

INSTITUTE  
OF ECONOMICS



Scuola Superiore  
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics  
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy  
ph. +39 050 88.33.43  
institute.economics@sssup.it

# LEM

## WORKING PAPER SERIES

### **Human-Mobility Networks, Country Income, and Labor Productivity**

Giorgio Fagiolo \*  
Gianluca Santoni \*

\* Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy

**2014/08**

**March 2014**

**ISSN (online) 2284-0400**

# Human-Mobility Networks, Country Income, and Labor Productivity

Giorgio Fagiolo\* and Gianluca Santoni†

March 2014

## Abstract

This paper asks whether the level of integration of world countries in the international network of temporary human mobility can explain differences in their per-capita income and labor productivity. We disentangle the role played by global country centrality in the network from traditional openness measures, which only account for local, nearest-neighbor linkages through which ideas and knowledge can flow. Using 1995-2010 data, we show that global country centrality in the temporary human-mobility network enhances both per-capita income and labor productivity. Our results hold cross-sectionally, as well as in a dynamic-panel estimation, and take into account potential endogeneity issues. Our findings imply that how close a country is to the theoretical technological frontier, depends not only on how much she is open to temporary human mobility, but mostly on whether she is embedded in a web of relationships connecting her with other influential partners in the network. Our exercises also suggest that most of the gain in income and productivity can be attained if country centrality in the network comes mostly from influential partners that lie not too far away from, but neither too close to them in the network.

**Keywords:** Temporary human-mobility network; International technological diffusion; Per-capita income; Productivity; Openness to mobility and trade; Centrality.

**JEL Classification:** O33, O47, C21, C26, F22.

---

\*Corresponding Author. Address: Istituto di Economia, Scuola Superiore Sant'Anna, Piazza Martiri della Libertà, 33. 56127 Pisa, Italy. Email: [giorgio.fagiolo@sssup.it](mailto:giorgio.fagiolo@sssup.it)

†Istituto di Economia, Scuola Superiore Sant'Anna, Pisa, Italy. Email: [gianluca.santoni@sssup.it](mailto:gianluca.santoni@sssup.it)

# 1 Introduction

In this paper, we study the network structure of human mobility across the globe, defined in terms of temporary cross-border flows of people, and we explore how such a structure affects cross-country differences in per-capita income and labor productivity (income and productivity, hereafter). We are in particular interested in understanding if world countries that enjoy a globally stronger integration in the human-mobility network are also characterized by higher income productivity, net of traditional explanations including openness to mobility, trade and capital to proxy for international technological diffusion.

Our starting hypothesis, in line with previous studies (Frankel and Romer, 1999), is that cross-border trade of goods and physical capital only imperfectly captures income- and productivity-enhancing interactions between countries in the international technological-diffusion process. Everything else being constant, such a process may also be facilitated the larger a country exposure to foreign influence, through temporary people mobility inducing more communication, exchange of ideas and sharing of tacit knowledge (Arrow, 1969, Lucas, 1993, Keller, 2004).

Here, we argue that cross-country income and productivity differences can be explained not only by the *local* openness of a country to temporary cross-border flows of people (Andersen and Dalgaard, 2011), but also by its *global* integration (or *centrality*) in the international human-mobility network, defined as the graph where nodes are world countries and edges represent cross-country temporary flows of people (possibly directed and weighted). Indeed, traditional country-openness measures only account for direct interactions with neighboring countries. For example, country openness to trade is proportional to a country's total imports and exports (Frankel and Rose, 2002), whereas the openness of a country to the influence of ideas through temporary human mobility is taken to be proportional to the total flow of arrivals and departures in and from a country (Andersen and Dalgaard, 2011).

However, it may be the case that local measures of connectivity such as openness to trade or mobility are not perfectly correlated with indicators accounting for the global embeddedness of a node in the network, which instead fully account for the overall position of a country within the complex web of interconnections in which it is entrenched. This is because nodes may be strongly connected locally, but with partners that are in turn relatively less connected with other important nodes, such as hubs or members of the network core. Conversely, it may be the case that some nodes are poorly connected locally, but they interact with very globally important partners, and therefore enjoy a high centrality score because of a few but relevant preferential links.

Building on this idea, we ask whether the global positioning of countries in the network of temporary human mobility can explain cross-country differences in income and productivity, beyond what openness measures already can do. We hypothesize that in addition to local

spillovers induced by openness to foreign influence, income and productivity can be enhanced also the higher country global centrality in the international network of temporary human mobility. We argue that, for any given level of local openness to mobility, being more globally central means being more globally embedded in a web of relationships that are likely to convey stronger income or productivity-enhancing spillovers. Indeed, *coeteris paribus*, sharing small flows of people with influential countries should be more income and productivity enhancing than interacting with less central nodes.

To test the hypothesis that country centrality in the international network of temporary human mobility matters for country income and productivity, we use 1995-2010 data on the flow of people travelling through the globe for tourism and business purposes. We run a number of econometric exercises where, in addition to openness to mobility and trade, and standard controls, we use as a covariate country centrality in the network of international temporary human mobility. We perform both cross-section and dynamic panel data analyses. Following [Frankel and Romer \(1999\)](#) and [Alcalá and Ciccone \(2004\)](#), in our cross-section exercises we address endogeneity issues by instrumenting openness to mobility, as well as weighted-centrality network indicators built using mobility data, with a gravity equation for bilateral mobility flows ([Keum, 2010](#)). In addition, to control for the possibility that movements of people are simply a substitute for movements of goods, we also include among the regressors trade openness, properly instrumented using again a trade gravity-model specification. Furthermore, we test the robustness of cross-section exercises with dynamic panels using internal instruments for the endogenous variables.

Our main result is that country centrality in the international temporary-mobility network positively and significantly affects per-capita income and labor productivity, even after one accounts for country openness to mobility and trade. This result is robust to a number of alternative econometric specifications and to different measures of country centrality. We also find that most of the gain in income and productivity can be attained if country centrality in the network stems mostly from influential partners that lie not too far away, but neither too close, from them in the network. This suggests the existence of an interesting non-linearity in the relationship between centrality and distance in the human mobility network, and hints to the presence of some complementarity between knowledge exchanged with nearest neighbors as opposed with that shared with more distant countries.

The remaining part of the paper is organized as follows. In [Section 2](#) we review the relevant literature. In [Section 3](#) we describe the data we employ in our analyses and we formally define the international network of temporary human mobility, discussing some of its topological properties (more details on this is in [Appendix A](#)). We back-up our main hypotheses with a theoretical framework discussed in [Section 4](#). The econometric setup and estimation strategy are presented in [Section 5](#). [Section 6](#) contains our main results, whereas [Section 7](#) concludes.

## 2 Literature Review

Our paper contributes to a flourishing literature that emphasizes how the behaviors of individuals, households, industries and countries can be explained not only by what their *direct* neighbors do (e.g., via externalities, peer effects, and strategic interdependencies), but also by the extent to which they are globally embedded in a network of relationships *indirectly* connecting them with all the other agents in the relevant population through chains or paths of links (Acemoglu et al., 2012, 2013, Ductor et al., 2013). This implies that it becomes crucial to get a deep understanding of the topological properties of the network connecting the agents, and identify how these properties can in turn affect the behavior of the agents (Schweitzer et al., 2009).

Such a research avenue has been boosted by a rising number of papers exploring the structure and consequences of economics networks, including trade (Fagiolo et al., 2009), production (Hausmann and Hidalgo, 2011), input-output interdependencies (McNerney et al., 2012) and finance (May et al., 2008, Haldane and May, 2011, NP, 2013). Furthermore, in the last years, the topics of international migration and human temporary mobility have received an increased attention in the field of network science, for their straightforward connections with crucial large-scale phenomena like global-epidemic spreading (Balcan et al., 2009).

The basic idea is that the ways in which people move across the planet at different spatio-temporal scales can be characterized using a complex-network approach, where nodes and links represent locations and flows between them. Using data ranging from survey and census databases, all the way to circulation of banknotes, commuter flows, mobile phone users, geo-located tweets, Facebook profiles and GPS records, a number of contributions have uncovered statistical regularities and anomalies concerning the networks of short-range people mobility (Brockmann et al., 2006, Brockmann and Theis, 2008, Simini et al., 2012, Csaji et al., 2013, Hawelka et al., 2013, Lu et al., 2013, Wang et al., 2014), international human migration (Tranos et al., 2012, Fagiolo and Mastrorillo, 2013a, Davis et al., 2013, Hofleitner et al., 2013), inter-urban traffic (De Montis et al., 2007) and infrastructures facilitating how people move across space (e.g., air, cargo and maritime transportation networks; see Barrat et al., 2004, Hua and Zhu, 2009, Kaluza et al., 2010, Woolley-Meza et al., 2011, Kaluza et al., 2011).

Understanding how human-mobility networks are shaped and evolve is important not only from a descriptive perspective, but also because it helps to predict the properties of diffusion processes occurring on them, e.g. global epidemics (Hufnagel et al., 2004, Eubank et al., 2004, Colizza et al., 2006a, Viboud et al., 2006, Balcan et al., 2009, Bajardi et al., 2011, Brockmann and Helbing, 2013) and computer viruses (Kleinberg, 2007), and to improve the design of traffic systems and urban planning (Horner and O’Kelly, 2001), traffic forecasting (Kitamura et al., 2000) and localized recommendations (Clements et al., 2011, Scellato et al., 2011).

Here, we employ data on people international travels for business and tourism purposes

to build a sequence of international cross-border temporary mobility networks. These data have been previously studied from a complex-network perspective in [Miguéns and Mendes \(2008\)](#) and [Brahmbhatt and Menezes \(2013\)](#), who highlight the non-trivial structure of the international temporary-mobility network. Their main finding is that the binary backbone of the network is almost random, whereas after accounting for heterogeneity of mobility-flow intensity by weighting the links with the number of people traveling on each corridor, one observes scale-free behavior. Interestingly, these contributions show that cross-border temporary mobility networks look more like other technological and economic networks, than like social and cultural ones. This suggests that an important role in shaping these networks is played by technology and knowledge diffusion across international borders.

The present paper exploits this intuition, in line with previous studies that attempted to gauge the impact of mobility on country income and productivity in the aggregate. In particular, we expand upon ([Andersen and Dalgaard, 2011](#)), who show that country openness to mobility facilitates the international spread of ideas and therefore enhances income and productivity.<sup>1</sup> More generally, the impact of the flow of goods (international trade) and people (migration) on country income and productivity has been investigated by several studies. For example, [Frankel and Romer \(1999\)](#) and [Alcalá and Ciccone \(2004\)](#) have employed trade openness —measured as the ratio of international over domestic economic activity— to evaluate the importance of cross-border flows of goods in the process of international technological diffusion. Similarly, [Ortega and Peri \(2013\)](#) build measures of openness to migration and show the existence of a positive effect of country integration on long-run income per capita.<sup>2</sup>

Here, we posit that openness measures do not fully account for the importance of interaction-related effects in the international diffusion process, as they do not capture spillovers coming from indirect linkages.<sup>3</sup> In other words, direct relationships in the network may not be sufficient to gauge the overall importance of a country in the whole web of interactions, which should be instead evaluated after having factored in the computation the importance of each partner, that of the partners of a partner, and so on. The role played by indirect network effects in explaining country dynamics has been investigated in the context of international trade by [Abeyasinghe and Forbes \(2005\)](#), who show that bilateral imports and exports can only explain a small fraction of the impact that an economic shock originating in a given country can have on another one, which is not among its direct-trade partners. Similarly, [Dees and Saint-Guilhem \(2011\)](#) report that countries that do not trade very much with the U.S. are largely influenced by its dominance over other trade partners linked with the U.S.

---

<sup>1</sup>On the direct effects of cross-border movements of people and productivity in European NUTS regions, see [Gambardella et al. \(2003\)](#).

<sup>2</sup>[Alesina et al. \(2013\)](#) also show that within-country heterogeneity in skilled immigration positively affects income and productivity.

<sup>3</sup>The importance of cross-country spillovers in explaining country performance, at both a global and local levels, has been explored in [Conley and Ligon \(2002\)](#) and [Moreno and Trehan \(1997\)](#).

More generally, [Ward et al. \(2013, p. 96\)](#) argue that “bilateral trade is not independent of the production, consumption, and trading decisions made by firms and consumers in third countries”. [Lee et al. \(2011\)](#) show that indirect links in the international-trade network are central in explaining how crises diffuse.

More generally, the fact that node connectivity and centrality in a network may explain node-related attributes has been widely documented in both theoretical and empirical contributions. For example, [Kali et al. \(2007\)](#), [Kali and Reyes \(2010\)](#), [Duernecker et al. \(2012\)](#) show that the level of country connectivity and integration in the international-trade network explains country growth and the patterns of diffusion of financial crises. [Reyes et al. \(2010\)](#) provide evidence supporting the idea that country centrality in the trade network can explain divergence in country development profiles better than traditional measures such as trade openness do. From an international finance perspective, country connectedness and centrality in financial networks have been used to predict the performance of countries during recessions and business cycle downturns ([Chinazzi et al., 2013](#), [Hale, 2012](#), [Minoiu et al., 2013](#)). Finally, [Fagiolo and Mastrorillo \(2013b\)](#) show that country centrality in the international-migration network can explain bilateral trade flows.

### 3 International Mobility: Data and Network Topology

We study international human mobility using data on temporary cross-border flows of people collected by the World Tourism Organization (UNWTO, <http://www2.unwto.org>) for the period 1995-2010.

Data records arrivals and departures of people traveling to a different country with respect to their usual place of residence, and stay there for less than one consecutive year. Outbound data are based on incoming visitors registered by the destination country and encompass both leisure and professional travelers, excluding border and seasonal workers as well as long term student.<sup>4</sup> As discussed in [Andersen and Dalgaard \(2011\)](#), disaggregated data reporting travels for business purposes only should be preferred. Indeed, such data should better proxy for the mechanisms through which sharing of ideas and diffusion of tacit knowledge actually occur. However, UNWTO data report tourism-business breakdowns solely for country-aggregate figures, and not for outbound bilateral flows. Therefore, following [Andersen and Dalgaard \(2011\)](#), we have chosen to employ in our analyses total bilateral travel flows (business plus tourism), although our results are quite robust to alternative measures of bilateral flows computed by multiplying all outbound total bilateral flows by constant business-travel shares reported for country aggregates.

---

<sup>4</sup>Complete bilateral data are available only for a limited subset of countries. In some cases, therefore, only arrival data from the most relevant origins are recorded, whereas the remaining flows are treated as being equal to zero.

After having removed countries that do not report either arrivals or departures in at least one year between 1995 and 2010, we are left with a final mobility dataset with yearly origin-destination outbound flows for  $N = 213$  world countries.<sup>5</sup> For each year  $t = 1995, \dots, 2010$ , we then build a  $N \times N$  weight matrix  $W^t$ , whose generic entry  $w_{ij}^t$  records the number of travellers who left from country  $i$  and arrived in country  $j$  during year  $t$ . More formally, we define the International Human Mobility Network (IHMN) as the  $N$ -sized graph where nodes are countries and links are weighted by the entries of the matrices  $W^t = \{w_{ij}^t\}$ ,  $i, j = 1, \dots, N$  and  $t = 1995, \dots, 2010$ . Note that  $w_{ii} = 0$  for all  $i$  and in general  $w_{ij} \neq w_{ji}$ , i.e. the network is directed. Accordingly, we define the binary projection of the IHMN as the network where a link exists if and only if the correspondent weight is strictly positive, resulting in adjacency matrices  $A^t = \{a_{ij}^t\}$ . In our econometric exercises, we will mostly employ symmetrized (undirected) versions of the IHMN, characterized by weight matrices  $\bar{W}^t = \{\bar{w}_{ij}^t\}$ , where  $\bar{w}_{ij}^t = w_{ij} + w_{ji}$ , and by adjacency matrices  $\bar{A}^t = \{\bar{a}_{ij}^t\}$ , where  $\bar{a}_{ij}^t = \max\{a_{ij}, a_{ji}\}$ . In other words, in the symmetrized version of the IHMN we do not distinguish between the direction of travel corridors and we aggregate in-coming and out-going link weights.

