offer a broader perspective on the phenomenon as a country's level of integration is assessed not only through its direct trade connections with the rest of the world but also uses the whole architecture induced by its second and higher-order connections.

We use trade data from the United Nations Commodity Trade Statistics Database and apply our methodology to a sample of 204 countries spanning the period from 1962 to 2016. A first descriptive analysis of the data reveals that our measure of integration is largely uncorrelated with the conventional indicators of openness (such as the trade share or the Sachs-Warner openness index). It also shows that, across the period being considered, the world as a whole has become more integrated. It has also become more unequal in this respect because the group of rich and most integrated countries has shown a persistent tendency to increase their integration, while the majority of poor and less integrated countries have been stagnating or

falling behind.

Then we pursue a systematic econometric analysis that revisits the long-standing debate in the empirical literature concerning the relationship between countries' outward orientation and their different growth experiences. To address model-selection concerns, we do it through a comprehensive Bayesian model-averaging analysis that considers all possible specifications involving any subset of 34 different variables as candidate regressors. The key finding is that our network-based measure of trade integration is strongly correlated with cross-country income differences, while the traditional indicators of country openness is only marginally so. In fact, trade integration stands out from all other regressors (except own lagged GDP) with a substantially larger posterior inclusion probability than for all those that are statistically significant. To check the robustness of our conclusions, we perform an extensive battery of sensitivity analyses and find that our baseline findings are largely unaffected if we use other data sets or rely on different variants for computing trade integration.

To sum up, we suggest that our analysis sheds new light on the nexus between openness and growth, pointing as well to a possible explanation for why the long debate that it sparked has remained largely inconclusive. The reason may be that trade-based integration in the world market – a natural and theoretically founded measure of a country's openness – requires a systemic evaluation of higher-order trade connections that goes well beyond (and tends to be only weakly related to) the direct trading magnitudes exclusively considered by the received openness indicators.

References

[1] Acharya, R.C. and W. Keller (2009): "Technology transfer through imports," *Canadian Journal of Economics* **42**, 1411-1448

[2] Alcala, F. and A. Ciccone (2004): "Trade and productivity", *Quarterly Journal of Economics* **119**, 613-646.

[3] Alvarez, F., F. Buera and R.E. Lucas (2017): "Idea Flows, Economic Growth, and Trade", working paper, University of Chicago and Federal Reserve Bank of Chicago **19667**.

[4] Arribas, I., F. Pérez, and E. Tortosa-Ausina (2009): "Measuring Globalization of International Trade: Theory and Evidence", *World Development* **37**, 127-145.

[5] Baumol, W., (1986): "Productivity Growth, Convergence, and Welfare: What the Long-run Data Show", *American Economic Review* **76**, 1072-85.

[6] Barro, R. J. (1991): "Economic Growth in a Cross Section of Countries, *Quarterly Journal of Economics* **106**, 407-433.

[7] Barro, R. (1996). Democracy and growth". *Journal of Economic Growth* **1 (1)**, 1-27.

[8] Barro, R., Lee, J.-W. (1994). "Sources of economic growth". *Carnegie-Rochester Conference Series on Public Policy*, **40**, 1-57.

[9] Bloch, F., Jackson, M., Tebaldi, P. (2017): "Centrality measures in networks," working paper.

[10] Buera, F. and E. Oberfield (2016): "The global diffusion of ideas," *Econometrica* **88**, 83-114.

[11] Casellit, F., G. Esquivel and F. Lefort (1996): "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics", *Journal of Economic Growth* **1**, 363-89.

[12] Caselli, F., and W. J. Coleman (2001): "Cross-country technology diffusion: the case of computers." *American Economic Review* 9**1**, 328-335.

[13] Chaney, Thomas (2014): "The network structure of international trade," *American Economic Review*, **104(11)**, 3600-3634.

[14] Ciccone, A. and M. Jarocinski (2010): "Determinants of economic growth: Will Data Tell?", *American Economic Journal: Macroeconomics* **2**, 222-46.

[15] Coe, D. and Helpman, E. (1995): "International R&D spillovers," *European Economic Review* **39**, pp. 859-87.

[16] Coe, D., E Helpman, and A. W. Hoffmaister (1997): "North-South R&D spillovers," *Economic Journal* **107**, pp. 134-49.

[17] Coe, D., E Helpman, and A. W. Hoffmaister (2009): "International R&D spillovers and institutions." *European Economic Review* **53**, 723-741

[18] Dollar, D. (1992): "Outward Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976-1985", *Economic Development and Cultural Change* **40**, 523-544.

[19] Dollar, D. and A. Kraay (2003): "Institutions, Trade, and Growth", *Journal of Monetary Economics* **50**, 133-162.

[20] Duernecker, G. and F. Vega-Redondo (2018): "Social networks and the process of globalization", *Review of Economic Studies*, **85(3)**, 1716-1751.

[21] Duernecker, G., M. Meyer, and F. Vega-Redondo (2020): "The network origins of economic growth," *IGIER working paper*, Bocconi University.

[22] Durlauf, S. N., P. A. Johnson, and J. R. W. Temple (2005): "Growth Econometrics", *Handbook of Economic Growth*, ed by P. Aghion and S. Durlauf.

[23] Easterly, W. (1993). "How much do distortions affect growth?". *Journal of Monetary Economics* **32(2)**, 187-212.

[24] Easterly, W., Levine, R. (1997). "Africa's growth tragedy: Policies and ethnic divisions". *Quarterly Journal of Economics* **112(4)**, 1203-1250.

[25] Easterly,W., Levine, R. (2001). "It's not factor accumulation: Stylized facts and growth models". *World Bank Economic Review* **15**, 177-219.

[26] Eberhardt, M., C. Helmers, and H. Strauss (2013): "Do spillovers matter when estimating private returns to R&D?," *The Review of Economics and Statistics* **95**, 436-448.

[27] Estevadeordal, A. and A. M. Taylor (2013): "Is the Washington Consensus Dead? Growth, Openness, and the Great Liberalization, 1970s-2000s", *Review of Economics and Statistics*, **95(5)**, 1669-1690.

[28] Etur, C., and A. Musolesi (2017): "Weak and strong cross-sectional dependence: a panel data analysis of international technological diffusion," *Journal of Applied Economics* **32**, 477-503.

[29] Feenstra, R.C., R.E. Lipsey, H. Deng, A.C. Ma, and H. Mo (2005): "World Trade Flows: 1962-2000", *NBER Working Papers* **11040**.

[30] Feyrer, J. (2019): "Trade and Income – Exploiting Time Series in Geography," *American Economic Journal: Applied Economics*, **11(4)**, 1-35.

[31] Frankel, J. A., and D. Romer (1999): "Does Trade Cause Growth?", *American Economic Review* **89**, 379-399.

[32] Gallup, J., Mellinger, A., Sachs, J. (2000). "The economic burden of malaria". Harvard University, Center for International Development Working Paper No. 1.

[33] George, E.I. and R.E. McCulloch (1997): "Approaches for Bayesian Variable Selection", *Statistica Sinica* **7**, 339-373.

[34] Hall, B.,H. and Jaffe, A.B. and Trajtenberg, M. (2001), "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools", *NBER Working Paper 8498*.

[35] Harrison, A., (1996), "Openness and Growth: A time-series, cross-country analysis for developing countries", *Journal of Development Economics*, **48**, 419-447.

[36] Harrison, A. and G. Hanson, (1999), "Who gains from trade reform? Some remaining puzzles". *Journal of Development Economics*, **59(1)**, 125-154.

[37] Hausmann, R., J. Hwang and D. Rodrik (2007): "What you export matters", *Journal of Economic Growth* **12**, 1-25.

[38] Jaffe, A., Trajtenberg, M., and Henderson, R. (1993): "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *The Quarterly Journal of Economics*, **108(3)**, 577-598.

[39] Keller, W. (1998): "Are international R&D spillovers trade-related?: Analyzing spillovers among randomly matched trade partners." *European Economic Review*, **42(8)**, **1469-1481**.

[40] Krugman, Paul (1991): "Increasing returns and economic geography," *Journal of Political Economy* **99**, 483-499.

[41] Leamer, E. (1978): *Specification Searches*, John Wiley & Sons, New York.

[42] Ley, E. and M. Steel (2009): "On the effect of prior assumptions in Bayesian model averaging with applications to growth regression", *Journal of Applied Econometrics* **24**, 651-674.

[43] Madigan, D. and J. York (1995): "Bayesian Graphical Models for Discrete Data", *International Statistical Review* **63**, 215-232.

[44] Magnus, J.R., O. Powell and P. Prüfer (2010): "A comparison of two model averaging techniques with an application to growth empirics", *Journal of Econometrics*, **154**, 139-153.

[45] Melitz, M. J. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, **71**, 1695-1725.

[46] Melitz, M.J. and S.J. Redding (2014), "Heterogeneous Firms and Trade", in Gopinath–Helpman–Rogoff eds., *Handbook of International Economics*, **4**, 1-54

[47] Moral-Benito, E. (2013): "Likelihood-based Estimation of Dynamic Panels with Predetermined Regressors", *Journal of Business and Economic Statistics* **31**, 451-472.

