Received: 2 April 2019 | Revised: 16 March 2022 | Accepted: 24 March 2022

DOI: 10.1002/jae.2920

WILEY Journal of APPLIED ECONOMETRICS

RESEARCH ARTICLE

Trade openness and growth: A network-based approach

Georg Duernecker[1] | **Moritz Meyer**[2] | **Fernando Vega-Redondo**[3]

[1]Goethe-University Frankfurt, CEPR, Frankfurt am Main, Germany

[2]World Bank, Washington, DC, USA

[3]Bocconi University, Milan, Italy

Correspondence
Georg Duernecker, Goethe-University Frankfurt, CEPR, Frankfurt am Main, Germany.
Email: duernecker@econ.uni-frankfurt.de

Abstract

We propose a novel approach to the study of international trade based on a theory of country integration that embodies a broad systemic viewpoint on the relationship between trade and growth. Our model leads to an indicator of country openness that measures a country's level of integration through the full architecture of its connections in the trade network. We apply our methodology to a sample of 204 countries and find a sizable and significant positive relationship between our integration measure and a country's growth rate, while that of the traditional measures of outward orientation is only minor and statistically insignificant.

KEYWORDS

Bayesian model averaging, dynamic-panel model, economic growth, globalization, network analysis, trade integration

1 | INTRODUCTION

A long-standing theme in the literature on economic growth has revolved around the question of what factors underlie the sharply contrasting growth performance of different countries. Inspired by early work of Baumol (1986) and Barro (1991), numerous studies have been devoted to establishing whether a given variable does or does not help explain cross-country growth differences. A central variable that has attracted considerable research interest in this literature is a country's openness to international trade. However, despite countless efforts and a long debate on the issue, no shared understanding has been reached on the relationship between trade openness and economic growth.[1] Indeed, the thorough reinvestigation of existing evidence that was undertaken by Rodriguez and Rodrik (2001) suggested that *in the absence of a supporting theory*, the received indicators used to measure a country's outward orientation did not convincingly capture truly relevant dimensions of economic growth.[2]

The objective of this paper is to revisit the debate by adopting a *systemic* (network-based) viewpoint that focuses on how effectively a country is *integrated* (through trade) in the world market. By doing so, we shall argue, new and

[Correction added on 12 August 2022, after first online publication: The article was published as Replication due to administrative error and has been corrected to Research Article in this version.]

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

[1]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 (see the early influential work by Dollar, (1992); Frankel & Romer, (1999); Sachs & Warner, (1995) or the more recent studies by Alcala & Ciccone, (2004); Dollar & Kraay, (2003); and Feyrer, (2019), supporting this same position).

[2]For a more recent review of the issue, see Winters (2004), Rodriguez (2007), and Estevadeordal and Taylor (2013).

valuable insights arise. The main novelty of our paper derives from the idea that a country's access to innovation opportunities depends on how well that country is integrated into the *whole* world trade network. Our approach is both dynamic and global in that not only direct trade connections but also *indirect* ones are taken to channel valuable (i.e., growth-generating) information. More specifically, we build on the theory of globalization developed by Duernecker and Vega-Redondo (2018) to formulate a model where intercountry knowledge flows are supported by trade. Under the assumption that foreign ideas are complementary to domestic ones in fueling innovation, our model predicts that countries can sustain high rates of economic growth only by having a rich pattern of direct and indirect trade-based connections to other countries.

To test the model's prediction, we use a dynamic-panel data set including 204 countries 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 several interesting patterns. One of them is that, while countries have become more integrated on average, the distribution of trade links across countries has become more unequal. In particular, we find that while the group of most integrated countries shows a persistent tendency to increase their integration, most less integrated countries stagnate or even display an opposite trend towards lower integration. An additional finding is that our measure of trade-based integration is essentially uncorrelated with the classical trade share (TS) variable for openness used by the literature. This observation provides support to the notion that a systemic/network perspective to understanding outward orientation is qualitatively distinct from the local one traditionally considered in the literature.

The paper then proceeds with a systematic analysis of the relationship between trade integration and economic growth. More specifically, we posit an econometric model where the GDP per capita of each country (in logs) is conceived as a linear projection of its lag and a set of covariates. The latter include our measure of trade-based integration and other 33 variables highlighted by the growth literature as potentially important factors correlated with country growth. These variables comprise the most commonly used openness indicators (such as the TS and the Sachs–Warner Index) along with a wide set of geographical, political and economic country characteristics.