Table 1 summarizes the evolution over time of the main IHMN topological properties. Note that the volume of the network almost doubles during the period under observation. The link-weight distribution appears to be extremely skewed to the right, and over the years it seems to have accentuated its skewness. The number of links in the network, and thus its density, substantially increases, due to a stronger integration and globalization of travels. This resulted in an increased binary symmetry, as unilateral corridors were reciprocated over time. The increasing weighted-asymmetry indicator (Fagiolo, 2006) also suggests that arrivals and departure flows have become more similar over time. Whereas the undirected IHMN is fully connected, its directed version features a giant component (with increasing size) and a large number of small disconnected subsets of countries, often made of singletons. The increasing connectivity of the IHMN over time can be inferred also by the decreasing average-path length, in both the undirected and directed case.

In Appendix A we present a detailed statistical analysis of the IHMN structure. Here, we focus instead on the patterns of country centrality in the IHMN. In particular, for the sake of the subsequent analysis, we move now to assess whether different node-centrality and openness indicators, computed on both the binary and weighted version of the *symmetrized* network, are correlated both among themselves and with traditional openness variables.

We begin by defining country openness to mobility (OPMOB) as the ratio between a country total (inward and outward) traveller flows in the IHMN and country workforce. Similarly, country openness to trade (OPTRA) is defined, as usual, as the ratio between total trade (country imports plus exports) and country (PPP) real GDP.<sup>6</sup>

<sup>5</sup>This covers about 96% of total bilateral flows in the database in each year.

<sup>6</sup>To double check our results, we have computed country total travel inflows and outflows using both IHMN



To proxy country integration within the IHMN, we calculate a number of *local* and *global* node centrality indicators, on both the binary and weighted versions of the network. We focus on *undirected* centrality measures, i.e. computed on the matrices  $(\bar{W}^t, \bar{A}^t)$ . This choice stems from the fact that in our subsequent analyses we do not want to discriminate between in-coming and out-going flows of people. Furthermore, undirected flows are more in line with basic openness indicators, which do not distinguish between in- and out-flows of people and commodities. Local centrality measures are proxied by node degree (total number of links) and node strength (sum of node link weights). As to global centrality indicators, we use Bonacich eigenvector centrality (Bonacich, 1987, Bonacich and Lloyd, 2001) and Katz centrality (Katz, 1953).<sup>7</sup> Bonacich centrality measures the influence of a node in a network, based on the idea that a few connections to most central nodes contribute more to the centrality of the node in question than many homogeneous connections to low-scoring nodes. Katz centrality modifies Bonacich centrality by introducing an attenuating factor  $\xi < 1$  that penalizes connections with distant neighbors.<sup>8</sup>

Table 2 reports our results for year 2010 (similar results hold also for all the other years in the sample). First note that centrality indicators, irrespective of whether they are local or global, or they are computed on binary and weighted versions of the IHMN, are positively and often very highly correlated between each other. This suggests that all these centrality measures can be taken as a proxy of the level of integration of a country in the network without much loss of information. The only exception is the correlation coefficient between node degree and strength, which is positive but not very high, indicating that countries holding many travel corridors in the IHMN do not necessarily feature large total flows. Second, openness measures are only weakly correlated with centrality indicators, indicating that the latter are conveying fresh and perhaps useful information, beyond that of openness, as far as measurement of country embeddedness in the network is concerned. Finally, the correlation between mobility and trade openness is positive but not extremely high. This militates in favor of the idea that movement of goods is not a perfect substitute for people-driven interactions.

---

link weights and also employing data from World Development Indicators (WDI), World Bank, available at <http://data.worldbank.org/data-catalog/world-development-indicators>. Results are robust to such alternative. OPTRA has been computed using BACI trade data (Gaulier and Zignago, 2010) and real (PPP) GDP from WDI.

<sup>7</sup>Similar results hold using the Page Rank (Brin and Page, 1998) indicator and, in directed versions of the IHMN, hub and authority scores (Kleinberg, 1999)

<sup>8</sup>The value of the attenuation factor is chosen such that it is smaller than the reciprocal of the absolute value of the largest eigenvalue of the adjacency or weighted matrix.

## 4 Theoretical Framework

This section discusses some theoretical foundations for the econometric setup that we employ to test our main hypotheses.

We begin by extending the conceptual framework in [Andersen and Dalgaard \(2011\)](#) to include network-based centrality measures in the explanation of country productivity and income differentials. Consider a human-capital augmented growth model ([Solow, 1956](#)), with a Cobb-Douglas constant-return-to-scale production function. The steady state level of per-capita (or per-worker) GDP in each country is given by:

$$y_t^* = \left( \frac{I}{n + \delta + g_t} \right)^{\frac{\beta}{1-\beta}} \Phi_t H, \quad (1)$$

where  $\Phi_t$  is the technology parameter (which grows at the rate  $g_t$ ),  $H$  is human capital,  $I$  is the investment rate,  $n$  is labor-force growth,  $\delta$  is capital depreciation and  $\beta$  represents the share of capital. Following [Nelson and Phelps \(1966\)](#), the rate of growth of the technological parameter can be assumed to be increasing in the gap with world-efficiency frontier  $\Phi_t^w$ :

$$\hat{\Phi}_t = \frac{1}{\Phi_t} \frac{\partial \Phi_t}{\partial t} = \gamma \left( \frac{\Phi_t^w - \Phi_t}{\Phi_t} \right) \quad (2)$$

where  $\gamma$  is a country-specific parameter. Assuming that world technology evolves at a constant rate, one gets that the frontier is determined exogenously by a “theoretical technology”<sup>9</sup>. In the long-run, the rate of technological progress will converge to the exogenous world efficiency growth. As a consequence, growth differences across countries will be temporary. Instead, cross-country differences in income (and productivity) levels will be determined by  $\gamma$ . In the steady state, country efficiency grows at the same rate as that of the world frontier, whereas the equilibrium level is given by  $\Phi_t^* = \frac{\gamma}{\gamma+g} \Phi_t^w$ .

The parameter  $\gamma$  captures a country’s ability to adopt the “theoretical technology”: a higher ability implies that a country will be closer to the frontier. However,  $\gamma$  can also be interpreted as a proxy for the rate of technology diffusion and knowledge spillovers across countries. Indeed, along with country specific characteristics (e.g., R&D expenditure), the speed of technological adoption may be influenced by cross-country interactions (such as cross-border travel flows), which help the circulation of ideas and innovation through a positive-spillover knowledge effect.

More formally, suppose that each country  $i$  engages in activities devoted to maximize its  $\gamma_i$ , exerting an effort  $e_i$ . In presence of strategic complementarities,  $\gamma_i$  may additively depend on two terms. The first one gives the benefits and costs of exerting the effort  $e_i$ . The second

<sup>9</sup>[Nelson and Phelps \(1966\)](#) define the “theoretical level of technology” as the most efficient technology that would prevail in case of instantaneous diffusion.

one reflects the strategic-complementarity effect and fully depends on country  $i$ 's position in the network, via its link weights. Assume for simplicity that the network is symmetric, so that one does not distinguish between in-coming and out-going interaction channels. Hence, following [Jackson and Zenou \(2014\)](#), we can write:

$$\gamma_i(\underline{e}, \underline{w}_i) = \alpha e_i - \frac{1}{2} e_i^2 + \psi \sum_{j=1}^N w_{ij} e_i e_j \quad (3)$$

where  $\alpha$  and  $\psi$  are positive parameters,  $N$  is the number of countries,  $w_{ij}$  is the link weight between country  $i$  and  $j$  in the IHMN for a given year,  $\underline{w}_i$  is the  $i$ -th row of the weight matrix  $W = \{w_{ij}\}$ , and  $\underline{e}$  is the vector of efforts. Note that Eq. (3) holds also if we replace link weights with binary adjacency-matrix entries  $a_{ij}$ .<sup>10</sup> In this quadratic benchmark model, players are ex-ante homogeneous in terms of observable characteristics  $(\alpha, \psi)$  and their heterogeneity is fully due to their position in the network. When two countries interact through cross-border travels ( $w_{ij} > 0$ ), the cross derivative of  $\gamma_i$  in Eq. (3), capturing spillover effects, is positive and equal to  $\psi$ . If all countries choose their effort playing a one-shot game with strategic complementarities, [Ballester et al. \(2006\)](#) have shown that (provided that  $\psi$  is not too large) the unique pure-strategy Nash equilibrium is given by  $\mathbf{e}^* = \alpha(\mathbf{I} - \psi\mathbf{W})^{-1}\mathbf{1}$ , that is the vector of country Bonacich (eigenvector) centrality ([Bonacich, 1987](#), [Bonacich and Lloyd, 2001](#)) in the IHMN.

Therefore, Bonacich (eigenvector) centrality, and more generally centrality measures taking into account the global importance of countries in the network, embody interaction effects in the equilibrium behavior. Net of country-specific characteristics, more central countries in the IHMN should be closer to the theoretical technological frontier and should exhibit higher income levels and productivity. Before formally testing this hypothesis, it is worth noting that in our sample the unconditional correlation among logged income per capita and country Bonacich centrality is typically positive and marginally decreasing. For relatively peripheral countries, an increase in the level of international interactions is indeed associated with a greater increase in real income (as the consequence of moving closer to the technological frontier).

---

<sup>10</sup>The network structure is exogenously-given. In other words, one assumes that the country-distance from the theoretical frontier (and this its centrality, income or productivity) is the long-run outcome of the underlying game.

# 5 Econometric Setup and Estimation Strategy

## 5.1 Cross-Section Exercises

We test the hypothesis that country centrality affects income and aggregate productivity, net of country openness to both mobility and trade, in both cross-section and dynamic setups. We begin by a cross-section exercise where our dependent variable  $y_j$  is year- $t$  GDP per-capita (income) and per-worker (labor productivity), both measured in PPP, of the destination country  $j$ . Notice that we suppress, for the sake of simplicity, time superscripts in cross-section exercises.

In a given year  $t$ , our preferred specification reads:

$$\ln(y_j) = \alpha_0 + \beta_0 \log(OPMOB_j) + \beta_1 \log(OPTRA_j) + \beta_2 \log(CENTR_j) + \beta_3 X_j + \epsilon_j \quad (4)$$

where  $OPMOB_j$  and  $OPTRA_j$  are country openness to mobility and trade;  $CENTR_j$  is country-centrality in the IHMN (see Section 3); and  $X_j$  is a vector of country-specific controls, including geography and environmental variables, and institutional-quality indicators. More precisely, we employ the proportion of land area in tropics and sub tropics (Gallup et al., 1999); the distance from the equator; a dummy controlling for whether the country is landlocked; rent from natural resources over GDP; a measure of the quality of institutions; and a full set of regional dummies.<sup>11</sup> Finally,  $\epsilon_j$  represents an idiosyncratic shock.

Estimating Eq. (4) using plain-vanilla OLS may cause biased results. This may be due to endogeneity issues, as richer countries may attract more travellers and become more open in both trade and travel networks, and more central in the IHMN.<sup>12</sup> This means that  $OPMOB$ ,  $OPTRA$  and  $CENTR$  may all end up being endogenous in Eq. (4). We now discuss how we address this potential issue in our exercises.

### 5.1.1 Endogeneity of Country Openness to Mobility and Trade

Following the identification strategy proposed in Frankel and Romer (1999), we use a 2SLS approach. We employ a gravity-model specification to instrument bilateral traveller flows divided by year- $t$  country  $j$  workforce, defined as  $\omega_{ij} = w_{ij}/N_j$ . We do so employing bilateral interaction effects controlling for geographical, political, cultural and economic factors, as it is

---

<sup>11</sup>Coordinates of country capitals and landlock dummies are from Mayer and Zignago (2011). Quality of institutions is measured using the institution-quality index Polity-2 from Marshall and Jaggers (1999). This variable scores governments on a range from -10 (most repressive) to +10 (most democratic). See Appendix B for more details on data sources and variable definitions. Note that we do not control here for differences in the intensity with which different technologies are used in different countries. As shown in Comin et al. (2008), this can be an additional income or productivity enhancer.

<sup>12</sup>Note that an omitted-variable bias may also plague our regressions, e.g. because transport and communication costs may drive both traveller flows, income and productivity.

customary in the trade and migration literature (Frankel and Romer, 1999, Ortega and Peri, 2013, Alesina et al., 2013, Beine et al., 2011). To account for the fact that our dependent variable in Eq. 4 is income (or productivity) in the destination country, we set up the gravity specification including population and area of destination country (as a size measure), geographical distance between the two countries, and dummies accounting for common borders, colonial ties and common language. As to “multilateral resistance” (Anderson and van Wincoop, 2003), we include only origin-country dummies to control for all push factors that may influence travel patterns.<sup>13</sup> Indeed, inclusion of destination-country multilateral terms would introduce endogenous variation in the fitted values, since it would account for all “pull” factors of location decisions, including income and productivity.<sup>14</sup>

Therefore, we fit the following instrumental gravity equation for workforce-rescaled bilateral traveller flows:

$$\begin{aligned} \omega_{ij} = & \alpha_0 \cdot POP_j^{\beta_1} \cdot AREA_j^{\beta_2} \cdot DIST_{ij}^{\beta_2} \cdot \\ & \exp\{\beta_5 BORDER_{ij} + \beta_5 LANGUAGE_{ij} + \beta_6 COLONY_{ij} + \\ & \beta_7 LANDLOCKED_{ij} + \beta_8 \Phi_{ij} + \delta_i\} \cdot \eta_{ij}, \end{aligned} \quad (5)$$

where  $\omega_{ij}$  is the share of travelers from country  $i$  to country  $j$  to the workforce of country  $j$ ;  $POP_j$  and  $AREA_j$  are population and area of the destination country  $j$ ;  $DIST_{ij}$  is geographical (geodesic) distance;  $BORDER_{ij}$ ,  $LANGUAGE_{ij}$ ,  $COLONY_{ij}$  and  $LANDLOCKED_{ij}$  are dummies for common border, language, colonial ties, and land-locking effects. Furthermore, to increase the explicative power of the instruments, we include the vector  $\Phi_{ij}$  featuring interaction effects between the border dummy and other regressors. The origin-specific dummy  $\delta_i$  controls instead for multilateral resistance.

We estimate cross-sectionally Eq. (5) with three methods: (i) plain OLS for strictly-positive  $\omega_{ij}$  only; (ii) Poisson pseudo maximum-likelihood (PPML) estimator (Santos Silva and Tenreyro, 2006) using both zero and strictly positive flows; (iii) PPML estimator using only strictly positive flows. While the first method is applied after log-linearizing Eq. (5), the second and third method uses the non-linear specification. Indeed, PPML estimators are known to perform better in case of heteroscedasticity and inflation of zeros (in our dataset the share of zero flows is 66%, a relatively large figure).

