Enhanced Gravity Model of trade: reconciling macroeconomic and network models

Assaf Almog, Rhys Bird, Diego Garlaschelli*

March 17, 2017

Abstract

The structure of the International Trade Network (ITN), whose nodes and links represent world countries and their trade relations respectively, affects key economic processes worldwide, including globalization, economic integration, industrial production, and the propagation of shocks and instabilities. Characterizing the ITN via a simple yet accurate model is an open problem. The traditional Gravity Model successfully reproduces the volume of trade between connected countries, using macroeconomic properties such as GDP, geographic distance, and possibly other factors. However, it predicts a network with complete or homogeneous topology, thus failing to reproduce the highly heterogeneous structure of the ITN. On the other hand, recent maximum-entropy network models successfully reproduce the complex topology of the ITN, but provide no information about trade volumes. Here we integrate these two currently incompatible approaches via the introduction of an Enhanced Gravity Model (EGM) of trade. The EGM is the simplest model combining the Gravity Model with the network approach within a maximum-entropy framework. Via a unified and principled mechanism that is transparent enough to be generalized to any economic network, the EGM provides a new econometric framework wherein trade probabilities and trade volumes can be separately controlled by any combination of dyadic and country-specific macroeconomic variables. The model successfully reproduces both the global topology and the local link weights of the ITN, parsimoniously reconciling the conflicting approaches. It also indicates that the probability that any two countries trade a certain volume should follow a geometric or exponential distribution with an additional point mass at zero volume.

Introduction

The International Trade Network (ITN) is the complex network of trade relationships existing between pairs of countries in the world. The nodes (or vertices) of the ITN represent nations and the edges (or links) represent their (weighted) trade connections. In a global economy extending across national borders, there is increasing interest in

*Instituut-Lorentz for Theoretical Physics, University of Leiden, Niels Bohrweg 2, 2333 CA, Leiden, The Netherlands

understanding the mechanisms involved in trade interactions and how the position of a country within the ITN may affect its economic growth and integration [1, 2, 3, 4, 5]. Moreover, in the wake of recent financial crises the interconnectedness of economies has become a matter of concern as a source of instability [6]. As the modern architecture of industrial production extends over multiple countries via geographically wider supply chains, sudden changes in the exports of a country (due e.g. to unpredictable financial, environmental, technological or even political circumstances) can rapidly propagate to other countries via the ITN. The assessment of the associated trade risks requires detailed information about the underlying network structure [7]. In general, among the possible channels of interaction among countries, trade plays a major role [2, 3, 4].

The above considerations imply that the empirical structure of the ITN plays a crucial role in increasingly many economic phenomena of global relevance. It is therefore becoming more and more important to characterize the ITN via simple but accurate models that identify both the basic ingredients and the mathematical expressions required to accurately reproduce the details of the empirical network structure. Reliable models of the ITN can better inform economic theory, foreign policy, and the assessment of trade risks and instabilities worldwide.

In this paper, we emphasize that current models of the ITN have strong limitations, and that none of them is satisfactory, from either a theoretical or a phenomenological point of view. We point out equally strong (and largely complementary) problems affecting on one hand traditional macroeconomic models, which focus on the local *weight* of the links of the network, and on the other hand more recent network models, which focus on the *existence* of links, i.e. on the global topology of the ITN. We then introduce a new model of the ITN that preserves all the good ingredients of the models proposed so far, while at the same time improving upon the limitations of such approaches. The model can be easily generalized to any (economic) network and provides an explicit specification of the full probability distribution that a given pair of countries is connected by a certain volume of trade, fixing an otherwise arbitrary choice in previous approaches. This distribution is found to be either geometric (for discrete volumes) or exponential (for continuous volumes), with an additional point mass at zero volume. This feature, which is different from all previous specifications of international trade models, is shown to replicate both the local trade volumes and the global topology of the empirical ITN remarkably well.

Preliminaries: building blocks of the model

Before we fully specify our model, we preliminarily identify its building blocks by reviewing the strengths and weaknesses of the two main modelling frameworks adopted so far.

Gravity models of trade

We start by discussing traditional macroeconomic models of international trade. These models have mainly focused on the *volume* (i.e. the value e.g. in dollars) of trade between countries, largely because economic theory perceives trade volumes as being *a priori* more informative than the topology of the ITN: the striking heterogeneity of trade volumes observed between different pairs of countries is clearly not captured by a purely 'binary' description where all connections are effectively given the same weight. Based on this argument, emphasis has been put on explaining the (expected) volume of trade between two countries, given certain dyadic and country-specific macroeconomic properties.

Jan Tinbergen, the physics-educated[1] Dutch economist who was awarded the first Nobel memorial prize in economics, introduced the so-called Gravity Model (GM) of trade [8]. The GM aims at inferring the volume of trade from the knowledge of Gross Domestic Product, mutual geographic distance, and possibly additional dyadic factors of macroeconomic relevance [9, 10]. In one of its simplest forms, the gravity model predicts that, if i and j label two different countries ($i, j = 1, N$ where N is the total number of countries in the world), then the expected volume of trade from i to j is

$$\langle w_{ij} \rangle = c \ \mathrm{GDP}_i^{\alpha} \ \mathrm{GDP}_j^{\beta} \ R_{ij}^{-\gamma} \qquad c, \alpha, \beta, \gamma > 0, \tag{1}$$

where GDP_k is the Gross Domestic Product of country k, R_{ij} is the geographic distance between countries i and j, and c, α, β, γ are free global parameters to be estimated. In the above *directed* specification of the GM, the flows w_{ij} and w_{ji} can be different. An analogous *undirected* specification exists, where the volumes of trade from i to j and from j to i are added together into a single value $w_{ij} = w_{ji}$ of bilateral trade. In the latter case, 1 still holds but with the symmetric choice $\alpha = \beta$. With this in mind, we will keep our discussion entirely general throughout the paper and, unless otherwise specified, allow all quantities to be interpreted either as directed or as undirected. Only in our final empirical analysis we will adopt an undirected description for simplicity.

More complicated variants of 1 use additional factors (with associated free parameters) either favoring or resisting trade [9, 10]. Like the GDP and geographic distances, these factors can be country-specific (e.g. population) or dyadic (e.g. common currency, trade agreements, shared borders, common language, etc.). In general, if we collectively denote with $\mathbf{n}_i$ the set of *node-specific* factors and with $\mathbf{D}_{ij}$ the set of *dyad-specific* factors used, 1 can be generalized to

$$\langle w_{ij} \rangle = F_{\phi}(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij}) \qquad F > 0, \tag{2}$$

where the functional form of $F_{\phi}(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij})$ need not be of the same type as in 1, and ϕ is a vector containing all the free parameters of the model (like c, α, β, γ for the particular case above). Indeed, although in this paper we focus on the GM applied to

[1] Jan Tinbergen studied physics in Leiden, where he carried out a PhD under the supervision of the theoretical physicist Paul Ehrenfest. Tinbergen defended his thesis in 1929, and then became a leading economist. He was awarded the first Nobel memorial prize in economics in 1969.

the international trade network, our discussion equally applies to many other models of (socio-economic) networks as well. For instance, the recently proposed Radiation Model (RM) [11] is also described by 2, where $\mathbf{n}_i$ and $\mathbf{D}_{ij}$ include certain geographical and demographical variables. Our following discussion applies to both the GM and the RM, as well as any other model described by 2. Similarly, it does not only apply to trade networks, since both the GM and the RM have been successfully applied to other systems as well, including mobility and traffic flows [11, 12, 13, 14], communication networks [15], and migration patterns [16] (the latter representing - to our knowledge - the earliest application of the GM to a socio-economic system, dating back to 1889 [17]).

It is generally accepted that the expected trade volumes postulated by the GM, already in its simplest form given by 1, are in good agreement with the observed flows between trading countries. To illustrate this result, in Fig. 1 we show a typical log-log plot comparing the empirical volume of the realized (bilateral) international trade flows with the corresponding expected values calculated under the GM as defined in 1. The figure shows the typical good agreement between the GM and the empirical non-zero trade volumes. However, it should be noted that, while 1 and 2 define the expected value of w_{ij}, the full probability distribution from which this expected value is calculated is not specified, and actually depends on how the model is implemented in practice. In the GM case, the distribution is chosen to be either Gaussian (corresponding to additive noise, in which case the expected weights can be fitted to the observed ones via a simple linear regression [18, 19]), log-normal (corresponding to multiplicative noise and requiring a linear regression of log-transformed weights [20]), Poisson [20], or more sophisticated [21] (see [22] for a review). The arbitrariness of the weight distribution already highlights a fundamental weakness of the traditional formulation of the model. Moreover, for both additive and multiplicative Gaussian noise, the model can produce undesired negative values.