Considering 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 that should be used to evaluate the implied conditional correlations. To tackle the problem, we adopt the Bayesian model-averaging (BMA) approach. This procedure associates to every possible model specification a posterior probability that the selected model is the correct (or "true") model. As a key finding of our empirical analysis, we establish that the trade-integration measure derived from our theory is strongly positively related to economic growth whereas the traditional measures of country openness are essentially uncorrelated with growth. Importantly, we show that this finding is robust to the use of different data sets or to the application of several variant formulations of our integration measure.

Lastly, we undertake considerable efforts to uncover potential channels underlying the positive relation between trade integration and economic growth. As a result of these efforts, we provide empirical support for the two essential features of our theory: (a) the notion of trade as a channel of information flows and (b) the focus on the overall architecture of the trading network and hence the role of indirect links.

Concerning (a), we consider two different (complementary) routes. First, we show that while trading flows in capital goods are positively related to growth, trading flows in raw commodities and economic growth are largely unrelated. Because, 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). We also provide more direct support for it by analyzing data on cross-country patent citations. We find that if we approximate the flow of ideas across two countries with the volume of patents in one country that are cited in the other country, such a variable is closely related to the corresponding trade-based distance from the former to the latter.

Concerning the feature described by (b), 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 are relegated to a very subsidiary role. We arrive at a different but similarly motivated point by constructing a pseudo-index of "globalization" that replaces the direct links reflecting bilateral trade with links that weigh geographic proximity alone. We then conclude that such a measure of "trade integration," which abstracts from direct trade links, enjoys support from the data that is very close to that provided by our benchmark measure.

Our paper relates and contributes to several strands in the literature. In the theoretical realm, the theory we develop is most closely related to those of Chaney (2014) and Alvarez et al. (2017), which posit that cross-border trade flows are a key conduit for economically valuable ideas—for example, information that enhances the technology and operational know-how of domestic firms. While Chaney's model assumes that international trade provides domestic firms with information that allows them to access foreign markets, the Alvarez et al. model postulates that “trade puts domestic producers in contact with (…) foreign and domestic producers from which they can learn and improve their technologies” (see also Buera & Oberfield, 2016).

There has also been a substantial empirical literature devoted to testing the informational implications of trade from a diverse set of perspectives. A seminal contribution to it was provided by Coe and Helpman (1995), who showed—for a sample of the 21 OECD countries plus Israel, during the period 1971–1990—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 only machinery and equipment, and it is also robust to controlling for institutional factors and human capital.[3] 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 adds to this empirical literature along several dimensions. A key feature is that we go well beyond the first-order effects associated to *direct* trade flows. As explained, this turns out to provide a significantly wider view of the problem, in which the *overall pattern* of international trade arises as 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*. Hence, at an admittedly high level, our analysis provides a trade-based foundation for the type of cross-country spillovers that have been found to be a key component of technological change (see, e.g., Eberhardt et al., 2013; Etur & Musolesi, 2017). In fact, we also find that these effects are not merely strong but also highly heterogeneous across countries because they are largely dependent on the position each country occupies in the overall trade network.[4]

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 (GI) 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 Appendix S1.

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 et al. (2017).

The model views the world economy as a directed network defined over a fixed set of nodes, $N = \{1, 2, \dots, n\}$, 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

[3]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 in Appendix S1, where we perform a spurious analysis that considers random permutations of the trade network.

[4]We thank an anonymous referee for suggesting this point.

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 is 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:

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

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, that is, $y_j(t)$. Then, the question arises of how that information is transferred from j to i. In line with existing literature (both theoretical and empirical),[5] 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 at every other j. The *value* of this information, however, is postulated to decrease with the time it takes to arrive due to several possible complementary factors, such as 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 $v_{ij}(t)$ 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

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

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

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

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

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 country i is 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 $i = 1, 2, \dots, n$, the number of projects z_i active in every country remains unchanged, so that

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

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:

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

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 into 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, we still need to articulate a useful operationalization of the theoretical framework to test this prediction. 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 GI Φ_i. The construction of these objects amounts to the following steps.