Starting from estimated coefficients from Eq. (5), we employ the predicted values  $\hat{\omega}_{ij}$  to

<sup>13</sup>Bertoli and Fernandez-Huertas Moraga (2013) discuss sufficient conditions under which the inclusion of one-way fixed effects allows to control for multilateral resistance (e.g., travelers have homogenous preferences across locations).

<sup>14</sup>Results do not dramatically change if we fit a gravity structural specification with two-side country dummies, where in constructing the instrument we do not consider destination fixed effects.

build an instrumented version of  $OPMOB_j$ , defined as:

$$\widehat{OPMOB}_j = \sum_{i \neq j} \hat{\omega}_{ij}, \quad (6)$$

which we then use in the cross-section estimation of Eq. (4)

In addition to openness to mobility, also country openness to trade (OPTRA) can generate endogeneity issues. To address this problem, we employ two alternative strategies. First, following [Frankel and Romer \(1999\)](#) and [Alcalá and Ciccone \(2004\)](#), we use the same gravity-equation approach that we used for mobility, and we fit Eq. (5) using as dependent variable the share of total trade (exports plus imports) to (PPP) real GDP of the destination country, i.e.  $\tau_{ij} = (\exp_j + \text{imp}_j)/\text{rGDP}_j$ . We then use predicted values  $\hat{\tau}_{ij}$  from the gravity equation to build an instrumented version of  $OPTRA_i$ , which reads:

$$\widehat{OPTRA}_j = \sum_{i \neq j} \hat{\tau}_{ij}. \quad (7)$$

Secondly, we double check our results using as an instrument for year- $t$  openness to trade its prior 10-year average.

### 5.1.2 Endogeneity of Network Centrality Measures

Identification of the effect of country centrality on income and productivity can be subject to endogeneity problems as well, as countries can travel their way to more central positions in the IHMN while they become richer or more productive. Again, we follow two strategies to limit endogeneity issues related to network centrality.

First, we employ binary centrality measures (computed on  $\bar{A}$  matrices), instead of weighted ones, and we instrument them with their values at a 5-year lag. Indeed, binary centrality measures may be thought to be less subject to endogeneity, as it is increasingly difficult that higher levels of country income or productivity can end up in the deletion of existing links or the creation of new links in the IHMN, and thus affect country-centrality scores. Furthermore, instrumenting binary centrality with their 5-year lag should further reduce endogeneity issues, while preserving a good proxy for contemporaneous centrality.

Second, we include weighted centrality indicators (computed on  $\bar{W}$  matrices) in our regressions and we instrument them with gravity predictions. To do so, we fit Eq. (5) replacing the dependent variable with actual IHMN travel flows  $w_{ij}$ , as taken from the original weight matrix  $W$ . We then use the predicted values  $\tilde{w}_{ij}$  from a PPML estimation to build a predicted (directed) weight matrix  $\tilde{W}$ , which we employ to compute weighted country-centrality indicators. These are eventually used as instruments in the second-stage estimation of Eq. (4). Notice that PPML gravity estimation is done on strictly-positive flows only. This choice al-

lows one to keep fixed the underlying binary structure (i.e. the adjacency matrix  $A$ ) and to improve the quality of the instrument.

## 5.2 Dynamic-Panel Exercises

We check the robustness of cross-section exercises in a dynamic setup, which fully exploits the time dimension of the data (1995-2010). We estimate our preferred income and labor productivity specification (Eq. 4) using internal instruments for the endogenous variable, namely lagged dependent variables of mobility and trade openness, and network centrality indicators. We employ a panel-data framework including country-fixed effects to control for all time-invariant determinants of income and labor productivity (e.g., geography and institution quality). Note that accounting for lagged dependent variables also helps one to control for the high persistence of structural variables, such as real income and labor productivity, and implicitly for the influence of the other endogenous variables in shaping the current level of  $y_j^t$ . We estimate the following equation:

$$\begin{aligned} \log(y_j^t) = & \alpha_0 + \rho \log(y_j^{t-1}) + \beta_1 \log(OPMOB_j^t) + \\ & + \beta_2 \log(OPTRA_j^t) + \beta_3 \log(CENTR_j^t) + \delta^t + \mu_j + \epsilon_j^t, \end{aligned} \quad (8)$$

where  $\mu_j$  controls for country-specific determinants,  $\delta^t$  is a set of time dummies and  $\epsilon_j^t$  is the error term. Since the dependent variable is likely to be very persistent, we employ a System-GMM estimator (Arellano and Bover, 1995, Blundell and Bond, 1998, Bond, 2002) in order to consistently estimate our parameters of interest and possibly overcome the problem of simultaneous-equation bias. We perform standard checks about our identifying assumptions and exogeneity of the instruments. Furthermore, we address endogeneity issues using a set of instruments  $Z_{it}$ , including two-year lags of the potentially endogenous covariates, i.e. the lagged dependent variable, mobility and trade openness, and network centrality indicators.<sup>15</sup>

# 6 Results

## 6.1 Cross-Section Estimation

This section reports results of cross-section estimations for year 2000.<sup>16</sup> We firstly discuss first-stage gravity estimations for instrumenting OPMOB and OPTRA. Next, we estimate a second-stage equation where our dependent variables are country income and productivity.

<sup>15</sup>Appendix C discusses the appropriateness and performance of a System-GMM estimator vs. existing alternatives.

<sup>16</sup>This is done in order to avoid dealing with post-crisis issues. Results are however robust in all other years. Data, codes and additional results are available from the Authors upon request.

Initially, we do so without introducing centrality measures. This allows us to compare our results with those in [Andersen and Dalgaard \(2011\)](#). Then we introduce network variables to see if country centrality can explain income and productivity beyond what country openness already does.

### 6.1.1 First-Stage Gravity Models for Mobility and Trade

Table 3 reports the results of the gravity equation from which we instrument mobility openness. In order to mitigate as much as possible measurement errors, we employ 1995-2000 averages for all our variables.

To begin with, note that the difference between OLS and full-sample PPML estimated coefficients are relatively small. Wald tests suggest that such differences are statistically significant (at 5%) for contiguity, language, colony and, marginally, for landlocked variables (at 10%). Instead, coefficients estimated with PPML on strictly-positive flows only (column 3) are very similar to those estimated on the full sample. This suggests that the divergence among OLS and PPML is mainly due to heteroscedasticity, and not sample truncation ([Santos Silva and Tenreyro, 2006](#)). To get a higher consistency, we therefore decide to use full-sample PPML estimates to construct our instruments for OPMOB.

Notice also that all coefficients exhibit the expected sign and significance levels: mobility flows increase if the two countries share the same language and colonial ties, while decrease with distance and if the destination country is landlocked. Noteworthy, the elasticity of travel to distance (-1.58) is remarkably similar to the -1.54 figure found by [Andersen and Dalgaard \(2011\)](#), who however employ a smaller sample and information set for year 2005.

We then plug predicted values from the gravity estimation into Eq. (7) to build our instrument for OPMOB. The correlation between the original and the instrumented variable turns out to be strong (0.74). Furthermore, regressing the log of observed mobility openness on its predicted values returns a highly significant slope of 1.3, with a  $R^2 = 0.55$ . This militates in favor of the soundness of our instrument for OPMOB.

Results for the trade-gravity equation are reported in Table 3. Note that geographical determinants of bilateral trade (i.e., distance and land-locking) are significant and with the expected sign and magnitude. The elasticity to geographical distance is significantly smaller in the PPML estimation than with OLS, thus confirming our preference for the non-linear estimator. Furthermore, the distance coefficient almost halves as compared to the gravity model for mobility.

As we did for mobility, we employ full-sample PPML estimates (second column) to build our instrument for OPTRA. We find that the correlation between observed and predicted OPTRA values is slightly lower than in case of mobility (0.65). However, a regression of OPTRA observed values vs. their predicted values gives a highly-significant slope coefficient



equal to 0.86 ( $R^2 = 0.42$ ).

Overall, a gravity approach seems to give strong instruments for both OPMOB and OPTRA. This is confirmed in both cases by very high F-tests for the first stages. However, it is worth noting that the two instruments are highly correlated with each other. After removing the effect of the exogenous covariates  $X_j$  appearing in Eq. (4), the residual correlation among predicted OPMOB and OPTRA is still very high (0.83). This suggests that the simultaneous inclusion of the two predicted measures in the second stage of the estimation may weaken identification and make inference less reliable. We will go back to this point below.

### 6.1.2 Estimation of Income and Productivity Equations

We now present our main results for the second stage of our 2SLS estimations of the effect of mobility on country income and productivity. We begin by estimating Eq. (4) without introducing network effects (CENTR), see Table 5. We instrument OPMOB with predictions from the gravity model, and OPTRA with our two alternative strategies (gravity predictions vs. prior 10-year averages).

We find that openness to mobility positively and significantly affects income (column 1) and labor productivity (column 3) if we do not control for trade openness. If we control for trade using its prior 10-year average, the estimated OPMOB coefficient is smaller and less significant for both income and productivity (columns 2 and 5). Furthermore, if we instrument trade openness using the gravity-based procedure, openness to mobility becomes even smaller and statistically not significant, cf. columns (3) and (6). The watering-down effect induced by openness to trade is evident also when one studies the quality of the instruments. The Kleibergen-Paap (KP) test (Stock and Yogo, 2005) is always above 15% critical values, but the inclusion of both instruments reduce drastically the KP statistics (columns 3 and 6). As mentioned above, this suggests that the high correlation among mobility and trade instruments may weaken identification. This is in line with Andersen and Dalgaard (2011), who find that introducing a gravity-based instrument for trade openness results in statistically-insignificant effects of openness to mobility. Note also that OPTRA is never significantly different from zero, both when measured as a prior 10-year average (columns 2 and 4) or as a gravity-instrumented variable (columns 3 and 6). As to the controls, they all appear to have expected signs and significance. Institutional quality has a positive (but statistically weak) effect on both income and labor productivity. Availability of natural resources has also a positive effect, stable and strongly significant over all specifications. Conversely, the percentage land in the tropics and the landlocked dummy have both a negative impact on income, whereas country distance from equator does not seem to play any role in explaining income and productivity.

Let us now turn to estimation of Eq. (4) with country centrality indicators. We begin with binary centrality, i.e. we employ adjacency matrices  $\bar{A}^t$  for the symmetrized version of the

IHMN. In order to deal with potential endogeneity, we instrument contemporaneous binary centrality values with their prior 5-year value. As mentioned above, the binary network is much more robust to endogeneity issues, as the effect of country income and productivity on mobility corridors should be weaker than the one on mobility flows (i.e., removing or adding links is less likely than changing link-weight levels). In other words, paralleling the trade literature, it is more likely that country income and productivity feedback on the *intensive* margins of mobility, than on its *extensive* margins. That said, we use a 5-year lag instrumentation to make even less relevant endogeneity concerns.

Results for the second stage of 2SLS estimation are reported in Tables 6 (income) and 7 (productivity).

As far as country income is concerned, it is easy to see that network centrality indicators exert a positive and significant effect, even if OPMOB is included among the regressors and trade openness is controlled for using 1990-2000 averages (Table 6, columns 1, 3, and 5). The estimated coefficient for OPMOB is positive and significant, even if, in terms of magnitude, (global) centrality measures seem to have a stronger effect. The impact of country network centrality remains positive even if we instrument trade openness using gravity predictions instead of lagged values (columns 2, 4, and 6). In such cases, however, OPMOB is no longer significant and the coefficient decreases. As already noted, the KP test weakens with the simultaneous inclusion of mobility and trade gravity-predicted instruments.

Results for labor productivity (Table 7) confirm what we have found for country income. Centrality measures remain positive and significant, whereas the effect of mobility becomes again insignificant when gravity-instrumented trade openness is included in the regression.

What happens instead when one considers *weighted* country centrality? Table 8 shows our 2SLS results for the second stage of the country-income equation (similar results hold for labor productivity). In this case, we instrument country centrality measures with those computed using the predicted IHMN, defined as the network where we keep fixed the observed binary structure and we replace observed weights with predicted ones from a gravity estimation.

It is easy to see that our main results robustly hold across all specifications. This implies that the IHMN plays a role in diffusing knowledge and ideas, and boosting country income and productivity, not only at the extensive level (i.e. in terms of mobility corridors) but also at the intensive level (i.e. in terms of the intensity of flows channeled via existing corridors). At the same time, the similar impact that binary and weighted centrality have on country income and productivity suggests that the heterogeneity of IHMN flows does not seem to be relevant in explaining the relation between country centrality and income or productivity. Indeed, the fact that some countries are related to other ones via preferential corridors conveying large number of people is as important as forming those connections in the first place.

It is also interesting to note how different centrality measures affect our dependent vari-

ables. In both income and productivity regressions, the elasticity of the dependent variable to node degree and strength is the smallest one, whereas Bonacich and Katz centrality seem to boost much more country performance. This can be explained by the fact that node degrees or strengths are local measures of connectivity, counting just the number of direct links held by a country in the IHMN (or the sum of their weights). Conversely, BON and KATZ take into account the overall position of a node in the entire network. Our results indicate that, net of openness to mobility, global centrality plays a more important role than local centrality: being connected with important partners, who are in turn connected with other relevant nodes, allows a country to benefit from the flow of ideas and knowledge attained via human temporary mobility. Furthermore, note that the elasticity of income and productivity to KATZ centrality is larger than the one for Bonacich centrality. Recall that Katz centrality discounts the importance of partners farther away from the node under consideration, even if they are very central, whereas the BON indicator treats the contribution of any country to centrality of the ones under consideration as they were equally distant. This suggests a sort of inverse U-shaped, non linear relation between centrality and distance in the IHMN. Income and productivity can be boosted by exchanges of ideas via direct partners, but most of the benefit may be generated by intermediate countries, that are not nearest neighbors in the IHMN but neither they are too distant. This is possibly because people from those countries are still able to convey fresh knowledge and ideas, which is complementary to that mastered in the country under consideration. Conversely, the flows coming from direct partners may become at some point redundant, or can substitute for the one already present in the destination. Similarly, being indirectly connected with very distant but central countries may imply exchanging knowledge and ideas that are too complementary and distant, and thus less likely to boost income and productivity. This non-linear effect reminds the interplay between strong and weak ties in labor-market networks ([Granovetter, 1973](#)).

## 6.2 Dynamic-Panel Estimation

In this section, we fully exploit the time dimension of our database (1995-2010), and estimate income and labor productivity specifications as in Eq. (8) employing internal instruments for the endogenous variables. Using a dynamic-panel setting allows one to include country fixed effects, which control for all time-invariant determinants of income and labor productivity, such as geographical and historical factors influencing institution quality.

Furthermore, the inclusion of dependent-variable lags in the right-hand side of the equation helps one to control for the high persistence of income and labor productivity, and, implicitly, for the influence of other endogenous variables in shaping the current level of  $y_j^t$ . In particular, the small econometric-sample size, and the high time persistency of our dependent variables, suggest to employ a System-GMM estimator, in order to overcome the problem of

simultaneous-equation bias (Bond, 2002).<sup>17</sup> As mentioned, endogeneity is worked around by including a vector of instruments  $\mathbf{Z}_j^t$ , which in our case features two-year lags of the potentially endogenous covariates (i.e. lagged dependent variable, mobility and trade openness, and network centrality indicators).