[48] Moral-Benito, E. (2015), "Model Averaging in Economics: An Overview". *Journal of Economic Surveys*, **29**, 46-75.

[49] Moral-Benito, E. (2016) Growth Empirics in Panel Data Under Model Uncertainty and Weak Exogeneity. *Journal of Applied Econometrics*, **31**, 584-602.

[50] Raftery, A. (1995): "Bayesian Model Selection in Social Research", *Sociological Methodology* **25**, 111-163.

[51] Raftery, A.E. and S.M. Lewis (1992): "How many iterations in the Gibbs sampler?", *Bayesian Statistics* **4**, 763-773.

[52] Rodriguez, F. (2007): "Openness and growth: what have we learned?", *United Nations, Department of Economics and Social Affairs*, Working Papers **51**.

[53] Rodriguez, F., and D. Rodrik (2001): "Trade Policy and Economic Growth: A Skeptics Guide to the Cross-National Evidence", *NBER Macroeconomics Annual 2000*, ed. by B. Bernanke and K. Rogoff.

[54] Sachs, J. D., and A. Warner (1995): "Economic Reform and the Process of Global Integration", *Brookings Papers in Economic Activity* **1**, 1-118.

[55] Sala-i-Martin, X. (1997a). "I just ran 4 million regressions". *NBER, Working Paper No. 6252.*

[56] Sala-i-Martin, X. (1997b). "I just ran 2 million regressions". *American Economic Review* **87 (2)**, 178-183.

[57] Sala-I-Martin, X., G. Doppelhofer and R.I. Miller (2004): "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach", *American Economic Review* **94**, 813-835.

[58] Samuelson, Paul (1954): "The Transfer problem and transport costs, II: Analysis of effects of trade impediments". *Economic Journal* **64**, 264-289.

[59] Winters, L.A. (2004): "Trade Liberalization and Economic Performance: An Overview", *Economic Journal* **114**, F4-F21.

Appendix

A Computation of the Globalization index

Our starting point is an $(n \times n)$-matrix A, which is row-stochastic, as the one constructed in Section 2.2. We think of it as the adjacency matrix of a weighted directed network over n nodes. Thus each entry a_{ij} is the relative weight with which node i connects to node j. Viewing such normalized weights as probabilities, the directed distance φ_{ij} from i to j is then identified as the expected number of steps required to reach j from i when, at every node $k = 1, 2, ..., n$, each possible link kl is chosen with probability a_{kl}. In our model, those paths reflect the transfer of information (or know-how) from one country to another, which occur with intensities that are proportional to the trades in the goods and services that embody that information.

To compute such expected magnitude, it is useful to consider the $(n-1) \times (n-1)$ matrix A_{-j} obtained from A by deleting the jth row and the jth column. (This matrix, of course, is no longer a stochastic matrix.) Then, it can be easily seen that the probability that a path that started at i is at $k \neq j$ after r steps is simply $[(A_{-j})^r]_{ik}$, where $(A_{-j})^r$ is the rth-fold composition of A_j with itself and $[\cdot]_{ik}$ stands for the ik-entry of the matrix $[\cdot]$. Thus, the probability that it visits node j for the first time in step $r+1$ is simply

$$\gamma_{ij}(r+1) = \sum_{k \neq j} [(A_{-j})^r]_{ik} \ a_{kj}.$$

Therefore, the expected number of steps φ_{ij} can be obtained as follows:

$$\begin{aligned} \varphi_{ij} &= \sum_{r=1}^{\infty} r \ \gamma_{ij}(r) = \sum_{r=0}^{\infty} (r+1) \ \sum_{k \neq j} [(A_{-j})^r]_{ik} \ a_{kj} \\ &= \sum_{k \neq j} \sum_{r=1}^{\infty} r \left[(A_{-j})^{r-1}\right]_{ik} \ a_{kj} = \left[\left(\sum_{r=1}^{\infty} r \ (A_{-j})^{r-1} \right)_{ik} \right]_{\substack{k=1,2,...,n \\ k \neq j}} \left(a_{kj} \right)_{\substack{k=1,2,...,n \\ k \neq j}} \end{aligned} \tag{17}$$

Using now a standard formula from linear algebra we have:

$$\sum_{r=1}^{\infty} r \ (A_{-j})^{r-1} = (I - A_{-j})^{-2}$$

so that, in an integrated matrix form, the (column) vector $\left(\varphi_{ij}\right)_{\substack{i=1,2,...,n \\ i \neq j}}$ can be written as follows

$$\left(\varphi_{ij}\right)_{\substack{i=1,2,...,n \\ i \neq j}} = (I - A_{-j})^{-2} \ \left(a_{ij}\right)_{\substack{i=1,2,...,n \\ i \neq j}}.$$

Finally, note that, because A is a row-stochastic matrix, it follows that

$$a_{ij} = 1 - \sum_{k \neq j} a_{ik}$$

and therefore

$$\Big(a_{ij}\Big)_{\substack{i=1,2,...,n \\ i \neq j}} = (I - A_{-j})\ e$$

where e is the column vector $(1, 1, ..., 1)^\top$. Hence the vector $\Big(\varphi_{ij}\Big)_{\substack{i=1,2,...,n \\ i \neq j}}$ can be computed from the following simple expression:

$$\begin{aligned} \Big(\varphi_{ij}\Big)_{\substack{i=1,2,...,n \\ i \neq j}} &= (I - A_{-j})^{-2} (I - A_{-j})\ e \\ &= (I - A_{-j})^{-1}\ e. \end{aligned}$$

B Empirical model

We follow the approach developed in Moral-Benito (2013, 2016) and augment the dynamic panel model of Section 5 by a feedback process which relates the predetermined variables to all lags of the explained variable, all lags of the predetermined variables, and the exogenous variables. Moreover, we transform the augmented model to obtain a simultaneous-equation representation. This representation has proven useful because it facilitates the estimation of the model by allowing a concentration of the parameters of the model's log-likelihood. Thus, for each country i, the model consists of a system of $T + (T-1)k$ equations, where T is the total number of time periods. Using matrix notation, we can write the model compactly as:

$$A\boldsymbol{R}_i = B\boldsymbol{Z}_i + \boldsymbol{U}_i \tag{18}$$

where the following definitions are used:

$$\begin{aligned} &\boldsymbol{R}_i = (\boldsymbol{y}_i, \mathbf{x}_i)' && \boldsymbol{y}_i = (y_{i1}, y_{i2}, ..., y_{iT})' \\ &&& \mathbf{x}_i = (\mathbf{x}_{i2}, \mathbf{x}_{i3}, ..., \mathbf{x}_{iT})' && \mathbf{x}_{it} = \left(x_{it}^1, x_{it}^2, ..., x_{it}^k\right)' \\ &\boldsymbol{Z}_i = (y_{i0}, \mathbf{x}_{i1}, \mathbf{z}_i)' && \mathbf{z}_i = \left(z_i^1, z_i^2, ..., z_i^m\right)' \\ &\boldsymbol{U}_i = (\epsilon_i + \mathbf{v}_i, \boldsymbol{\xi}_i)' && \mathbf{v}_i = (v_{i1}, v_{i2}, ..., v_{iT})' \\ &&& \boldsymbol{\xi}_i = (\boldsymbol{\xi}_{i2}, \boldsymbol{\xi}_{i3}, ..., \boldsymbol{\xi}_{iT})' && \boldsymbol{\xi}_{it} = \left(\xi_{it}^1, \xi_{it}^2, ..., \xi_{it}^k\right)' \end{aligned}$$

$$A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & \mathbf{I} \end{pmatrix} \qquad A_{11} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\alpha & 1 & 0 & \cdots & 0 \\ 0 & -\alpha & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -\alpha & 1 \end{pmatrix} \qquad A_{12} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ -\boldsymbol{\beta} & 0 & \cdots & 0 \\ 0 & -\boldsymbol{\beta} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\boldsymbol{\beta} \end{pmatrix}$$

where $\mathbf{I}$ is an identity matrix of dimension $(T-1)k \times (T-1)k$, ϵ_i can be interpreted as an individual-specific effect and $\boldsymbol{\xi}_{it}$ is a $k \times 1$ vector of prediction errors. Furthermore, we have:

$$B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \qquad B_1 = \begin{pmatrix} \alpha + \gamma_y & \boldsymbol{\beta} + \boldsymbol{\gamma} & \boldsymbol{\delta} \\ \gamma_y & \boldsymbol{\gamma} & \boldsymbol{\delta} \\ \vdots & \vdots & \vdots \\ \gamma_y & \boldsymbol{\gamma} & \boldsymbol{\delta} \end{pmatrix} \qquad B_2 = \begin{pmatrix} \boldsymbol{\pi}_{2y} & \boldsymbol{\pi}_{2x} & \boldsymbol{\pi}_{2z} \\ \boldsymbol{\pi}_{3y} & \boldsymbol{\pi}_{3x} & \boldsymbol{\pi}_{3z} \\ \vdots & \vdots & \vdots \\ \boldsymbol{\pi}_{Ty} & \boldsymbol{\pi}_{Tx} & \boldsymbol{\pi}_{Tz} \end{pmatrix}$$

$$\boldsymbol{\beta} = \left(\beta^1, \beta^2, ..., \beta^k\right) \qquad \boldsymbol{\gamma} = \left(\gamma^1, \gamma^2, ..., \gamma^k\right) \qquad \boldsymbol{\delta} = \left(\delta^1, \delta^2, ..., \delta^m\right)$$

$$\boldsymbol{\pi}_{ty} = \begin{pmatrix} \pi_{ty}^1 \\ \pi_{ty}^2 \\ \vdots \\ \pi_{ty}^k \end{pmatrix} \qquad \boldsymbol{\pi}_{tx} = \begin{pmatrix} \pi_{tx}^{11} & \pi_{tx}^{12} & \ldots & \pi_{tx}^{1k} \\ \pi_{tx}^{21} & \pi_{tx}^{22} & \ldots & \pi_{tx}^{2k} \\ \vdots & \vdots & & \vdots \\ \pi_{tx}^{k1} & \pi_{tx}^{k2} & \ldots & \pi_{tx}^{kk} \end{pmatrix} \qquad \boldsymbol{\pi}_{tz} = \begin{pmatrix} \pi_{tz}^{11} & \pi_{tz}^{12} & \ldots & \pi_{tz}^{1m} \\ \pi_{tz}^{21} & \pi_{tz}^{22} & \ldots & \pi_{tz}^{2m} \\ \vdots & \vdots & & \vdots \\ \pi_{tz}^{k1} & \pi_{tz}^{k2} & \ldots & \pi_{tz}^{km} \end{pmatrix}.$$

Under normality of the random disturbances, the model in (18) gives rise to the following log-likelihood function:

$$\mathcal{L}(\mathbf{y}, \mathbf{X}|\mathbf{Z}, \boldsymbol{\theta}) \propto -\frac{N}{2} \log |\boldsymbol{\Omega}| - \frac{1}{2} tr\left(\boldsymbol{\Omega}^{-1} \mathbf{U}\mathbf{U}'\right) \tag{19}$$

where $\mathbf{y}$, $\mathbf{X}$ and $\mathbf{Z}$ are the observations on $\mathbf{y}_i$, $\mathbf{x}_i$ and $\mathbf{z}_i$ for all N countries in the sample, $\boldsymbol{\theta}$ is the vector of model parameters, and $\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2, ..., \mathbf{U}_N]$. Moreover, $\boldsymbol{\Omega}$ is the variance-covariance matrix of $\mathbf{U}$ and $tr(\cdot)$ denotes the trace of the corresponding matrix. Notice that the following simplification was made $\sum_{n=1}^{N} \mathbf{U}_n' \boldsymbol{\Omega}^{-1} \mathbf{U}_n = tr(\boldsymbol{\Omega}^{-1} \mathbf{U}\mathbf{U}')$. Also notice that the determinant of A is equal to unity.

C Integrated likelihood

The integrated likelihood used in Equation (13) is defined as follows:

$$p(\mathbf{y}|M_j) = \int p(\mathbf{y}|M_j, \boldsymbol{\theta}) f(\boldsymbol{\theta}|M_j) d\boldsymbol{\theta} \tag{20}$$

where $p(\mathbf{y}|M_j, \boldsymbol{\theta})$ is the conditional likelihood of the data. The expression in (20) is typically hard to evaluate, but there exists a simple and accurate approximation of it, the so-called BIC approximation which makes use of Laplace's method. Let $m(\boldsymbol{\theta}) = \log(p(\mathbf{y}|M_j, \boldsymbol{\theta}) f(\boldsymbol{\theta}|M_j))$ denote the posterior mode, and construct a Taylor-series expansion of $m(\cdot)$ around $\tilde{\boldsymbol{\theta}}$, where $\tilde{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} m(\boldsymbol{\theta})$:

$$m(\boldsymbol{\theta}) = m(\tilde{\boldsymbol{\theta}}) + (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})' m'(\tilde{\boldsymbol{\theta}}) + \frac{1}{2}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})' m''(\tilde{\boldsymbol{\theta}})(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}) \tag{21}$$

where m' and m'' are the first and second derivative of m, respectively. $m(\boldsymbol{\theta})$ reaches its maximum at $\tilde{\boldsymbol{\theta}}$, therefore $m'(\tilde{\boldsymbol{\theta}}) = 0$, and Equation (21) becomes

$$m(\boldsymbol{\theta}) = m(\tilde{\boldsymbol{\theta}}) + \frac{1}{2}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})' m''(\tilde{\boldsymbol{\theta}})(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}) \tag{22}$$

Inserting (22) into the integral gives:

$$p(\mathbf{y}|M_j) = \int e^{m(\tilde{\boldsymbol{\theta}}) + \frac{1}{2}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})' m''(\tilde{\boldsymbol{\theta}})(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})} d\boldsymbol{\theta} = e^{m(\tilde{\boldsymbol{\theta}})} \int e^{\frac{1}{2}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})' m''(\tilde{\boldsymbol{\theta}})(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})} d\boldsymbol{\theta} \tag{23}$$

The integral is a Gaussian integral and, therefore we get the following expression:

$$p(\mathbf{y}|M_j) = e^{m(\tilde{\boldsymbol{\theta}})} (2\pi)^{\frac{k}{2}} | - m''(\tilde{\boldsymbol{\theta}})|^{-\frac{1}{2}} \tag{24}$$

where k and $| - m''(\tilde{\boldsymbol{\theta}})|$ are, respectively, the rank and the determinant of $-m''(\tilde{\boldsymbol{\theta}})$. In large sample $\tilde{\boldsymbol{\theta}} \approx \hat{\boldsymbol{\theta}}$, where $\hat{\boldsymbol{\theta}}$ is the maximum likelihood estimator of $\boldsymbol{\theta}$. By taking logs, we obtain:

$$\log p(\mathbf{y}|M_j) = \log p(\mathbf{y}|M_j, \hat{\boldsymbol{\theta}}) + \log f(\hat{\boldsymbol{\theta}}|M_j)) + \frac{k}{2}\log(2\pi) - \frac{1}{2}\log| - m''(\tilde{\boldsymbol{\theta}})| \tag{25}$$

Following Raftery (1995), in large samples, $-m''(\hat{\boldsymbol{\theta}}) \approx N\mathbf{I}$, where N is the number of observations and $\mathbf{I}$ is the expected Fisher information matrix. Using that, we get $| - m''(\hat{\boldsymbol{\theta}})| \approx N^k|\mathbf{I}|$ and:

$$\log p(\mathbf{y}|M_j) = \log p(\mathbf{y}|M_j, \hat{\boldsymbol{\theta}}) + \log f(\hat{\boldsymbol{\theta}}|M_j)) + \frac{k}{2}\log(2\pi) - \frac{k}{2}\log N - \frac{1}{d}\log|\mathbf{I}| \tag{26}$$

The first and the fourth term on the right-hand side of this expression are of order N and $\log N$ respectively, whereas all other terms are of order 1 or less. When we remove these terms we arrive at the following expression for the (approximated) integrated likelihood:

$$\log p(\mathbf{y}|M_j) = \log p(\mathbf{y}|M_j, \hat{\boldsymbol{\theta}}) - \frac{k}{2}\log N \tag{27}$$

This expression is well-known and it is very similar to the Akaike information criterion. With this expression at hand, we are almost ready to compute the posterior model probability given in (13). One more step is required since the model in (18) does not give us $p(\mathbf{y}|M_j, \hat{\boldsymbol{\theta}})$ but rather $p(\mathbf{y}, \mathbf{X}_j|M_j, \hat{\boldsymbol{\theta}})$, which is the joint conditional likelihood of $(y, \mathbf{X}_j)$, with M_j containing the relevant $\mathbf{Z}$-regressor variables.

In the BMA, we consider different models each consisting of a particular combination of regressor variables. If we were to use the joint likelihood $p(y, \mathbf{X}_j|\cdot)$ we would compare different likelihoods, for instance, $p(\mathbf{y}, \mathbf{X}^1, \mathbf{X}^2, ...\mathbf{X}^k|\cdot)$ and $p(\mathbf{y}, \mathbf{X}^4, \mathbf{X}^5, ...\mathbf{X}^k|\cdot)$ which are, in fact, not comparable. Thus, instead, we proceed as follows. For a given model M_j, we, first, maximize (19) to obtain the maximum likelihood estimate of θ_j. Then we compute the likelihood of the outcome variable y conditional on the estimated model, that is $p(\mathbf{y}|M_j, \hat{\boldsymbol{\theta}}_j)$. Most importantly, this statistic is comparable across the different models and hence we can use this expression to compute the posterior probability of the underlying model. The conditional likelihood $p(\mathbf{y}|M_j, \hat{\boldsymbol{\theta}}_j)$ can be obtained in a relatively straightforward manner by transforming the model given in (18) as follows:

Given $\hat{\boldsymbol{\theta}}$, we, first, substitute the feedback process into the outcome-equation which yields:

$$y_{n,1} = (\hat{\alpha} + \hat{\gamma}_0)y_{n,0} + \left(\hat{\boldsymbol{\gamma}} + \hat{\boldsymbol{\beta}}\right)\mathbf{x}_{n,1} + \hat{\boldsymbol{\delta}}\mathbf{z}_n + \epsilon_n + v_{n,1} \tag{28}$$

and for $t = 2, ..., T$, we get:

$$y_{n,t} = \hat{\alpha}y_{n,t-1} + \left[\hat{\gamma}_0 + \hat{\boldsymbol{\beta}}\hat{\boldsymbol{\pi}}_{t0}\right]y_{n,0} + \left[\hat{\boldsymbol{\gamma}} + \hat{\boldsymbol{\beta}}\hat{\boldsymbol{\pi}}_{t1}\right]\mathbf{x}_{n,1} + \left[\hat{\boldsymbol{\delta}} + \hat{\boldsymbol{\beta}}\hat{\boldsymbol{\pi}}_{t2}\right]\mathbf{z}_n + \hat{\boldsymbol{\beta}}\boldsymbol{\xi}_{n,t} + \epsilon_n + v_{n,t} \tag{29}$$

For each country observation i, the model in (28)-(29) is a system of T equations which can be compactly written as:

$$\mathcal{A}\mathbf{y}_i = \mathcal{B}\mathbf{Z}_i + \mathcal{C}\mathbf{U}_i \tag{30}$$

where the following definitions are applied:

$$\mathcal{A} = \hat{A}_{11} \qquad \mathcal{B} = \begin{bmatrix} 0 \\ \left(\mathbf{I}_{T-1} \otimes \hat{\boldsymbol{\beta}}\right)\hat{B}_2 \end{bmatrix} + \hat{B}_1 \qquad \mathcal{C} = \left[\mathbf{I}, -\hat{A}_{12}\right].$$

$\mathbf{I}_{T-1}$ is an identity matrix of order $T-1$. The variables $\mathbf{y}_i$, $\mathbf{Z}_i$ and $\mathbf{U}_i$ are defined as above in (18) together with the matrices $\hat{A}_{11}$, $\hat{\boldsymbol{\beta}}$, $\hat{B}_2$, $\hat{B}_1$, $\hat{A}_{12}$ which are evaluated at the ML-estimate $\hat{\boldsymbol{\theta}}$. Finally, we write the log-likelihood of observation $\mathbf{y}$, conditional on $\mathbf{Z}$ and $\hat{\boldsymbol{\theta}}$ as follows:

$$\log p(\mathbf{y}|M_j, \hat{\boldsymbol{\theta}}) \propto -\frac{N}{2}\log|\mathcal{C}\hat{\boldsymbol{\Omega}}\mathcal{C}'| - \frac{1}{2}tr(\hat{\boldsymbol{\Omega}}^{-1}\mathbf{U}\mathbf{U}'). \tag{31}$$

The expression in (31) is substituted into (27) to obtain the approximated integrated likelihood.

D Data

		Mean	Median	Std	Min	Max
1.	Logarithm of real GDP per capita	8.35	8.39	1.30	5.19	10.72
	Total population (mill.)	38.9	7.98	124	0.35	1148
	Annual growth rate of population	0.02	0.02	0.01	-0.01	0.06
	Price level of investment	0.93	0.64	2.06	0.11	21.5
	Exports plus imports as a share of GDP	0.53	0.46	0.36	0.04	2.90
	Consumption share of GDP	0.72	0.72	0.15	0.23	1.32
	Investment share of GDP	0.22	0.21	0.09	0.03	0.57
	Government share of GDP	0.10	0.08	0.06	0.02	0.39
	Labor force participation rate	0.39	0.39	0.08	0.19	0.57
2.	Life expectancy at birth, in years	59.9	61.8	12.0	30.3	78.8
	Population density, people per km^2	120	37	391	1.4	4547
	Urban population, % of total	0.45	0.43	0.24	0.02	1.00
	Population aged 0-14, % of total	0.38	0.41	0.09	0.16	0.50
	Population share aged 65+, % of total	0.06	0.04	0.04	0.01	0.18
3.	Sachs & Warner openness measure	0.50	0.50	0.50	0.00	1.00
4.	Democracy index	0.58	0.70	0.38	0.00	1.00
5.	1/0 dummy for former Spanish colony	0.21	0.00	0.41	0.00	1.00
6.	1/0 dummy for armed conflict	0.16	0.00	0.37	0.00	1.00
7.	Land area in km^2 (thousand)	1026	272	2115	0.61	9590
	Land share in the geographic tropics	0.55	0.95	0.48	0.00	1.00
	Population share in the geographic tropics	0.51	0.78	0.49	0.00	1.00
	Land area within 100km of navigable water	0.48	0.38	0.37	0.00	1.00
	Land share in Koeppen-Geiger tropics	0.38	0.06	0.42	0.00	1.00
	1/0 dummy for landlocked country	0.16	0.00	0.37	0.00	1.00
	Air distance to NYC, Rotterdam, Tokyo	4205	4065	2594	140	9590
	Timing of national independence	0.96	1.00	0.97	0.00	2.00
	1/0 dummy for socialist rule in 1950-95	0.10	0.00	0.30	0.00	1.00
8.	Average years of primary schooling	2.87	2.63	1.79	0.02	7.51
	Average years of secondary schooling	1.06	0.72	1.05	0.01	5.09
9.	Globalization Index	0.56	0.55	0.08	0.39	0.76
10.	1/0 dummy for Western European country	0.18	0.00	0.39	0.00	1.00
	1/0 dummy for Latin-American country	0.26	0.00	0.44	0.00	1.00
	1/0 dummy for East Asian country	0.11	0.00	0.31	0.00	1.00
	1/0 dummy for Sub-Saharan country	0.26	0.00	0.44	0.00	1.00

Data sources: 1. Penn World Tables, 2. World Development Indicators, 3. Sachs and Warner: "Trade Openness Indicators", Dataset: sachswarneropen.xls, 4. Polity IV Project: Regime Authority Characteristics and Transitions Datasets: p4v2010.xls, 5. Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) geo_cepii.xls, 6. Uppsala Conflict Data Program (UCDP), Dataset: 64464_UCDP_PRIO_ArmedConflictDataset_v42011.xls, 7. Gallup, Mellinger, Sachs, Harvard University Center for International Development, Datasets: physfact_rev.csv (Physical geography and population), kgzones.csv (Köppen-Geiger Climate zones), geodata.csv (Geography and Economic Development), 8. Barro and Lee 2000, Dataset: appendix_data_tables_in_panel_set_format.xls, 9. UN Comtrade

Table 10: Data: Sources and descriptive statistics.

Asia: Afghanistan, Armenia, Azerbaijan, Bahrain, Bangladesh, Bhutan, Brunei, Cambodia, *China*, Georgia, Hong Kong, *India*, *Indonesia*, *Iran*, Iraq, *Israel*, *Japan*, *Jordan*, Democratic Republic of Korea, *Republic of Korea*, Kuwait, Kyrgyzstan, Laos, Lebanon, Macao, *Malaysia*, Maldives, Mongolia, Myanmar, *Nepal*, Oman, *Pakistan*, *Philippines*, Qatar, Saudi Arabia, Singapore, *Sri-Lanka*, *Syria*, Tajikistan, *Thailand*, *Turkey*, Turkmenistan, United Arab Emirates, Uzbekistan, Vietnam, Yemen, Former Yemen
Europe: Albania, Andorra, *Austria*, Belarus, *Belgium*, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Former Czechoslovakia, Czech Republic, *Denmark*, Estonia, *Finland*, *France*, Germany, East Germany, Former USSR, Gibraltar, *Greece*, Hungary, Iceland, *Ireland*, *Italy*, Kazakhstan, Latvia, Lithuania, Luxembourg, Macedonia, Malta, Moldova, *Netherlands*, *Norway*, Poland, *Portugal*, Romania, Russia, San Marino, Serbia-Montenegro, Slovakia, Slovenia, *Spain*, *Sweden*, *Switzerland*, Ukraine, *United Kingdom*, Former Yugoslavia
Africa: *Algeria*, Angola, Benin, Botswana, Burkina Faso, Burundi, *Cameroon*, Cape Verde, *Central African Republic*, Chad, Comoros, *Democratic Republic of Congo*, *Republic of Congo*, Cote d'Ivoire, Djibouti, *Egypt*, Equatorial Guinea, Eritrea, Ethiopia, Gabon, *Gambia*, *Ghana*, Guinea, Guinea-Bissau, *Kenya*, Kiribati, Lesotho, Liberia, Libya, Madagascar, *Malawi*, *Mali*, Mauritania, *Mauritius*, Morocco, *Mozambique*, Namibia, *Niger*, Nigeria, *Rwanda*, *Senegal*, Seychelles, *Sierra Leone*, Somalia, *South Africa*, Sudan, Swaziland, Tanzania, *Togo*, *Uganda*, *Tunisia*, *Zambia*, *Zimbabwe*
North America: Antigua & Barbuda, Bahamas, Barbados,*Belize*, Bermuda, *Canada*, *Costa Rica*, Cuba, Dominia, *Dominican Republic*, *El Salvador*, Grenada, Greenland, *Guatemala*, *Haiti*, *Honduras*, *Jamaica*, *Mexico*, Netherlands Antilles, *Nicaragua*, Former Panama, *Panama*, Saint Kitts-Nevis, Saint Lucia, Saint Vincent, *Trinidad-Tobago*, *United States*
South America: *Argentina*, Aruba, *Bolivia*, *Brazil*, *Chile*, *Colombia*, *Ecuador*, *El Salvador*, Guyana, *Paraguay*, *Peru*, Suriname, *Uruguay*, *Venezuela*
Australia: *Australia*, Fiji, French Polynesia, Marshall Islands, Micronesia, New Caledonia, *New Zealand*, Palau, *Papua New Guinea*, Solomon Islands, Samoa, Tonga, Tuvalu
Countries in *italic* are included in the Bayesian model averaging analysis.