The ***first step*** of the procedure involves the construction of the matrix of trade flows $X \equiv \left(x_{ij}\right)_{i,j=1}^n$ 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 $\tilde{X} \equiv \left(\tilde{x}_{ij}\right)_{i,j=1}^n$, 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, that is, $\beta_i = y_i/Y$, 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, that is, $(y_i/Y)y_i$. 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—that is, $(1/y_i)\sum_{j\neq i} x_{ij}$—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 $A = \left(a_{ij}\right)_{i,j=1}^{n}$ is nonnegative and row-stochastic $\left(\sum_{j=1}^{n} a_{ij} = 1\right)$, 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 Section A in Appendix S1. There, we show that such expected path lengths can be computed as follows:

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

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 "iceberg costs."[6] 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 GI 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 are taken from the United Nations Commodity Trade Statistics Database (UN Comtrade), which covers 204 countries on an annual basis for the period from 1962 to 2016. See Table 2 in Appendix S1 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 to 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 data set allows for an accurate description of the world trade network.[7]

[6]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).

[7]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.

FIGURE 1 World trade discrete network in 2015

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

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

Figure 1 provides a schematic visualization of the world trade flows in 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. Several observations arise. Most interestingly, the figure shows that the world trade network is far from complete. That is, many countries are connected to only 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 in-degree and out-degree of any given country, where the former reflects its imports and the latter its exports. Because both perspectives yield an equivalent representation of the network,[8] we choose the import-based one and define simply the *degree* of a country as the number of other countries from which it receives imports.

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 25th, 50th, and the 75th percentile of the distribution of countries according to their degree. The numbers in the first row indicate that, at the 25th percentile, the countries imported from only 26% of all countries in 1965, whereas at the 75th percentile, they imported from more than half of all countries.

[8] 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.

		$\bar{\lambda}$	$\bar{v}^{10}$	$\bar{v}^{3}$	$\bar{v}^{1}$
	All	3.3	79.6	55.9	34.0
1965	Low income	2.9	81.3	57.6	34.2
	High income	4.4	62.3	36.5	20.0
	All	1.3	77.3	54.0	31.2
2015	Low income	1.1	80.9	57.3	33.4
	High income	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

Table 1 also shows a marked contrast between high-income and low-income countries. For concreteness, we classify countries as high income (low income) if, in 2015, their GDP per capita was above (below) 50% of the US level. Then, the second row of the table shows that in 1965, low-income countries imported, on average, from only 38.8% of all countries, while high-income countries did so from an average of 74.3% of countries. Even the most connected low-income countries at the 75th percentile imported from less than 50% of countries whereas at that same percentile, the high-income countries imported from almost every country in the world. Overall, the pattern described for 1965 is essentially maintained for 2015, although the extent of average connectivity grows, with the increase being especially significant for the low-income 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, in 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 the weight of the median m_j. We denote by v_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 $\bar{v}^u$, where the averages are taken either at the whole-world level or are separately computed for high-income or low-income 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.

The results show a strong concentration of countries' imports on only a few links for both 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 of 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 high-income and low-income countries. The import connections for low-income countries are more highly concentrated than for high-income 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 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 Section 2.2, to compute the GI, Φ_{it}, for every country i in our sample and every year $t = 1962, \dots, 2016$, then contrast it with two of the leading openness measures proposed in the literature: TS—the sum of exports and imports of a country as a fraction of its GDP—and the Sachs–Warner index (SWI)—a widely used five-criterion index of openness.[9] 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 descending order according to their 2015 value of the GI. Quite interestingly, we find

[9] See below for a detailed explanation of it.

TABLE 3 The globalization index—summary statistics and comparison with other openness indices: trade share and the Sachs–Warner index