We organize our dynamic-panel results in three tables. The first two (Tables 9 and 10) reports results for binary centrality indicators for country income and productivity. In the third one (Table 11), we present our findings for weighted-network indicators (in the case of country income). Furthermore, we employ different lag structures to analyze the robustness of our results. In the first three columns, we use the more parsimonious specification, with only the first two lags for all the endogenous variables. Then, in columns (4)-(6), we extend the lag structure for  $\log(y_j^t - 1)$  until lag 4. Finally, in columns (6)-(9), we employ all available lags as instruments.

We evaluate the robustness of our dynamic-panel regressions with a Hansen J-test, to ensure that instruments are not correlated with the residuals. In particular, the employed dynamic specification suggests to use an Arellano-Bond test. This allows one to check if the residuals of the first-difference equations are not correlated at the second order. An accurate choice of the instruments solves the problem of instrument proliferation (Roodman, 2009a,b). Note also that the number of instruments that we have used does not weaken the Hansen-test statistic. Indeed, if we enlarge the number of instruments, Hansen tests increase and approach 1 (see columns 6-9), indicating that more parsimonious specifications are likely to be more reliable.<sup>18</sup>

Overall, cross-section results are confirmed. The coefficients associated to our centrality indicators are always positive and highly significant. Moreover, country centrality clearly outweighs openness to mobility in impacting the dependent variable in all specifications. In the case of productivity, however, their magnitude becomes weaker and OPMOB becomes statistically insignificant. In addition, the effect of trade openness, consistently with Andersen and Dalgaard (2011), is never statistically different from zero. Weighted-centrality results for income (Table 11) are in line with binary ones, with OPMOB that keeps its significance only when one inserts all available lags as instruments (similar results hold also for productivity, not shown here for the sake of brevity). Finally, note that the ranking in the impact that different centrality indicators have on income and productivity, which we noticed in cross-section exercises, is preserved also in dynamic-panel estimation, and across binary vs. weighted network variables. Indeed, Katz centrality exerts the strongest influence of both income and productivity, followed by Bonacich centrality and node degree/strength.

---

<sup>17</sup>In Appendix C, we discuss the pros and cons of using a System-GMM estimator vis-à-vis difference GMM estimators (Arellano and Bond, 1991).

<sup>18</sup>See Appendix C for additional robustness checks on the System-GMM estimator.

## 7 Concluding Remarks

Country income and productivity are known to be positively influenced by the process of technological diffusion, which is in turn boosted by cross-border trade of goods and physical capital (Frankel and Romer, 1999). In addition, exchange of ideas and knowledge facilitated by country openness to temporary human mobility can also enhance country per-capita, or per-worker, GDP (Andersen and Dalgaard, 2011).

In this paper, we have argued that, net of such traditional explanations, country income and productivity can be positively affected also by the level of its global integration in the international network of temporary human mobility (IHMN), as measured by its *centrality* in such a graph at each point in time.

Our analysis has proceeded along three main avenues. First, we have defined the IHMN and explored its topological properties. Second, we have presented a simple theoretical framework, whose prediction is that countries that are more central in the IHMN should be closer to the theoretical technological frontier and therefore exhibit higher income and productivity. Third, we have taken this prediction to the data, using both cross-section and dynamic-panel approaches. Our main concern in estimation regards potential endogeneity issues, stemming from reverse-causation linkages from country income and productivity to openness and integration. These problems have been dealt with using an IV gravity-based approach in cross-section exercises, and System-GMM estimators in dynamic-panel setups.

Our main result is that, in addition to country openness to mobility and trade, country income and productivity increases the more countries are central in the IHMN. This implies that the impact of human mobility in the international technological-diffusion process depends not only on how many direct partners a country has (and how strongly it is connected with them), but mostly on whether such a country is embedded in a web of relationships that connect her with other influential partners in the network. Interestingly, our exercises seem to indicate that most of the gain in income and productivity can be attained if their centrality in the network comes mostly from influential partners that lie not too far away, but neither too close, from them in the network. This points to intriguing non-linearities in the relationship between centrality and distance in the IHMN, and suggests the existence of some complementarity between the nature of the knowledge exchanged with nearest neighbors as opposed to that shared with more distant countries.

Our findings consistently hold across a large number of econometric setups. For example, we have employed alternative instrumentation approaches in cross-section estimation; or different estimators and lag structures in dynamic-panel exercises. Furthermore, our main qualitative result does not depend on the network-centrality measure actually used, and are quite robust to the inclusion of both openness to mobility and trade in the regressions, as well as of a wealth of geographical and institutional controls.

The present work can be extended in many ways. Perhaps the most interesting one, and certainly the one ranking first in our agenda, concerns a more thorough analysis of the role and impacts of country integration in the macroeconomic network as a whole. Indeed, countries are embedded in a web of many different types of relationships (or interaction channels) in addition to temporary mobility, ranging from trade and investment to finance and international migration. Such a structure can be characterized in terms of multiplexes (or multi-networks, cf. [De Domenico et al., 2013](#)), where at each point in time a country can be connected to others through many different links, whose presence/absence and intensity may be correlated. One could therefore apply recently-developed techniques to measure the overall centrality of countries in such networks ([Halu et al., 2013](#)), and put to a test several hypotheses, concerning e.g. the relative importance of centrality in different layers in boosting country income, productivity, and performance in terms of GDP growth.

## References

- Abeysinghe, T. and Forbes, K. (2005). Trade linkages and output-multiplier effects: a structural var approach with a focus on asia. *Review of International Economics*, 13:356–375.
- Acemoglu, D., Carvalho, V., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 8:1977–2016.
- Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2013). The network origins of large economic downturns. NBER Working Paper No. 19230.
- Alcalá, F. and Ciccone, A. (2004). Trade and productivity. *The Quarterly Journal of Economics*, 119(2):612–645.
- Alesina, A., Harnoss, J., and Rapoport, H. (2013). Birthplace diversity and economic prosperity. NBER Working Papers 18699, National Bureau of Economic Research, Inc.
- Andersen, T. and Dalgaard, C.-J. (2011). Flows of people, flows of ideas, and the inequality of nations. *Journal of Economic Growth*, 16(1):1–32.
- Anderson, J. E. and van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, 93(1):170–192.
- Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58(2):277–97.
- Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29 – 51.
- Arrow, K. J. (1969). Classificatory notes on the production and transmission of technological knowledge. *American Economic Review*, 59(2):29–35.
- Bajardi, P., Poletto, C., Ramasco, J. J., Tizzoni, M., Colizza, V., and Vespignani, A. (2011). Human Mobility Networks, Travel Restrictions, and the Global Spread of 2009 H1N1 Pandemic. *PLoS ONE*, 6(1):e16591+.
- Balcan, D., Colizza, V., Goncalves, B., Hu, H., Ramasco, J. J., and Vespignani, A. (2009). Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*, 106(21484).
- Ballester, C., Calvo-Armengol, A., and Zenou, Y. (2006). Who’s who in networks. wanted: The key player. *Econometrica*, 74(5):1403–1417.

- Barrat, A., Barthélemy, M., Pastor-Satorras, R., and Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, 101(11):3747–52.
- Beine, M., Docquier, F., and Ozden, C. (2011). Diasporas. *Journal of Development Economics*, 95(1):30–41.
- Bertoli, S. and Fernandez-Huertas Moraga, J. (2013). Multilateral resistance to migration. *Journal of Development Economics*, 102(C):79–100.
- Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115 – 143.
- Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology*, 92(5):1170–1182.
- Bonacich, P. and Lloyd, P. (2001). Eigenvector-like measures of centrality for asymmetric relations. *Social Networks*, 23(3):191–201.
- Bond, S. (2002). Dynamic panel data models: a guide to microdata methods and practice. Cemmap working papers, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.
- Bond, S., Nauges, C., and Windmeijer, F. (2002). Unit roots and identification in autoregressive panel data models: A comparison of alternative tests. 10th International Conference on Panel Data, Berlin, July 5-6, 2002 C5-4, International Conferences on Panel Data.
- Bond, S. R., Hoeffler, A., and Temple, J. (2001). GMM Estimation of Empirical Growth Models. CEPR Discussion Papers 3048, Center for Economic Policy Research.
- Brahmbhatt, J. and Menezes, R. (2013). On the relation between tourism and trade: A network experiment, network science workshop (nsw), 2013 ieee 2nd.
- Brin, S. and Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. In *Proceedings of the seventh international conference on World Wide Web, WWW7*, pages 107–117, Amsterdam, The Netherlands. Elsevier Science Publishers.
- Brockmann, D. and Helbing, D. (2013). The Hidden Geometry of Complex, Network-Driven Contagion Phenomena. *Science*, 342(6164):1337–1342.
- Brockmann, D., Hufnagel, L., and Geisel, T. (2006). The scaling laws of human travel. *Nature*, 439(7075):462–465.



- Brockmann, D. and Theis, F. (2008). Money circulation, trackable items, and the emergence of universal human mobility patterns. *IEEE Pervasive Computing*, 7(4):28–35.
- Chinazzi, M., Fagiolo, G., Reyes, J. A., and Schiavo, S. (2013). Post-mortem examination of the International Financial Network. *Journal of Economic Dynamics and Control*, 37(8):1692–1713.
- Clements, M., Serdyukov, P., de Vries, A. P., and Reinders, M. J. T. (2011). Personalised travel recommendation based on location co-occurrence. physics.soc-ph 1106.5213, arXiv.org.
- Colizza, V., Barrat, A., Barthélemy, M., and Vespignani, A. (2006a). The role of the airline transportation network in the prediction and predictability of global epidemics. *PNAS*, 103(7):2015–2020.
- Colizza, V., Flammini, A., Serrano, M. A., and Vespignani, A. (2006b). Detecting rich-club ordering in complex networks. *Nature Physics*, 2:110–115.
- Comin, D., Hobijn, B., and Rovito, E. (2008). Technology usage lags. *Journal of Economic Growth*, 13(4):237–256.
- Conley, T. G. and Ligon, E. (2002). Economic Distance and Cross-Country Spillovers. *Journal of Economic Growth*, 7(2):157–87.
- Csaji, B., Browet, A., Traag, V., Delvenne, J.-C., Huens, E., Dooren, P. V., Smoreda, Z., and Blondel, V. D. (2013). Exploring the mobility of mobile phone users. *Physica A*, 392(6):1459–1473.
- Davis, K., D’Odorico, P., Laio, F., and Ridolfi, L. (2013). Global spatio-temporal patterns in human migration: A complex network perspective. *PLoS ONE*, (8):e53723.
- De Domenico, M., Solè-Ribalta, A., Cozzo, E., Kivela, M., Moreno, Y., Porter, M. A., Gómez, S., and Arenas, A. (2013). Mathematical Formulation of Multilayer Networks. *Physical Review X*, 3:041022.
- De Montis, A., Barthélemy, M., Chessa, A., and Vespignani, A. (2007). The structure of inter-urban traffic: A weighted network analysis. *Environment and Planning B: Planning and Design*, 34(5):905–924.
- Dees, S. and Saint-Guilhem, A. (2011). The role of the united states in the global economy and its evolution over time. *Empirical Economics*, 41:573–591.
- Ductor, L., Fafchamps, M., Goyal, S., and van der Leij, M. (2013). Social Networks and Research Output. *Review of Economics and Statistics*, forthcoming.

- Duernecker, G., Meyer, M., and Vega-Redondo, F. (2012). Being Close to Grow Faster: A Network-Based Empirical Analysis of Economic Globalization. Economics Working Papers ECO2012/05, European University Institute.
- Eubank, H., Guclu, S., Kumar, V. S. A., Marathe, M., Srinivasan, A., Toroczkai, Z., and Wang, N. (2004). Controlling Epidemics in Realistic Urban Social Networks. *Nature*, 429.
- Fagiolo, G. (2006). Directed or undirected? a new index to check for directionality of relations in socio-economic networks. *Economics Bulletin*, 3:1–12.
- Fagiolo, G. and Mastrorillo, M. (2013a). International Migration Network: Topology and modeling. *Physical Review E*, 88:012812.
- Fagiolo, G. and Mastrorillo, M. (2013b). Migration and trade: A complex-network approach. physics.soc-ph 1309.5859v1, arXiv.org.
- Fagiolo, G., Schiavo, S., and Reyes, J. (2009). World-Trade Web: Topological properties, dynamics, and evolution. *Physical Review E*, 79(036115).
- Frankel, J. and Rose, A. (2002). An estimate of the effect of common currencies on trade and income. *The Quarterly Journal of Economics*, 117(2):437–466.
- Frankel, J. A. and Romer, D. H. (1999). Does trade cause growth? *American Economic Review*, 89(3):379–399.
- Gallup, J. L., Mellinger, A. D., and Sachs, J. D. (1999). Geography datasets. Center for International Development at Harvard University.
- Gambardella, A., Mariani, M., and Torrisci, S. (2003). How provincial is your region? effects on labour productivity and employment in europe. LEM Papers Series 2003/06, Laboratory of Economics and Management (LEM), Sant’Anna School of Advanced Studies, Pisa, Italy.
- Gaulier, G. and Zignago, S. (2010). Baci: International trade database at the product-level. the 1994-2007 version. Working Papers 2010-23, CEPII research center.
- Granovetter, M. (1973). The Strength of Weak Ties. *The American Journal of Sociology*, 78(6):1360–1380.
- Haldane, A. and May, R. (2011). Systemic risk in banking ecosystems. *Nature*, 469:351–355.
- Hale, G. (2012). Bank relationships, business cycles, and financial crises. *Journal of International Economics*, 88(2):312–325.

- Halu, A., Mondragòn, R., Panzarasa, P., and Bianconi, G. (2013). Multiplex PageRank. *PLoS ONE*, 8(10):e78293.
- Hausmann, R. and Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, 16:309–342.
- Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., and Ratti, C. (2013). Geo-located twitter as the proxy for global mobility patterns. cs.SI 1311.0680, arXiv.org.
- Hofleitner, A., Chiraphadhanakul, T. V., and State, B. (2013). Coordinated migration. Facebook Data Science Team.
- Horner, M. W. and O’Kelly, M. E. (2001). Embedding economies of scale concepts for hub network design. *Journal of Transport Geography*, 9(4):255–265.
- Hua, Y. and Zhu, D. (2009). Empirical analysis of the worldwide maritime transportation network. *Physica A*, 388:2061–71.
- Hufnagel, L., Brockmann, D., and Geisel, T. (2004). Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences of the United States of America*, 101(42):15124–15129.
- Im, K. S., Pesaran, M. H., and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115(1):53–74.
- Jackson, M. O. and Zenou, Y. (2014). *Games on Networks*. forthcoming.
- Kali, R., Méndez, F., and Reyes, J. (2007). Trade structure and economic growth. *Journal of International Trade & Economic Development*, 16:245–269.
- Kali, R. and Reyes, J. (2010). Financial contagion on the international trade network. *Economic Inquiry*, 48:1072–1101.
- Kaluza, P., Kölzsch, A., Gastner, M., and Blasius, B. (2011). The complex network of global cargo ship movements. *Journal of the Royal Society*, 388:1093–1103.
- Kaluza, P., Kolzsch, A., Gastner, M. T., and Blasius, B. (2010). The complex network of global cargo ship movements. *Journal of The Royal Society Interface*.
- Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika*, 18:39–43.
- Keller, W. (2004). International technology diffusion. *Journal of Economic Literature*, 42(3):752–782.