A related and more important limitation of the GM is that, at least in its simplest and most natural implementations, it cannot generate zero volumes – thereby predicting a fully connected network [23, 22, 24]. This means that the GM can be fitted only to the *non-zero* weights, i.e. the volumes existing between pairs of *connected* countries. If used in this way, the model effectively disregards the empirical structure of the network, both as input (thus making predictions on the basis of incomplete data) and as output (thus failing to reproduce the topology). Operatively, the GM can be used only *after* the presence of a trade link has been established independently [22]. As observed in [21], *"Omitting zero-flow observations implies that we loose information on the causes of (very) low trade"*, because any fit to positive-only flows would significanlty underestimate the effects of factors that diminish trade. This problem is particularly critical since roughly half of the possible links are found to be *not* realized in the real ITN [25, 26, 27, 28]. Clearly, the same problem holds for the RM and any more general model of the form specified in 2.

While there are variants and extensions of the GM that do generate zero weights and a realistic link density (e.g. the so-called Poisson pseudo-maximum likelihood models [20] and 'zero-inflated' gravity models [21]), these variants systematically fail in reproducing the observed topology [22, 10]. In other words, while these models can generate the correct number of connections, they tend to put many of the latter in the 'wrong place' in the network. Indeed, even in its generalized forms, the GM predicts a largely homogeneous network structure, while the empirical topology of the ITN is much more heterogeneous and complex [23, 22]. Established empirical signatures of this heterogeneity include a broad distribution of the degree (number of connections) and the strength (total trade volume) of countries [25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35], the rich-club phenomenon (whereby well-connected countries are also connected to each other) [36, 37], strong clustering and (dis)assortative patterns [26, 27]. These highly skewed structural properties are remarkably stable over time. However, their are not replicated by any current version of the GM [22].

Network models of trade

As we mentioned at the beginning, many processes of great economic relevance crucially depend on the large-scale topology of the ITN. In light of this result, the sharp contrast between the observed topological complexity of the ITN and the homogeneity of the network structure generated by the GM (including its extensions) calls for major improvements in the modelling approach. In particular, in assessing the performance of a model of the ITN, emphasis should be put on how reliably the (global) empirical network structure, besides the (local) volume of trade, is replicated. In the network science literature, successful models of the ITN have been derived from the *maximum entropy* principle [25, 26, 27, 28, 24, 38, 39, 40, 41, 42]. These models construct ensembles of random networks that have some desired topological property (taken as input from empirical data) and are maximally random otherwise. Typically, the constrained properties are chosen to be the degrees and/or the strenghts of all nodes. In this way the models can perfectly replicate the observed strong heterogeneity of these purely local properties, and at the same time illustrate its immediate (i.e. prior to invoking any other more complicated network formation mechanism) structural effects on any higher-order topological property of the network.

In general, different choices of the constrained properties lead to different degrees of agreement between the model and the data. This can generate intriguing and counter-intuitive insight about the structure of the ITN. For instance, contrary to what naive economic reasoning would predict, it turns out that the knowledge of purely binary local properties (e.g. node degrees) can be *more informative* than the knowledge of the corresponding weighted properties (e.g. node strengths). Indeed, while the binary network reconstructed only from the knowledge of the degrees of all countries is found to be topologically very similar to the real ITN, the weighted network reconstructed only from the strengths of all countries is found to be much denser and very different from the real network [26, 27, 28]. This is somewhat surprising, given that economic theory postulates that weighted properties are *per se* more informative than the corresponding binary ones.

The solution to this apparent paradox lies in the fact that, while the knowledge of the *entire* weighted network is necessarily more informative than that of its binary projection (in accordance with economic postulates), the knowledge of certain *marginal* properties of the weighted network can be unexpectedly less informative than the knowledge of the corresponding marginal properties of the binary network. In fact, it turns out that if the degrees of countries are (not) specified in addition to the strengths of countries, the resulting maximum-entropy model can(not) reproduce the empirical weighted network of international trade satisfactorily [27, 40, 41].

An important take-home message is that, in contrast with the mainstream literature, models of the ITN should aim at reproducing not only the strength of countries (as the GM automatically does by approximately reproducing all non-zero weights), but also their degree (i.e. the number of trade partners) [26, 27, 28, 41]. In order to devise improved models of the ITN, one should therefore include the degrees, which are purely topological properties, among the main target quantities to replicate. This is the guideline we will follow in this paper.

Unlike the GM, maximum-entropy models of trade are *a priori* non-explanatory, i.e. they take as input structural properties (as opposed to explanatory economic factors) to explain other structural properties. However, they can in fact be used to select *a posteriori* an explicit, empirically validated functional dependence of the structure of the ITN on underlying explanatory factors. For models with country-specific constraints, this operation can be carried out as follows. Mathematically, controlling for node-specific properties is realized by assigning one or more Lagrange multipliers, also known as 'hidden variables' or 'fitness parameters' $\mathbf{x}_i$, to each node. If a certain choice of local constraints is found to replicate the higher-order properties of the real-world network satisfactorily, then one can look for an empirical relationship between the values of the associated hidden variables and those of candidate non-topological, country-specific factors of the type $\mathbf{n}_i$, like the GDP or total import/export. If the hidden variables are indeed (at least approximately) found to be functions of some country-specific factors (i.e. if $\mathbf{x}_i \approx \mathbf{f}(\mathbf{n}_i)$), then one can replace $\mathbf{x}_i$ with $\mathbf{f}(\mathbf{n}_i)$ in the maximum-entropy model, thus reformulating the latter as a model with explanatory variables (i.e. 'regressors') of trade, precisely like the GM. Already more than a decade ago, the approach outlined above led to the definition of a GDP-driven model for the binary topology of the ITN, where $\mathbf{x}_i \propto \mathrm{GDP}_i$. The model, which is a reformulation of a maximum-entropy model for binary networks with given degrees, predicts that the probability of a trade connection existing from country i to country j is

$$p_{ij} = \frac{\delta\ \mathrm{GDP}_i\ \mathrm{GDP}_j}{1 + \delta\ \mathrm{GDP}_i\ \mathrm{GDP}_j} \qquad \delta > 0, \tag{3}$$

where δ is a free parameter that allows to reproduce the empirical link density. The model has been tested successfully in multiple ways [24, 25, 30, 32, 38].

The GM in 1 and the maximum-entropy model in 3 have complementary strengths and weaknesses, the former being a good model for non-zero volumes (while being a bad model for the topology) and the latter being a good model for the topology (while providing no information about trade volumes). An attempt to reconcile these two complementary and currently incompatible approaches has been recently proposed via the definition of an extension of the maximum-entropy model to the case of weighted networks [42]. Since, as we mentioned, a maximum-entropy model of weighted networks with given strengths and degrees [40] can correctly replicate many structural properties of the ITN [41], it makes sense to reformulate such model as an economically inspired model of the ITN. Indeed, like in the binary case, the hidden variables enforcing the constraints are found to be strongly correlated with the GDP, thus allowing to express both p_{ij} and $\langle w_{ij} \rangle$ as functions of the GDP [42]. The resulting model is confirmed to be in good accordance with both the topology and the volumes observed in the real ITN.

Unfortunately, in the above approach the choice of country-specific constraints (degrees and strengths) only allows for regressors that have a corresponding country-specific nature. This makes the model incompatible with the inclusion of dyadic variables of the type $\mathbf{D}_{ij}$ and represents a strong limitation for (at least) two reasons. Firstly, one of the main lessons learnt from the traditional GM is that the addition of geographic distances improves the fit to the empirical volumes significantly. Indeed, in the light of the large body of knowledge accumulated in the international economics literature, it is hard to imagine a realistic and economically meaningful model of international trade that does not allow for simple pair-wise quantities controlling for trade 'costs' and 'incentives', including geography [9, 10]. Secondly, even if the structure of the ITN can be replicated satisfactorily in terms of the 'GDP-only' model defined in 3 [25, 30, 32], recent analyses have found evidence that certain metric (although not necessarily geographic[2]) distances do also play a role in determining the topology of the ITN [43]. Together, these two pieces of evidence call for an inclusion of dyadic factors in $\langle w_{ij} \rangle$ and p_{ij}, and highlight a limitation of current maximum-entropy models based only on country-specific constraints.