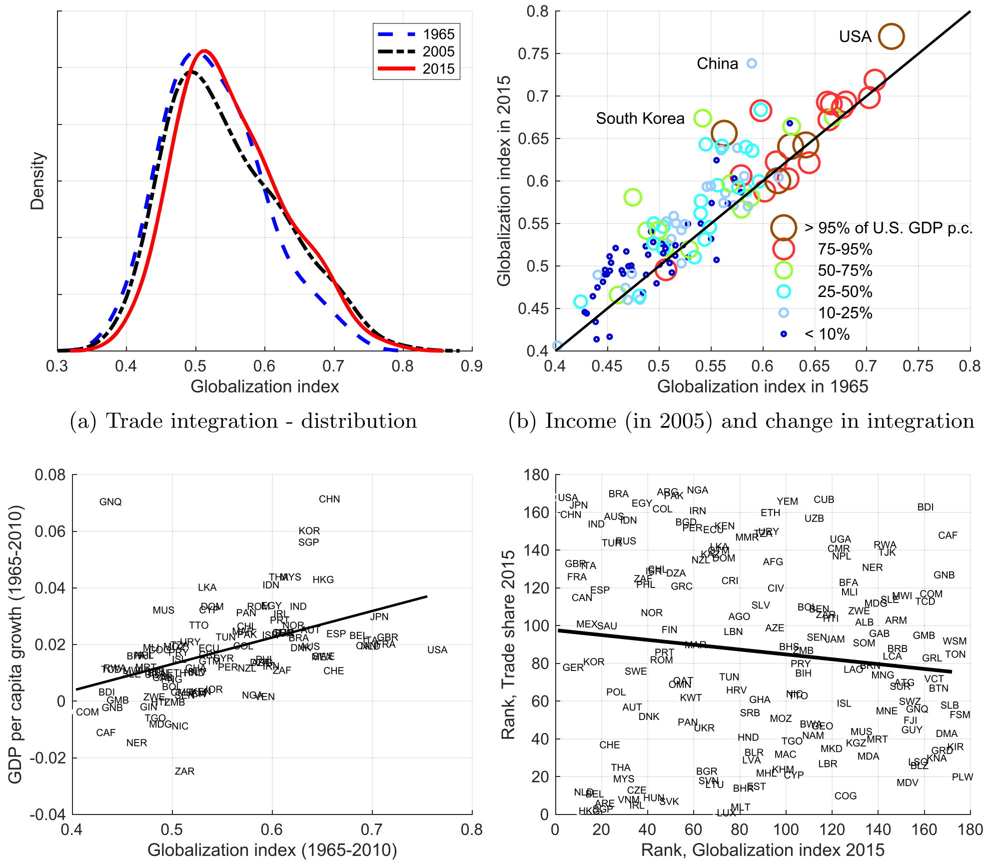
	Trade integration			Δ	Rank			Rank_{TS}	SWI
	1965	1990	2015	65–15	1965	1990	2015	2015	1990
United States	0.72	0.76	0.77	0.05	1	1	1	121	1
China	0.59	0.64	0.74	0.15	24	18	2	112	0
Germany	0.71	0.74	0.72	0.01	2	2	3	41	1
United Kingdom	0.70	0.71	0.70	0.00	3	4	4	91	1
France	0.68	0.72	0.69	0.01	4	3	5	84	1
Mexico	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
DRep. 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

Note: Rank: Ranking of each country in terms of the GI in a given year. 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.

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. Among countries lying in the middle range, the pattern is less clear cut, with some countries becoming more integrated while others are 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 TS. 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, there has been a general shift towards more integration at the world level between 1965 and 2015.
- 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 more integration than poorer ones. (Each circle represents a country and the size of a circle is proportional to the country's per capita GDP relative to US per capita GDP.)
- Relatedly, panel (c) displays a strong relationship between a country's 1965–2010 average of the GI (x-axis) and the annual GDP per capita growth rate (y-axis). That is, countries that 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 three-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 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 TS is very low and equal to −0.06. Furthermore, some of the most integrated economies, 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)—that is, 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. Reciprocally, one third of the bottom 50% of countries are classified as open.

Such a lack of correlation between our GI 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%; (ii) nontariff barriers cover more than 40% 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%. In view of criteria (i)–(v), a useful basis to understanding 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). As these authors show, most of the explanatory power of the SWI comes from the two nontrade 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, like 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. To explore the low correlation between the two measures, we first present a stylized numerical example of a hypothetical world composed of only three countries that are linked by exports and imports. This example will be useful to demonstrate what features of trade determine a country's

TABLE 4 A simple example for a world with three countries

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

level of integration and its TS and how a low TS 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 the three-country example again but suppose that while countries 2 and 3 are as before, country 1 exports less, that is, $x_{1j} = 1/10$ for $j = 2, 3$.