- Keum, K. (2010). Tourism flows and trade theory: a panel data analysis with the gravity model. *The Annals of Regional Science*, 44(3):541–557.
- Kitamura, R., Chen, C., Pendyala, R. M., and Narayanan, R. (2000). Micro-simulation of daily activity-travel patterns for travel demand forecasting. *Transportation*, V27(1):25–51.
- Kleinberg, J. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 465:604–632.
- Kleinberg, J. (2007). Computing: The wireless epidemic. *Nature*, 449(7160):287–288.
- Lee, K., Yang, J., Kim, G., Lee, J., Goh, K., and Kim, I. (2011). Impact of the topology of global macroeconomic network on the spreading of economic crises. *PLoS One*, 6(3):e18443.
- Lu, X., Wetter, E., Bharti, N., Tatem, A. J., and Bengtsson, L. (2013). Approaching the Limit of Predictability in Human Mobility. *Scientific Reports*, 3.
- Lucas, Robert E, J. (1993). Making a miracle. *Econometrica*, 61(2):251–72.
- Marshall, M. and Jagers, K. (1999). "Polity iv project: Political regime characteristics and transitions, 1800-2007". Center for international development and conflict management, university of maryland.
- May, R. M., Levin, S. A., and Sugihara, G. (2008). Complex systems: ecology for bankers. *Nature*, 451(7181):893–895.
- Mayer, T. and Zignago, S. (2011). Notes on CEPII's distances measures: The GeoDist database. Working papers, CEPII research center.
- McNerney, J., Fath, B. D., and Silverberg, G. (2012). Network structure of inter-industry flows. *Physica A: Statistical Mechanics and its Applications*, 392(24):6427–6441.
- Miguéns, J. I. L. and Mendes, J. F. F. (2008). Travel and tourism: Into a complex network. *Physica A: Statistical Mechanics and its Applications*, 387(12):2963–2971.
- Minoiu, C., Kang, C., Subrahmanian, V., and Berea, A. (2013). Does financial connectedness predict crises? IMF Working Paper 13/267, International Monetary Fund.
- Moreno, R. and Trehan, B. (1997). Location and the Growth of Nations. *Journal of Economic Growth*, 2(4):399–418.
- Nelson, R. R. and Phelps, E. S. (1966). Investment in humans, technological diffusion and economic growth. *American Economic Review*, 56(2).

- Nickell, S. J. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49(6):1417–26.
- NP (2013). Special issue on Complex Networks in Finance. *Nature Physics*. 9(3): 119–197.
- Opsahl, T., Colizza, V., Panzarasa, P., and Ramasco, J. J. (2008). Prominence and control: the weighted rich-club effect. *Physical Review Letters*, 101(16):168702.
- Ortega, F. and Peri, G. (2013). Migration, trade and income. IZA Discussion Papers 7325, Institute for the Study of Labor (IZA).
- Reyes, J., Schiavo, S., and Fagiolo, G. (2010). Using complex networks analysis to assess the evolution of international economic integration: The cases of east asia and latin america. *The Journal of International Trade and Economic Development*, 19(2):215–239.
- Roodman, D. (2009a). How to do xtabond2: An introduction to difference and system gmm in stata. *Stata Journal*, 9(1):86–136.
- Roodman, D. (2009b). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1):135–158.
- Santos Silva, J. and Tenreyro, S. (2006). The Log of Gravity. *Review of Economics and Statistics*, 88(4):641–658.
- Scellato, S., Noulas, A., and Mascolo, C. (2011). Exploiting place features in link prediction on location-based social networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '11*, pages 1046–1054, New York, NY, USA. ACM.
- Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A., and White, D. R. (2009). Economic networks: The new challenges. *Science*, 325(5939):422–425.
- Simini, F., González, M. C., Maritan, A., and Barabasi, A.-L. (2012). A universal model for mobility and migration patterns. *Nature*, 484:96–100.
- Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1):65–94.
- Stock, J. and Yogo, M. (2005). *Testing for Weak Instruments in Linear IV Regression*, pages 80–108. Cambridge University Press, New York.
- Tranos, E., Gheasi, M., and Nijkamp, P. (2012). International migration: A global complex network. Tinbergen Institute Discussion Papers 12-123/VIII, Tinbergen Institute.

- Viboud, C., Bjornstad, O. N., Smith, D. L., Simonsen, L., Miller, M. A., and Grenfell, B. T. (2006). Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science*, 312:447–451.
- Wang, X.-W., Han, X.-P., and Wang, B.-H. (2014). Correlations and scaling laws in human mobility. *PLoS ONE*, 9(1):e84954.
- Ward, M. D., Ahlquist, J. S., and Rozenas, A. (2013). Gravity’s rainbow: A dynamic latent space model for the world trade network. *Network Science*, 1(1):95–118.
- Woolley-Meza, O., Thiemann, C., Grady, D., Lee, J., Seebens, H., Blasius, B., and Brockmann, D. (2011). Complexity in human transportation networks: a comparative analysis of worldwide air transportation and global cargo-ship movements. *The European Physical Journal B*, 84:589–600.

**Table 1:** The International Human Mobility Network: Descriptive statistics

years	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Travellers (Mln)	408.3	447.2	487.8	503.2	526.2	568.5	577.2	592.7	583.4	632.2	673.4	720	770.1	789.7	752.3	796.6
Max link weight (Mln)	19.2	20.3	18.2	17.7	17.5	19.3	18.6	18.5	17.6	19.4	20.3	19.7	19.4	20.3	19.5	20.1
Average link weight	52154	52383	52831	53964	54034	54534	55299	56682	55442	58582	59865	63654	65264	67348	63501	66743
Median link weight	1060	909	775	792	738	646	679	710	743	787	803	923	944	1063	1034	1060
# Links	7828	8537	9234	9325	9739	10424	10437	10457	10522	10792	11249	11311	11800	11725	11847	11935
Density	0.173	0.189	0.204	0.207	0.216	0.231	0.231	0.232	0.233	0.239	0.249	0.25	0.261	0.26	0.262	0.264
# Bilateral Links	2918	3250	3640	3750	3972	4438	4484	4496	4546	4686	4966	5032	5336	5318	5414	5436
Bilateral density	0.373	0.381	0.394	0.402	0.408	0.426	0.43	0.43	0.432	0.434	0.441	0.445	0.452	0.454	0.457	0.455
Weighted asymmetry <sup>a</sup>	0.408	0.448	0.462	0.46	0.458	0.458	0.451	0.453	0.469	0.415	0.402	0.383	0.377	0.384	0.381	0.368
# SCC <sup>b</sup>	33	33	33	33	34	34	31	33	33	32	27	25	23	27	27	25
# WCC <sup>c</sup>	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Size of largest SCC	181	181	181	181	180	180	183	181	181	182	187	189	191	187	187	189
Size of largest WCC	213	213	213	213	213	213	213	213	213	213	213	213	213	213	213	213
Diameter (directed)	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Diameter (undirected)	3	3	3	3	2	2	2	3	2	2	2	2	2	2	2	2
APL <sup>d</sup> (directed)	1.922	1.833	1.805	1.803	1.789	1.765	1.767	1.762	1.751	1.749	1.748	1.744	1.735	1.729	1.726	1.732
APL (undirected)	1.721	1.694	1.672	1.67	1.657	1.637	1.637	1.636	1.635	1.626	1.612	1.61	1.596	1.598	1.595	1.592

<sup>a</sup>See Fagiolo (2006)

<sup>b</sup>Strongly connected components

<sup>c</sup>Weakly connected components

<sup>d</sup>Average path-length

B \ W	OPMOB	OPTRA	STR	BON	KATZ
OPMOB	–	0.6339	0.1862	0.3218	0.3019
OPTRA	0.6339	–	0.1189	0.2229	0.2065
DEG	0.3235	0.1743	0.4706	0.8969	0.9078
BON	0.3483	0.1924	0.9824	0.9054	0.9017
KATZ	0.3361	0.1835	0.9962	0.9949	0.9093

**Table 2:** Correlation among centrality measures and openness variables. Year: 2010. OPMOB: Openness to mobility. OPTRA: Openness to trade. Note: the lower-diagonal quadrant refers to binary centrality indicators. The upper-diagonal quadrant to weighted centrality indicators. DEG = node degree; STR = node strength; KATZ: Katz centrality (Katz, 1953); BON = Bonacich (Eigenvector) Centrality (Bonacich, 1987, Bonacich and Lloyd, 2001).

Dep. Var.	OLS $\log(\omega_{ij})$	PPML $\omega_{ij} \geq 0$	PPML $\omega_{ij} > 0$
$\log(POP)_j$	-0.119*** (0.016)	-0.195* (0.114)	-0.256* (0.144)
$\log(AREA)_j$	-0.191*** (0.013)	-0.238*** (0.078)	-0.202** (0.098)
$\log(DIST)_{ij}$	-1.680*** (0.023)	-1.540*** (0.097)	-1.427*** (0.115)
$BORDER_{ij}$	0.295 (1.033)	-6.094*** (2.018)	-6.205*** (2.012)
$LANGUAGE_{ij}$	1.339*** (0.050)	0.819*** (0.149)	0.776*** (0.165)
$COLONY_{ij}$	2.108*** (0.114)	0.770*** (0.144)	0.777*** (0.148)
$LANDLOCKED_j$	-1.231*** (0.056)	-0.975*** (0.136)	-0.972*** (0.292)
No. of observations	16,169	36,312	16,169
R <sup>2</sup>	0.615	0.554	0.511

**Table 3:** Gravity-model fit of bilateral travel flows (as a share of population workforce). All regressions include country of origin fixed effects, an intercept and interactions among regressors and contiguity dummy, not reported here. The full sample covers 196 countries. All variables computed as 1995-2000 averages. Significance levels: \* = 0.10, \*\* = 0.05, \*\*\* = 0.01. Standard errors clustered at the dyadic level in parentheses.



Dep. Var.	OLS $\log(\tau_{ij})$	PPML $\tau_{ij} \geq 0$	PPML $\tau_{ij} > 0$
$\log(POP)_j$	0.036*** (0.010)	-0.224** (0.104)	-0.230** (0.103)
$\log(AREA)_j$	-0.125*** (0.008)	0.066 (0.094)	0.067 (0.093)
$\log(DIST)_{ij}$	-1.255*** (0.016)	-0.785*** (0.147)	-0.769*** (0.148)
$BORDER_{ij}$	2.249*** (0.789)	1.116 (1.088)	1.037 (1.061)
$LANGUAGE_{ij}$	0.805*** (0.033)	0.586*** (0.077)	0.574*** (0.076)
$COLONY_{ij}$	1.092*** (0.070)	0.464** (0.205)	0.462** (0.206)
$LANDLOCKED_j$	-0.496*** (0.029)	-0.851*** (0.282)	-0.843*** (0.280)
No. of observations	25,880	32,448	25,880
R <sup>2</sup>	0.698	0.167	0.272

**Table 4:** Gravity-model fit of country total trade (imports plus exports) as a share of destination country real (PPP) GDP. Year: 2000. All regressions include country of origin fixed effects, an intercept and interactions among regressors and contiguity dummy, not reported here. The full sample covers 196 countries. All variables computed as 1995-2000 averages. Significance levels: \* = 0.10, \*\* = 0.05, \*\*\* = 0.01. Standard errors clustered at the dyadic level in parentheses.

Dep. Var.	log(real GDP per-capita)			log(real GDP per-worker)		
	(1)	(2)	(3)	(4)	(5)	(6)
log(OPMOB)	0.297*** (0.082)	0.211** (0.100)	0.162 (0.154)	0.334*** (0.080)	0.229** (0.104)	0.098 (0.160)
log(OPTRA) <sub>1990–2000</sub>		0.162 (0.108)			0.204 (0.132)	
log(OPTRA)			0.051 (0.283)			0.231 (0.314)
INST	0.020* (0.011)	0.025** (0.012)	0.024** (0.011)	0.018* (0.011)	0.023* (0.012)	0.023** (0.011)
TROPICS	-0.473*** (0.158)	-0.502*** (0.170)	-0.400** (0.170)	-0.667*** (0.166)	-0.713*** (0.174)	-0.676*** (0.187)
LATITUDE	-0.274 (0.469)	-0.051 (0.535)	-0.100 (0.547)	-0.650 (0.470)	-0.436 (0.528)	-0.464 (0.539)
LANDLOCKED	-0.257** (0.113)	-0.314*** (0.120)	-0.410*** (0.118)	-0.265** (0.123)	-0.336*** (0.127)	-0.452*** (0.130)
log(RENT)	0.186*** (0.041)	0.174*** (0.039)	0.190*** (0.040)	0.192*** (0.046)	0.172*** (0.043)	0.192*** (0.042)
Constant	9.792*** (0.230)	9.073*** (0.563)	9.818*** (0.367)	10.781*** (0.235)	9.913*** (0.646)	11.049*** (0.443)
No. of Observations	135	135	131	134	134	130
Kleibergen-Paap F stat	29.34	10.87	4.83	29.74	10.60	4.62
Stock-Yogo critical values <sup>†</sup>	16.38/8.96	16.38/8.96	7.03/4.58	16.38/8.96	16.38/8.96	7.03/4.58

**Table 5:** The effect of Mobility Openness on Income and Labor Productivity. Cross-Section 2SLS estimation, second stage. Year 2000. All regressions contain regional dummies.  $\log(\text{OPTRA})_{1990-2000}$  refers to the log of trade openness averaged over the period 1990-2000 (data are from Penn World Tables 7.1), INST is the average Polity-2 score in the same period. OPTRA and OPMOB are instrumented using gravity predictions. Robust Standard Errors in parenthesis. \*, \*\*, \*\*\* significance at the 10, 5 and 1 percent level. *Note.* (<sup>†</sup>): Critical values for 10% and 15% max IV size.

Dep. Var.	log(real GDP per-capita)					
	(1)	(2)	(3)	(4)	(5)	(6)
log(OPMOB)	0.302*** (0.082)	0.176 (0.158)	0.311*** (0.083)	0.196 (0.150)	0.309*** (0.086)	0.200 (0.147)
log( <i>ND</i> )			0.374*** (0.142)	0.395** (0.195)		
log( <i>BON</i> )	0.530*** (0.194)	0.603** (0.291)				
log( <i>KATZ</i> )					0.939** (0.368)	0.926* (0.474)
log(OPTRA) <sub>1990–2000</sub>	0.168 (0.109)		0.160 (0.109)		0.162 (0.110)	
log(OPTRA)		0.257 (0.354)		0.215 (0.331)		0.191 (0.319)
INST	0.011 (0.010)	0.009 (0.012)	0.011 (0.010)	0.009 (0.011)	0.012 (0.010)	0.011 (0.011)
TROPICS	-0.438*** (0.152)	-0.428** (0.173)	-0.445*** (0.152)	-0.428** (0.171)	-0.454*** (0.151)	-0.429** (0.167)
LATITUDE	-0.547 (0.478)	-0.534 (0.474)	-0.537 (0.489)	-0.498 (0.487)	-0.511 (0.492)	-0.454 (0.493)
LANDLOCKED	-0.147 (0.113)	-0.221* (0.118)	-0.151 (0.114)	-0.237** (0.119)	-0.156 (0.114)	-0.252** (0.119)
log(RENT)	0.166*** (0.042)	0.181*** (0.041)	0.166*** (0.042)	0.180*** (0.041)	0.167*** (0.042)	0.181*** (0.040)
Constant	10.566*** (0.608)	11.758*** (1.130)	7.523*** (0.964)	8.339*** (0.686)	8.552*** (0.678)	9.444*** (0.339)
No. of Observations	135	131	135	131	135	131
Kleibergen-Paap F stat	7.56	3.18	7.58	3.42	7.45	3.52
Stock-Yogo critical values <sup>†</sup>	7.03/4.58		7.03/4.58		7.03/4.58	

**Table 6:** The Effect of Country *Binary* Centrality on Income. Cross-Section 2SLS estimation, second stage. Year 2000. log(OPTRA)<sub>1990–2000</sub> refers to the log of trade openness averaged over the period 1990-2000 (data are from Penn World Tables 7.1), INST is the average Polity-2 score in the same period. OPTRA and OPMOB are instrumented using gravity predictions. Robust Standard Errors in parenthesis. \*, \*\*, \*\*\* significance at the 10, 5 and 1 percent level. *Note.* (†): Critical values for 10% and 15% max IV size.