Combining all the above considerations, it is clear that an improved model of the ITN should aim at retaining the realistic trade volumes postulated by models based on 2 (including the GM, the RM, and possibly many more), while combining them with a realistic network topology generated by (extensions of) maximum-entropy models. Such a model should also aim at providing the full probability distribution, and not only the expected values as in 1, of trade flows and, unlike the GDP-only model in 3 [25] or its current weighted extension [42], allow for the inclusion of both dyadic and node-specific macroeconomic factors.

The Enhanced Gravity Model of international trade

In this Section, we introduce what we call the Enhanced Gravity Model (EGM) of trade. The EGM mathematically formalizes the two ingredients that, in the light of

[2]Building on the hypothesis of the existence of underlying hidden metric spaces in which real-world networks are embedded, Ref. [43] models the ITN by looking for an optimal embedding of countries in some abstract metric space. The resulting inferred distances are interpreted as incorporating all possible sources of empirically revealed trade costs, possibly including geographic distances as well. However, since the postulated embedding space is either a unidimensional circle or a hyperbolic plane, these distances are necessarily different from the usual geographic distances R_{ij} appearing in the GM and measured as geodesics on our spherical tridimensional world.

the previous discussion, any 'good' model of economic networks should feature: namely, realistic (trade) volumes and a realistic topology, both controllable by macroeconomic factors.

A single model for topology and weights

The first lesson we have learnt is that 2 is successful in reproducing link weights only *after* the existence of the links themselves has been preliminarly established. This implies that 2, as a model of real-world trade flows, is actually unsatisfactory and should rather be reformulated as a *conditional expectation* of the weight w_{ij}, given that $w_{ij} > 0$. In other words, if a_{ij} denotes the entry of the adjacency matrix of the ITN (defined as $a_{ij} = 1$ if $w_{ij} > 0$ and $a_{ij} = 0$ if $w_{ij} = 0$), an improved model should be such that 2 is replaced by

$$\langle w_{ij} | a_{ij} = 1 \rangle = F_\phi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij}) \qquad F > 0, \tag{4}$$

where $\langle w_{ij} | a_{ij} = 1 \rangle$ is the *conditional expected weight* of the trade link from country i to country j, *given that such link exists.* This operation ensures that, whatever the new model looks like, its predictions for the expected trade volume between connected pairs of countries remain identical to the ones proposed in more traditional macroeconomic models. For instance, choosing $F_\phi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij}) = c\ \mathrm{GDP}_i^\alpha\ \mathrm{GDP}_j^\beta\ R_{ij}^{-\gamma}$ as in 1 allows us to retain (in almost intact form) all the empirical knowledge that has accumulated in the econometrics literature since Jan Tinbergen's introduction of the GM. An important difference, however, is that in our model the trade volumes will be drawn from a different probability distribution.

The second lesson we have learnt is that, in analogy with 4, 3 should be generalized to allow for both dyadic ($\mathbf{D}_{ij}$) and node-specific ($\mathbf{n}_i$) factors as follows:

$$p_{ij} = \langle a_{ij} \rangle = \frac{G_\psi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij})}{1 + G_\psi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij})} \qquad G > 0, \tag{5}$$

where a crucial requirement is that G can in general be different from F in 4 and, correspondingly, the vector ψ of parameters can be different from ϕ. Note that, since p_{ij} is monotonic in G, the above expression is entirely general, i.e. we have put no restriction on the functional form of p_{ij}. It is also worth noticing that the explanatory factors used in 4 and 5 need not coincide. However, to avoid using different symbols for the arguments of the two functions, we adopt the convention that $\mathbf{D}_{ij}$ and $\mathbf{n}_i$ denote the sets of all factors used as arguments of either F or G, and that these functions can have flat (i.e. no) dependence on some of their arguments. For instance, 5 reduces to 3 by setting $\mathbf{n}_i = \mathrm{GDP}_i$ and assuming flat dependence on $\mathbf{D}_{ij}$, or it reduces to the hyperbolic model in Ref. [43] by setting $\mathbf{D}_{ij}$ equal to the hyperbolic distance and assuming flat dependence on $\mathbf{n}_i$.

We want our model to produce both 4 as the desired (gravity-like) conditional expectation for link weights and 5 as a realistic expected topology. To do so, we introduce the full probability $P(\mathbf{W})$ that the model produces a weighted network specified by the $N \times N$ matrix $\mathbf{W}$ with entries (w_{ij}). Without loss of generality, we assume that w_{ij} is a non-negative integer number (allowing for non-negative real numbers is straightforward, upon taking the limit of a vanishing unit of weight, as we briefly discuss later). The probability $P(\mathbf{W})$ is the key quantity that fully specifies the model and determines both the topology and the link weights of the ITN. From $P(\mathbf{W})$, we can define the dyadic (marginal) probability $q_{ij}(w)$ that w_{ij} takes the particular value w. Note that the event $w = 0$ indicates the absence of a trade link and is included as a possible outcome in $q_{ij}(w)$. The normalization condition is therefore $\sum_{w\geq 0} q_{ij}(w) = 1$ for all i, j. Note that we are *not* assuming independence of the trade volumes w_{ij} and w_{kl} between two distinct country pairs, or equivalently the factorization of $P(\mathbf{W})$ into the product $\prod_{i,j} q_{ij}(w_{ij})$ of dyadic probabilities. However, we will later find that the desired model has precisely this independence property. Importantly, unlike the traditional GM, in our approach dyadic independence is a consequence and not a postulate.

We now look for the form of $q_{ij}(w)$ that enforces both 4 and 5. Let us consider the latter first. In terms of $q_{ij}(w)$, the probability p_{ij} that i and j are connected (irrespective of the volume of trade) is given by the complement of the probability $q_{ij}(0)$ that they are not connected, i.e.

$$p_{ij} = \sum_{w>0} q_{ij}(w) = 1 - q_{ij}(0). \tag{6}$$

Imposing that 6 has the form dictated by 5 leads to the following unique choice for $q_{ij}(0)$:

$$q_{ij}(0) = \frac{1}{1 + G_\psi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij})} \qquad G > 0. \tag{7}$$

We now relate $q_{ij}(w)$ to 4 in a similar manner. The expected trade volume, irrespective of whether a link exists, is

$$\langle w_{ij}\rangle \equiv \sum_{w>0} w\ q_{ij}(w), \tag{8}$$

while the conditional expectation, given that the link exists, is

$$\langle w_{ij}|a_{ij} = 1\rangle \equiv \sum_{w>0} w\ q_{ij}(w|a_{ij} = 1) = \frac{\langle w_{ij}\rangle}{p_{ij}} \tag{9}$$

where

$$q_{ij}(w|a_{ij} = 1) = \frac{q_{ij}(w)}{\sum_{u>0} q_{ij}(u)} = \frac{q_{ij}(w)}{p_{ij}} \qquad w > 0 \tag{10}$$

is the *conditional* probability that w_{ij} equals w, given that the link is realized. Setting 9 equal to 4 leads to

$$\langle w_{ij}\rangle = \frac{F_\phi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij})\ G_\psi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij})}{1 + G_\psi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij})} \qquad F, G > 0. \tag{11}$$

11 carries an important message. It reveals us that, while a superficial inspection of 8 might suggest that the expected trade volume $\langle w_{ij}\rangle$ is independent on the topology of the ITN, i.e. on $q_{ij}(0)$ or equivalently G, this is actually not the case. In fact, $q_{ij}(0)$ is coupled to the other values $q_{ij}(w)$ (with $w > 0$) through the normalization condition manifest in

6. This necessarily implies that *the topology of the ITN must have an immediate effect on the expected volume of trade between any two countries.* This effect is rigorously quantified in 11, which shows that $\langle w_{ij} \rangle$ depends on both F and G. This result confirms the inconsistency of the traditional GM defined in terms of 1 and of any of its extensions of the form 2. By contrast, *the expected topology of the ITN is independent on the expected volumes of trade*, since p_{ij} depends on G but not on F. This simple but, to the best of our knowledge, previously unrecognized result highlights a nontrivial asymmetry between weights and topology in the ITN and, by extension, in any (economic) network described by our generic expressions involving F and G. This basic finding provides a natural explanation for the aforementioned empirical observation that the topology of the ITN and several other networks can be satisfactorily reconstructed from aggregate local constraints [26, 40], while the same result does not hold for the weighted structure of the same network(s) [27, 28], unless topological information is explicitly included as an additional constraint [40, 41].