In contrast, 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 GDPs for each country;
- in (3), their TSs;
- in (4), their individual openness;
- in (5), the diffusion matrix;
- in (6), the expected path lengths involved in joining every pair of countries; and
- 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 also be worth noting that their individual openness $\theta_i = 1$ for all of them because 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 TS 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, as 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. Moreover, so are the *indirect* (multiple-link) paths flowing *into* Country 2 or 3 because 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 the volume of a country's exports and imports matters for the determination of a country's TS; it is its pattern of import links (direct and indirect) that plays the key role in determining the GI. This means that, in general,

TABLE 5 Trade share, globalization index, and higher order trading connections

		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
	DRep. 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

Note: 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.

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 the full pattern of combinations in the relationship between the TS and the GI is 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 in 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 arise from the table is the following: the countries with a high GI are characterized by many import links. Thus, in the column labeled "#" 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 fewer and, in many cases, less than half of all countries.

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 data set, such as the United States, China, or France. Nevertheless, the number of import links of these countries is substantially lower, which is why these countries rank low according to our GI.

But why is the number—not only 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 this idea empirically, in the last five columns of Table 5, we list the total diffusion weight reached by each country through direct import links and the import links of orders 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, that is, A^s. We observe that it is not only the value of $\bar{a}_i$ for the least globalized countries in panels (c) and (d) that is a small fraction of the value of the most globalized countries in panels (a) and (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 (33) of alternative variables 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, 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. In contrast, 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 that 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 that 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.

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 empirical models in a more or less ad hoc fashion. Over the years, this practice has led to many variables being proposed 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 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 is 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 ($j = 1, \ldots, 2^K$) 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.

[10]BMA 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), 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 like ours. 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.

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

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

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 binomial-beta prior structure (named after the implied model-size prior distribution) that 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 covariates. This is computationally unfeasible when the number K of regressors is large—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. When this Markov chain is simulated for a sufficiently long time—so that the model-to-model transition probabilities become stationary—it can be taken to have converged to the desired posterior distribution. In Section E in Appendix S1, 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 1 in Appendix S1 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. 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 we take the within-period average value for the regressor variables. We use 10-year intervals in the benchmark case, and as a robustness check, we also use 5-year intervals.[11]

Finally, 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 2 in Appendix S1 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. We report the data source and some descriptive statistics in Table 1 in Appendix S1.

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. In contrast, concerning the seven columns in the table, for the moment, we focus only on the first four that correspond to our benchmark globalization measure, while the last three columns will be discussed in Section 7.4 when we consider an alternative globalization measure relying only 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.

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

[12]Ideally, we would have 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 are not available after 1992.

TABLE 6 Results of the Bayesian model-averaging analysis

Description of variable	Benchmark $E(\theta_k\|\mathbf{y})$	*PIP*	$\%_{sig}$	Beta	Higher order trade $E(\theta_k\|\mathbf{y})$	*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 100 km 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 1950–1995	−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

- The fraction of models, $\%_{sig}$, where coefficient estimates are significant at the 5% level.
- The standardized coefficients, "beta," obtained by using standardized data in the analysis.[13]

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 both positive and significant at the 1% level. To reinforce the latter point, we also note that the estimated 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 in line with the estimated posterior model size of 8.7. More importantly for our purposes, our GI scores a very high inclusion probability of 85%.

[13]The standardization is achieved by demeaning and normalizing the original data so that each variable has a mean of zero and a unit standard deviation. 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% Bayesian coverage region that excludes zero. Using this "pseudo-t" statistic, we associate the levels of significance of 10%, 5%, and 1% to the ratios of posterior mean to standard deviation of 1.645, 1.960, and 2.576, respectively.

3. The standardized 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 correlation of the GI with GDP per capita is both statistically and economically significant.

Jointly considered, the above three points provide substantial support to the existence of a strong positive relationship between trade integration and income per capita. In fact, note that this applies both to the level and to the growth rate of per capita GDP because 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, 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 (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. in Appendix S1, 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 because Section 4 already showed that GI is practically uncorrelated with TS and SWI. Nevertheless, we find analyzing this issue in the broader context of the BMA worthwhile. Concretely, we include different combinations of the different openness measures into the BMA to check whether excluding some variables significantly alters the results for the others. As shown in Table 6 in Appendix S1, by and large, the analysis does not uncover any notable dependencies between the different measures.

Finally, 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 is an additional finding of some relevance to growth empiricists. This applies, for example, to the variables representing the *Government share*, the *Average years of primary schooling*, the SWI, or the *Annual population growth rate*. These are variables characterized by low values of the *PIP*—indicating that the models that 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 Section G in Appendix S1, we explore this discordance in some detail and explain 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. The latter identifies robust estimates with those that are statistically significant within the models that include the corresponding variables; the former also considers the support/likelihood that those models receive from the data in the first place (compared to the models that do not include the variables in question).

7 | THE KEY FEATURES OF THE GI

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

7.1 | Good-specific trade flows

First, we investigate to what extent the relationship between trade integration and growth is associated to varying intensities of trade in different types of goods. Recall that our measure of country integration, the GI index, has been computed using the ***total*** bilateral trade flows. Consequently, it treats Brazilian coffee exports to Japan and Japanese computer equipment shipped to Brazil equivalently (conditional on having the same dollar value). Arguably, all kinds of trade are not equally meaningful and should not have the same relation to a country's economic performance. In other words, if trade involves sophisticated goods (e.g., 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, such as raw materials. As a first step towards exploring this question, we consider here trade flows at the one-digit product level and separate them into the following two broad categories:

TABLE 7 Statistics derived by BMA analysis for the various product-based globalization indices

$E(\theta_k \| y)$	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 manufactured 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

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

Several 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 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] Taken together, this empirical finding is consistent with the interpretation 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.

7.2 | Trade and the flow of ideas

In this section, we aim at directly testing 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

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

[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 that specialize in the type of goods that rich countries typically export grow faster than countries that specialize in other goods. 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 a strong positive relationship 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.

TABLE 8 Country distance and patent citations

	Model 1: Avg_{ij}			Model 2: $Prob^{inv}_{ij}$		
	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

Note: 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: $Prob^{inv}_{ij}$ 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.

> "[…] *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 patent database that contains detailed information on all US patents granted between 1963 and 1999 (roughly three million patents) and all citations made to these patents between 1975 and 1999 (over 16 million citations).[18] Furthermore, the data set also includes, for each patent, the identity and the address (country, city, zip code and street) of every inventor who was involved in it. Based on the aforementioned information, we construct two variables, Avg_{ij} and $Prob^{inv}_{ij}$ (see Section I in Appendix S1 for the details). The variable Avg_{ij} measures how many patents of country j are cited, on average, by patents of country i. The variable $Prob^{inv}_{ij}$ specifies the fraction of cross-country co-patenting (bilateral) relationships of inventors from country i that 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 the following model by OLS:

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

where $y_{ij} \in \left\{ Avg_{ij}, Prob^{inv}_{ij} \right\}$, 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 measured by their relative GDP and the geographical distance between countries expressed in kilometers.

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 that 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 ($Prob^{inv}_{ij}$) and are more likely to cite each other's 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 when, at the same time, the bilateral network distance is controlled for.

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

[18] See Hall (2001) for a comprehensive description of the data set.

TABLE 9 Direct versus indirect links

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

In the spirit of our analysis of Section 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. Whereas 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 measures. In particular, we find a robust and positive relationship between countries' closeness and the bilateral knowledge exchange for capital goods. In contrast, the coefficient estimates suggest no significant relationship for raw materials.

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, the results show that close proximity between countries is significantly related to the bilateral exchange of ideas. However, an important qualification, also suggested by the analysis of Section 7.1, is that such a phenomenon arises only when the goods traded are of the type we have generically labeled as capital goods, that is, when they are likely to embody valuable information and know-how.

7.3 | The role of direct and indirect trade links

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 allow us to shed light on the relative importance of direct and indirect links. In the first experiment, we calculate the network distance between any pair (j, i) from a modified adjacency matrix A where we remove the direct connection between j and i by setting $a_{ji} = 0$. (Note that we eliminate *only* that link and therefore all other direct links remain untouched.) Then, if we compute the induced modified distance $\hat{\varphi}_{ji}$, the only way that one can reach i from j is indirectly – that is, by some path of length $m+1$ of the sort $j \to k_1 \to k_2 \to \dots \to k_m \to i$ where $m \geq 1$ and the link $j \to i$ is not used along the path. Doing this *separately* for each pair (j, i), we arrive at the matrix of modified distances $\left(\hat{\varphi}_{ji}\right)_{i,j=1}^{n}$ where, for each pair, the corresponding direct link from j to i is not used. Therefore, the modified GI given by $\hat{\Phi}_i \equiv \sum_{j \neq i} \beta_j f(\hat{\varphi}_{ji})$ involves distances for every pair of nodes that, by construction, rely only on indirect connections between that pair. The second row in Table 9 shows that the results for this modified GI are almost the same as those for the original measure. In other words, the correlation of our baseline GI with growth turns out not to depend on the direct trade connections between countries.

In contrast, we use only the direct trade connections of countries to compute the GI in the second experiment. 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:

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

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,

[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 the 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 the maximum distance across countries and years. None of these had any significant quantitative effect on the results.

we compute the GI as described in Section 2.2 and include it into BMA 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 relationship 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 | GI 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 GI that uses only the higher order trade connections of a country and replaces the first-order connections with 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) 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, that is,

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

This is a directed-distance measure that computes the expected length of all trade-weighted paths arriving at country i through the countries m that export to it, directly and indirectly, then assesses 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 GI as before, then included it into the BMA analysis to investigate its relationship to growth. The results, 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 Section 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 S1, providing here merely a succinct advance of that work. First, we test the robustness to different data inputs by considering alternative data sources and utilizing different waves of data sets. 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, we have considered other prominent notions of centrality such as PageRank centrality in our robustness analysis. 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.

[20]In Section H.2 in Appendix S1, 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 that 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.

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-order 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 TS 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 most 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 this work through a comprehensive BMA 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 per capita differences, while the traditional indicators of country openness are only marginally so. In fact, trade integration stands out from all other regressors (except own lagged GDP per capita) with a substantially larger posterior inclusion probability than for all regressors 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, 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.

OPEN RESEARCH BADGES

This article has earned an Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results.

DATA AVAILABILITY STATEMENT

The data and computer codes used in this paper are available in the JAE Data Archive http://qed.econ.queensu.ca/jae/datasets/duernecker001/.

REFERENCES

Acharya, R. C., & Keller, W. (2009). Technology transfer through imports. *Canadian Journal of Economics*, *42*, 1411–1448.

Alcala, F., & Ciccone, A. (2004). Trade and productivity. *Quarterly Journal of Economics*, *119*, 613–646.

Alvarez, F., Buera, F., & Lucas, R. E. (2017). Idea flows, economic growth, and trade: University of Chicago and Federal Reserve Bank of Chicago.

Arribas, I., Pérez, F., & Tortosa-Ausina, E. (2009). Measuring globalization of international trade: Theory and evidence. *World Development*, *37*, 127–145.

Barro, R. J. (1991). Economic growth in a cross section of countries. *Quarterly Journal of Economics*, *106*, 407–433.

Baumol, W. (1986). Productivity growth, convergence, and welfare: What the long-run data show. *American Economic Review*, *76*, 1072–1085.

Buera, F., & Oberfield, E. (2016). The global diffusion of ideas. *Econometrica*, *88*, 83–114.

Caselli, F., & Coleman, W. J. (2001). Cross-country technology diffusion: The case of computers. *American Economic Review*, *91*, 328–335.

Caselli, F., Esquivel, G., & Lefort, F. (1996). Reopening the convergence debate: A new look at cross-country growth empirics. *Journal of Economic Growth*, *1*, 363–389.

Chaney, T. (2014). The network structure of international trade. *American Economic Review*, *104*(11), 3600–3634.

Coe, D., & Helpman, E. (1995). International R & D spillovers. *European Economic Review*, *39*, 859–887.

Coe, D., Helpman, E., & Hoffmaister, A. W. (1997). North-south RandD spillovers. *Economic Journal*, *107*, 134–149.
Coe, D., Helpman, E., & Hoffmaister, A. W. (2009). International RandD spillovers and institutions. *European Economic Review*, *53*, 723–741.
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.
Dollar, D., & Kraay, A. (2003). Institutions, trade, and growth. *Journal of Monetary Economics*, *50*, 133–162.
Duernecker, G., Meyer, M., & Vega-Redondo, F. (2020). The network origins of economic growth: Bocconi University IGIER Working Paper.
Duernecker, G., & Vega-Redondo, F. (2018). Social networks and the process of globalization. *Review of Economic Studies*, *85*(3), 1716–1751.
Durlauf, S. N., Johnson, P. A., & Temple, J. R. W. (2005). Growth econometrics. In P. Aghion, & S. Durlauf (Eds.), *Handbook of economic growth*. North-Holland.
Eberhardt, M., Helmers, C., & Strauss, H. (2013). Do spillovers matter when estimating private returns to RandD? *The Review of Economics and Statistics*, *95*, 436–448.
Estevadeordal, A., & Taylor, A. M. (2013). Is the Washington Consensus dead? Growth, openness, and the great liberalization, 1970s–2000s. *Review of Economics and Statistics*, *95*(5), 1669–1690.
Etur, C., & Musolesi, A. (2017). Weak and strong cross-sectional dependence: A panel data analysis of international technological diffusion. *Journal of Applied Economics*, *32*, 477–503.
Feyrer, J. (2019). Trade and income—Exploiting time series in geography. *American Economic Journal: Applied Economics*, *11*(4), 1–35.
Frankel, J. A., & Romer, D. (1999). Does trade cause growth? *American Economic Review*, *89*, 379–399.
Hall, B. H. (2001). The NBER patent citation data file: Lessons, insights and methodological tools NBER Working Paper 8498.
Harrison, A. (1996). Openness and growth: A time-series, cross-country analysis for developing countries. *Journal of Development Economics*, *48*, 419–447.
Harrison, A., & Hanson, G. (1999). Who gains from trade reform? Some remaining puzzles. *Journal of Development Economics*, *59*(1), 125–154.
Hausmann, R., Hwang, J., & Rodrik, D. (2007). What you export matters. *Journal of Economic Growth*, *12*, 1–25.
Jaffe, A., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, *108*(3), 577–598.
Keller, W. (1998). Are international RandD spillovers trade-related?: Analyzing spillovers among randomly matched trade partners. *European Economic Review*, *42*(8), 1469–1481.
Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, *99*, 483–499.
Leamer, E. (1978). *Specification searches*: John Wiley & Sons.
Ley, E., & Steel, M. (2009). On the effect of prior assumptions in Bayesian model averaging with applications to growth regression. *Journal of Applied Econometrics*, *24*, 651–674.
Madigan, D., & York, J. (1995). Bayesian graphical models for discrete data. *International Statistical Review*, *63*, 215–232.
Magnus, J. R., Powell, O., & Prüfer, P (2010). A comparison of two model averaging techniques with an application to growth empirics. *Journal of Econometrics*, *154*, 139–153.
Moral-Benito, E. (2013). Likelihood-based estimation of dynamic panels with predetermined regressors. *Journal of Business and Economic Statistics*, *31*, 451–472.
Moral-Benito, E. (2015). Model averaging in economics: An overview. *Journal of Economic Surveys*, *29*, 46–75.
Moral-Benito, E. (2016). Growth empirics in panel data under model uncertainty and weak exogeneity. *Journal of Applied Econometrics*, *31*, 584–602.
Raftery, A. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.
Rodriguez, F. (2007). Openness and growth: What have we learned?: United Nations, Department of Economics and Social Affairs Working Paper 51.
Rodriguez, F., & Rodrik, D. (2001). Trade policy and economic growth: A skeptic's guide to the cross-national evidence. In B. Bernanke, & K. Rogoff (Eds.), *NBER Macroeconomics Annual 2000*: MIT Press.
Sachs, J. D., & Warner, A. (1995). Economic reform and the process of global integration. *Brookings Papers in Economic Activity*, *1*, 1–118.
Sala-i-Martin, X., Doppelhofer, G., & Miller, R. I. (2004). Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. *American Economic Review*, *94*, 813–835.
Samuelson, P. (1954). The transfer problem and transport costs, II: Analysis of effects of trade impediments. *Economic Journal*, *64*, 264–289.
Winters, L. A. (2004). Trade liberalization and economic performance: an overview. *Economic Journal*, *114*, F4–F21.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of the article.

How to cite this article: Duernecker, G., Meyer, M., & Vega-Redondo, F. (2022). Trade openness and growth: A network-based approach. *Journal of Applied Econometrics*, 37(6), 1182-1203. https://doi.org/10.1002/jae.2920