Dep. Var.	log(real GDP per-worker)					
	(1)	(2)	(3)	(4)	(5)	(6)
log(OPMOB)	0.315*** (0.089)	0.119 (0.163)	0.321*** (0.090)	0.138 (0.156)	0.314*** (0.091)	0.136 (0.152)
log( <i>ND</i> )			0.338** (0.153)	0.409** (0.196)		
log( <i>BON</i> )	0.489** (0.209)	0.653** (0.293)				
log( <i>KATZ</i> )					0.800** (0.387)	0.879* (0.473)
log(OPTRA) <sub>1990–2000</sub>	0.195 (0.139)		0.190 (0.139)		0.195 (0.140)	
log(OPTRA)		0.430 (0.372)		0.383 (0.352)		0.353 (0.344)
INST	0.010 (0.011)	0.007 (0.012)	0.011 (0.011)	0.009 (0.012)	0.012 (0.011)	0.011 (0.012)
TROPICS	-0.629*** (0.164)	-0.676*** (0.191)	-0.642*** (0.164)	-0.684*** (0.188)	-0.657*** (0.163)	-0.691*** (0.184)
LATITUDE	-0.761 (0.474)	-0.810* (0.470)	-0.767 (0.479)	-0.787* (0.477)	-0.749 (0.479)	-0.746 (0.481)
LANDLOCKED	-0.179 (0.129)	-0.248* (0.134)	-0.187 (0.129)	-0.273** (0.135)	-0.200 (0.127)	-0.303** (0.132)
log(RENT)	0.176*** (0.046)	0.192*** (0.044)	0.174*** (0.047)	0.189*** (0.044)	0.172*** (0.047)	0.188*** (0.043)
Constant	11.272*** (0.702)	13.049*** (1.140)	8.498*** (1.088)	9.444*** (0.739)	9.458*** (0.788)	10.650*** (0.405)
No. of Observations	134	130	134	130	134	130
Kleibergen-Paap F stat	7.36	3.12	7.28	3.29	7.12	3.52
Stock-Yogo critical values <sup>†</sup>	7.03/4.58		7.03/4.58		7.03/4.58	

**Table 7:** The Effect of Country *Binary* Centrality on Productivity. Cross-Section 2SLS estimation, second stage. Year 2000. log(OPTRA)<sub>1990–2000</sub> refers to the log of trade openness averaged over the period 1990-2000 (data are from Penn World Tables 7.1), INST is the average Polity-2 score in the same period. OPTRA and OPMOB are instrumented using gravity predictions. Robust Standard Errors in parenthesis. \*, \*\*, \*\*\* significance at the 10, 5 and 1 percent level. *Note.* (†): Critical values for 10% and 15% max IV size.

Dep. Var.	log(real GDP per-capita)					
	(1)	(2)	(3)	(4)	(5)	(6)
log(OPMOB)	0.275*** (0.085)	0.180 (0.149)	0.261*** (0.089)	0.086 (0.187)	0.296*** (0.085)	0.204 (0.141)
log(NS)			0.093*** (0.036)	0.111** (0.048)		
log(BON)	0.209** (0.090)	0.239** (0.118)				
log(KATZ)					0.573** (0.235)	0.618** (0.305)
log(OPTRA) <sub>1990–2000</sub>	0.171 (0.105)		0.215** (0.109)		0.174 (0.110)	
log(OPTRA)		0.185 (0.309)		0.395 (0.383)		0.182 (0.304)
INST	0.016 (0.010)	0.014 (0.010)	0.018* (0.010)	0.014 (0.011)	0.015 (0.010)	0.012 (0.010)
TROPICS	-0.449*** (0.148)	-0.402** (0.158)	-0.434*** (0.142)	-0.477*** (0.170)	-0.452*** (0.145)	-0.416*** (0.155)
LATITUDE	-0.307 (0.463)	-0.309 (0.464)	-0.457 (0.463)	-0.405 (0.478)	-0.365 (0.459)	-0.355 (0.453)
LANDLOCKED	-0.205* (0.110)	-0.281** (0.110)	-0.230** (0.108)	-0.322*** (0.113)	-0.187* (0.108)	-0.268** (0.107)
log(RENT)	0.181*** (0.038)	0.198*** (0.037)	0.186*** (0.037)	0.205*** (0.039)	0.184*** (0.038)	0.200*** (0.038)
Constant	9.633*** (0.527)	10.638*** (0.606)	7.561*** (0.872)	8.644*** (0.575)	8.702*** (0.630)	9.606*** (0.328)
No. of Observations	135	131	135	131	135	131
Kleibergen-Paap F stat	7.48	3.39	6.21	2.97	7.89	3.72
Stock-Yogo critical values <sup>†</sup>	7.03/4.58		7.03/4.58		7.03/4.58	

**Table 8:** The Effect of Country *Weighted* Centrality on Income. Cross-Section 2SLS estimation, second stage. Year 2000. Weighted network centrality measures are instrumented using centrality computed on the predicted IHMN. log(OPTRA)<sub>1990–2000</sub> refers to the log of trade openness averaged over the period 1990-2000 (data are from Penn World Tables 7.1), INST is the average Polity-2 score in the same period. OPTRA and OPMOB are instrumented using gravity predictions. Robust Standard Errors in parenthesis. \*, \*\*, \*\*\* significance at the 10, 5 and 1 percent level. *Note.* (†): Critical values for 10% and 15% max IV size.

Dep. Var.	$\log(y_j^t) = \log(\text{real GDP per capita})$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log(y_j^{t-1})$	0.974*** (0.008)	0.979*** (0.008)	0.980*** (0.008)	0.967*** (0.008)	0.973*** (0.007)	0.973*** (0.007)	0.970*** (0.007)	0.976*** (0.006)	0.975*** (0.007)
$\log(ND_j^t)$		0.043*** (0.011)			0.048*** (0.010)			0.051*** (0.010)	
$\log(BON_j^t)$	0.065*** (0.015)			0.069*** (0.014)			0.071*** (0.013)		
$\log(KATZ_j^t)$			0.132*** (0.034)		0.144*** (0.033)				0.150*** (0.031)
$\log(OPMOB_j^t)$	0.011** (0.005)	0.013** (0.005)	0.012** (0.005)	0.013*** (0.005)	0.014*** (0.005)	0.014*** (0.005)	0.009** (0.004)	0.009** (0.004)	0.010** (0.005)
$\log(OPTRA_j^t)$	-0.003 (0.013)	-0.014 (0.013)	-0.015 (0.013)	0.002 (0.011)	-0.006 (0.010)	-0.008 (0.011)	0.002 (0.010)	-0.007 (0.009)	-0.008 (0.010)
No. of Observations	2,226	2,226	2,226	2,226	2,226	2,226	2,226	2,226	2,226
No. of ID	160	160	160	160	160	160	160	160	160
AR(1) p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AR(2) p-value	0.187	0.180	0.174	0.199	0.192	0.185	0.195	0.188	0.179
Hansen p-value	0.046	0.038	0.038	0.170	0.136	0.172	0.988	0.992	0.990
No. of Instruments	127	127	127	152	152	152	218	218	218

**Table 9:** The Effect of Country *Binary* Centrality on Income. Dynamic Panel: 1995-2010. System GMM Estimates. All regressions contain country and time dummies. Endogenous variables: lagged dependent variable, centrality indicator, openness to mobility and trade. In all regressions we use two-year lags as instrument for endogenous regressors. Columns 4-6: we employ also  $t - 4$  lag for the lagged dependent variable. Columns 7-9: we use all available instruments for the lagged dependent variable. Robust SE in parenthesis. Significance levels:  $*p < 0.1$ ,  $**p < 0.05$ ,  $***p < 0.01$ .

Dep. Var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$\log(y_j^t) = \log(\text{real GDP per-worker})$								
$\log(y_j^{t-1})$	0.984*** (0.009)	0.989*** (0.008)	0.989*** (0.009)	0.976*** (0.007)	0.981*** (0.007)	0.982*** (0.007)	0.977*** (0.006)	0.983*** (0.007)	0.982*** (0.007)
$\log(ND_j^t)$		0.025** (0.011)		0.036*** (0.010)			0.037*** (0.010)		
$\log(BON_j^t)$	0.042*** (0.014)			0.056*** (0.013)			0.056*** (0.013)		
$\log(KATZ_j^t)$			0.081** (0.036)			0.108*** (0.032)			0.103*** (0.028)
$\log(OPMOB_j^t)$	0.001 (0.005)	0.003 (0.005)	0.003 (0.006)	0.003 (0.005)	0.004 (0.005)	0.005 (0.005)	-0.000 (0.004)	0.001 (0.005)	0.002 (0.005)
$\log(OPTRA_j^t)$	0.009 (0.012)	0.001 (0.013)	0.001 (0.013)	0.014 (0.009)	0.007 (0.010)	0.007 (0.011)	0.015* (0.009)	0.009 (0.009)	0.009 (0.009)
No. of Observations	2,130	2,130	2,130	2,130	2,130	2,130	2,130	2,130	2,130
No. of ID	153	153	153	153	153	153	153	153	153
AR(1) p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AR(2) p-value	0.987	0.986	0.997	0.972	0.999	0.979	0.962	0.990	0.975
Hansen p-value	0.050	0.049	0.050	0.218	0.194	0.193	0.999	0.998	0.998
No. of Instruments	127	127	127	152	152	152	218	218	218

**Table 10:** The Effect of Country *Binary* Centrality on Productivity. Dynamic Panel: 1995-2010. System GMM Estimates. All regressions contain country and time dummies. Endogenous variables: lagged dependent variable, centrality indicator, openness to mobility and trade. In all regressions we use two-year lags as instrument for endogenous regressors. Columns 4-6: we employ also  $t - 4$  lag for the lagged dependent variable. Columns 7-9: we use all available instruments for the lagged dependent variable. Robust SE in parenthesis. Significance levels:  $*p < 0.1$ ,  $**p < 0.05$ ,  $***p < 0.01$ .

Dep. Var.	$\log(y_j^t) = \log(\text{real GDP per-capita})$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log(y_j^{t-1})$	0.973*** (0.009)	0.979*** (0.010)	0.984*** (0.008)	0.966*** (0.008)	0.967*** (0.009)	0.977*** (0.008)	0.967*** (0.007)	0.969*** (0.008)	0.977*** (0.007)
$\log(NS_j^t)$		0.022*** (0.005)		0.023*** (0.005)				0.023*** (0.005)	
$\log(BON_j^t)$	0.052*** (0.011)			0.052*** (0.010)			0.053*** (0.010)		
$\log(KATZ_j^t)$			0.078*** (0.030)			0.091*** (0.027)			0.101*** (0.024)
$\log(OPMOB_j^t)$	0.006 (0.006)	0.002 (0.008)	0.012*** (0.005)	0.009* (0.005)	0.006 (0.006)	0.013*** (0.005)	0.006 (0.005)	0.003 (0.006)	0.010*** (0.005)
$\log(OPTRA_j^t)$	0.000 (0.013)	-0.004 (0.015)	-0.017 (0.012)	0.007 (0.011)	0.005 (0.014)	-0.010 (0.011)	0.009 (0.010)	0.008 (0.011)	-0.009 (0.009)
No. of Observations	2,226	2,226	2,226	2,226	2,226	2,226	2,226	2,226	2,226
No. of ID	160	160	160	160	160	160	160	160	160
AR(1) p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AR(2) p-value	0.239	0.163	0.170	0.251	0.172	0.184	0.249	0.172	0.188
Hansen p-value	0.022	0.018	0.028	0.179	0.124	0.142	0.993	0.994	0.994
No. of Instruments	127	127	127	152	152	152	218	218	218

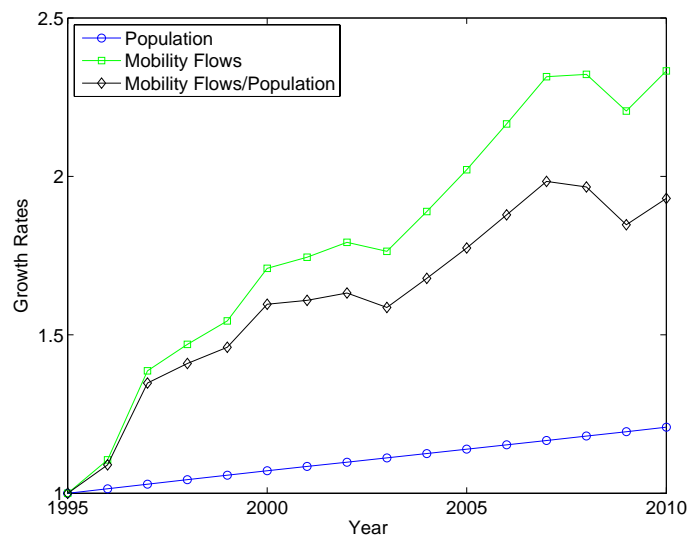
**Table 11:** The Effect of Country *Weighted* Centrality on Income. Dynamic Panel: 1995-2010. System GMM Estimates. All regressions contain country and time dummies. Endogenous variables: lagged dependent variable, centrality indicator, openness to mobility and trade. In all regressions we use two-year lags as instrument for endogenous regressors. Columns 4-6: we employ also  $t - 4$  lag for the lagged dependent variable. Columns 7-9: we use all available instruments for the lagged dependent variable. Robust SE in parenthesis. Significance levels:  $*p < 0.1$ ,  $**p < 0.05$ ,  $***p < 0.01$ .



# A Topological Properties of the International Human-Mobility Network

## A.1 Basic Facts about International Human-Mobility Data

The total flow of travellers has undergone a dramatic increase in the period 1995-2010. As Figure A1 shows for our sample of 213 countries, the total volume of the network has risen by a factor of 2.4. This pattern holds also for the share of total travellers over world population. In 2010 nearly 780 millions of people (11.2% of world population) has temporarily moved worldwide, while in 1995 mobility flows involved only 5.8% of the overall population.



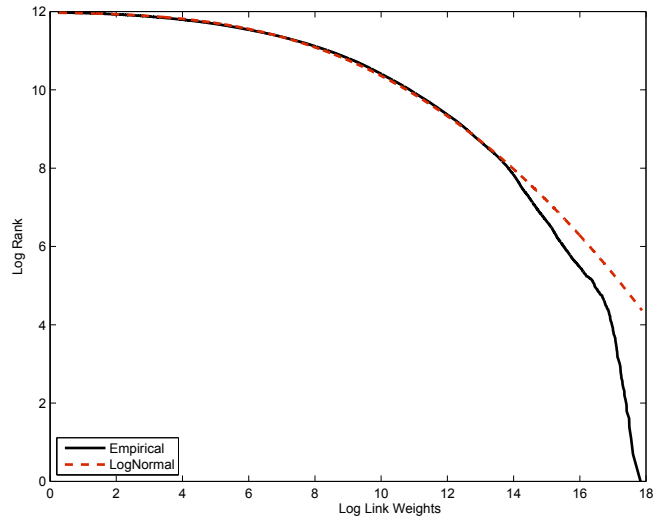
**Figure A1:** Growth rates of total mobility flows, world population and share of travellers over world population. Base year: 1995 = 1.