Maximum entropy construction

7 and 11 fix two important properties we require for $q_{ij}(w)$ and ultimately $P(\mathbf{W})$, but they do not specify these probability distributions uniquely. To do so, we take the maximum-entropy approach and look for the form of $P(\mathbf{W})$ that maximizes the entropy

$$S = -\sum_{\mathbf{W}} P(\mathbf{W}) \ln P(\mathbf{W}), \tag{12}$$

where the sum extends over all weighted graphs with N nodes and non-negative integer edge weights, subject to the constraints specified by 7 and 11. The maximum-entropy method ensures that the resulting functional form of $P(\mathbf{W})$ is maximally random, given the desired constraints. As well known, this procedure is guaranteed to lead to the least biased inference, i.e. to introduce no unjustified 'hidden' assumption in picking a specific form of $P(\mathbf{W})$.

Since 7 is equivalent to 5, we select $\langle a_{ij} \rangle$ and $\langle w_{ij} \rangle$ as the two sets of constraints specifying our model. In this way, if we introduce α_{ij} and β_{ij} as the (real-valued) Lagrange multipliers required to enforce the expected value of $a_{ij} = \Theta(w_{ij})$ and w_{ij} respectively (where $\Theta(x) = 1$ if $x > 0$ and $\Theta(x) = 0$ otherwise), then the maximum-entropy problem becomes equivalent to one solved exactly in Ref. [44]. There, it was shown that upon introducing the so-called *Hamiltonian*

$$H(\mathbf{W}) = \sum_{i,j} [\alpha_{ij}\Theta(w_{ij}) + \beta_{ij} w_{ij}], \tag{13}$$

(representing a linear combination of the quantities whose expected value is being constrained) and the *partition function* $Z = \sum_{\mathbf{W}} e^{-H(\mathbf{W})}$, the maximum-entropy probability $P^*(\mathbf{W})$ with constraints $\langle a_{ij} \rangle$ and $\langle w_{ij} \rangle$ is found to be

$$P^*(\mathbf{W}) = \frac{e^{-H(\mathbf{W})}}{Z} = \prod_{i,j} q^*_{ij}(w_{ij}), \tag{14}$$

where, given $x_{ij} \equiv e^{-\alpha_{ij}} \in (0, +\infty)$ and $y_{ij} \equiv e^{-\beta_{ij}} \in (0, 1)$,

$$q^*_{ij}(w) \equiv \frac{x_{ij}^{\Theta(w)} y_{ij}^w (1 - y_{ij})}{1 - y_{ij} + x_{ij} y_{ij}}, \qquad w \geq 0 \tag{15}$$

is the resulting (maximum-entropy) probability that the link from node i to node j carries a weight w. This probability is called the Bose-Fermi distribution, as it unifies the Bose-Einstein and Fermi-Dirac distributions encountered in quantum statistical physics [44]. We stress again that all our formulas apply to both directed and undirected representations of the network and, correspondingly, the sums and products over i, j should be interpreted as $i \neq j$ in the directed case (where the pairs i, j and j, i are different) and as $i < j$ in the undirected one (where the pair i, j is the same as the pair j, i). As we had anticipated, the factorization of $P^*(\mathbf{W})$ in terms of products of $q^*_{ij}(w)$ shows that, for this particular choice of the constraints, pairs of nodes turn out to be statistically independent as in the standard GM approach, even if we have not assumed this independence as a postulate in our approach.

Importantly, while the constraints used in the maximum-entropy models of the ITN considered so far in the literature are observed topological properties (e.g. the degrees and/or the strengths of nodes), the constraints considered here are economically-driven expectations, namely 5 and 11. This key step allows us to reconcile macroeconomic and network approaches within a generalized framework, and represents an important difference with respect to previous models. In particular, we use 6, 8 and 9 to express p_{ij}, $\langle w_{ij} \rangle$ and $\langle w_{ij} | a_{ij} = 1 \rangle$ in terms of x_{ij} and y_{ij} [44]:

$$p_{ij} = 1 - q^*_{ij}(0) = \frac{x_{ij} y_{ij}}{1 - y_{ij} + x_{ij} y_{ij}}, \tag{16}$$

$$\langle w_{ij} \rangle = \sum_{w>0} w \; q^*_{ij}(w) = \frac{p_{ij}}{1 - y_{ij}}, \tag{17}$$

$$\langle w_{ij} | a_{ij} = 1 \rangle = \frac{\langle w_{ij} \rangle}{p_{ij}} = \frac{1}{1 - y_{ij}}. \tag{18}$$

Now, equating 16 to 5 and 17 to 11 (or, equivalently, 18 to 4) allows us to find the values of x_{ij} and y_{ij} solving the original problem:

$$x_{ij} = \frac{G_\psi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij})}{F_\phi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij}) - 1}, \tag{19}$$

$$y_{ij} = \frac{F_\phi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij}) - 1}{F_\phi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij})}. \tag{20}$$

Inserting 19 and 20 into 15, we finally get the explicit probability of any two countries trading a volume w, as a function of any choice of the factors $\mathbf{n}_i$ and $\mathbf{D}_{ij}$.

In terms of conditional probabilities, our model becomes extremely simple: establishing a link from country i to country j is a Bernoulli trial with success probability p_{ij} given by 5; if realized, this link acquires a weight w with probability

$$q^*_{ij}(w | a_{ij} = 1) = \frac{[F_\phi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij}) - 1]^{w-1}}{[F_\phi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij})]^w} \qquad w > 0, \tag{21}$$

which is a geometric distribution representing the chance of $w-1$ consecutive successes, each with probability y_{ij}, followed by a failure with probability $1-y_{ij}$. The above result provides an insightful interpretation of the realized volumes in the model in terms of processes of link establishment and link reinforcement (see Discussion).

Maximum-likelihood parameter estimation

We now take an econometric perspective and discuss how the model parameters can be chosen to optimally fit a specific empirical instance of the network. To this end, we use the Maximum Likelihood principle applied to network models [38]. If $\mathbf{W}^*$ denotes the weight matrix (with entries w^*_{ij}) of the empirical network, our model generates this particular matrix with probability $P^*(\mathbf{W}^*)$. We therefore define the log-likelihood function as

$$\mathcal{L}(\phi,\psi)=\ln P^*(\mathbf{W}^*)=\sum_{i,j}\ln\frac{(G_\psi)^{a^*_{ij}}(F_\phi-1)^{w^*_{ij}-a^*_{ij}}}{(1+G_\psi)(F_\phi)^{w^*_{ij}}} \tag{22}$$

(where we have dropped the dependence of F and G on $\mathbf{n}_i$, $\mathbf{n}_j$ and $\mathbf{D}_{ij}$) and look for the parameter values ϕ^*,ψ^* that maximize $\mathcal{L}$ by requiring that all the first derivatives with respect to ϕ and ψ vanish simultaneously:

$$\vec{\nabla}_\phi\mathcal{L}(\phi,\psi) = \sum_{i,j}\left[\frac{w^*_{ij}-a^*_{ij}}{F_\phi-1}-\frac{w^*_{ij}}{F_\phi}\right]\vec{\nabla}_\phi F_\phi=\vec{0} \tag{23}$$

$$\vec{\nabla}_\psi\mathcal{L}(\phi,\psi) = \sum_{i,j}\left[\frac{a^*_{ij}}{G_\psi}-\frac{1}{1+G_\psi}\right]\vec{\nabla}_\psi G_\psi=\vec{0}. \tag{24}$$

Upon checking that the second derivatives have the correct sign, i.e. that $\mathcal{L}(\phi^*,\psi^*)$ is indeed a maximum for $\mathcal{L}$, the solution (ϕ^*,ψ^*) to the above equations yields the optimal parameter values in our model. Selecting these values into 19 and 20 yields the values x^*_{ij} and y^*_{ij} that, when inserted into 15, fully specify the model.

The above expressions, which are valid *any* specification of the EGM, show that the estimation of the parameter ϕ nicely separates from that of ψ. This result solves, in a single shot, two major problems encountered in previous econometric approaches: on one hand, in most alternative models the estimation of the parameters determining the expected weights is badly affected by the presence of the zeroes; on the other hand, the expected number of zeroes may paradoxically depend on the (arbitrary) units of measure for the weights. For instance, if $q_{ij}(w)$ is a Poisson distribution as in zero-inflated GMs [20, 21, 22], then its only parameter (the mean) determines both the magnitude of link weights and the connection probability p_{ij}. As the units of money in the data are changed arbitrarily (e.g. from dollars to thousands of dollars), so will the estimated mean and the resulting expected number of zeroes. By contrast, in our model the zeroes affect ψ but not ϕ and the money units affect ψ but not ϕ.

Empirical analysis

We can finally test the predictions of our model against empirical international trade data. These predictions are found to improve dramatically upon those of the traditional GM.

Model specification

We adopt an undirected network description (where the connection between two countries carries a weight equal to the total trade in either direction) to facilitate the definition of the topological properties characterizing the ITN. Previous work has shown that, given the highly symmetric structure of the ITN, the undirected representation retains all the basic properties of the network [30, 26, 27].

We choose $F_\phi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij})$ in such a way that the expected non-zero trade flow $\langle w_{ij} | a_{ij} = 1 \rangle$ is the same as in the GM defined by 1 (now interpreted as a conditional expectation). This means choosing $\mathbf{n}_i = \mathrm{GDP}_i$, $\mathbf{D}_{ij} = R_{ij}$, $\phi = (c, \alpha, \gamma)$ and

$$F_\phi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij}) = c \; (\mathrm{GDP}_i \; \mathrm{GDP}_j)^\alpha \; R_{ij}^{-\gamma}, \tag{25}$$

where we have set $\beta = \alpha$ due to undirectedness. Similarly, we choose $G_\psi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij})$ in such a way that the probability p_{ij} is the same as in the model defined in 3, i.e. $\psi = \delta$ and

$$G_\psi(\mathbf{n}_i, \mathbf{n}_j, \mathbf{D}_{ij}) = \delta \; \mathrm{GDP}_i \; \mathrm{GDP}_j. \tag{26}$$

With the above specification, the expected topology does not depend on any dyadic factor. This is the simplest choice that is found to reproduce the topology of the ITN very well [25, 30, 32, 38] and is supported by empirical evidence that dyadic factors like geographic distances [45] and trade agreements [43] have a much weaker effect on the purely binary topology of the ITN than on trade volumes. Of course our formalism has been designed in such a way that we can immediately add dyadic factors, and is therefore much more general. For instance, we might easily add 'hidden' metric distances inferred via an optimal geometric embedding [43] (although they would not be identifiable with some empirically measurable, 'external' macroeconomic factors like those used elsewhere in our model).

Given the above model specification, for a given instance $\mathbf{W}^*$ of the empirical network we find the optimal parameter values c^*, α^*, γ^* and δ^* through the maximum likelihood conditions given by 23 and 24. Importantly, 24 reads in this case $\partial \mathcal{L} / \partial \delta = 0$ and yields a value δ^* that ensures that the expected number of links $\sum_{i,j} p_{ij} = \sum_{i,j} G_\psi / (1 + G_\psi)$ is exactly equal to the empirical number $L^* = \sum_{i,j} a_{ij}^*$, irrespective of the volumes of trade. This result, which is equivalent to what is found for the purely binary model defined by 3 [38], shows that, unlike the standard GM, our model always generates the correct number of links and, unlike some more complicated variants of the GM, it does so independently of the money units chosen for the volumes.

Testing the model against real data

We first test the performance of the EGM in replicating the empirical trade volumes, i.e. the purely local structure of the ITN. In Fig. 1, superimposed to the previous results for the standard GM given by 1, the empirical non-zero link weights w^*_{ij} are also compared with their conditional expected value $\langle w_{ij}|a_{ij}=1\rangle$ under the EGM given by 25. As expected, the two sets of points largely overlap, confirming that, in terms of trade volumes, the EGM cannot do worse than the GM. Moreover, the EGM is more parsimonious than the GM as it achieves a narrower scatter of points while having no dedicated free parameter to tune the variance.

We then check the performance of the EGM in replicating the purely binary topology of the ITN. As a first illustration, in Fig. 2 we show all the trade links of the country with maximum degree (USA), the one with minimum degree (Western Sahara) and one with intermediate degree (Vanuatu). We also show the corresponding predictions under the standard GM (where 1 is first fitted to the non-zero flows and then extended to all pairs of countries) and the EGM. Since the traditional GM predicts a fully connected network, i.e. an expected degree $\langle k_i\rangle_{GM}=N-1$ for all i, its prediction is correct only for the country with maximum degree and deteriorates dramatically as the degree decreases. By contrast, the EGM gives an expected degree $\langle k_i\rangle_{EGM}=\sum_{j\neq i}p_{ij}$ (see Materials and Methods) which is in good agreement with the empirical one for the entire range of connectivity.

We now consider higher-order topological properties as a more stringent and quantitative test. In the top left panel of Fig. 3 we plot the average degree (k_i^{nn}) of the trade partners of each country i versus the number of such partners, i.e. the degree (k_i) of country i itself. Similarly, in the top right panel of Fig. 3 we plot the clustering coefficient (c_i), i.e. the fraction of trade partners of country i that trade with each other, again versus the number (k_i) of such partners. The empirical quantities are compared with the expected quantities under the GM and the EGM. The exact expressions for both empirical and expected quantities are provided in Materials and Methods. The decreasing empirical trends observed in both plots shows that the trade partners of poorly connected countries (small k_i) are on average highly connected, both to the rest of the world (large k_i^{nn}) and among themselves (large c_i). By contrast, countries that trade with a high-degree country (large k_i) are on average poorly connected, both to the rest of the world (small k_i^{nn}) and among themselves (small c_i). For both properties, we find that the EGM is in excellent agreement with the empirical ITN, as opposed to the classical GM which systematically generates nearly constant and much higher values, as a result of predicting a complete network.

Having checked that the EGM does very well in separately replicating both the local link weights and the global topology of the ITN, we now perform a last and most severe test monitoring properties that combine topological and weighted information together. In the bottom left panel of Fig. 3 we plot the average strength (s_i^{nn}), i.e. the average traded volume, of the trade partners of each country i versus the strength (s_i) of country i itself. In the bottom right panel, we plot a weighted version of the clustering coefficient (c_i^w) of country i, again versus the strength (s_i) of country i. The empirical trends are compared with the predictions of the GM and EGM (see Materials and Methods for all definitions). These two plots are in some sense the weighted counterparts of the purely binary plots considered above. We find that, on average, countries connected to countries with a low trade activity (small s_i) trade a lot with the rest of the world (large s_i^{nn}) but relatively less among themselves (small c_i^w). Countries connected to countries with a large volume of trade (large s_i) have instead a small trade activity with the rest of the world (small s_i^{nn}), but trade relatively strongly with each other (large c_i^w). Again, we find that both trends are replicated very well by the EGM, while the standard GM fails systematically.

Discussion

In this paper we have introduced the EGM as a novel, advanced model for the ITN and economic networks in general. Phenomenologically, the EGM allows us to reconcile two very different approaches that have remained incompatible so far: on one hand, the established GM which successfully reproduces non-zero trade volumes in terms of GDP and distance, while failing in predicting the correct topology [22]; on the other hand, network models which have been successful in reproducing the topology [25] but are more limited in predicting link weights [42]. To our knowledge, the EGM is the first model that can successfully reproduce the binary and the weighted empirical properties of the ITN simultaneously. Just like the standard GM, the RM [11] or similar models, the EGM can accomodate additional economic factors in terms of extra dyadic and country-specific properties.

The agreement between the EGM and trade data calls for an interpretation of the process generating the network in the model. In this respect, we notice that 15 and 21 allow us to interpret the realized trade volumes in the EGM as the outcome of two equivalent processes (a serial and a parallel one) of link creation and link reinforcement. In the serial process, for a given pair of countries i, j a trade link of unit weight is first established with success probability p_{ij} and its volume is then incremented in unit steps, each with success probability y_{ij}. After the first failure, the process stops and starts again for a different pair of countries, and so on for all pairs. In the equivalent parallel process, all pairs of countries simultaneously explore the mutual benefits of trade and engage in a first connection, each with its probability p_{ij}. Then, all pairs of nodes for which the previous event has been successful reinforce their existing connection by a unit weight, each with its probability y_{ij}. The process stops as soon as there are no more successful events. In either case, 15 is the resulting probability that the realized volume is w.

Importantly, 15 implies that $q_{ij}^*(w)$ is a geometric distribution with an extra point mass $q_{ij}^*(0)$ at zero volume, i.e. the first event has a probability p_{ij} which is in general different from the probability y_{ij} of all subsequent events. This distinguishing property of the Bose-Fermi distribution [44] ensures a realistic network formation mechanism where the establishment of a trade connection for the first time is intrinsically different (and therefore associated to a different 'cost') from the reinforcement of an already existing trade connection. This desirable distinction, interpretable for instance in terms of profitability of trade, has been advocated in previous studies [9, 10, 21]. Here, it is implemented naturally within the maximum-entropy framework via 13, where the (expected) binary topology is enforced separately from the (expected) link weights. Notice that the distinction disappears if the parameter α_{ij} in 13 is set to zero, i.e. if the constraint on the expected value of $\Theta(w_{ij})$ (the expected topology) is removed as in the standard GM. In such a case, p_{ij} becomes equal to y_{ij} (i.e. link creation and link reinforcement become equally likely) and therefore $q^*_{ij}(w)$, not only $q^*_{ij}(w|a_{ij}=1)$, becomes a geometric distribution.

Consistently with the fact that trade volumes are always reported as integer multiples of some indivisible money unit (e.g. dollars), we have assumed non-negative integer weights. It is easy to show that, if we take the limit of a vanishing money unit, trade volumes become non-negative real numbers and $q^*_{ij}(w)$ becomes an exponential density with an extra point mass at zero volume, while $q^*_{ij}(w|a_{ij}=1)$ becomes a purely exponential density. Crucially, the extra point mass $q^*_{ij}(0)$ ensures that, even in this continuous limit, p_{ij} is unchanged and the topology is still described by 5. In absence of α_{ij}, in this limit the network would become fully connected as in all specifications of the GM with continuous volumes [39].

Our results may have strong implications for the theoretical foundations of trade models and for the resulting policy implications. It is known that the traditional GM is consistent with a number of (possibly conflicting) micro-founded model specifications [46, 47, 48, 49]. For instance, a gravity-like relation can emerge as the equilibrium outcome of models of trade specialization and monopolistic competition with intra-industry trade [10, 50]. The empirical failure of the standard GM highlights a previously unrecognized limitation of these micro-founded models, at least in their current form, and indicates the need for an appropriate reformulation that makes these model consistent with the EGM, i.e. with a realistic topology of the ITN. How policy implications change as the result of such a reformulation of current micro-founded models is an important point to add to the future research agenda. We therefore believe that the EGM can represent a novel benchmark supporting improved theories of trade and refined policy scenarios.

Materials and Methods

Data

We have used international trade and GDP data from the database curated by Gleditsch [51] for the years 1950, 1960, 1970, 1980, 1990 and 2000. This database includes yearly trade volumes w_{ij} (which we have symmetrized by taking the sum of $w_{ij}+w_{ji}$), yearly GDP values, and the (time-independent) distance matrix R_{ij}. The numer N of countries increases over time from roughly 85 in 1950 to approximately 200 in 2000. Both GDP and trade date are reported in U.S. dollars. To produce Fig. 2, we have used the BACI database [52], which reports imports and exports between $N=208$ countries in 2011. The BACI data were originally in disaggregated form, where total trade was resolved into 96 different non-overlapping commodity classes. We have aggregated all these commodity classes together, and again symmetrized, to obtain a dataset consistent with the Gleditsch data used for the earlier years.

Observed network properties

Given a weighted undirected network with weight matrix $\mathbf{W}$ and adjacency matrix $\mathbf{A}$, with entries related through $a_{ij} = \Theta(w_{ij})$, the *degree* of node i is defined as

$$k_i = \sum_{j \neq i} a_{ij}, \tag{27}$$

the *average nearest-neighbor degree* of node i is defined as

$$k_i^{nn} = \sum_{j \neq i} \frac{a_{ij} k_j}{k_i} = \frac{\sum_{j \neq i} \sum_{k \neq j} a_{ij} a_{jk}}{\sum_{j \neq i} a_{ij}}, \tag{28}$$

and the *(binary) clustering coefficient* of node i is defined as

$$c_i = \frac{\sum_{j \neq i} \sum_{k \neq i,j} a_{ij} a_{jk} a_{ki}}{\sum_{j \neq i} \sum_{k \neq i,j} a_{ij} a_{ki}}. \tag{29}$$

The *average nearest neighbor strength* of node i is defined as

$$s_i^{nn} = \sum_{j \neq i} \frac{a_{ij} s_j}{k_i} = \frac{\sum_{j \neq i} \sum_{k \neq j} a_{ij} w_{jk}}{\sum_{j \neq i} a_{ij}} \tag{30}$$

(where $s_i = \sum_{j \neq i} w_{ij}$ is the *strength* of node i) and the *weighted clustering coefficient* of node i is defined as

$$c_i^w = \frac{\sum_{j \neq i} \sum_{k \neq i,j} (w_{ij} w_{jk} w_{ki})^{\frac{1}{3}}}{\sum_{j \neq i} \sum_{k \neq i,j} a_{ij} a_{ki}}. \tag{31}$$

Expected network properties

The expected value (under the EGM) of each of the network properties defined above can be calculated either numerically, by averaging over many network realizations sampled independently from the probability $P^*(\mathbf{W})$ in 14, or analytically, using the following approach. First of all, in this model the expected value of all ratios can be approximated by the ratio of the expected values [40, 41]. Secondly, all numerators and denominators involve only products over distinct pairs of nodes, which are statistically independent in the model. Using 15, the expected values of such products can therefore be calculated exactly in terms of x_{ij} and y_{ij} as follows:

$$\langle \sum_{i,j,k,...} a_{ij} \cdot a_{jk} \cdot ... \rangle = \sum_{i,j,k,...} \langle a_{ij} \rangle \cdot \langle a_{jk} \rangle \cdot \langle ... \rangle, \tag{32}$$

$$\langle \sum_{i,j,k,...} w_{ij}^{\alpha} \cdot w_{jk}^{\beta} \cdot ... \rangle = \sum_{i,j,k,...} \langle w_{ij}^{\alpha} \rangle \cdot \langle w_{jk}^{\beta} \rangle \cdot \langle ... \rangle, \tag{33}$$

where $\langle a_{ij} \rangle = p_{ij}$, as given by 16, and

$$\langle w_{ij}^{\gamma} \rangle \equiv \sum_{w=0}^{\infty} w^{\gamma} q_{ij}(w) = \frac{x_{ij}(1-y_{ij})\mathrm{Li}_{-\gamma}(y_{ij})}{1-y_{ij}+x_{ij}y_{ij}}, \tag{34}$$

$\mathrm{Li}_n(z) = \sum_{l=1}^{\infty} \frac{z^l}{l^n}$ denoting the so-called $n-$th polylogarithm of z. From the above two considerations, it follows that the expected properties of all quantities of interest can be approximated with entirely analytical expressions obtained by simply replacing a_{ij} with p_{ij} and w_{ij}^{γ} with $\langle w_{ij}^{\gamma} \rangle$ in 27, 28, 29, 30 and 31. Via x_{ij} and y_{ij}, the expected values are ultimately a function of only the GDPs and distances. In our analysis, after preliminary checking that the analytical expressions matched extremely well with the numerical averages over realizations, we have systematically adopted the analytical approach, which requires no sampling of networks and is therefore extremely efficient.

Acknowledgements

AA and DG acknowledge support from the Dutch Econophysics Foundation (Stichting Econophysics, Leiden, the Netherlands). This work was also supported by the Netherlands Organization for Scientific Research (NWO/OCW).

References

[1] F. Schweitzer, G. Fagiolo, D. Sornette, F. Vega-Redondo, A. Vespignani, D.R. White. *Economic networks: The new challenges*, Science **325(5939)**, 422 (2009).

[2] R. Kali and J. Reyes,*The architecture of globalization: a network approach to international economic integration*, Journal of International Business Studies **38**, 595 (2007).

[3] R. Kali and J. Reyes, *Financial contagion on the international trade network*, Economic Inquiry **48**, 1072 (2010).

[4] S. Schiavo, J. Reyes, G. Fagiolo, *International trade and financial integration: a weighted network analysis*, Quantitative Finance **10(4)**, 389-399 (2010).

[5] J. Spitz, T. Kastelle, *Gains from Trade: the Impact of International Trade on National Economic Convergence; A Complex Network Analysis Approach*, Eastern Economic Association Annual Meeting, 2010.

[6] S. Battiston et al., *Complexity theory and financial regulation*, Science 351:6275, 818-819 (2016).

[7] P. Klimek, M. Obersteiner and S. Thurner, *Systemic trade risk of critical resources*, Science Advances, **1**, e1500522 (2015).

[8] J. Tinbergen, *Shaping the World Economy: Suggestions for an International Economic Policy*, (The Twentieth Century Fund, New York, 1962).

[9] P. van Bergeijk, S. Brakman (eds.) *The gravity model in international trade* (Cambridge University Press, Cambridge, 2010).

[10] L. De Benedictis, D. Taglioni, *The gravity model in international trade*, in The Trade Impact of European Union Preferential Policies, pp. 55-89 (Springer Berlin Heidelberg, 2011).

[11] F. Simini, M.C. González, A. Maritan, A. Barabási *A universal model for mobility and migration patterns* Nature **484**, 96–100 (2012).

[12] G. K. Zipf, *The* P_1P_2/D *hypothesis: on the intercity movement of persons*, Am. Sociol. Rev. **11**, 677-686 (1946).

[13] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, A. Vespignani, *Multiscale mobility networks and the spatial spreading of infectious diseases*, Proc. Natl. Acad. Sci. USA **106**(51), 21484-21489 (2009).

[14] W.-S. Jung, F. Wang, H. E. Stanley, *Gravity model in the Korean highway*, EPL **81**, 48005 (2008).

[15] G. Krings, F. Calabrese, C. Ratti, V. D. Blondel, *Urban gravity: a model for intercity telecommunication flows*, J. Stat. Mech. **2009**, L07003 (2009).

[16] G. Fagiolo, M. Mastrorillo, *International migration network: Topology and modeling*, Phys. Rev. E **88**, 012812 (2013).

[17] E. G. Ravenstein, *The Laws of Migration*, Journal of the Royal Statistical Society **52**(2), 241-305.

[18] R. Glick and A.K. Rose *Does a Currency Union affect Trade? The Time Series Evidence*, NBER Working Paper No. 8396 (2001).

[19] A. K. Rose, M. M. Spiegel *A Gravity Model of Sovereign Lending: Trade, Default, and Credit* IMF Staff Papers, Vol. 51, Special Issue, International Monetary Fund (2004)

[20] J. Silva and S.Tenreyro *The Log of Gravity* The Review of Economics and Statistics, **88**, issue 4, pages 641-658 (2006).

[21] G.J. Linders and H.L.F. de Groot *Estimation of the Gravity Equation in the Presence of Zero Flows* Tinbergen Institute Discussion Paper No. 06-072/3 (2006).

[22] M. Duenas, G. Fagiolo, *Modeling the International-Trade Network: A Gravity Approach*, Journal Of Economic Interaction And Coordination **8** 155-178 (2013).

[23] G. Fagiolo, *The international-trade network: gravity equations and topological properties*, Journal of Economic Interaction and Coordination **5(1)**, 1-25 (2010).

[24] T. Squartini and D. Garlaschelli, *Jan Tinbergen's legacy for economic networks: from the gravity model to quantum statistics*, in Econophysics of Agent-Based Models, pp. 161-186 (Springer International Publishing, 2014).

[25] D. Garlaschelli, M.I. Loffredo, *Fitness-Dependent Toplogical Properties of the World Trade Web*, Phys. Rev. Lett **93**, 188701 (2004).

[26] T. Squartini, G. Fagiolo, D. Garlaschelli, *Randomizing world trade. I. A binary network analysis*, Phys. Rev. E **84**, 046117 (2011).

[27] T. Squartini, G. Fagiolo, D. Garlaschelli, *Randomizing world trade. II. A weighted network analysis*, Phys. Rev. E **84**, 046118 (2011).

[28] G. Fagiolo, T. Squartini, D. Garlaschelli, *Null Models of Economic Networks: The Case of the World Trade Web*, J. Econ. Interac. Coord. **8**(1), 75-107 (2013).

[29] A. Serrano and M. Boguna,*Topology of the world trade web*, Phys. Rev. E **68**,015101(R) (2003).

[30] D. Garlaschelli and M. Loffredo, *Structure and Evolution of the World Trade Network*, Physica A **355**, 138 (2005).

[31] A. Serrano, M. Boguna, and A. Vespignani, *Patterns of dominant flows in the world trade web*, J. Econ. Interact. Coord.**2**, 111 (2007).

[32] D. Garlaschelli, T. Di Matteo, T. Aste, G. Caldarelli, and M. Loffredo, *Interplay between topology and dynamics in the World Trade Web*, Eur. Phys. J. B **57**, 1434 (2007).

[33] G. Fagiolo, J. Reyes, S. Schiavo, *On the topological properties of the world trade web: A weighted network analysis*, Physica A **387(15)**, 3868-3873 (2008).

[34] G. Fagiolo, J. Reyes, S. Schiavo, *World-trade web: Topological properties, dynamics, and evolution*, Physical Review E **79(3)**, 036115 (2009).

[35] G. Fagiolo, J. Reyes, S. Schiavo, *The evolution of the world trade web: a weighted-network analysis*, Journal of Evolutionary Economics, **20(4)**, 479-514 (2010).

[36] V. Colizza, A. Flammini, M. A. Serrano, and A. Vespignani, *Detecting rich-club ordering in complex networks*, Nature Physics **2**, 110 - 115 (2006).

[37] V. Zlatic, G. Bianconi, A. D. Guilera, D. Garlaschelli, F. Rao, and G. Caldarelli, *On the rich-club effect in dense and weighted networks* Eur. Phys. J. B **67**, 271-275 (2009).

[38] D. Garlaschelli, M.I. Loffredo, *Maximum Likelihood: Extracting Unbiased Information from Complex Networks*, Phys. Rev. E **78**, 015101(R) (2008).

[39] A. Fronczak, P. Fronczak, J.A. Holyst, (2012) 'Statistical mechanics of the international trade network', *Phys. Rev. E*, Vol. 85, pp.056113

[40] R. Mastrandrea, T. Squartini, G. Fagiolo, D. Garlaschelli, *Enhanced reconstruction of weighted networks from strengths and degrees*, New J. Phys. **16**, 043022 (2014).

[41] R. Mastrandrea, T. Squartini, G. Fagiolo, D. Garlaschelli, *Reconstructing the world trade multiplex: the role of intensive and extensive biases*, Phys. Rev. E **90**, 062804 (2014).

[42] A. Almog, T. Squartini, and D. Garlaschelli, *A GDP-driven model for the binary and weighted structure of the International Trade Network*, New J. Phys. **17**, 013009 (2015).

[43] G. García-Pérez, M. Boguñá, A. Allard, M. Serrano, *The hidden hyperbolic geometry of international trade: World Trade Atlas 1870–2013*, Scientific Reports, **6**, 33441, (2016).

[44] D. Garlaschelli, M.I. Loffredo, *Generalized Bose-Fermi Statistics and Structural Correlations in Weighted Networks*, Phys. Rev. Lett. **102**, 038701 (2009)

[45] F. Picciolo, T. Squartini, F. Ruzzenenti, R. Basosi, D. Garlaschelli, *The role of distances in the World Trade Web*, Proceedings of the Eighth International Conference on Signal-Image Technology & Internet-Based Systems (SITIS 2012), pp. 784-792 (edited by IEEE) (2013).

[46] J. E. Anderson, *A Theoretical Foundation for the Gravity Equation*, American Economic Review, **69** (1), 106-16 (1979).

[47] J. Bergstrand, *The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence*, The Review of Economics and Statistics, vol. 67, issue 3, pages 474-81 (1985).

[48] A.V. Deardorff, *Determinants of Bilateral Trade:Does Gravity Work in a Neoclassical World?* NBER Working Paper, No. 5377, (1995).

[49] J.E. Anderson, E. Wincoop, *Gravity with Gravitas: A Solution to the Border Puzzle*, NBER Working Paper, No. 8079, (2001).

[50] M.Fratianni, *Expanding RTAs, trade flows, and the multinational enterprise*, Journal of International Business Studies, Vol 40, **7**, 1206-1227, (2009).

[51] Gleditsch, K. S. (2002). *Expanded trade and GDP data.* Journal of Conflict Resolution, 46(5), 712-724.

[52] G. Gaulier and S. Zignago, CEPII Working Paper 23 (2010).

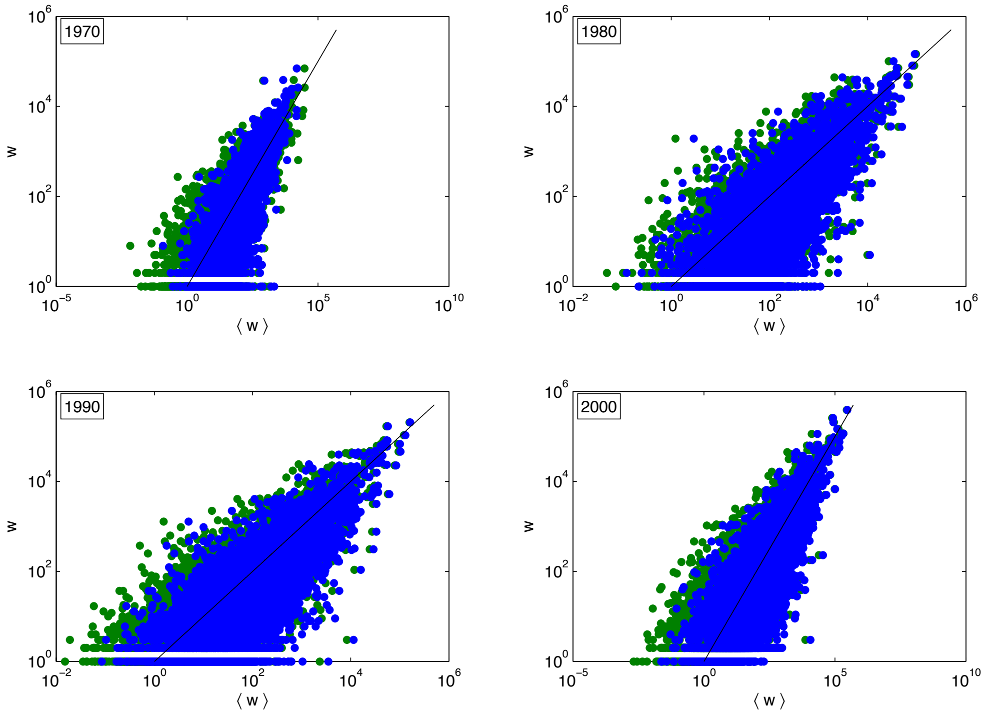

Figure 1: **Empirical versus model-generated trade flows.** Log-log plot comparing the empirical volume (y-axis) of all non-zero bilateral trade flows in the ITN with the corresponding (conditional) expected volume (x-axis) predicted by the Gravity Model defined in 1 (green, parameters estimated via OLS of log-transformed weights) and by the Enhanced Gravity Model defined in 4 and 25 (blue, parameters estimated via Maximum Likelihood as in 23). Top left: 1970, top right: 1980, bottom left: 1990, bottom right: 2000. The black line is the identity line corresponding to the ideal, perfect match that would be achieved if empirical weights were exactly equal to their (conditional) expected values, i.e. in absence of randomness. The results for the two models largely overlap, the main difference being in the fitting procedure and resulting from the fact that the EGM predicts an exponential (plus delta at zero) volume distribution, while the GM predicts a multiplicative gaussian distribution. Note that, even though the expression for the conditional volumes are the same and the same macroeconomic factors have been chosen in both models, the EGM is more parsimonious as it achieves a narrower scatter of points while *not* requiring an extra free parameter (which is instead required in the GM) to estimate the variance.

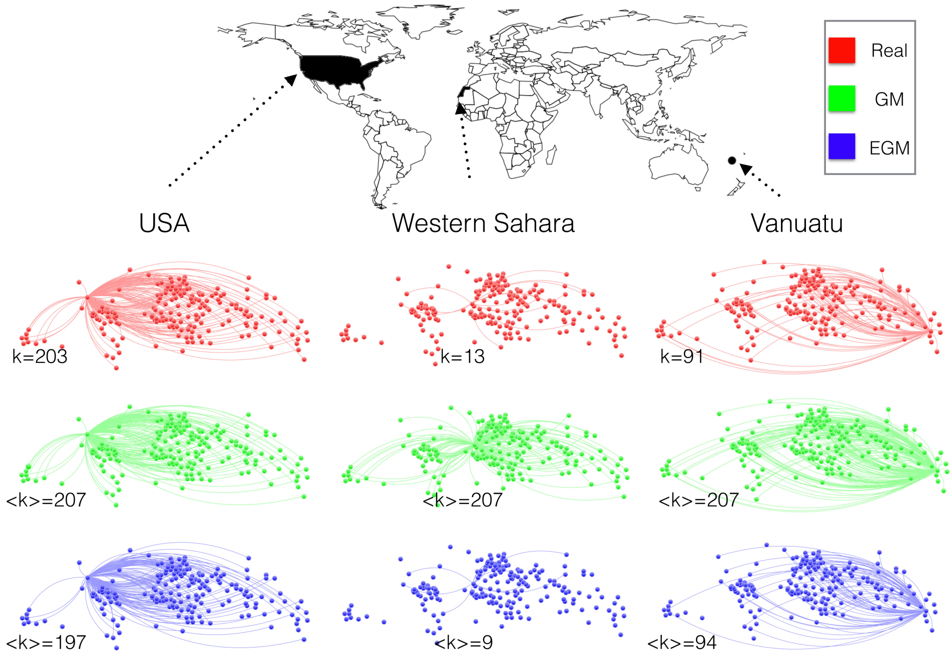

Figure 2: **Country-based network configurations for year 2011 in the real ITN (red), the GM (green) and the EGM (blue).** For three representative countries, we show the connections to all trade partners in the world. The total number of countries in the data (see Materials and Methods) is $N = 208$. The three countries are selected on the basis of their empirical degree k: the country with maximum degree (USA, $k = N - 1 = 207$), the one with minimum degree (Western Sahara, $k = 13$) and one with intermediate degree (Vanuatu, $k = 91$). The GM produces always the maximum number $(N - 1)$ of connections. By contrast, the EGM produces connections randomly with probability p_{ij}, so links change from realization to realization. The expected degree is however independent of the individual realizations and is close to the empirical one for all countries. We have selected a typical realization that produces a degree equal to the expected degree for each of the three countries.

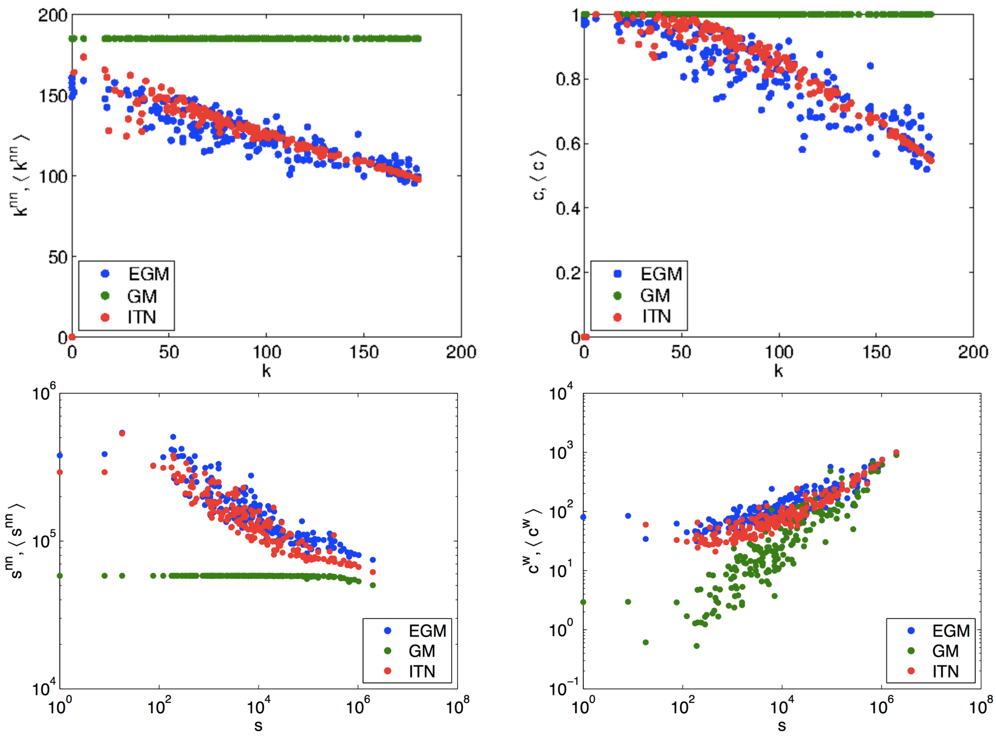

Figure 3: **Network properties in the real ITN (red), the EG (green) and the EGM (blue).** Top left: average nearest neighbor degree k_i^{nn} versus degree k_i for all nodes. Top right: clustering coefficient c_i versus degree k_i for all nodes. Bottom left: average nearest neighbor strength s_i^{nn} versus strength s_i for all nodes. Bottom right: weighted clustering coefficient c_i^w versus strength s_i for all nodes. All results are for the 2000 shapshot of the ITN. For all the other years in the analysed sample, we systematically obtained very similar results. See Materials and Methods for information about the data and all definitions of empirical and observed quantities.