Table 11: Sample of countries

E Markov chain - Monte Carlo - Model Composition

Here we describe the first-order Markov that, as explained in Section 5, approximates the posterior probability distribution induced by our BMA analysis. This Markov chain evolves according to the following transition kernel. Suppose the current state of the chain is M_j. Then, a candidate model is sampled from the neighborhood of M_j, where the neighborhood consists of the set of models with either one variable more or one variable less than in M_j. The candidate model, denoted by $M_{j'}$, is then "compared" to M_j and it is accepted with probability $\min\{1, \frac{P(M_{j'}|\mathbf{y})}{P(M_j|\mathbf{y})}\}$. If the candidate model is accepted then the Markov chain moves to $M_{j'}$, otherwise it stays at M_j. The ratio $\frac{P(M_{j'}|\mathbf{y})}{P(M_j|\mathbf{y})}$ is the posterior odds (= prior odds $\times$ Bayes Factor) and it measures how much the data supports one model over the other. The posterior odds for M_j and $M_{j'}$ is given by:

$$\frac{p(M_{j'}|\mathbf{y})}{p(M_j|\mathbf{y})} = \frac{p(\mathbf{y}|M_{j'})}{p(\mathbf{y}|M_j)} \times \frac{p(M_{j'})}{p(M_j)}$$

Here, $p(\mathbf{y}|M.)$ and $p(M.)$ are the integrated likelihood and the prior probability of a given model, respectively.

To check the mixing and convergence properties of the simulated chain, we compute the following diagnostic statistics. First, we compute the statistic $Corr(\Pi, Freq)$ tests for convergence of the Markov Chain, which consists of the following steps: (1) discard the first S_0 steps of the simulated Markov chain to eliminate possible effects from influential starting values; (2) split the remaining chain into two parts: the first S_1 steps and the subsequent S_2 steps; (3) compute the transition matrix T_1, where an element of T_1, say t_{ij}, records how many times the chain has moved from model m_i to model m_j. The dimension of T_1 is equal to the number of different models in S_1; (4) convert T_1 into the transition probability matrix P_1. An element of P_1, say p_{ij}, is determined as $t_{ij}/\sum_{k=1}^{\dim(T)} t_{ik}$ and it measures the probability of the chain to move from m_i to m_j, conditional on being in m_i; (5) calculate the ergodic probability of being in m_i (from P_1^{∞}), which gives the unconditional probability of observing model m_i; (6) derive, for every $m_i \in S_1$, the empirical frequency in S_2 as $c_i/\dim(S_2)$, where c_i counts how often model m_i is visited in S_2; (7) denote by $Corr(\Pi, Freq)$ the correlation coefficient between the ergodic probabilities of all models in S_1 and their empirical frequencies in S_2. $Corr(\Pi, Freq)$ approaches one when the Markov chain reaches stationarity. This is because any two subsets of a stationary chain give rise to the same stationary distribution, and the stationary distribution is (in a large sample) identical to the empirical frequency of each state.

Second, we also compute the statistic $Corr(Bayes, Freq)$ which is another stationarity test that involves the following steps: (1) eliminate a burn-in period from the simulated Markov chain and identify the model with the highest posterior probability, denoting it by $\bar{m}$; (2) compute the empirical frequency for each model in the chain and denote it by f_i; (3) calculate the relative frequency for each model with respect to the best model: $f_i/f_{\bar{m}}$; (4) determine the Bayes factor for each model with respect to the best model: $b_i/b_{\bar{m}}$ [the Bayes factor is the ratio of the posterior probabilities of two models]; (5) compute the correlation coefficient $Corr(Bayes, Freq)$ between $f_i/f_{\bar{m}}$ and $b_i/b_{\bar{m}}$. $Corr(Bayes, Freq)$ approaches 1 as the chain reaches stationarity. This is because the model selection along the chain is based upon the Bayes factor (the probability that the chain accepts to move to a candidate model is equal to the Bayes factor between the current model and the candidate model), and as a result, the chain visits those models more often which have a high posterior probability.

Thirdly, we derive the **Raftery-Lewis** dependence factor which is a measure for the mixing behavior of the Markov chain. Dependence factors above 5 are critical and indicate bad mixing of the chain or influential starting values – see Raftery and Lewis (1992) for details (the parameter values required in the test are as in Raftery and Lewis (1992) and given by $q = 0.025$, $r = 0.005$, $s = 0.95$, $\epsilon = 0.001$). To obtain an accurate representation of the posterior distribution, it is important that the chain explores those areas in the model space which have a high probability mass. We follow George and McCulloch (1997) and use a capture-recapture algorithm to estimate what fraction of the total posterior probability mass the Markov Chain has visited.

In Table 12, we report a number of statistics describing the properties of the simulated Markov chain. *Markov Steps* refers to the total number of steps (in 1000) of the simulated

chain. *Posterior model size* refers to the posterior model size. *Models covering* 50% is the number of models with the highest posterior model probability which, in sum, account for 50% of the posterior model probability. *P(max)* is the maximum posterior model probability achieved by a single model. *Visited probability* refers to the estimated fraction of the total posterior probability mass that the Markov Chain has visited. This number is computed by using the capture-recapture algorithm described in George and McCulloch (1997). The remaining statistics describe the convergence and mixing properties of the simulated chain. Generally, the values of these indicators indicate very good mixing and convergence properties of the simulated Markov chain. For example, the values of Corr(Π,Freq) and Corr(Bayes,Freq) are very close to unity, suggesting that the simulated Markov chain has reached stationarity. Furthermore, we obtain a Raftery-Lewis factor equal to 3.38 which indicates fast mixing of the process. Factors above 5 are critical and indicate bad mixing of the chain or influential starting values. Lastly, the estimate for the total posterior probability mass that the Markov Chain has visited is very high and equal to 98%. The high value is reassuring because an accurate representation of the posterior distribution requires that the Markov chain reaches the areas in the model space with high probability mass.

	Benchmark	**Higher-order trade**
Markov steps (× 1000)	836	996
Posterior model size	8.7	8.3
Models covering 50%	58	97
Pr(best model)	7.20	4.78
Visited probability	98.0	95.7
Corr(Π,Freq)	0.997	0.909
Corr(Bayes,Freq)	0.998	0.966
Raftery-Lewis factor	3.38	3.42

Table 12: MC^3 statistics.

F Robustness

As advanced in Section 8, here we explore the sensitivity of the findings in Section 6 to variations in the data input, as well as to modifications of the underlying model assumptions, and to alternative measures of network centrality.

F.1 Data

Raw data vs. cleaned data: In our baseline approach, we use the raw trade data from the UN Comtrade to compute the GI. There exists, however, a National Bureau of Economic Research project led by Robert Feenstra that has systematically cleaned the UN Comtrade data from a number of inconsistencies. The resulting data set is available from the Center for International Data and a detailed description of it is provided in Feenstra et al. (2005). As a

robustness check, we use these data instead of the raw trade data to compute our Globalization Index. Then, we perform a BMA analysis where we include this new measure. The row labeled *Feenstra* in Table 13 shows the resulting findings are very similar to the baseline results.

	$E(\theta_k\|\mathbf{y})$	PIP	$\%_{sig}$
Benchmark	**6.289*****	**85**	**99**
Feenstra	5.497***	77	97
IMF DOTS	6.084***	88	96
PWT 6.2	5.589***	83	99
PWT 6.3	4.387**	95	88
PWT 7.0	7.318***	80	97
PWT 7.1, 60-09	5.823**	71	96
PWT 7.1, 5 yrs	2.169***	79	97

Table 13: Robustness: Data.

IMF DOTS: The International Monetary Fund (IMF) publishes the Direction of Trade Statistics (DOTS) which provides detailed data on bilateral trade flows. We use these data instead of the UN Comtrade data to compute the GI. Again, the results are largely unchanged – see row "IMF DOTS" in Table 13.

Penn World Tables: A number of the variables included in the empirical analysis are constructed from data taken from the Penn World Tables (PWT). Ciccone and Jarocinski (2010) raise the important concern that the results of growth empirics are often sensitive to revisions in the PWT data. We address this concern by using different releases of the PWT to compute the relevant variables. Table 13 compares the results. By and large, our baseline finding are robust to revisions of the PWT. An advantage of the recent releases of the PWT is that they extend the time period covered by the data, which allows us to consider a longer period in the BMA. Specifically, we can use the period from 1960-2010, which gives us a total of five observations for each country. Again, the results are very similar to the baseline findings. As yet another check, we also organize the data into five year time intervals (instead of using 10-year intervals) which gives us a total of 10 country observations. As can be seen in Table 13, the higher-frequency data do not lead to noteworthy changes in the sign and significance levels of the results

F.2 Model specification

In the baseline approach we use a Binomial-Beta structure as the model prior distribution. We test the robustness of this choice by using a uniform prior as in Moral-Benito (2016). Accordingly, all models are equally likely a priori, and $P(M_j) = 2^{-K}$, where K is the number of potential regressor variables. The results of the BMA analysis with the uniform prior are very close to the results of the baseline case, as can be seen from Table 14. Hence we conclude that

the assumption on the model prior distribution does not have a significant effect on the results.

	$E(\theta_k\|\mathbf{y})$	PIP	$\%_{sig}$
Benchmark	**6.29*****	**0.85**	**0.99**
Uniform prior	6.46**	0.95	0.97
Rich	9.52***	0.98	0.99
Poor	4.18**	0.79	0.84
15 best covariates	6.41***	0.99	0.99
10 best covariates	6.37***	0.99	1.00
15 worst covariates	4.93***	0.99	0.98
10 worst covariates	4.61***	0.99	1.00

Table 14: Robustness: Model.

Next, we note that there is a large degree of heterogeneity among the countries in our sample. To account for this heterogeneity, we include into the empirical analysis a large set of covariates, as well as a full set of dummy variables to control for country-, region-, and time-fixed effects.

Despite these efforts we cannot exclude the possibility of additional dependencies that we have not properly controlled for. Two particularly relevant concerns are the existence of spatially correlated shocks that affect geographically-proximate countries, and the potentially differential effect of the covariate for developed and undeveloped countries. We address these concerns in various ways. First, we cluster the data by splitting the sample into rich and poor countries. More concretely, we consider countries as poor (rich) if their GDP per person in 1960 was less (more) than 1/5 of the U.S.-level. The resulting samples consist of 48 poor countries and 34 rich countries. Then, we perform the BMA analysis on both samples separately and report the results in Table 14. Importantly, the positive relation between openness and growth is found to be very robust for both rich and poor countries but it seems somewhat stronger for the rich.

In a similar vein, we have also interacted in a separate experiment the trade share with the region fixed effect to assess whether the relation between the traditional openness measure and growth varies across regions. Interestingly, we find that the insignificant relationship between the trade share and growth that arises in the baseline case also holds across all five regions that we consider. Moreover, to account for spatially correlated shocks we interacted the region-fixed effect with time-fixed effects. And again we arrive at the conclusion that the results are very similar to the baseline results, especially for the estimate of the Globalization measure. For conciseness we do not report the results of the last two experiments here but they are available upon request.

As an additional robustness test, we check whether our main findings are sensitive to the number of regressor variables included in the empirical model. In the baseline case, we consider 34 candidate regressors. Now we include only a subset of these variables into the model. In particular, we pick those 10 (15) variables which had the highest posterior inclusion probability

in the baseline case. As an additional experiment, we select - together with the GI and initial GDP per person – those 10 (15) variables which had the lowest posterior inclusion probability. The results of the BMA analysis are in Table 14, and again we observe no significant change with respect to the baseline findings.

Finally, we recall that the analysis in Section 6 reveals a weak relationship between the traditional measures of openness – such as the trade share and the Sachs-Warner indicator – and economic growth, as indicated by low values of their posterior inclusion probability. Now we want to address the concern that this result may driven by a potential dependence between our GI and the traditional measures. Table 15 shows the posterior mean and the inclusion probability (rows) of the openness measures for the baseline case and for different combinations of included variables (column). The results in the table do not reveal any notable dependencies between the different measures.

		Baseline	Include			
			GI S&W	GI TS	GI	TS S&W
GI	$E(\theta_k\|\mathbf{y})$	6.29***	6.21***	6.15***	6.11***	
	PIP	0.85	0.83	0.81	0.80	
TS	$E(\theta_k\|\mathbf{y})$	-0.07		-0.04		0.10
	PIP	0.04		0.04		0.05
S&W	$E(\theta_k\|\mathbf{y})$	0.18***	0.18***			0.13**
	PIP	0.16	0.36			0.07

Table 15: Robustness: Openness measures.

F.3 Alternative measures of network centrality

Our Globalization index reflects a notion of network centrality that is known as closeness centrality. Other prominent notions of centrality considered in the literature are PageRank, Bonacich, Eigenvalue, or Betweenness centralities – see e.g. Bloch, Jackson, and Tebaldi (2017). Since they all behave quite similarly for the relevant parameter ranges, in order to avoid unnecessary redundancies we focus on PageRank centrality. According to PageRank centrality, a central/influential node is identified as one that is largely connected to central/influential nodes. If we denote by $\nu = (\nu_i)$ the vector specifying such an "impact" for every node i, the centrality condition can then be written as

$$\nu = (\tilde{A})^T \nu,$$

where $\tilde{A}$ is a perturbation of the adjacency matrix A defined by $\tilde{A} = \alpha A + (1-\alpha)U$, where $0 < \alpha < 1$, and U is a (stochastic) matrix with entries all equal $1/n$. The matrix $\tilde{A}$ can still be formally interpreted as the transition probability matrix of a Markov process. Such a Markov process is clearly ergodic and thus has a unique invariant distribution. This allows PageRank

to identify the centrality of any given node i as its weight in that invariant distribution, so that we may write

$$\nu = \frac{1-\alpha}{n}(I - \alpha A^{\top})^{-1}\mathbf{e} \tag{32}$$

where n is the dimension of A and $\mathbf{e}$ is a column vector of all 1's. The notion of centrality given by (32) implicitly presumes that all nodes in the network are symmetric and command the same value. But, of course, just as we did for our baseline measure introduced in Subsection 2.2, we want to account for the fact that countries are very different in relative size within the world economy. Again, this can be captured by replacing the uniform weighting embodied by the vector $\mathbf{e}$ by the alternative vector $\boldsymbol{\beta}$ (also used by our baseline measure) where each β_i captures the fraction of country i's GDP in world economy. This leads to the following modified notion of PageRank centrality:

$$\nu = \frac{1-\alpha}{n}(I - \alpha A^{\top})^{-1}\boldsymbol{\beta}, \tag{33}$$

which is the measure of integration we apply to our full sample of 200 countries and all years from 1962 to 2012. Table 16 below reports the outcome of the BMA exercise for different values of α. There we observe that the magnitude of the posterior mean estimate of the PageRank coefficient, the corresponding inclusion probability, and the $\%_{sig}$ statistic all grow monotonically with α, only achieving truly high values when this parameter is also high. These results are very much in line with those obtained for our benchmark measure of country integration, since the parameter α plays in the present case a role analogous to δ for our benchmark integration measure. Here, α determines how much PageRank is dependent on the network architecture, hence depending on the full set of paths that, directly and indirectly, join each pair of nodes. The results of Table 16, therefore, are again a manifestation of the importance that long-range indirect connections on growth even if integration were measured by the notion of PageRank centrality.

	$E(\theta_k\|\mathbf{y})$	PIP	$\%_{sig}$
PageRank centrality			
$\alpha = 0.95$	2.6332**	0.63	76
$\alpha = 0.75$	2.0135*	0.34	71
$\alpha = 0.50$	0.7264	0.11	52
$\alpha = 0.25$	0.0506	0.08	44

Table 16: Global vs. local connections: PageRank centrality.

In addition to PageRank centrality, we have experimented with a number of other integration measures that belong to none of the aforementioned centrality concepts. Most noteworthy among those is the approach suggested by Arribas et al. (2009). One of the indicators they use to assess a country's integration is what they call *Degree of Connection* (DTC), which compares the trade of a given country in the actual world with what would prevail in an ideal and perfectly integrated one. More specifically, DTC measures whether a country has its international flows

match the weight of the other countries, being equal to 1 in case of a perfect match. Clearly, this approach is conceptually very different from ours. Arribas et al. (2009) also consider the *Degree of Openness* (DO) which, for each country, is equivalent to 1 minus its corresponding diagonal element in our adjacency matrix A. These two different indicators capture a country's aggregate trade flows but not its architecture of first- and higher-order trade connections. Consequently, it is not surprising that the correlation between our integration indicator and DO and DTC is generally very low (just as we showed to be the case with the traditional measures of openness in Section 4). For example, in the year 2004, it is equal to -0.03 and -0.05, respectively. We also find an insignificant role for these indicators when included in the BMA. For instance, the posterior mean associated with the indicator DO is not significant (even at the 10% level) and the posterior inclusion probability is only 6%.[21]

Lastly, we consider three different versions of random perturbations to the diffusion matrix A in order to address the criticism expressed in Keller (1998) that a spurious version of the trade network is likely to have the same implications for the global transmission of information than the actual trade network. At the same time, the analysis below allows to assess the importance of different dimensions of the network structure for the relationship between the GI and growth. In the first case, we keep the structure of the original matrix A as in the baseline case - in terms of the number of each country's links and the set of its partners - and we just perturb the weight of existing links. In particular, we randomly assign a weight between 0 and 1 to each existing link and re-normalize the resulting matrix so that it is row-stochastic. This approach implies only a small modification to the original transition matrix A because the structure of the matrix is preserved. Using this modified version of the transition matrix, we compute the GI according to the approach described in Section 2.2. Clearly, the values of the GI depend on the realization of the random draws of the link weights. To eliminate the variation in the GI that is due to this randomness, we compute the GI for 100 different sets of realizations and average over the outcomes. Lastly, we include the resulting GI into the BMA analysis. The estimated coefficient of the GI comes out significant only at the 10 percent level and the posterior inclusion probability drops from 85% in the baseline case to 42%.

The second case that we consider involves a more substantial modification of the matrix A. For each country we keep the number and the weight of existing links but we assign the links to a randomly selected set of trading partners. That is, we reshuffle the existing links of a given country. As before, we use the perturbed transition matrix to compute the GI, then we average over 100 different realizations and include the resulting GI in the BMA. The estimated coefficient of the GI becomes insignificant and the posterior inclusion probability of only 2 percent is significantly below the baseline value. In the last case we allocate the total weight of each

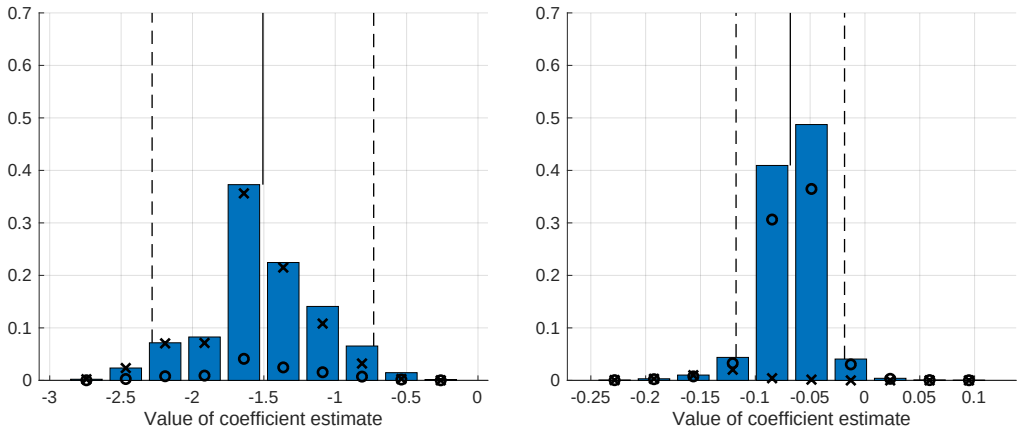
[21] In another experiment, we identify the first principle component (FPC) of trade openness and compare it to the GI. To conduct this comparison, we compute the correlation between the two variables and, in addition, include the FPC instead of the GI into the BMA. We find a correlation coefficient of -0.36 which is slightly higher (in absolute terms) than that for trade openness and the GI of -0.10. Still, the value is rather low, indicating a relatively weak relation between the two variables. When including the FPC into the BMA we find a posterior inclusion probability for this variable of less than 1 percent.

country's links to a randomly selected set of trading partners. That is, we keep the outward orientation of countries as in the baseline case but perturb the number of links. Also in this case we obtain in the BMA an insignificant coefficient estimate and very low posterior inclusion probability for the GI of 2 percent.

We interpret these findings as reflecting the importance of both, the structure of the trade network - in terms of the number of links of a country and the set of its trading partners - as well as the intensity of trade connections between countries for the explanatory power of the GI. If we keep the structure but modify the intensity of trade connections (as in the first case) then the posterior inclusion probability of the GI declines substantially but it is still higher than for 26 out of 34 the included covariates. Instead, if we perturb the set of trading partners (second case) and, in addition, also the number of links (third case), then the relation between the modified GI and growth becomes very weak.

G Explaining discordance within the BMA analysis

As can be seen from Table 6, for some of the regressor variables there is a marked misalignment between the posterior inclusion probability and the $\%_{sig}$-statistic for several of the variable included in the BMA analysis. For example, the *Government share* has a PIP of only 11% but the estimated coefficient is significant in 91% of the models. To understand this pattern it is useful to consider Figure 3, which focuses on the variables *Government share*, and *Armed conflict*. It shows the posterior probability mass over the whole range of coefficient estimates (bars) as well as, for each value of the estimated coefficient, the share of models where the estimation is significant at the 5% level (crosses) and the posterior inclusion probabilities of the respective models (circles). The solid line and the broken lines represent the posterior mean and the 95% confidence bounds, respectively.



(a) Government share ($PIP = 0.11$; $\%_{sig} = 91\%$) (b) Armed conflict, ($PIP = 0.75$; $\%_{sig} = 6\%$)

Figure 3: Discordance between PIP and $\%_{sig}$.

A comparison of the two panels yields some useful insights. It illustrates, in particular, that the posterior inclusion probability of a variable and the share of significant coefficient estimates can be very different. In Panel (a) we observe that the coefficient associated with the *Government share* is typically estimated very precisely across models (crosses are close to the top of each bar), while the models which contained this variable are generally characterized by a low goodness-of-fit (circles close to the bottom). As a result, the posterior inclusion probability of these models is rather low across the entire range of the estimated coefficient. The opposite can be observed in Panel (b), which shows the same set of statistics for the variable *Armed conflict*. We find that the coefficient for *Armed conflict* is generally very imprecisely estimated, whereas the models which include it have a high goodness-of-fit and thus provide this variable with a high posterior inclusion probability.

Thus, in sum, the point here is that the identification of robust covariates according to the posterior inclusion probability (as done by model averaging approach) can lead to conclusion that are very different from the traditional (single-equation) growth empirics that typically evaluates variables on the basis of the significance level of the estimated coefficient for a certain model specification. As a result of this practice, much of the empirical growth literature considers the variable *Government share* as robustly related to growth (see for example, the work by Barro (1991, 1996) and Caselli et al (1996)) whereas the results above lead to conclude the exact opposite. The same applies (but in reverse order) to the variable *Armed conflict*. With a posterior inclusion probability of 75% this variable is found to be strongly related to growth. This result is in stark contrast to much of the existing empirical work which interprets the mostly insignificant coefficient estimates for this variable as evidence for a limited explanatory role. See, for example, Barro and Lee (1994) and Easterly and Levine (1997). Such contradictory assessment can be established also for several other candidate regressors, such as the *Investment price* (Easterly (1993)), the *Life expectancy* Barro and Lee (1994), *Democracy* (Barro (1996), Dollar and Kraay (2003)), *Landlocked* (Easterly and Levine (2001)), or *Former Spanish colony* (Barro 1996), all of which have been suggested to be important for economic growth. Instead, according to our results, these variables are characterized by low values of the posterior inclusion probability, hence indicating a weak relationship to growth. For yet other variables, our results are in line with the findings of the traditional empirical growth literature. This includes, for example, the *Investment share* and the dummy variable for *Sub-Saharan countries*.[22]

H Geography and the Globalization Index

H.1 Modified Globalization index

The computation of the modified GI presented in Section 7.4 for a given country i involves the following steps. First, we denote by $\varphi_{m,j,-i}$ the expected number of steps required to reach j

[22] See Barro (1991, 1996), Barro and Lee, (1994), Caselli et al. (1996), Easterly and Levine (1997) and Sala-i-Martin (1997a, 1997b).

from any country $m \neq i$, conditional on **not** utilizing any of the links that involve country i. $\varphi_{m,j,-i}$ can be derived as follows:

$$\varphi_{m,j,-i} = \sum_{k \neq i,j} \sum_{r=1}^{\infty} r \left[\left(A_{-i,-j} \right)^{r-1} \right]_{m,k} a_{k,j} \tag{34}$$

Here $A_{-i,-j}$ is a $(n-2) \times (n-2)$ matrix obtained from the original adjacency matrix A by deleting the ith and the jth column, and the ith and the jth row. $[\cdot]_{m,k}$ indicates the elements of the mth row and the kth column of the array $[\cdot]$. Rearranging equation (34) yields the following expression:

$$\varphi_{m,j,-i} = \left[\left(\sum_{r=1}^{\infty} r \left(A_{-i,-j} \right)^{r-1} \right)_{m,k} \right]_{k=1,2,...,n;\, i \neq k \neq j} \left(a_{k,j} \right)_{k=1,2,...,n;\, i \neq k \neq j} \tag{35}$$

where $\left(a_{k,j} \right)_{k=1,2,...,n; k \neq i,j}$ is an $(n-2) \times 1$ vector that is obtained from the jth column of matrix A by deleting the ith and the jth element. We use $\sum_{r=1}^{\infty} r \left(A_{-i,-j} \right)^{r-1} = \left(I - A_{-i,-j} \right)^{-2}$ and substitute it into (35), to obtain

$$\varphi_{m,j,-i} = \left[\left(I - A_{-i,-j} \right)^{-2}_{m,k} \right]_{k=1,2,...,n;\, i \neq k \neq j} \left(a_{k,j} \right)_{k=1,2,...,n;\, i \neq k \neq j} \tag{36}$$

We compute $\varphi_{m,j,-i}$ for all combinations of (m, j), where $m = 1, 2, ..., n$ and $j = 1, 2, ..., n$, with $m \neq i$, $j \neq i$. This yields the $(n-1) \times (n-1)$ dimensional matrix $\left(\varphi_{m,j,-i} \right)^{n}_{m=1,j=1;\, m \neq i \neq j}$. An element of which specifies the expected number of steps from any country j to each of country i's potential trading partners $m = 1, 2, ..., n, m \neq i$. The key difference to the related matrix in the benchmark case, i.e. $\left(\varphi_{m,j} \right)^{n}_{m=1,j=1}$, is that here all connections from and to country i are disregarded. The remaining steps of the calculations involve the aggregation of $\varphi_{m,j,-i}$ using the distance-related weighting factors as described in the main text.

H.2 The Frankel-Romer approach

The volume of trade of a country is potentially affected by its rate of economic growth, which renders the matrix $A_t = \left(a_{ijt} \right)^{n}_{i,j=1}$ induced by the trade flows of year t and the resulting GI, Φ_{it}, possibly endogenous to growth. In this section, we take a step to alleviate this endogeneity issue. More concretely, in the spirit of the approach pursued by Frankel and Romer (1999), we construct a modified GI measure that is based on bilateral geographical distance alone, and rely on it to instrument for Φ_{it}.

More concretely, the procedure implements the following steps. Let geo_{ij} denote the geographical distance (measured in kilometers) between countries i and j. In the first step, we replace the elements of the transition matrix, a_{ijt}, with the inverse of the geographical distance, $1/geo_{ij}$, between countries i and j. Naturally, after this step, the sum of each row is no longer

equal to one. Thus, to make the matrix row-stochastic, we normalize the elements of each row by the sum of each row. Let $\tilde{A}_t$ denote this modified transition matrix. An element of this matrix, denoted by $\tilde{a}_{ijt}$ is given by $\frac{1/geo_{ij}}{\sum_k 1/geo_{ik}}$. Clearly, $\tilde{a}_{ijt}$ is exogenous to growth. Next, we use the modified transition matrix, $\tilde{A}_t$, to compute the GI as described in Equations (9) and (10). Let by $\tilde{\Phi}_{it}$ denote the value of the modified GI for country i in period t. A key step of our approach is to use $\tilde{\Phi}_{it}$ as an instrument for the potentially endogenous GI, Φ_{it}. Specifically, we estimate by OLS the following first-stage regression:

$$\Phi_{it} = \alpha + \gamma\tilde{\Phi}_{it} + \mu_i + \zeta_t + \epsilon_{it}$$

where μ_i and ζ_t represent country and time fixed effects. We also consider a version where we do not include fixed effects. We compute $\tilde{\Phi}_{it}$ for all countries in our sample and for all years, which gives us a total of 9553 observations. The estimated value of γ obtained from the regression is equal to 0.45 and is highly significant with a 95% confidence interval of $[0.42, 0.47]$. Moreover, the F-statistic of this regression is equal to 768.1 and, thus, it far exceeds the value of 10 which is typically considered the critical value for indicating weak instruments. Let by $\widehat{\Phi}_{it}$ denote the predicted values of the regression. In the final step, we include $\widehat{\Phi}_{it}$ instead of the baseline GI measure, Φ_{it}, into the BMA. Importantly, the estimated coefficient of the modified GI is highly significant and the posterior inclusion probability of 62 percent is only slightly below that of the baseline GI.

Two remarks are in order. First, even though the geographical distance between countries is time invariant, the values of the modified GI are not necessarily constant over time. This is because, the number and the distribution of links in the trade network can change from year to year. Second, and relatedly, while the modified GI alleviates the endogeneity issue - by using geographical distance as a measure of bilateral trade intensity - it does not completely remove it. For, arguably, we cannot exclude the possibility that the number of a country's links is endogenous to its growth performance. That is, our approach does not tackle the endogeneity of whether two given countries engage in bilateral trade at all (extensive margin of trade) but only how much they trade (intensive margin). As a result of the latter observation, we do not interpret the results of the BMA with the modified GI as causal per se.

I Analysis of the patent data

In our analysis in Subsection 7.2, we focus on the patents originating in a sample of $n = 149$ countries that cite at least one other patent from a foreign country. That is, we disregard patents which (i) cite no other patent, or (ii) cite only patents of the same country. The latter condition derives from the fact that we are interested on the flow of ideas between countries and, naturally, own-country citations are not taken to contribute to it. The analysis has centered on two variables, Avg_{ij} and $Prob_{ij}^{inv}$, that measure, respectively, the average number of cited

patents from j cited in every citing patent from i, and the fraction of cross-country patenting relationships that connect an inventor from i with another in j. Here we provide a precise description of how these variables are derived.

First, we explain the computation of each Avg_{ij}. It is based on two matrices, P and C, of the following form:

$$P = \begin{pmatrix} 0 & p_{12} & p_{13} & ... & p_{1n} \\ p_{21} & 0 & p_{23} & ... & p_{2n} \\ . & . & . & ... & . \\ p_{n1} & p_{n2} & p_{n3} & ... & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & c_{12} & c_{13} & ... & c_{1n} \\ c_{21} & 0 & c_{23} & ... & c_{2n} \\ . & . & . & ... & . \\ c_{n1} & c_{n2} & c_{n3} & ... & 0 \end{pmatrix}.$$

The elements p_{ij} in matrix P represent the number of country-i patents that cite at least one country-j patent. Notice that, in general, $p_{ij} \neq p_{ji}$ and, of course, we may also find that many elements in P for which $p_{ij} = 0$. That is, cross-country patenting need not be symmetric and the cross-citing patent network could be quite sparse. In fact, in our case the total number of elements for which $p_{ij} > 0$ is equal to 3376 (thus much lower than the maximum $n(n-1)$) while the total number of patents that cite a foreign patent is equal to $\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij} = 2.98MM$.[23]

On the other hand, the elements c_{ij} in matrix C count how many country-j patents are cited in total by country-i patents. Notice that this is a conditional statement as we include only those country-i patents in c_{ij} which cite at least one country-j patent. $\sum_{j=1}^{n} c_{ij}$ is the total number of foreign patents cited by country-i patents. For our sample, we obtain that the total number of citations to foreign patents is equal to $\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} = 5.82mill.$ The element-by-element division of both matrices C and P gives $Avg_{ij} = c_{ij}/p_{ij}$ which is the average number of country-j patents cited per country-i patents.

Next, we explain how the variables $Prob_{ij}^{inv}$ are obtained. Their computation relies on the following matrix:

$$T = \begin{pmatrix} 0 & t_{12} & t_{13} & ... & t_{1n} \\ t_{21} & 0 & t_{23} & ... & t_{2n} \\ . & . & . & ... & . \\ t_{n1} & t_{n2} & t_{n3} & ... & 0 \end{pmatrix}.$$

An element t_{ij} in matrix T specifies the total number of bilateral co-patenting relationships between inventors from countries i and j. To fix ideas, consider two patents: Patent 1 was created by a team of 4 U.S. inventors, 2 French inventors and 2 German inventors. Patent 2 was created by 2 U.S. inventors and 3 French inventors. Then, for this example, we would get $t_{US,FRA} = t_{FRA,US} = 8 + 6 = 14$, $t_{US,GER} = t_{GER,US} = 8$, $t_{FRA,GER} = t_{GER,FRA} = 4$. In our sample, the number of entries in the matrix T for which $t_{ij} > 0$ is equal to 1918 and the total number of collaborations between international inventors is $\sum_{i=1}^{n}\sum_{j=1}^{n} t_{ij} = 286,168$. Computing the fraction $t_{ij}/\sum_{j=1}^{n} t_{ij}$ for each $i, j = 1, 2, ..., n$ we arrive at the corresponding

[23]Note that if a country-i patent cites country-j and country-k patents, then this country-i patent will be counted in both p_{ij} and p_{ik}. Due to this multiple counting of patents, we get that the row-sum $\sum_{j=1}^{n} p_{ij}$ is higher than the total number of country-i patents.

$Prob_{ij}^{inv}$.