We now turn to investigate the topological properties of the weighted-directed network of international human-mobility (IHMN), see Section 3.

Figure A2 plots the pooled link-weight distribution (using a size-rank characterization) and the correspondent log-normal fit. Link weights appear to follow a log-normal distribution (as it happens for the international trade), with only a very small fraction of the data in the right tail that could be approximated by a Pareto density.

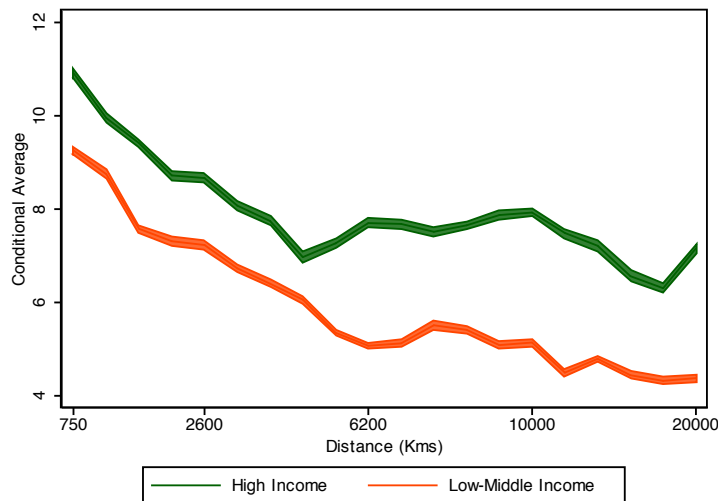
We now explore the behavior of IHMN link weights conditional on geographical distance travelled. Note that, in the pooled sample, people have travelled an average bilateral distance of around 6600. As it happens for migration stocks (Fagiolo and Mastrorillo, 2013a), average mobility flows decrease with geographical distance (see Figure A3). Interestingly, the decline is sharper for mobility flows originated from low and middle income countries<sup>19</sup>, suggesting that

<sup>19</sup>Income group is defined according to the World Bank classification (<http://data.worldbank.org/about/>



**Figure A2:** Size-rank plot of the pooled mobility-flow distribution.

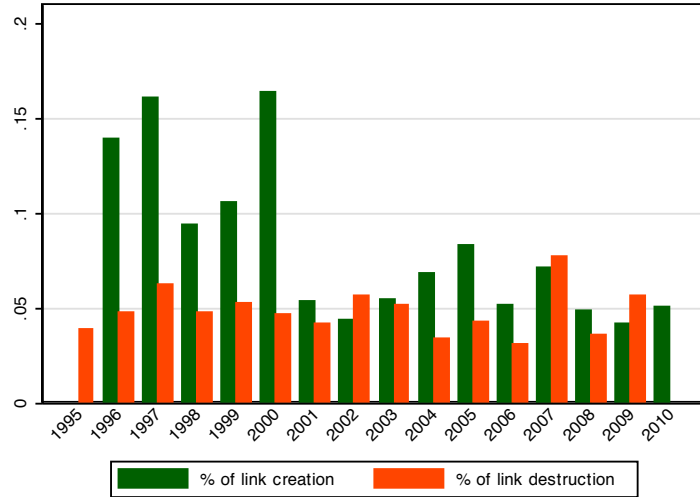
income also plays a major role in shaping the international human-mobility flow distribution. It is worth noting that in 2010, 70% of the total flows originated from high-income countries and almost half of world international human mobility involves travels between high-income countries (49% of total flows in 2010). A large share of mobility, about one fifth (21%), occurs between low and middle income countries, whereas the remaining 9% is from low/middle-income countries to high-income ones.



**Figure A3:** Average of mobility flows conditional to geographical distance between countries. X-axis employs 20 geographical-distance bins, i.e. 5-percentiles of the geographical-distance distribution.

The IHMN is quite geographically concentrated. For example, over 8.3% of total world [country-classifications/world-bank-atlas-method](#)).

mobility flows is accounted by mobility pattern among NAFTA countries (Canada, Mexico and the U.S.). The U.S. are also the largest recipient of incoming flows (7.5%), followed by Spain, France and Italy (together accounting for 17.1%) and China (4.6%). Germany, instead, accounted for 10% of world outgoing flows in 2010, followed by the U.S. (9.4%) and the U.K. (6.7%). It is also worth mentioning the sharp increase in the number of travellers from the BRICs: indeed, about 7.8% of total flows in 2010 originated from those countries, a figure that almost doubled since 1995 (3.8%).



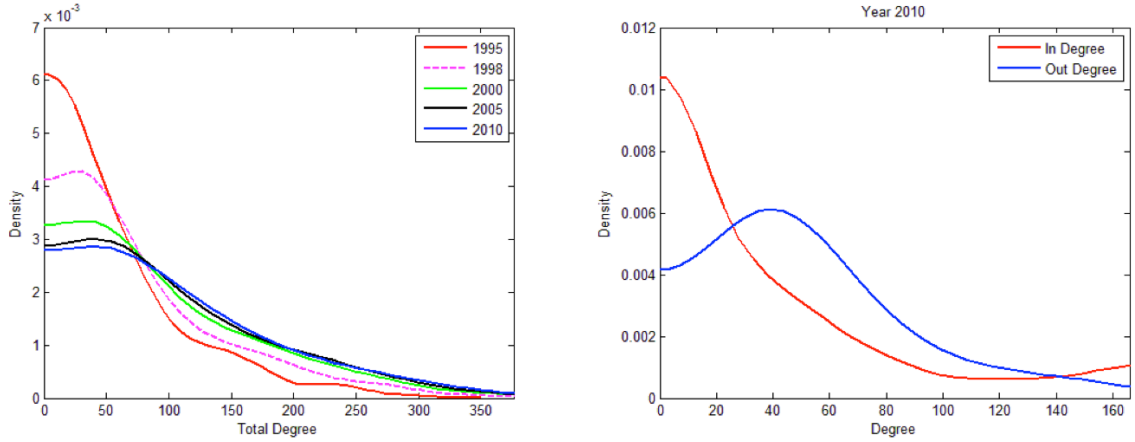
**Figure A4:** Percentage of links in the IHMN whose weight changes from zero to strictly positive (link creation) and from positive to zero (link destruction).

The IHMN is characterized by a relative large turbulence as far as the emergence and disappearing of mobility corridors. Regarding the emergence of links, the fraction of flows which switch from zero to a positive value is relatively high, on average 6.2% over the whole period (with a sharp decrease after year 2000), see Figure A4. This relatively high number of newly-formed links couples with the relatively low number of disappearing links. Over the whole period, in fact, the fraction of links that switched from a positive flow to zero has remained quite low and stable (around 4%). This explains the large increase in network density.

## A.2 Node Connectivity

We now explore node-degree and strength distributions. As the left panel of Figure A5 shows, node total-degree distributions have shifted to the right and their tails have become fatter. This is explained by the increase in the overall connectivity of the network. The average number of link (both incoming and outgoing ones), has grown from 73 to about 112, while the median number of mobility corridors increased by 40% (from 57 to 93), thus entailing a rise

in the overall density of the network. Statistical tests show that the pooled, total node-degree distribution follows a log-normal density, except for the highest 22% of the observation for which the decay is Pareto.



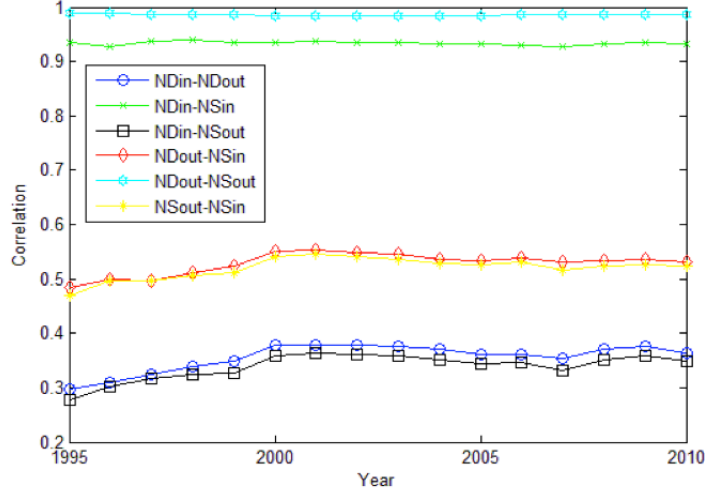
**Figure A5:** Total degree (left) and year-2010 in/out degree (right) distributions.

The right panel of Figure A5 depicts the distribution of in and out degree for year 2010. Interestingly, the two distributions exhibit a different behavior. In particular, out degree is characterized by a well-defined modal value (around 38 links), whereas the in-degree distribution is much more skewed to the left, with a fat right tail. This is mainly due to the fact that in the IHMN there exists a much more uniform number of sending countries than of destination countries. This is reflected in a coefficient of variation of the number of incoming mobility path 40% higher than for outgoing corridors (which remain very stable over time).

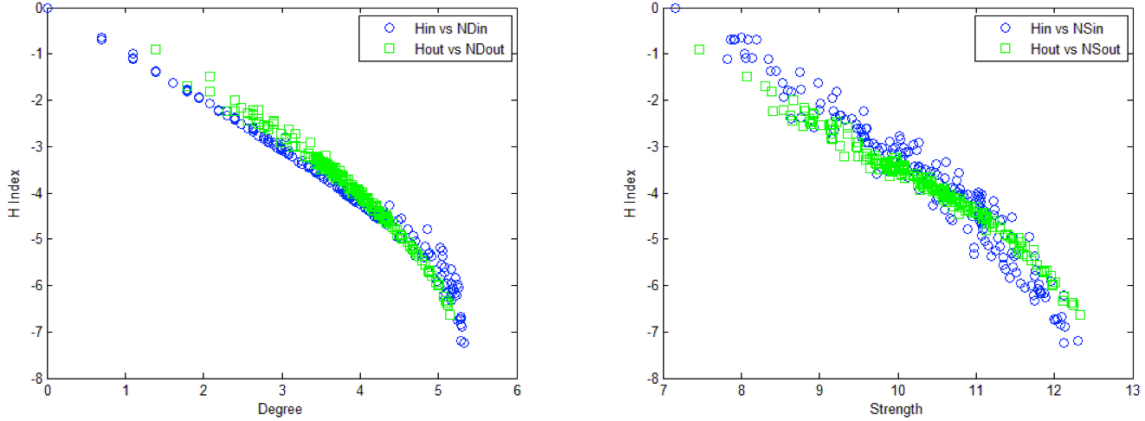
This is not true for in- and out-strength distributions, which are both well-approximated by log-normal densities. This skewed behavior, as it happens for migration networks (Fagiolo and Mastroiello, 2013a), is robust to rescaling for country population, indicating that the presence of hubs in the weighted network does not depend on country size.

Figure A6 plots Pearson correlation coefficients between node in/out degree and strength over time. Note that countries with more inward or outward traveling corridors also hold the largest in and out mobility flows. On the contrary, countries that are chosen as preferred destinations (both from a binary and weighted perspective), are not necessarily holding many outward corridors or send many travellers abroad.

In order to evaluate the level of diversification of in and out mobility flows at a country level, we compute the Herfindahl concentration index of in and out link weight portfolios. Results are reported in Figure A7, where the H-index is plotted against node degree and strength for year 2010 in a log-log scale. A power-law relation emerges, implying that countries with a larger share of mobility routes and total in and out flows are associated with less concentrated portfolios of origin/ destinations. This parallels what has been already found for migration



**Figure A6:** Pearson correlation coefficients between node in/out degree and strength over time.



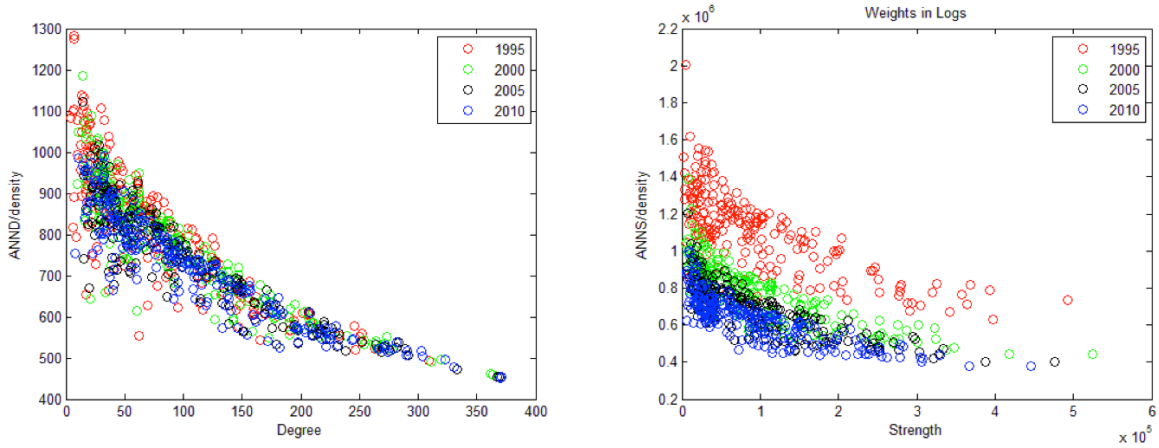
**Figure A7:** Herfindahl (H) concentration index vs node degree (left) and strength (right). Year 2010. Log-log scale.

networks (Fagiolo and Mastrorillo, 2013a).

### A.3 Assortativity, Clustering, and Rich-Club Behavior

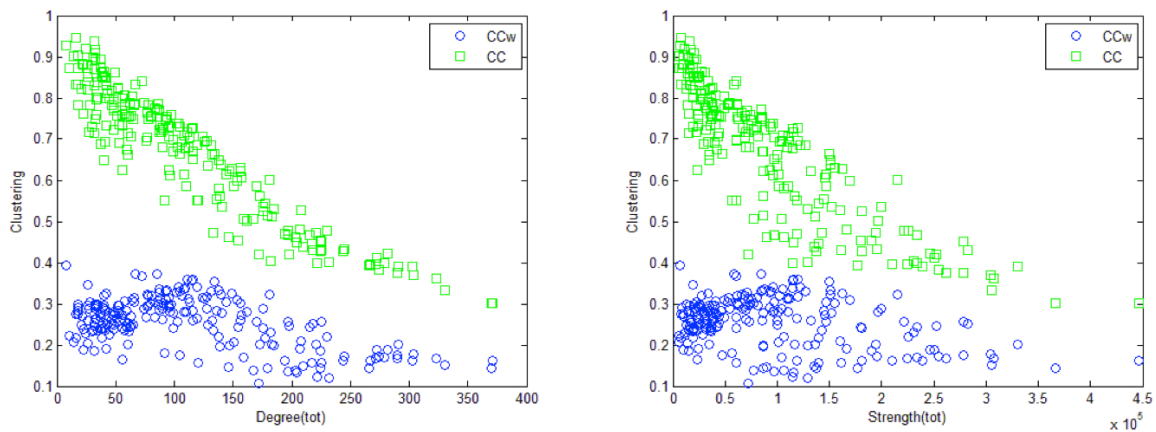
We evaluate the assortativity (disassortativity) of the IHMN by computing Pearson correlation coefficients between total node degree (strength) and average nearest-neighbor degree (strength) for each given country across time. Results are reported in Figure A8. As it happens for international migration, the IHMN appears to be weakly disassortative: more connected players are usually connected to countries with a relatively small number of mobility routes and mobility-flow intensity.

Moving to network clustering, we observe that the HIMN is characterized a relatively high



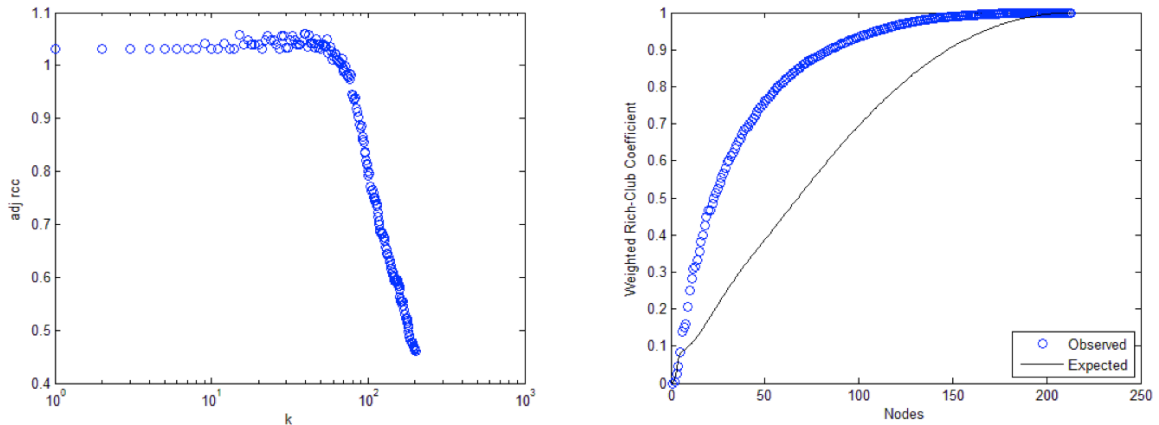
**Figure A8:** Disassortative behavior in the IHMN. Total node degree (left) and strength (right) vs. average nearest-neighbor degree (ANND) and strength (ANNS). In the weighted case, logged link weights have been employed.

binary clustering coefficient (CC), with average CC that grows over time: from 0.60 in 1995 to 0.67 in 2010. In weighted terms, the average weighted clustering coefficient (CCw) grows from 0.23 to 0.25 in the same period. Coupled with the evidence about a decreasing APL over time (see Table 1), this suggests that the HIMN shows some small-world behavior. To further analyze this issue, we have computed average CC and APL values, along with their expected counterparts in Erdos-Renyi (ER) random graphs that keep fixed the observed density of the network. The observed APL is indeed higher than the expected one, while the CC is substantially higher (about three times) than its ER counterpart (about three times). This implies that over time, new mobility channels progressively close open triplets, shrinking the overall geodesic distance and increasing the clustering coefficient of the network.



**Figure A9:** Binary clustering coefficient (CC) and weighted clustering coefficient (CCw) vs node degree and strength in year 2010.

Figure A9 reports the correlation between binary (CC) and weighted (CCw) clustering coefficient vs total degree and total strength in year 2010. The binary coefficient appears to be negatively correlated with both degree and strength, suggesting a hierarchical structure in the network where most connected nodes tend to form few triangles with their neighbors (as it happens in the international migration network, see [Fagiolo and Mastrorillo, 2013a](#)). Conversely, nodes with a small number of mobility routes, or with marginal ones in term of intensity, are usually connected with countries that are more likely connected with each other. The absence of a clear positive correlation among CCw and NS implies that the weight of such triangles is relatively weak.



**Figure A10:** Binary (adjusted) rich-club coefficient (left, adj rcc) and weighted rich-club coefficient (right) in year 2010.

Finally, we explore whether the IHMN exhibits any rich-club behavior. Following [Colizza et al. \(2006b\)](#), [Opsahl et al. \(2008\)](#), we compute both binary (adjusted) and weighted coefficients, see Figure A10. It is easy to see that whereas the binary version of the network does not exhibit any rich-club ordering (the coefficient stays very close to one), the weighted counterpart displays a lot of concentration of mobility flows within the rich club made of the first 25 nodes (around 50%). To attain 90% of the volume of the network is sufficient to consider the “richest” 90 countries in the world, where richness is in terms of total strength (where link weights are computed in levels).

## B Data Sources and Variable Definitions

Name	Definition	Source
Real GDP per capita	Real GDP, PPP (constant 2005 US\$) divided by mid-year population	WDI
Real GDP per worker	Real GDP, PPP (constant 2005 US\$) divided by workforce (from employment over population ratio)	WDI
POP	Resident population	WDI
OPMOB	Openness to mobility: arrivals plus departures over country workforce	WDI
OPTRA	Openness to trade: imports plus exports over country real GDP (PPP)	Penn World Table 7.1 and CEPII-BACI
$w_{ij}^t$	Bilateral mobility flows: People moving from $i$ to $j$ in year $t$ (and staying there for less than one year)	UNWTO
$\omega_{ij}^t$	Bilateral mobility flows / country $j$ workforce	UNWTO and WDI
ND	Node degree: Number of (inward- and outward-) country links in the IHMN	UNWTO*
NS	Node strength: sum of (inward- and outward-) country link weights in the IHMN	UNWTO*
BON	Bonacich Centrality ( <a href="#">Bonacich, 1987</a> )	UNWTO*
KATZ	Katz Centrality ( <a href="#">Katz, 1953</a> )	UNWTO*
TROPICS	Proportion of land in tropics and subtropics	<a href="#">Gallup et al. (1999)</a>
LATITUDE	Distance from equator (absolute latitude / 90)	CEPII*
LANDLOCKED	Dummy if the country is landlocked	CEPII
RENT	Rent from natural resources percentage of GDP	WDI
INST	Quality of institutions (Polity-2)	<a href="#">Marshall and Jaggers (1999)</a>
AREA	Country area in Km <sup>2</sup>	CEPII
BORDER	Dummy equal to one if two countries share common borders	CEPII
DIST	Distance from most populated cities	CEPII
COLONY	Dummy equal to one for country pairs ever in colonial relationship	CEPII
LANGUAGE	Dummy equal to one if country pairs have common official primary language	CEPII
Regional Dummies	Sub-Saharan Africa, E Asia & Pacific, Latin America & Caribbean, S Asia, Europe & C Asia, M East & N Africa	WDI

**Table B1:** Variables Definition and Data Sources.

WDI: World-Development Indicators (<http://data.worldbank.org/data-catalog/world-development-indicators>); CEPII-BACI ([http://cepii.fr/CEPII/fr/bdd\\_modele/presentation.asp?id=1](http://cepii.fr/CEPII/fr/bdd_modele/presentation.asp?id=1)); UNWTO: World Tourism Organization (<http://www2.unwto.org>); Penn World Table 7.1 ([https://pwt.sas.upenn.edu/php\\_site/pwt71/pwt71\\_form.php](https://pwt.sas.upenn.edu/php_site/pwt71/pwt71_form.php)). (\*): Our own calculations.



## C The System-GMM Estimator: Robustness Checks

In this Appendix, we report on some robustness checks we have performed to ensure that the choice for a System-GMM estimator was the right one, given the structure of our data.

First of all, we note that the number of elements in the matrix of instruments grows quadratically in the time dimension. Therefore, we want to check if the sample we are using contains adequate information for estimation. In order to exclude a “small-sample bias” issue, we exploit the fact that estimating a dynamic model with OLS will produce an upward bias for the autoregressive coefficient, while if using an FE estimator will generate a downward bias.

We use the coefficients estimated via OLS vs. a fixed-effect (FE) model as upper and lower bounds for our preferred estimator, i.e. the System-GMM. As reported in Table C1 (income) and C2 (productivity), the autoregressive coefficient obtained through a System-GMM always lies in the interval defined by OLS and FE estimates. This confirms that our sample contains enough information to ensure consistency of our estimators, thus excluding the presence of a finite-sample bias (Bond et al., 2001), in both income and labor productivity estimation.

Second, we ask whether that a System-GMM is the most appropriate estimator given the high persistence of the dependent variable, and our relatively-small sample ( $N = 160$ ). We do so using the fact that if the process were  $AR(1)$ , then OLS estimation for  $\rho$  would have been upward biased, whereas a Within estimator for  $\rho$  would have delivered downward-biased estimates (Nickell, 1981), with a distortion inversely proportional to the econometric sample size. Such extreme cases can provide the boundaries for a consistent estimation of  $\rho$ . Tables C1 and C2 indicate that our estimates for  $\rho$  lie between OLS and Within bounds, suggesting that the System-GMM is an appropriate estimator in our case.

Third, we compare the performance of a System-GMM estimator with that of a Difference-GMM estimator. Figures in Column (3) of both Table C1 and Table C2 reveal that the autoregressive coefficient estimated using a Difference-GMM lies below the Within estimate. This is in line with Bond et al. (2001), who suggest that it is “likely that the GMM (diff) estimate is also biased downwards [...], perhaps due to weak instruments”. All this reassures us that a System-GMM estimator should be preferred in our case, for its superior finite-sample properties.

Finally, our estimates show that the first-order auto-regressive coefficient is very close to unity. Therefore, we check for the presence of a unit-root using the the Im-Pesaran-Shin (IPS) test (Im et al., 2003). The test returns a p-value of 0.028 for the labor productivity equation, which gives us a reasonable confidence in rejecting the null hypothesis that our panel exhibits a unit root. In the case of income, instead, the IPS test strongly advises against rejection. Even if the moment condition of the System-GMM should still remain valid in the case of a unit root (providing that there are not  $N$  specific drifts, see Bond et al., 2002), we re-estimate

our income regression (with binary centrality indicators) using a different measure of real per-capita GDP, namely the Penn World Tables (7.1) PPP-Converted GDP Per Capita at 2005 constant prices. For this series, the p-value of the IPS test is 0.0381, i.e. the null can safely be rejected. Results are reported in Table C3: even if the overall diagnostics of the GMM are worst in this case (e.g., AR(2) is only barely rejected), it is worth noting that the coefficients associated with binary centrality measures are very similar to those in Table 9, both in terms of magnitude and significance.

Dep. Var.	(1)	(2)	(3)	(4)
$\log(y_j^t)$	OLS	System-GMM	Difference-GMM	Fixed Effects
$\log(y_j^{t-1})$	0.992*** (0.001)	0.968*** (0.008)	0.679*** (0.052)	0.929*** (0.016)
$\log(BON_j^t)$	0.014*** (0.002)	0.069*** (0.014)	-0.039* (0.023)	0.013 (0.008)
$\log(OPMOB_j^t)$	0.003*** (0.001)	0.012*** (0.004)	0.064*** (0.020)	0.012*** (0.004)
$\log(OPTRA_j^t)$	-0.001 (0.002)	0.003 (0.011)	0.050** (0.024)	0.011 (0.007)
No. of Observations	2,226	2,226	2,057	2,226
R <sup>2</sup>	0.999			0.955
No. of ID		160	160	160
AR(1) p-value		3.47e-08	0.000104	
AR(2) p-value		0.199	0.356	
Hansen p-value		0.165	0.0540	
No. of Instruments		152	95	

**Table C1:** Coefficient boundaries for dynamic-panel estimation: Country Income with Binary Bonacich Centrality. All regressions contain country and time dummies. Robust SE in parenthesis. Significance Levels: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . Column (2): System-GMM, using lag 2 to lag 4 for lagged dependent variable, and lag 2 for network centrality, OPMOB and OPTRA as instruments.

Dep. Var.	(1)	(2)	(3)	(4)
$\log(y_j^t)$	OLS	System-GMM	Difference-GMM	Fixed Effects
$\log(y_j^{t-1})$	0.991*** (0.002)	0.980*** (0.007)	0.710*** (0.054)	0.926*** (0.017)
$\log(BON_j^t)$	0.015*** (0.003)	0.053*** (0.014)	-0.023 (0.022)	0.013 (0.009)
$\log(OPMOB_j^t)$	0.003** (0.001)	0.003 (0.005)	0.061*** (0.019)	0.011** (0.005)
$\log(OPTRA_j^t)$	0.001 (0.002)	0.009 (0.010)	0.025 (0.024)	0.009 (0.008)
No. of Observations	2,130	2,130	1,968	2,130
R <sup>2</sup>	0.999			0.936
No. of ID		153	153	153
AR(1) p-value		0.000	0.000	
AR(2) p-value		0.976	0.802	
Hansen p-value		0.202	0.0249	
No. of Instruments		152	95	

**Table C2:** Coefficient boundaries for dynamic-panel estimation: Country Productivity with Binary Bonacich Centrality. All regressions contain country and time dummies. Robust SE in parenthesis. Significance Levels:  $*p < 0.1$ ,  $**p < 0.05$ ,  $***p < 0.01$ . Column (2): System-GMM, using lag 2 to lag 4 for lagged dependent variable, and lag 2 for network centrality, OPMOB and OPTRA as instruments.

Dep. Var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$\log(y_j^t)$								
$\log(y_j^{t-1})$	0.988*** (0.009)	0.993*** (0.010)	0.995*** (0.009)	0.989*** (0.009)	0.994*** (0.009)	0.996*** (0.009)	0.990*** (0.009)	0.994*** (0.009)	0.994*** (0.008)
$\log(ND_j^t)$		0.029** (0.012)			0.040*** (0.011)			0.034*** (0.010)	
$\log(BON_j^t)$	0.052*** (0.017)			0.064*** (0.014)			0.055*** (0.014)		
$\log(KATZ_j^t)$			0.089** (0.036)			0.120*** (0.031)			0.098*** (0.028)
$\log(OPMOB_j^t)$	0.006 (0.006)	0.007 (0.006)	0.007 (0.006)	0.003 (0.005)	0.004 (0.005)	0.004 (0.006)	0.002 (0.005)	0.003 (0.005)	0.003 (0.005)
$\log(OPTRA_j^t)$	-0.005 (0.015)	-0.014 (0.014)	-0.016 (0.013)	-0.003 (0.015)	-0.011 (0.014)	-0.015 (0.014)	0.002 (0.014)	-0.003 (0.014)	-0.004 (0.013)
No. of Observations	2,233	2,233	2,233	2,233	2,233	2,233	2,233	2,233	2,233
No. of ID	161	161	161	161	161	161	161	161	161
AR(1) p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AR(2) p-value	0.015	0.015	0.015	0.016	0.016	0.016	0.016	0.016	0.016
Hansen p-value	0.063	0.066	0.077	0.174	0.222	0.228	0.996	0.990	0.995
No. of Instruments	127	127	127	152	152	152	218	218	218

**Table C3:** Using an alternative country-income definition to test the effect of country *binary* centrality on income. Country income: Per capita GDP (ppp) 2005 constant US\$ from PWT7.1. Dynamic Panel: 1995-2010. System GMM Estimates. All regressions contain country and time dummies. Endogenous variables: lagged dependent variable, centrality indicator, openness to mobility and trade. In all regressions we use two-year lags as instrument for endogenous regressors. Columns 4-6: we employ also  $t - 4$  lag for the lagged dependent variable. Columns 7-9: we use all available instruments for the lagged dependent variable. Robust SE in parenthesis. Significance levels: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .