

Flows of People, Flows of Ideas, and the Inequality of Nations*

Thomas Barnebeck Andersen Carl-Johan Dalgaard
University of Southern Denmark University of Copenhagen

Forthcoming: *Journal of Economic Growth*

February 1, 2011

Abstract

The present paper examines a neglected determinant of aggregate productivity: temporary cross-border flows of people. We hypothesize that interaction between people from different nations facilitates the international diffusion of ideas, thus stimulating aggregate productivity. In order to assess the causal impact of people flows on productivity, we construct an instrument for people flows. By analogy to the trade/growth literature, this instrument is derived from a fitted gravity equation involving geographic determinants of bilateral travel flows. Our cross-section analysis reveal that greater international interaction leads to higher productivity; a very similar result, qualitatively as well as quantitatively, is obtained when we employ dynamic panel data methods for the purpose of identification.

Keywords: Technology diffusion, International travel, Productivity, IV estimation.

JEL Classification codes: O33, O47, C21

*We would like to thank three anonymous referees, Phil Abbott, Antonio Ciccone, Alistar Dieppe, Kevin Fox, Justin Lin, David Weil, and seminar participants at the University of Copenhagen, the DEGIT conference in Jerusalem, Göteborg University, the University of Hannover, the IMF Institute, the IMF Research Department, the UNIDO conference “productivity and growth in Africa and Asia” and the conference “Understanding and Predicting Productivity Growth” organized by the central bank of Sweden for useful comments. The usual disclaimer applies. We also thank Francesco Alcalá and Antonio Ciccone for sharing their data. Contact: barnebeck@sam.sdu.dk or carl.johan.dalgaard@econ.ku.dk.

1 Introduction

It is well documented that the bulk of observed differences in GDP per worker can be accounted for by the variation in aggregate total factor productivity (TFP).¹ A leading theoretical explanation for aggregate TFP differences is that some countries are closer to the technological frontier than others, as a result of higher rates of technology adoption.² This theoretical work has been supplemented by empirical studies of the international diffusion process. In particular, it has been argued that cross-border flows of goods (international trade) and capital (FDI) are important vehicles for the diffusion of technology across the globe.³ The underlying logic of why international trade and FDI are expected to enable diffusion of technology is that (intermediate) goods and machinery embody leading-edge technological knowledge.

At the same time, however, it is well known that flows of goods and capital also proxy other aspects of global interaction. Indeed, in the influential study by Frankel and Romer (1999), which examines the relationship between trade and productivity, the authors observe that perhaps (p. 393) “[t]he literal shipment of goods between countries does not raise income. Rather, trade is a proxy for the many ways in which interactions between countries raise income – specialization, spread of ideas, and so on. Trade is likely to be highly, but not perfectly, correlated with the extent of such interactions. Thus, trade is an imperfect measure of income-enhancing interactions among countries.” In particular, Frankel and Romer emphasize (p. 381) “[the] exchange of ideas through communication and travel” as an important income-enhancing mechanism.

The contribution of the present paper lies in gauging the empirical importance of this latter technology diffusion mechanism in the aggregate. Specifically, we examine the hypothesis that societies more exposed to foreign influence, through temporary in- and outflows of travellers, will be able to obtain useful technologies, ideas, and organizational strategies from abroad. Once implemented, the acquired knowledge should stimulate aggregate productivity and GDP per worker.

To test this hypothesis, we measure the intensity of travel as the ratio of international arrivals plus departures to the size of the labor force, and examine its impact on TFP. In a simple OLS setting, the travel intensity variable can account for more than 45% of the variation in aggregate TFP in a diverse cross section of countries. Moreover, the correlation is robust to the inclusion of additional determinants of long-run productivity such as (measures of) institutional quality and geographic circumstances.⁴

¹See e.g. Klenow and Rodriguez-Clare (1997), Hall and Jones (1999), and Caselli (2006).

²The pioneering theoretical contribution is Nelson and Phelps (1966). Recent notable contributions include Lucas (1993), Howitt (2000), and Eaton and Kortum (2002).

³The literature is surveyed in Keller (2004).

⁴That the institutional make-up of a country empirically is a key determinant of long-run productivity was origi-

In order to examine whether the cross-sectional correlation between movements of people and productivity is causal, we propose the following identification strategy. Inspired by Frankel and Romer (1999) and Alcalá and Ciccone (2004), we estimate a gravity equation for bilateral *travel* flows. That is, instead of the sum of bilateral imports and exports to nominal GDP, our left-hand side variable is the sum of bilateral in- and outflows of people to the size of the labor force. We subsequently use fitted aggregate travel shares, derived from the gravity equation, as an instrument for aggregate travel. To our knowledge, this is the first time a gravity equation has been estimated in an effort to study travel flows.

Our 2SLS estimates suggest that increasing the intensity of travel by 1% increases the level of TFP and GDP per worker by roughly 0.2%. This finding is also robust to the inclusion of other fundamental determinants of productivity such as institutional quality and key climate related circumstances.

A concern is that the fitted aggregate travel share (i.e., our instrument) may be correlated with time-invariant determinants of productivity, in spite of our best efforts to control for such determinants. In light of this difficulty, we therefore further scrutinize the people/productivity nexus by employing dynamic GMM methods to attain identification in a panel data setting. The main virtue of the panel approach is that it allows us to introduce fixed effects, thus pruning the data from the influence of time invariant determinants of productivity. In addition, it allows us to expand the country coverage significantly, as we are no longer constrained by the availability of time invariant instruments for determinants of labor productivity. Reassuringly, using the Arellano-Bond procedure in this larger sample of countries we obtain point estimates that are almost identical to those obtained in the cross-section analysis. Hence, the main result is robust to this alternative (yet, admittedly, somewhat mechanical) identification strategy.

An issue of some importance is whether the impact we attribute to movements of people simply masks an impact from movements of goods. In the cross-sectional setting we can address this issue by introducing trade as a separate productivity determinant, instrumenting it using fitted trade shares along the lines of Frankel and Romer (1999) and Alcalá and Ciccone (2004). With two endogenous variables and two separate instruments we are in principle able to separately identify the impact from these two forces of globalization. In practise, however, movements of people and goods are highly correlated and, it turns out, difficult to disentangle. Still, we generally find in the cross-sectional setting that movements of people has a causal impact on productivity, conditional on movements of goods. This result is further corroborated in the panel data setting: using the

nally demonstrated in important contributions by Hall and Jones (1999) and Acemoglu et al. (2001). Sachs (2003) argues convincingly that some aspects of climate also matters directly for aggregate productivity.

Arellano-Bond procedure travel emerges as a robust determinant of labor productivity, conditional on trade.

Before we turn to the analysis itself, it is worth emphasizing that the notion that interaction of individuals matters for international diffusion of knowledge is not novel. Contributions to the theory of endogenous growth has already incorporated this mechanism. Lucas (1993) develops a model where human capital spillovers ensure international convergence in income per worker. Lucas emphasizes learning-by-doing as a form of human capital accumulation, which could be relevant in trying to understand the East Asian growth miracle. Irwin and Klenow (1994) provide empirical evidence on learning in the semi-conductor industry. In particular, they find evidence of international knowledge spillovers stemming from learning-by-doing, thus supporting the spillover mechanism in Lucas' model. It is hard to believe that diffusion of tacit knowledge like that acquired through learning-by-doing can occur without any personal interaction, which then almost inevitably involves cross-border flows of people.

Personal interaction may be crucial for diffusion of technological knowledge more generally. The reason is that even technological knowledge is not always fully codified since this would be prohibitively costly. Aspects of technological knowledge therefore remains tacit in nature and may then only be transferred fully via personal interaction, preferably face-to-face (Keller, 2004; Storper and Venables, 2004; Jones, 2007). This could especially be true for knowledge spillovers between researchers, commonly assumed to exist in idea-based endogenous growth models (Romer, 1990; Aghion and Howitt, 1992, and many others).

That personal contacts could be key in the diffusion process was pointed out early on by Arrow (1969); an assertion which more recently has received strong empirical support (Foster and Rosenzweig, 1995; Almeida and Kogut 1999; Agrawal et al., 2006).⁵ Furthermore, Arrow (1969) argues that a relative lack of personal contacts outside a country may be an important impediment to the adoption of valuable technological knowledge from abroad. Accordingly, international travel ought to facilitate the diffusion of technological knowledge by allowing people to build personal relationships across borders.

Specific cases of knowledge spillovers through personal (international) interaction can also

⁵Foster and Rosenzweig present survey results deriving from a questionnaire circulated in 186 Indian villages around the time of the "Green Revolution". Over a third of the respondents claimed to have received most of their information regarding cropping practices from friends and neighbours, rather than through campaigns orchestrated by government agencies. Similar results emerge from another survey conducted among farmers in the Philippines. Analyzing patent citations, Almeida and Kogut show that ideas in the U.S. semi-conductor industry are spread, in part, by the mobility of patent-holders. Ethnographic research of Silicon Valley and Route 128 corroborates this finding (Saxenian 1994). Extending this line of research, Agrawal et al. find that close geographic proximity is not necessary in order to transfer tacit knowledge; social relationships may act as conduits for knowledge flows. Using U.S. patent data, they find that knowledge flows go disproportionately to an inventor's prior location.

be provided. One example concerns the creation of the so-called “Desh Garment Company” in Bangladesh, which was founded on interaction with Korean Daewoo, and turned out to be a resounding success. As described by Rhee (1990, p. 336): “*The collaboration agreement, which was to run five years, involved several key elements: six months of training for Desh workers in Korea: start-up activities to involve certain purchases of machinery by Desh from Daewoo, which would then handle the installation, supervise and advise on the actual start-up; production, to be managed by Desh with consultation and supervision provided by Daewoo.*” Hence the international in- and outflow of people was an important ingredient in establishing this new endeavour, and thereby in its ultimate success.⁶

Some direct evidence on the association between cross-border movements of people and productivity does exist. Gambardella et al. (2003) find a significant relationship between the annual number of airplane passengers (embarked and disembarked) in regions of Europe and regional average labor productivity. Along a similar vein Dowrick and Tani (2008) document a positive impact of business travel on industry level productivity in Australia. To our knowledge, however, the present paper is the first to examine the same issue in the aggregate while addressing the inherent endogeneity of travel.

Finally, it is worth observing that although our analysis pertains to the 20th century the underlying mechanism – cross-border interaction of people enables technology diffusion – may well have been a persistent force of development for a long time. Moreover, contemporary societies that are more prone to interaction have likely also been so historically.

Recent theoretical and empirical work suggests that the pre-industrial interaction of societies may importantly contribute to an understanding of contemporary productivity differences. From a theoretical angle, Ashraf and Galor (2007) hypothesize that societies characterized by higher cultural diversity are more able to adjust to new technological paradigms; cultural homogeneity by contrast favours the intergenerational transmission of knowledge, which is an advantage in technologically stagnant societies. The authors construct a unified growth model which predicts that the culturally diverse society should undergo the industrial revolution earlier than the more culturally homogenous society, thus instigating income divergence. As cultural diversity (partly) originates in the interaction of people from different societies, it seems plausible to assert a connection between cultural diversity and (historical) cross-border movements of people.⁷ Empirically, Ashraf and Galor (2008) find that genetic diversity, which can be viewed as a proxy for cultural diversity,

⁶Another interesting case involves how Ryanair came to implement ticketless boarding after one of the company’s CEOs had observed the usefulness of the technology first hand on a transatlantic connection to Ireland, operated by a competitor (Crespi et al., 2008). For other cases and further references, see Rauch (2001).

⁷See also the discussion of historical evidence in Ashraf and Galor (2007, Section 3).

is non-monotonically related to population density in 1500 consistent with the underlying thesis that cultural diversity have spurred technological adoption historically, up to a point. In related research, Spolaore and Wacziarg (2009) hypothesize that cultural differences between individual countries and the frontier countries may act as a barrier to technology adoption, and proceed to show that genetic differences has a statistically significant effect on income differences. This can be taken to suggest that societies that historically have not interacted (for which reason they exhibit marked genetic differences) have not managed to benefit from their separate technological advances, for which reason they have failed to converge in income levels.

The paper proceeds as follows: Section 2 lays out a simple framework which conceptually forms the basis for our empirical analysis of cross-country technology diffusion. Aside from motivating our empirical specification this discussion also allows us to contrast our work with previous studies in a transparent way. Section 3 contains our cross-section analysis, whereas Section 4 turns to panel data. Section 5 concludes.

2 A Conceptual Framework

Consider the following Solow (1956) model, slightly modified to allow for human capital and technology transfer. The production function is Cobb-Douglas, exhibiting constant returns to capital and labor input. Labor is augmented by technology, A_t (which grows at the rate g_t), but also by human capital, h . Along a steady state trajectory GDP per worker, y_t^* , is then given by

$$y_t^* = \left(\frac{s}{n + \delta + g_t} \right)^{\frac{\alpha}{1-\alpha}} A_t h, \quad (1)$$

where s is the savings/investment rate, n is the rate of labor force growth, δ is the rate of capital depreciation, and α is capital's share. This equation, in its log form, was estimated by Mankiw, Romer and Weil (1992). Their cross-section analysis invoked the identifying assumption that A is a random variable with a constant mean.⁸ However, more recent work has deemed this assumption suspect. As a result, we will instead entertain the idea that countries, to varying extent, tap into a "world technology frontier", A_t^w . In the spirit of Nelson and Phelps (1966) we assume that local technology is characterized by the following first-order linear differential equation

$$\dot{A}_t = \phi \cdot (A_t^w - A_t), \quad (2)$$

⁸They also allowed human capital to be accumulated. To capture this, equation (1) could be restated as $(s/n + \delta + g_t)^{\frac{\alpha}{1-\alpha}} A_t h^*$, where h^* is the steady state stock of human capital per worker.

where $\phi > 0$ captures the rate of technological diffusion. The parameter ϕ can therefore be thought to capture adoption or knowledge spillovers from abroad. To complete the model we assume the world technology frontier expands over time at a constant rate of technological progress, $\dot{A}_t^w = g^w A_t^w$. The evolution of technology in the economy over time is fully described by $\dot{A}_t = g^w A_t^w$ and equation (2). A steady state for this system is a A_t^w/A_t -ratio such that the rate of (local) technological progress is equal to the frontier rate, $\dot{A}_t/A_t = g_t = \dot{A}_t^w/A_t^w = g^w$. This ratio, which is unique and can be interpreted as the steady state distance to the frontier, is determined by ϕ ,

$$A_t = \frac{\phi}{\phi + g^w} A_t^w. \quad (3)$$

If equation (3) is substituted into equation (1) (along with $g_t = g^w$) we are left with a complete solution for the long run level of GDP per worker as a function of the parameters of the model and the exogenous level of A_t^w .

While simple, this structure holds several implications which are useful for the empirical examination of international technology diffusion. As should be clear, the model implies that growth differences between countries should be temporary in nature. In the long run, differences in growth of labor productivity, y , as well as productivity, A , disappear. In addition, persistent differences in productivity are due to differences in ϕ .⁹

With this sort of framework in mind, explaining observed differences in the levels of A (and y), rather than differences in growth rates, is a sensible focus for the empirical analysis. The reason is that the former reflects persistent variation whereas the latter is a transitional phenomenon. Consequently, in the main part of our analysis we therefore examine, by way of cross-section regression analysis, the hypothesis that cross-border flows of people influences the size of ϕ . If so, it should impact on A and y .

In contrast, the existing literature on aggregate knowledge spillovers typically use panel data. In effect, these studies take a log-differenced version of equation (3) to the data, thereby examining the determinants of TFP growth (\dot{A}_t/A_t). The right hand side variable is R&D expenditures in leading economies, weighted by either trade shares (Coe and Helpman, 1995; Coe et al., 1997) or FDI shares (de la Potterie and Lichtenberg, 2001). A few observations regarding the difference in empirical strategy and specification, compared with the present paper, are worthwhile.

First, while the existing literature attempts to resolve the problem of endogeneity of trade/FDI (if at all) by invoking internal instruments, the levels specification allows us to invoke an external instrument for travel. Second, since we mostly perform cross-section analysis, adding accumulated

⁹Howitt (2000) develops a multi-country Schumpeterian growth model, which contains similar reduced form properties. In Howitt's model g^w is determined by global R&D effort and ϕ is endogenously determined as well.

global R&D investment (to proxy A^w) would not make much sense. Using bilateral travel flows, one could generate a variable involving weighted accumulated R&D investment for the purpose of estimation, thereby mimicking the independent variable from existing panel studies. We do not follow this track since we do not wish to limit attention to R&D spillovers *per se* but rather examine knowledge diffusion more generally; the study of R&D spillovers through travel is thus left for future research. Instead, we opt for a parsimonious specification which involves regressing A on travel (and other plausible determinants of the rate of diffusion) so as to capture the influence from ϕ . Third, panel data allows for country specific intercepts in the regression; this is infeasible in a cross-section analysis. Instead, in the main analysis, we add controls for other potentially “deep” determinants of ϕ , aside from global interaction via travel, and carefully try to instrument these when relevant.

While the main part of the analysis employ cross-section data, we also check the robustness of our results in a panel setting. In this context we regress labor productivity on the lagged level of labor productivity and the level of cross-border movements of people, while controlling for fixed effects and common time varying determinants of labor productivity. Notice that such a dynamic specification is more general than the cross-section specification as estimating equations such as (1) and (3) involves a steady state assumption, which can be questioned. Indeed, countries may rather be in the process of converging towards the steady state, which we can capture in the panel data setting. Invoking fixed effects in the panel setup (so as to eliminate the influence of time invariant determinants of ϕ), however, leads to biased estimates due to the presence of the lagged dependent variable on the list of regressors. To deal with this issue we invoke internal instruments to produce a consistent estimate of the impact of people flows on productivity.

Finally, an important implication of the model is that ϕ only affects y via A .¹⁰ If indeed our travel variable only captures the technology diffusion channel, we expect that the impact of travel intensity on y can be accounted for entirely by the association between travel and productivity.

3 Cross-Section Analysis

In this section we present the cross-section analysis. Section 3.1 considers the impact of travel on productivity, whereas Section 3.2 examines the impact of cross-border movements of people and goods simultaneously.

¹⁰In theory this is still true if s and n are endogenously determined in a standard growth framework. See e.g. the discussion in Mankiw, Romer and Weil (1992, p. 411).

3.1 The Impact of Travel on Productivity

3.1.1 Specification

On the basis of the conceptual framework outlined in Section 2, the empirical strategy runs as follows. Contingent on data for TFP (proxy for A), our empirical specification is a regression model of the form:

$$\log(A_i) = \gamma_1 + \gamma_2 TRAVEL_i + \mathbf{X}'_i \cdot \boldsymbol{\gamma}_3 + \varepsilon_i, \quad (A)$$

where i refers to countries, $TRAVEL_i$ is the cross-border flow of people, \mathbf{X}_i is a vector of controls, and ε_i captures omitted factors and noise. In choosing relevant controls, \mathbf{X}_i , we follow the approach taken in the literature on “deep determinants” of productivity, pioneered by Hall and Jones (1999) and Acemoglu et al. (2001). In particular, following Rodrik et al. (2004), the set of variables to be included in \mathbf{X}_i can be partitioned into three main subsets: “institutions”, “integration” (participation in the world economy) and “geography”. The rationale for adopting these controls in the specific context of technology diffusion is as follows.

At the fundamental level there are two sorts of reasons why countries do not adopt new technologies from abroad. First, there may be a lack of “willingness” to adopt. That is, the incentive structure may discourage adoption of foreign technologies. This would be the case absent well-established property rights. Related, powerful groups in society may attempt to block innovations if they stand to lose economic or political influence (Parente and Prescott, 1999; Acemoglu and Robinson, 2000). We try to take these kinds of barriers to adoption into account by controlling for the institutional environment. Second, there could be a lack of “access” to foreign technology. Such lack of access could be geographically founded. For example, Diamond (1999) argues convincingly for the difficulty in transferring (agricultural) technologies across climate zones. To capture considerations along these lines we include a set of geo-controls. But a lack of access could also be caused by a lack of global integration, which we control for via the TRAVEL variable. Hence, although TRAVEL is singled out in equation (A), and in the analysis to follow, we fully recognize that it conceptually captures “integration”.

Our central tests involve equation (A). But we also run regressions involving labor productivity:

$$\log(y_i) = \beta_1 + \beta_2 TRAVEL_i + \mathbf{X}'_i \cdot \boldsymbol{\beta}_3 + \nu_i. \quad (Y)$$

A particular reason why we run the regression displayed in equation (Y) is that by comparing the size of our estimates of γ_2 with those for β_2 we get a sense of whether the full effect of TRAVEL runs through TFP. This would be the case if TRAVEL only drives knowledge spillovers (matters for ϕ but not for investment rates).

3.1.2 Data

In calculating the key dependent variable, the level of TFP, we follow the literature on development accounting by invoking an aggregate Cobb-Douglas production function, which exhibits constant returns to the principal rival factors of production: capital and human input.¹¹ Specifically, following Klenow and Rodriguez-Clare (1997) we use that GDP per worker can be written as

$$y_i = A_i \left(\frac{k_i}{y_i} \right)^{\frac{\alpha}{1-\alpha}} h_i,$$

where A_i is productivity (TFP), $k_i \equiv K_i/L_i$ is capital per worker, and h_i is human capital per worker. Assuming markets are competitive, α can be identified by the share of capital in national accounts. We assume α is 1/3 for all countries.¹² Contingent on data for y , k and h we can then calculate A as the residual. Data for real (PPP) GDP per worker, y , the capital stock per worker, k , and the stock of human capital per worker, h , are those used by Caselli (2006). These data, which are based on Penn World Tables, version 6.1 (Heston et al., 2002), and on Barro and Lee (2001), enable us to construct a productivity measure dating from the mid-90s. Specifically, data on real GDP per worker and capital are from 1996, whereas human capital figures are from 1995. Focusing on mid-90s is important since data on cross-border flows of people is only available from 1990 onwards.

Our primary measure of cross-border flow of people is the sum of yearly international arrivals and departures, taken from World Development Indicators (2004) (WDI); the institution in charge of collecting the data is the World Tourism Organization (UNWTO). This variable measures arrivals and departures of people traveling to, and staying in, places outside their usual place of residence for no more than one consecutive year for business, leisure, and other purposes not related to an activity remunerated from within the place visited.¹³ In particular, we use (the log of) the average sum of yearly arrivals and departures over the 1990-1996 period divided by the size of the workforce. We denote this variable TRAVEL.

It is worth observing that TRAVEL is analogous to the conventional summary measure of trade, where total imports and exports are added up and divided by total economic activity in a country. In other words, much like in the case of measuring “trade” we do not take a stance as to whether in- or outflows are more important. Similarly, much like in the measurement of trade, where all

¹¹See Caselli (2006) for a survey of the literature.

¹²This is the conventional assumption in the literature. The study by Gollin (2002) suggests it is not an unreasonable approximation.

¹³The labelling of the underlying variables in WDI is slightly misleading in that it describes the raw data as “international tourist departures” and “international tourist arrivals”, respectively. As the WDI background notes make clear, however, tourism is only a subset of the annual departures and arrivals.

goods flows are admitted in the measure, we do not exclude particular types of travelling, like that attributed to tourism for instance. Naturally, this choice can be questioned. Hence, in checking the robustness of our findings we invoke a second travel variable which exclusively captures business travel.

For the years 1994-1998, UNWTO (2000) provides a breakdown of total international arrivals into tourist and business travellers, respectively. Unfortunately, we do not have similarly disaggregated data for international departures. Hence, the ratio of business arrivals to the size of the labor force, averaged over the period 1994-98, will have to serve as a proxy for “the business travel intensity”, which involves both business arrivals and departures. We call this second travel variable BTRAVEL. While BTRAVEL is a very useful complement to TRAVEL, it is not perfect. In part because it does not capture business departures, and in part because the number of business arrivals undoubtedly is measured with error.¹⁴

Finally, in constructing the instrument for TRAVEL we employ data on bilateral travel flows. These data have been obtained from UNWTO for the years 2003-2007.¹⁵ When estimating the gravity equation below we focus on 2005 to maximize the sample size; however, our results are not sensitive to this particular choice. We provide a detailed description of the bilateral travel data in Section 3.1.4 below, and the bilateral travel data are compared to the corresponding data for trade in Section 3.2.2

Our institutional measure is the index of government anti-diversion policies, proposed by Hall and Jones (1999). This variable combines indicators of bureaucratic quality, law and order, corruption, risk of expropriation, and the likelihood of government repudiation of contracts. Since we are not interested in examining the impact of any particular dimension of “institutions”, we opt for a kitchen sink approach when measuring this important structural characteristic. We denote this variable INSTITUTIONS.¹⁶

Turning to measures of geography, continent dummies (America, Asia & Oceania, Europe, and Africa) have been widely employed in the empirical growth literature. As argued by Sachs (2003), however, adding continent dummies to the regression is likely to be an incomplete way of controlling for the influence of geography on development. Indeed, most geographical analyses stress several other factors such as geographic isolation and the disease environment. To capture the impact

¹⁴UNWTO reports the total number of arrivals as well as the total number of business arrivals, leisure arrivals, and “other” arrivals. However, in some country cases the three last categories do not sum to the total arrivals number; a number of arrivals are thus not classified.

¹⁵See the Appendix for details on data retrieval.

¹⁶A further advantage of using the GADP index is that Hall and Jones (1999) provide instruments, which we shall invoke below. If we instead follow Rodrik et al. (2004) and employ the rule of law index (1996), we obtain results that are very similar to those reported below.

of geography more fully we therefore invoke three additional geo-controls in our regressions: (i) a dummy variable which takes on a value of 1 if the nation is landlocked (henceforth LANDLOCK); (ii) the absolute latitude of the country (LATITUDE); and (iii) the fraction of land area in the tropics (TROPICS). The source of LANDLOCK and TROPICS is Gallup et al. (1999); LATITUDE is taken from Hall and Jones (1999).

Finally, we need a measure of trade openness. For this purpose we shall rely on the real openness measure of Alcalá and Ciccone (2004). They define REAL OPENNESS as imports plus exports in exchange rate USD relative to GDP in purchasing power parity USD. The data source for this measure is WDI.

3.1.3 Partial Correlations

In order to examine the influence of TRAVEL on productivity a cross-plot is a natural starting point. As is visually obvious from Figure 1, the two variables have a high positive correlation. The figure has two clear outliers: Lesotho (LSO) and Romania (ROM). Lesotho, on the one hand, is a small monarchy completely encapsulated by South Africa. In particular, some 40% of all male Basotho are migrant workers in South Africa.¹⁷ Romania, on the other hand, has a large Romany population with a long tradition of crossing borders. The high positive correlation is confirmed in Table 1 (Panel B), which reports the correlation matrix associated with our key variables. The correlation between TRAVEL and productivity is 0.67.

> Figure 1 about here <

> Table 1 about here <

In Table 2 we report results from OLS regressions involving equation (A) (Columns 1-4, 6) and equation (Y) (Columns 5 and 7).

Column (1) shows that TRAVEL and a constant (the line in Figure 1) can account for as much as 46% of the variation in productivity across countries. As we move from left to right in Table 2 we progressively add additional controls to test the robustness of the partial correlation between TRAVEL and productivity. The important message from the regressions is that TRAVEL is estimated with high precision in all columns. This remains true when we employ BTRAVEL instead of TRAVEL (Columns 6-7). It is also worth noting that TRAVEL (/BTRAVEL), INSTITUTIONS, and geography account well for the variation in labor productivity, with a R^2 of 0.86.

¹⁷Guest workers are by definition not contained in the TRAVEL variable. So what explains the high travel activity is probably that Basotho guest workers living in South Africa are the direct cause of temporary flows of visitors (themselves visiting family at home, family visiting them in South Africa, and otherwise).

> Table 2 about here <

As a final check of the partial correlations reported above we have done a series of LAD (median) regressions to check robustness.¹⁸ The results, which are available upon request, show that the partial correlations between TRAVEL (/BTRAVEL) and productivity are not driven by outliers. Indeed, TRAVEL remains significant in all the specifications from Table 2 (typically at the 1% level) when re-estimated by LAD. Overall, these results document that TRAVEL is robustly correlated with productivity.

3.1.4 Identification of Travel

TRAVEL is likely to be endogenous in equations (A) and (Y). Countries whose productivity is high for reasons unrelated to international travel may simply experience a higher intensity of international travel. As a result, we need to invoke a suitable instrument for TRAVEL.

As argued in the seminal paper by Frankel and Romer (1999), certain geographic characteristics are important determinants of the extent to which a given country is engaged in international trade. Since a country's level of productivity does not affect these geographic characteristics, i.e. they are exogenous, the information contained in geography is a candidate instrument for international trade. This insight can also be used in the present context.

We adopt the Frankel and Romer (1999) approach in order to construct a geography based instrument for travel. The instrument is constructed in two steps. First, fitted bilateral travel shares are constructed by regressing bilateral travel shares (the sum of in- and outbound travel divided by the size of the labor force) on determinants such as distance between countries, country size, common border, etc. Second, the fitted bilateral travel shares are then aggregated in order to construct the geographic component of countries' overall travel shares.

Our preferred measure of people flows, TRAVEL, includes both business travellers and leisure travellers. In an effort to span as much of the business component as possible we adopt the *exact* same specification of the gravity equation as Frankel and Romer (and Alcalá and Ciccone, 2004). The notion is that by using geographic determinants of bilateral trade in the context of bilateral travel we primarily span exogenous variation in bilateral business travel, as related to bilateral trade. Consequently, we estimate a regression with the (log) of the sum of in- and outbound travel between countries i and j divided by the size of the labor force in country i (we denote this variable TRAVEL_{ij}) as dependent variable and "the Frankel-Romer controls" as independent variables. Country i 's predicted travel share is then obtained as the sum of the predicted geographic component of its bilateral travel with other countries in the world.

¹⁸This method of estimation is more resistant to outliers than OLS (Deaton, 1997).

Description of the Bilateral Travel Data Before turning to estimation of the travel gravity equation, a description of the travel data is in order. Data on bilateral travel flows between countries in 2005 were obtained from UNWTO. The travel data is constructed on the basis of data provided by the destination countries, and it therefore corresponds to arrivals in these countries. Very few countries actually publish data on outbound bilateral travel. If arrival data in destination countries is complete, one can construct data on departures by using the information on arrivals. However, some destination countries only provide arrival information on the most significant countries of origin. For “insignificant” countries, arrivals are therefore expected to be quite low. We thus treat a missing value as a zero. This does introduce a measurement error in the dependent variable, but under classical assumptions this is innocuous; it only increases the variance of the error term.

To get a sense of the information contained in the bilateral travel data we begin with some country examples. Consider Tables 3 and 4, which convey information on in- and outbound travel for the US and Nigeria, respectively; one being a large developed country and the other a large less developed country.

> Tables 3 and 4 about here <

For the US the vast majority of in- and outbound travel is from and to the two neighboring countries and fellow NAFTA members, Mexico and Canada. More broadly, travel flows to and from the US seem fairly symmetrical in the sense that six out of the 10 countries found on the “inbound top-10” are also among the top-10 outbound destinations.

The travel patterns for Nigeria are somewhat different in nature. Much like the US we do find that neighboring countries (Niger, Benin, Cameroon and Chad) are strongly represented in terms of inbound travel.¹⁹ But this is not true for outbound travel, where other countries are found at the top. In particular, outbound travel to United Kingdom and Saudi Arabia (about 238,000 trips) nearly matches outbound travel to the next 18 destinations taken together (roughly 244,000 trips). The popularity of these two destinations may have an explanation in past colonial ties (UK), and also in the fact that Nigeria is an OPEC member.

Another theme of Tables 3 and 4 is that the level of travel activity in the US exceeds travel activity in Nigeria by a factor of ten. This also translates into a large difference in per capita travel, since the US population only exceeds the Nigerian population by roughly a factor of two. In addition, it is worth observing that whereas outbound travel exceeds inbound travel in the US case, the opposite is true for Nigeria. These facts may simply reflect a considerably higher

¹⁹The high inbound travel intensity from Liberia is likely explained by the fact that Nigeria took thousands of refugees from Liberia during its civil war (1989-2003).

level of income per capita in the US (facilitating outbound travel), and of its immediate neighbors (facilitating higher inbound travel), which underlines the need for TRAVEL to be instrumented, and which suggests that spatial fixed effects are important to include in the regression analysis.

Another way to get a sense of the bilateral travel patterns is to move away from the country as the relevant unit of analysis and instead consider a regional breakdown. More specifically, in Table 5 we construct a “transition matrix”, with regions defined as in the World Development Indicators 2008. In the matrix, row i refers to outbound travel from region i , whereas column j refers to inbound travel to region j . Diagonal elements refers to intraregional travel. Finally, the bottom rows of Table 5 report regional average GDP per capita, total regional population, and the number of countries per region.

The first thing to observe from Table 5 is the importance of travel within regions. In every region, with Latin America & the Carribean as the exception, most travel is intraregional in nature. More specifically, summing the diagonal elements together we find that intraregional travel in 2005 consisted of about 569 million trips, out of a total of some 745 million. Accordingly, 3/4 of total (global) travel occurs within regions. This suggests that geographic proximity matters to travel flows, much like geographic proximity is known to influence trade flows.

A second observation is that income (per capita) is strongly associated with travel intensity. For example, outbound travel from Europe and the US (the two richest regions in the world) involved some 469 million trips. Hence, despite the fact that these two regions only accounted for about 19% of the global population in 2005, and hosts 21% of the worlds’ countries, they accounted for about 2/3 of global travel. Another way to appreciate the income gradient implicit in the travel data is to calculate the cross-regional correlation between travel flows and income. The correlation between outbound travel and income is positive, albeit modest, at about 0.3. But if one examines the correlation between outbound *interregional* travel and income (the total net of the diagonal) the correlation rises to an astonishing 0.96. Citizen’s of richer regions travel more, and especially farther. The fact that the richest places on earth also sees the most travel activity mirrors the insights from the comparison between the rich United States and poor Nigeria. Once again, this underscores the need to develop an instrument for travel activity in order to elicit information about the impact of travel on productivity.

> Table 5 about here <

Another way to examine the bilateral travel data is through the lens of a gravity equation, to which we now turn.

Estimation of Bilateral Travel Gravity Equation We use the bilateral travel data in the estimation of the following equation:

$$\log(\text{TRAVEL}_{ij}) = \mathbf{a}'\mathbf{X}_{ij} + e_{ij}, \quad (4)$$

where \mathbf{X}_{ij} is the Frankel-Romer set of regressors.²⁰ The data on the explanatory variables are taken from CEPII.²¹ CEPII has built and made publicly available two datasets. The first is a dyadic dataset (in the sense that it includes variables valid for pairs of countries) providing among other things different measures of bilateral distances and a contiguity variable. The second dataset (not dyadic) includes among other things a variable indicating whether the country is landlocked, etc. The two datasets cover 225 countries. Remaining variables such as labor force (used to construct the dependent variable) as well as population are taken from WDI.

Results from estimation of the bilateral equation are provided Table 6. The results from the travel gravity equation are largely as one would expect if travel is narrowly connected to trade in that distance has a large and significant negative effect on bilateral travel flows; the elasticity being around -1.5 if the country pair in question do not share a common border, -0.3 if they do. Travel between countries i and j is also decreasing (increasing) in the size of the population of country i (j), the elasticity being about -0.3 (0.7). In addition, bilateral travel decreases with the size of country j , the elasticity being slightly larger (in absolute value) than -0.1. Finally, travel between countries i and j decreases if one of them is landlocked. Holding all else fixed, a non-landlocked country has twice the amount of bilateral travel as compared to a landlocked country.

>Table 6 about here<

Constructing the Travel Instrument Country i 's predicted travel share, our instrument, is obtained as the sum of the predicted geographic component of its bilateral travel share with other countries in the world:

$$\widehat{\text{TRAVEL}}_i = \sum_{j \neq i} \exp(\widehat{\mathbf{a}}'\mathbf{X}_{ij}). \quad (5)$$

Figure 2 provides a cross plot of the predicted (aggregate) travel share, $\widehat{\text{TRAVEL}}$, against TRAVEL . The simple correlation between $\widehat{\text{TRAVEL}}$ and TRAVEL is 0.58, whereas regressing TRAVEL on $\widehat{\text{TRAVEL}}$ and a constant term yields a (robust) t -value of 7.2 and a R -squared of 0.33.

²⁰With the exception of the dependent variable, the specification in Table 6 is, as mentioned in the text, the exact same as the one used in Frankel and Romer (1999; their Table 1 on page 384). \mathbf{X}_{ij} includes a constant, the distance between countries i and j , population, country area, landlock dummies, and interaction terms (all terms are interacted with a contiguity/common-border dummy).

²¹<http://www.cepii.fr/anglaisgraph/bdd/distances.htm>

> Figure 2 about here <

3.1.5 Instrumental Variables Estimation

In Table 7 we re-estimate the specifications in Table 2 by 2SLS. In order to instrument the INSTITUTIONS variable, we rely on two language variables: the fraction of the population speaking English as first language and the fraction speaking one of the major European languages (English, French, German, Portuguese or Spanish) as first language. These language instruments, which were originally proposed by Hall and Jones (1999), allow us to retain the full OLS sample when employing IV methods. The set of instruments therefore consists of the predicted trade share and the two language variables.

> Table 7 about here <

Several features of the table stand out. First, TRAVEL is highly significant in all specifications. Second, the coefficient estimate associated with TRAVEL is remarkably stable despite the progressive inclusion of additional control variables. Third, the coefficient associated with TRAVEL and BTRAVEL essentially coincide.

To gauge the quality of our instruments we rely on the framework provided by Stock and Yogo (2004). This framework provides a multivariate generalization of the well-known Staiger-Stock “rule-of-thumb”. The null being tested is that instruments are weak in the sense that inference based on IV estimates is plagued with size distortions. The Kleibergen-Paap statistic reported in Table 7 must be above the relevant critical value reported in Stock and Yogo (2004) for instruments to be strong.

The critical values allow one to perform four tests, viz. that the size is at least 10, 15, 20 or 25 percent, respectively. Note that with one endogenous variable and three excluded instruments, the rule-of-thumb, which is widely used in the literature, deems instruments strong if the maximal size is less than 15 percent (see Stock and Yogo, 2004, Table 2).²²

If we look at Column (4) in Table 7, our preferred specification, the Kleibergen-Paap value is 7.75, which means that with a nominal size of 0.05 we are unable to reject that the actual size is at least 0.15, but we can reject that it is above 0.20. In other words, weak instruments result in an effective size between 0.15 and 0.20. To address this problem, Stock and Yogo recommend using the Limited Information Maximum Likelihood estimator (LIML), which is superior to 2SLS when instruments are weak.

²² An alternative way of viewing the rule of thumb (with one endogenous variable and three excluded instruments) is that it deems instruments strong if the null of a maximal bias (relative to OLS) of at least 10 percent is rejected.

Consequently, in Table 8 we report estimates using LIML.²³ With a Kleibergen-Paap value of 7.75 in column (3) we can now conclude that the maximal size is below ten percent, as the critical values are different under LIML estimation. The same is true when BTRAVEL is employed, as seen from Columns (6) and (7) of Table 8. Hence with LIML we are able to control the size distortion to a level acceptable by the standards of the rule-of-thumb. In addition, the fact that our instrument is about equally strong in explaining TRAVEL and BTRAVEL suggests that we indeed are able to span mainly the business component of TRAVEL. The LIML point estimates for TRAVEL and BTRAVEL are also very similar.

> Table 8 about here <

Importantly, the impact of TRAVEL on $\log(A)$ is not significantly different from the impact of TRAVEL on $\log(y)$. This is consistent with the hypothesis that TRAVEL induces knowledge spillovers, thereby increasing labor productivity.

Our results suggest that TRAVEL has a substantial impact on productivity: An (exogenous) increase in TRAVEL by 1% leads to an increase in productivity by roughly 0.2%. To get a sense of what this elasticity implies, in terms of predicted TFP differences, one may begin by observing that moving from the 10th percentile to the 90th percentile in the TRAVEL distribution (in the sample of Tables 2, 7 and 8) involves increasing TRAVEL by roughly 3 yearly departures/arrivals per worker. Using the elasticity of 0.2, this difference in TRAVEL predicts a difference in the level of TFP by roughly a factor of 2.2. Moving from the 10th to the 90th percentile in the TFP distribution (in the same sample) involves a difference of roughly a factor of 5. Hence the economic significance of TRAVEL is substantial, but not implausible.

3.2 The Impact of Travel conditional on Trade

An issue of interest is whether the results above are masking a productivity enhancing effect of trade. While our identification strategy involves a travel-specific instrument, it is clear that it could be picking up the effects of trade. If so, then the exclusion restriction could be questioned. Of course, power problems notwithstanding, the OID tests reported in Tables 7 and 8 give no particular cause for such concerns. Nevertheless, the issue seems worth pursuing further.

From a theoretical angle trade may have multiple effects on productivity. Exposure to international trade may, aside from stimulating technology diffusion as noted in the Introduction, induce an intensified state of competition and lead to specialization in production. It is plausible that

²³Note that with just-identification 2SLS and LIML coincide, which is why Column (1) in Table 7 is omitted from Table 8.

these latter effects also matters for A . However, in theory, specialization and competition will not necessarily *stimulate* productivity. Consider specialization: it is clear that if an economy ends up specializing in less innovative sectors overall productivity could suffer, at least in the medium term.²⁴ Naturally, specialization could also give rise to dynamic benefits, aside from the static efficiency gain. But *a priori* the impact is ambiguous. The same goes for competition. On the one hand, in a large market firms may try harder to innovate so as to “escape competition”. But, on the other hand, in more competitive markets profits tend to be lower, which in turn reduces the incentives to innovate. The net effect of competition is therefore also ambiguous *a priori*.²⁵ To be sure, whether trade should have an impact on productivity beyond what is captured by TRAVEL is an empirical question.

3.2.1 Partial Correlations

We start by comparing the predictive power of TRAVEL and REAL OPENNESS in the OLS setting. In Table 9, Columns (1) to (5), we add REAL OPENNESS to all columns of Table 2. REAL OPENNESS is insignificant throughout. Accordingly, TRAVEL continues to be a strong (log-linear) predictor of TFP and labour productivity, when REAL OPENNESS is introduced. Moreover, the OLS estimates are in economic size very similar to those reported in Table 2.

In Columns (6) to (10) of the table, the same set of regressions are reported where TRAVEL has been substituted for by BTRAVEL. As is apparent, the results are very similar to those reported in columns (1) to (5); in all cases BTRAVEL remains significant when REAL OPENNESS is introduced. Moreover, the point estimate for BTRAVEL is very similar to those obtained without the inclusion of REAL OPENNESS.

> Table 9 about here <

3.2.2 Identification of Trade

In an effort to obtain identification of both TRAVEL and TRADE, we once again invoke the Frankel-Romer (1999) approach. That is, we use predicted *trade* shares, deriving from a gravity equation where the *bilateral trade* share is the dependent variable, to instrument the trade variable.

²⁴See Matsuyama (1992) and Galor and Mountford (2008) for growth theories which explicitly shows how this outcome could be the result of trade liberalisation.

²⁵See Aghion and Griffith (2005) for a thorough theoretical and empirical discussion of this issue. In particular, the authors present evidence of a hump-shaped association between measures of competition and productivity growth using micro-data, thereby corroborating the theoretical prediction of an ambiguous effect of competition on productivity.

More specifically, we rely on the updated and expanded version of the “Frankel-Romer instrument” computed by Alcalá and Ciccone (2004).

Alcalá and Ciccone (2004) use the exact same gravity specification as Frankel and Romer, but they do not report the estimation results. In order to understand the differences between the travel and the trade gravity equation, we therefore report the Alcalá and Ciccone trade gravity equation in Table 10.²⁶

>Table 10 about here<

Comparing Tables 6 and 10, we immediately observe that the two estimated gravity equations are very similar; that is, the sign and significance of coefficients are the same. For purposes of obtaining identification of both travel and trade using the two gravity-based instruments, it would have been preferable if the geographical variables impacted travel and trade in different ways. This would have allowed us to pin-point which underlying geographical dimensions were responsible for separating the influences of respectively travel and trade. Nevertheless, as coefficients do differ in magnitude across the two regressions, we will obtain differences in predicted travel shares and predicted trade shares; and it is these differences in predicted shares upon which identification hinges.

The fact that Tables 6 and 10 deliver results which are so similar is of course understandable considering the scatter plot of the underlying bilateral data, cf. Figure 3. There is a very strong correlation between bilateral travel and bilateral trade (correlation coeff. = 0.67).

>Figure 3 about here<

Constructing the predicted trade share as described in Section 3.1.4 gives rise to the Alcalá-Ciccone instrument for REAL OPENNESS. Figure 4 provides a scatter plot of the two instruments. With a correlation of 0.87 in our estimation sample, the figure confirms the impression obtained from inspection of Tables 6 and 10 that disentangling an impact from people flows versus goods flows is likely to be tricky. We now turn to this issue.

>Figure 4 about here<

3.2.3 Instrumental Variables Estimation

Turning to the IV setup, Table 11 reports the first stage results when REAL OPENNESS is introduced into the framework of Table 7. In the interest of brevity we only report the setting that

²⁶We thank Francesco Alcalá and Antonio Ciccone for kindly providing us with their data as well as their code files.

involves geographic controls. The first six columns involves TRAVEL, whereas the last six relates to the case where BTRAVEL is introduced instead.

> Table 11 about here <

The results are mixed. When TRAVEL is used the fitted travel share instrument turns out to be insignificant. However, when BTRAVEL is used as travel indicator the fitted travel share is significant, conditional on fitted trade shares, the language variables, and the three geo-controls (cf. Columns 7-9). Moreover, the fitted trade shares predict trade, and the language variables matter mainly to INSTITUTIONS. Accordingly, when BTRAVEL is used the scope for separating the influence from trade and travel is greater than when TRAVEL enters the model.

Tables 12 reports the second stage results. As one may have anticipated from the first stage results, second stage results are somewhat clearer in the case where BTRAVEL is invoked as the travel variable of choice; in all columns BTRAVEL exerts a significant influence on total factor productivity (cf. Table 12, Columns 6-9). By contrast, REAL OPENNESS is never significant. The same pattern emerges from Columns (2) and (3). However, in light of the first stage results, identification must be considered weak.

> Table 12 about here <

Another feature of the results in Table 12 is that the point estimates for TRAVEL (/BTRAVEL, respectively) increases considerably in absolute value compared to the results from Table 7-9. Undoubtedly, this is caused by multicollinearity between the travel and trade variables, as well as between the instruments for TRAVEL (/BTRAVEL) and REAL OPENNESS.

To be sure, it would be foolhardy to make steadfast conclusions regarding the relative magnitude of the impact of travel versus trade; identification is not strong, and the IV estimates are therefore associated with wide confidence intervals. However, it does seem reasonable to conclude from the analysis that trade does not *dominate* travel as a determinant of productivity. Rather, the common theme that emerges from the analysis is that travel retains significance when trade is introduced.

4 Panel Data Analysis

In this section we turn to panel data. We have yearly data on $\log(y)$ and TRAVEL for the period 1990 to 2006, which we use in the regressions below. Unfortunately, panel data on $\log(A)$ are not available over this period. The key problem is that data on human capital levels are only available every 5th year, ending in 2000 (Barro and Lee, 2001). Hence, if we invoke data on TFP we would be

down to only 2 time-observations, rendering the use of GMM methods impossible for all practical purposes. As a result, we stick with GDP per worker as our measure of productivity.²⁷

An advantage of using panel data is that we can control for time invariant determinants of productivity more fully by introducing fixed effects. Absent time varying instruments for travel (or trade) we are, however, forced to rely on internal instruments (or, more appropriately, lagged values of the explanatory variables in levels as instruments) for the purpose of identification.

What is the identifying assumption in this setting? In the panel setting the identifying assumption is that past levels of travel and trade hold no impact on growth, *conditional* on current income (which controls for convergence effects), travel, and trade. A reasonable concern is whether the exclusion restriction is plausible. It may be, if we accept “integration” as a “fundamental determinant” of productivity alongside geography and institutions (cf. Rodrik et al, 2004). The reason is that by employing fixed effects estimation we should be able to eliminate the influence from both geography and institutions, where the latter may have been influenced by historical cross-country interaction. Despite this one may question the excludability of past trade and travel levels on the grounds that they may have helped shape the contemporary *level* of TFP, which likely influences growth in labor productivity. But TFP is also (implicitly) controlled for by the inclusion of the level of GDP per worker. Overall, we believe the requirements for valid identification are not outrageously implausible on *a priori* grounds in the present context. Naturally, we can perform the relevant over-identification test whether data rejects instrument validity.

We estimate the following dynamic fixed-effects panel data model,

$$\log(y_{it}) = \eta_1 + \eta_2 TRAVEL_{it} + \eta_3 \log(y_{it-1}) + \mathbf{X}'_i \cdot \boldsymbol{\eta}_4 + u_i + \nu_{it}. \quad (6)$$

where $\mathbf{X}'_i \cdot \boldsymbol{\eta}_4$ is the observed fixed effects. In equation (Y) these included continent dummies, INSTITUTIONS, LANDLOCK, LATITUDE, and TROPICS. Furthermore, u_i is an *unobserved* fixed effect. Specifically, instead of using instruments for time-invariant variables, we simply remove them (observed as well as unobserved) by a first difference transformation of (6):

$$\Delta \log(y_{it}) = \eta_2 \Delta TRAVEL_{it} + \eta_3 \Delta \log(y_{it-1}) + \Delta \nu_{it}. \quad (7)$$

This model can then be estimated using lagged values of the explanatory variables in levels as instruments in a GMM framework along the lines of Arellano and Bond (1991). The Arellano-Bond estimator is the standard estimator for dynamic panel data models with unobserved heterogeneity (Greene, 2008).

²⁷Updated data on travel and labor productivity are from the WDI (2007).

We also estimate versions of (6) where REAL OPENNESS is included among the right-hand side variables. This allows us to gauge the impact from cross-border flows of people, conditional on trade, while relying on the *within* variation as opposed to the *between* variation, and by way of an alternative identification strategy.

A useful aspect of the panel exercise is that it allows for an independent check of the cross-sectionally estimated elasticity of productivity with respect to travel. Observe that in the panel setting we allow for transitional dynamics, by conditioning on $\log(y_{it-1})$ in equation (6). As a consequence, the parameter η_2 represents an impact effect and is therefore not directly comparable to the cross sectional results (i.e., β_2 in equation (Y)). For a proper comparison we solve for “the steady state” (i.e., put $\log(y_{it}) = \log(y_{it-1}) = \log(y^*)$ in equation (6)) so as to obtain a “long-run” elasticity (recall, TRAVEL is logged): $\partial \log(y^*) / \partial \text{TRAVEL} = \eta_2 / (1 - \eta_3)$. This elasticity should represent the counterpart to β_2 in equation (Y). Using the delta method, we can subsequently perform a nonlinear test of the hypothesis that the long-run elasticity equals our cross-section estimate; that is, $H_0 : \beta_2 = \eta_2 / (1 - \eta_3) = 0.2$.

Accordingly, (6) represents a generalization of (Y) in the sense that it provides information about convergence as well as allowing for unobserved fixed effects. The central question is whether $\eta_2 / (1 - \eta_3)$ matches β_2 ; the absence of large discrepancies between panel and pure cross-section results would strengthen the case for a causal interpretation of our cross-sectional findings. Note, in addition, that there is a flavor of out-of-sample testing to this exercise. Since we do not have to rely on external instruments we can expand the country coverage dramatically; the panel estimations involve nearly twice as many countries as the cross-section analysis. Table 13 report estimation results.

The first thing to note is that TRAVEL is highly significant in all columns. Columns (1) and (5) are estimated using all instruments, whereas Columns (2) to (4) and (6) to (8) conserve on the use of instruments sequentially. Prolific use of internal instruments may cause problems in finite samples, hence it is important to check robustness of results to reducing the number of instruments (Roodman, 2008). Columns (2) to (4) and (6) to (8) reduce the maximum number of lags that can be used as instruments from four to two, and, as inspection of these columns reveal, nothing changes statistically when we limit the number of instruments in this manner. Of particular interest is Columns (5) to (8) where we have included REAL OPENNESS in addition to TRAVEL. The trade measure is only marginally significant in column (5) of the table, but it turns insignificant and changes sign when the number of lags that can be used as instruments are sequentially reduced from four (column (6)) to two (column (8)). TRAVEL, on the other hand, is unaffected by both the inclusion of REAL OPENNESS and the reduction in the number of allowable internal instruments. Consistent with the cross-section analysis we thus find little evidence to suggest that omitting trade

influences the estimates for travel.

> Table 13 about here <

Turning to the economic significance of travel, the obtained estimate basically coincides with the cross-section estimate. As can be seen from the table, a $\beta_2 = \eta_2/(1 - \eta_3) = 0.2$ hypothesis cannot be rejected in any of the specifications reported in Table 13; our remarks on economic significance from Section 3.1.5 therefore carry over.

Overall, it is rather remarkable that we obtain a long-run travel-productivity elasticity of about 0.2 whether we use: (i) OLS and cross-sectional data (Table 2); (ii) 2SLS, external instruments, and cross-sectional data (Tables 7-9 and 12); or (iii) internal instruments and panel data (Table 13). Taken together, these results make a strong case for the importance of cross-border flows of people for long-run productivity.

5 Concluding Remarks

Theoretically, observed differences in TFP levels can (in part) be motivated by differences in technology adoption rates. But why do some countries adopt technology more readily than others? In the present paper we have examined the hypothesis that integration, as manifested in cross-border flows of people, facilitates the international spread of ideas.

There are many reasons why the direct interaction between individuals may be crucial. Tacit knowledge almost inevitably needs to be communicated person-to-person. Even general technological knowledge is often not completely codified, which implies that it can only be transmitted fully if individuals meet in person. Many growth theories build on the idea that spillovers between individuals exist and are substantial.

Our empirical analysis shows that a compelling case can be made for the existence of cross-border spillovers using aggregate data. The intensity of travel is a strong predictor of aggregate productivity levels, even controlling for measures of the institutional infrastructure, trade, and key climate related circumstances. The estimated causal effect of travel on productivity is statistically and economically significant. Moreover, the association between international travel and GDP per worker can be accounted for by the association between international travel and TFP. Taken together, these findings suggest that cross-border flows of people is an important vehicle for knowledge diffusion.

Our results raises some new interesting policy issues. For instance, bilateral visa restrictions, which are arguable a major impediment to the travel opportunities of billions of people, may be a barrier to knowledge diffusion. The results presented in this paper point to costs of visa restrictions

in terms of lower productivity; these costs must be weighed against the benefits in terms of stemming the unwanted consequences of cross-border travel such as inflows of illegal immigration, criminals and potential terrorists. Quantifying these costs seem to be an interesting topic for future research.

References

- [1] Acemoglu, D. and J. Robinson, 2001. Political Losers As a Barrier to Economic Development. *American Economic Review*, 90, pp. 126-130.
- [2] Acemoglu, D., S. Johnson and J. Robinson, 2001. The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, 91, 1369-1401.
- [3] Aghion P. and R. Griffith, 2005. *Competition and Growth*. MIT Press, Cambridge.
- [4] Aghion P. and P. Howitt, 1992. A Model of Growth Through Creative Destruction. *Econometrica* 60, 323-51.
- [5] Agrawal, A., I. Cockburn and J. McHale, 2006. Gone but forgotten: knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography*, 6, 571-591.
- [6] Alcalá, F. and A. Ciccone, 2004. Trade and Productivity. *Quarterly Journal of Economics*, 119, 613-646.
- [7] Almeida, P. and B. Kogut, 1999. Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science* 45: 905-917.
- [8] Arellano, M. and S. Bond, 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58, 277-297.
- [9] Arrow, K., 1969. Classificatory notes on the production and transmission of technological knowledge. *American Economic Review*, 59, 29-35.
- [10] Ashraf, Q and O. Galor , 2007. "Cultural Assimilation, Cultural Diffusion and the Origin of the Wealth of Nations," Mimeo (Brown University).
- [11] Ashraf Q. and O. Galor O, 2008. "Human Genetic Diversity and Comparative Economic Development," Mimeo (Brown University).
- [12] Barro, R. J. and Lee J-W., 2001. International data on educational attainment: updates and implications. *Oxford Economic Papers*, 53, 541-63.
- [13] Caselli, F., 2006, Accounting for Cross-Country Income Differences, in Philippe Aghion and Steve Durlauf (eds.), *Handbook of Economic Growth*, Vol. 1A, North-Holland, Amsterdam.

- [14] Coe, D. and E. Helpman, 1995. International R&D Spillovers. *European Economic Review*, 39, p. 857-87.
- [15] Coe, D., E. Helpman and A. Hoffmaister, 1997. North-South R&D Spillovers. *Economic Journal*, 107, 134-49.
- [16] Crespi, G., C. Chriscuolo, J.E. Haskel and M. Slaughter, 2008. Productivity growth, knowledge flows, and spillovers. NBER working paper 13959.
- [17] Deaton, A., 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development. Policy*. The Johns Hopkins University Press.
- [18] Diamond, J., 1999. *Guns, Germs and Steel: The Fates of Human Societies*. W.W. Norton & Company Inc., New York.
- [19] de la Potterie B. and F. Lichtenberg, 2001. Does Foreign Direct Investment Transfer Technology Across Borders? *The Review of Economics and Statistics*, 83, p. 490-497.
- [20] Dowrick, S. and M. Tani, 2008. International Business Visits and the Technology Frontier. IZA DP No. 3417.
- [21] Eaton J. and S. Kortum, 2002. Technology, Geography, and Trade. *Econometrica* 70, 1741-79.
- [22] Foster, A., and M. Rosenzweig, 1995. Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture. *Journal of Political Economy*, 103, 1176-1209.
- [23] Frankel J. and D. Romer, 1999. Does trade cause growth? *American Economic Review*, 89, 379-99.
- [24] Gallup, J., J. Sachs, and A. Mellinger, 1999. Geography and Economic Development. *International Regional Science Review*, 22, 2, 179-232
- [25] Galor O, and A. Mountford, 2008. Trading Population for Productivity: Theory and Evidence. *Review of Economic Studies*, 75 , p. 1143-79.
- [26] Gambardella A, M. Mariani and S. Torrìsi, 2003. How 'Provincial' is your Region? Effects on Labour Productivity in Europe. Working Paper (Laboratory of Economics and Management Sant'Anna School of Advanced Studies).
- [27] Glaeser, E., R. La Porta, F. Lopes-de-Silanes, A. Shleifer, 2004. Do Institutions Cause Growth? *Journal of Economic Growth*, 9, 271-303.

- [28] Gollin, D., 2002. Getting Income Shares Right. *Journal of Political Economy*, 110, 458-74.
- [29] Greene, W., 2008. *Econometric Analysis*, Sixth Edition, Prentice Hall.
- [30] Hall, R. and C. Jones, 1999. Why do some countries produce so much more output than others? *Quarterly Journal of Economics*, 114, 83-116.
- [31] Heston A., Summers R. and B. Aten, 2002. Penn World Table Version 6.1. Center for International Comparisons at the University of Pennsylvania (CICUP).
- [32] Howitt, P., 2000. Endogenous Growth and Cross-Country Income Differences. *American Economic Review*, 90, 829-46.
- [33] Irwin D., and P. Klenow, 1994. Learning Spillovers in the Semi-Conductor Industry. *Journal of Political Economy*, 102, 1200-27.
- [34] Jones, A., 2007. More than 'managing across borders?' the complex role of face-to-face interaction in globalizing law firms, *Journal of Economic Geography*, 7, 223-246.
- [35] Klenow, P. and A. Rodriguez-Clare, 1997. The Neoclassical Revival: Has it gone too far? In B. Bernanke and J. Rotemberg (eds.) *NBER Macroeconomic Annual*, Cambridge, MA: MIT Press, 73-102.
- [36] Keller, W., 2004. International Technology Diffusion. *Journal of Economic Literature*, XLII, 752-82.
- [37] Lucas, R. Jr., 1993. Making a Miracle. *Econometrica*, 61, 251-72.
- [38] Mankiw, G., D. Romer and D. Weil, 1992. A contribution to the empirics of economic growth. *Quarterly Journal of Economics*, 107, p. 407-37.
- [39] Matsuyama, K., 1992. Agricultural productivity, comparative advantage and economic growth. *Journal of Economic Theory*, 58, 317-334.
- [40] Nelson, R. and E. Phelps, 1966. Investment in Humans, Technological Diffusion, and Economic Growth. *American Economic Review*, 56, p. 69-75.
- [41] Parente, S. and E. Prescott, 1999. Monopoly Rights: A Barrier to Riches. *American Economic Review*, 1216-1233.
- [42] Rauch, J.E., 2001. Business and Social Networks in International Trade. *Journal of Economic Literature*, 39, 1177-1203.

- [43] Rhee, Y.W., 1990. The Catalyst Model of Development: Lessons from Bangladesh's Success with Garment Exports. *World Development*, 18, p. 333-46.
- [44] Roodman, D., 2008. How to Do xtabond2: An introduction to "Difference" and "System" GMM in Stata, *Stata Journal*, vol. 9(1), p. 86-136.
- [45] Rodrik D., A. Subramanian and F. Trebbi, 2004. Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development. *Journal of Economic Growth*, 9, p. 131-165.
- [46] Romer, P., 1990. Endogenous Technical Change. *Journal of Political Economy* 98, 71-102.
- [47] Sachs, J., 2003. Institutions Don't rule: Direct Effects of Geography on Per Capita Income. NBER working paper no. 9490.
- [48] Saxenian, A., 1994. *Regional Advantage*. Harvard University Press.
- [49] Solow, R., 1956. A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70, p. 65-94.
- [50] Spolaore, E. and R. Wacziarg, 2009. The Diffusion of Development. *Quarterly Journal of Economics*, 124, p. 469-529
- [51] Staiger, D. and J. Stock, 1997. Instrumental variables with Weak Instruments. *Econometrica* 65, 557-586.
- [52] Stock, J. and M. Yogo, 2004. Testing for Weak Instruments in Linear IV Regression. Manuscript (Harvard University).
- [53] Storper, M. and A. Venables, 2004. Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography*, 4, 351-370.
- [54] UNWTO, 2000. *Compendium of Tourism Statistics*. The World Tourism Organization, Madrid.
- [55] WDI, 2004. *World Development Indicators CD-rom*. The World Bank.
- [56] WDI, 2007. *World Development Indicators CD-rom*. The World Bank.

A ON THE SOURCES OF THE BILATERAL TRAVEL DATA

The data that we use on bilateral travel flows are collected by the World Tourist Organization (UNWTO). They are freely available for the purpose of research but needs to be requested (see <http://www.unwto.org/statistics/data/note.pdf>). Also, the UNWTO requests to be alerted to research publications that result from the use of the data.

The data we thus obtained concerns the period 2003-2007; by now, updates may be available.

The data package that we received was in form of a large number of country specific files detailing outbound travel from each nation. However, the source of this information is arrival data. “Outbound travel” is, by the UNWTO, inferred from “inbound travel” since virtually no country keeps records about outbound travel.

To take a concrete example: the information on outbound trips from the US, reported above, derive from registrations (in connection with visa applications, or in airports, or even in hotel registers) in destination country. Since most of the countries in the world are part of the data set, this information allows the UNWTO to infer outbound travel.

After receiving the raw data from the UNWTO we have subsequently consolidated the resulting data in dyadic format, so as to estimate a gravity equation. The resulting data set is available from either one of the authors upon request.

FIGURES AND TABLES



Figure 1. TRAVEL versus. log(A).

Notes: Number of observations is 72. Country codes are: (AGO) Angola, (ARG) Argentina, (AUS) Australia, (AUT) Austria, (BDI) Burundi, (BEL) Belgium, (BEN) Benin, (BGD) Bangladesh, (BOL) Bolivia, (BRA) Brazil, (BWA) Botswana, (CAN) Canada, (CHE) Switzerland, (CHL) Chile, (CHN) China, (CIV) Cote d'Ivoire, (COL) Colombia, (CRI) Costa Rica, (CYP) Cyprus, (DNK) Denmark, (DOM) Dominican Republic, (DZA) Algeria, (ECU) Ecuador, (EGY) Egypt, (ESP) Spain, (ETH) Ethiopia, (FIN) Finland, (FJI) Fiji, (FRA) France, (GBR) United Kingdom, (GRC) Greece, (GTM) Guatemala, (HKG) Hong Kong, China, (HND) Honduras, (HUN) Hungary, (IDN) Indonesia, (IND) India, (IRL) Ireland, (IRN) Iran, (ISL) Iceland, (ISR) Israel, (ITA) Italy, (JOR) Jordan, (JPN) Japan, (KEN) Kenya, (KOR) Korea, Rep., (LKA) Sri Lanka, (LSO) Lesotho, (MAR) Morocco, (MDG) Madagascar, (MEX) Mexico, (MUS) Mauritius, (MYS) Malaysia, (NER) Niger, (NGA) Nigeria, (NIC) Nicaragua, (NLD) Netherlands, (NOR) Norway, (NPL) Nepal, (NZL) New Zealand, (PAN) Panama, (PER) Peru, (PHL) Philippines, (PNG) Papua New Guinea, (PRT) Portugal, (PRY) Paraguay, (ROM) Romania, (SGP) Singapore, (SLV) El Salvador, (SWE) Sweden, (SYR) Syrian Arab Republic, (TCD) Chad, (THA) Thailand, (TTO) Trinidad and Tobago, (TUN) Tunisia, (TUR) Turkey, (TZA) Tanzania, (USA) United States, (VEN) Venezuela, (ZAF) South Africa, (ZWE) Zimbabwe.

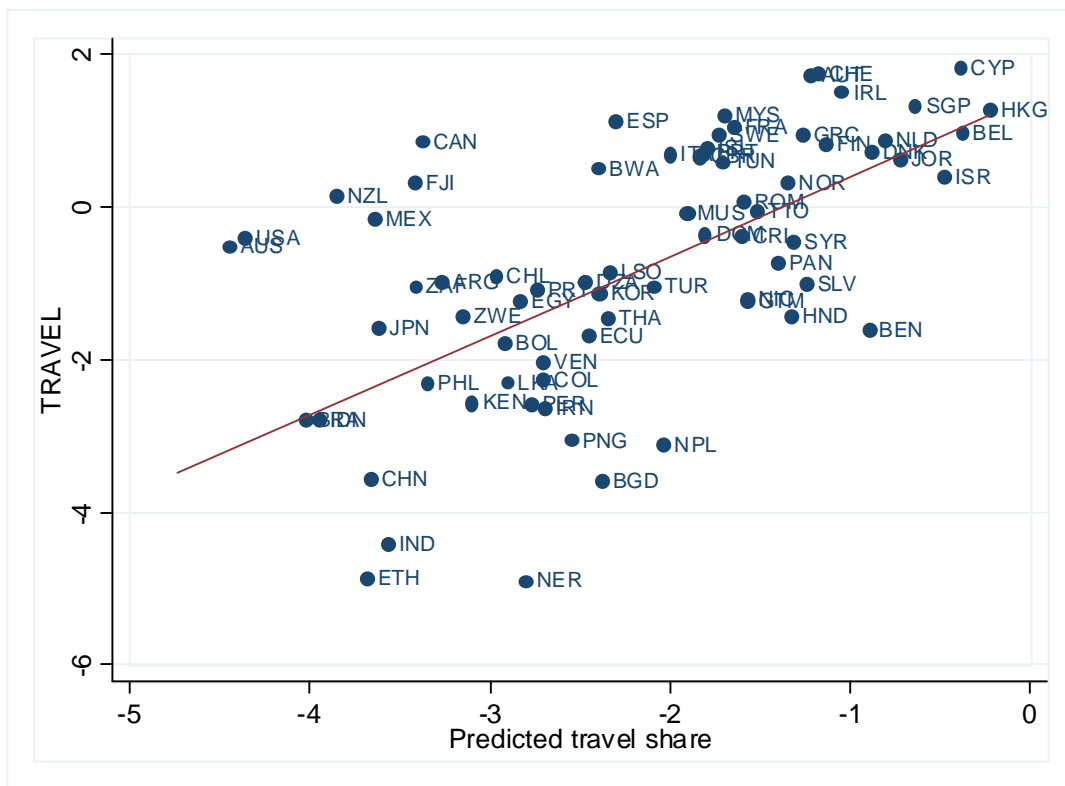


Figure 2. Predicted travel share versus TRAVEL.

Notes: Number of observations is 72. See Figure 1 for the key to country codes.

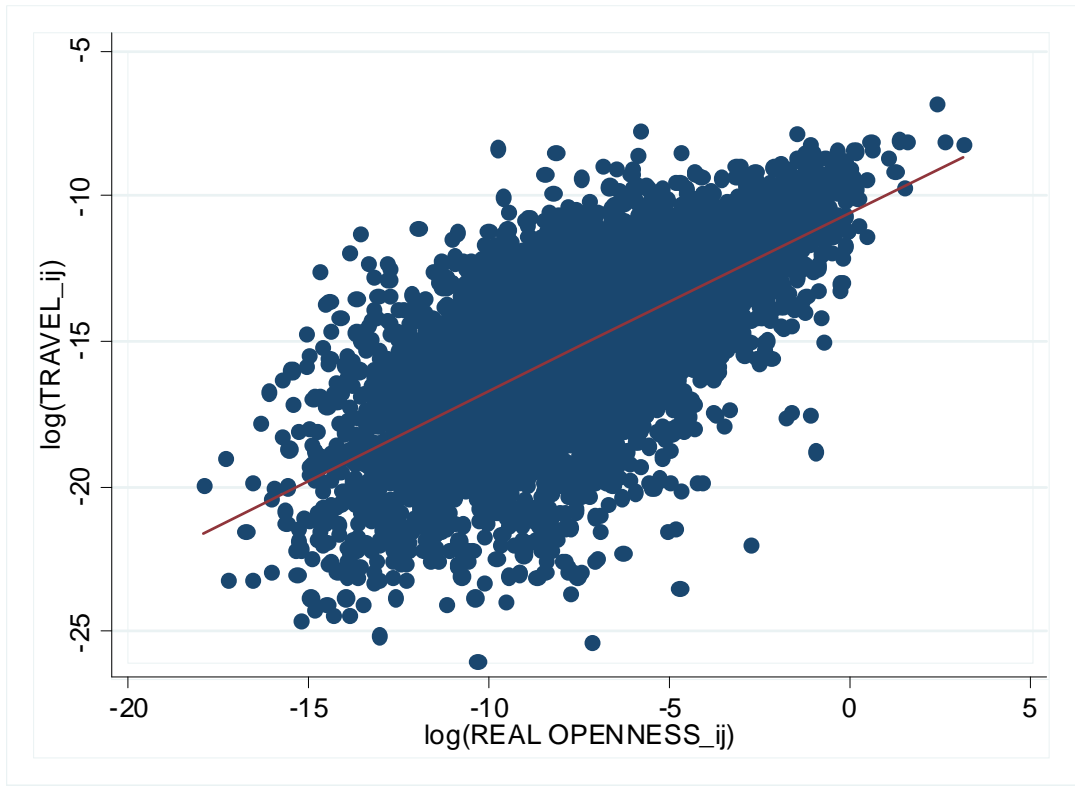


Figure 3. Bilateral trade versus bilateral travel.

Notes: Number of observations is 6303; correlation coefficient is 0.67.

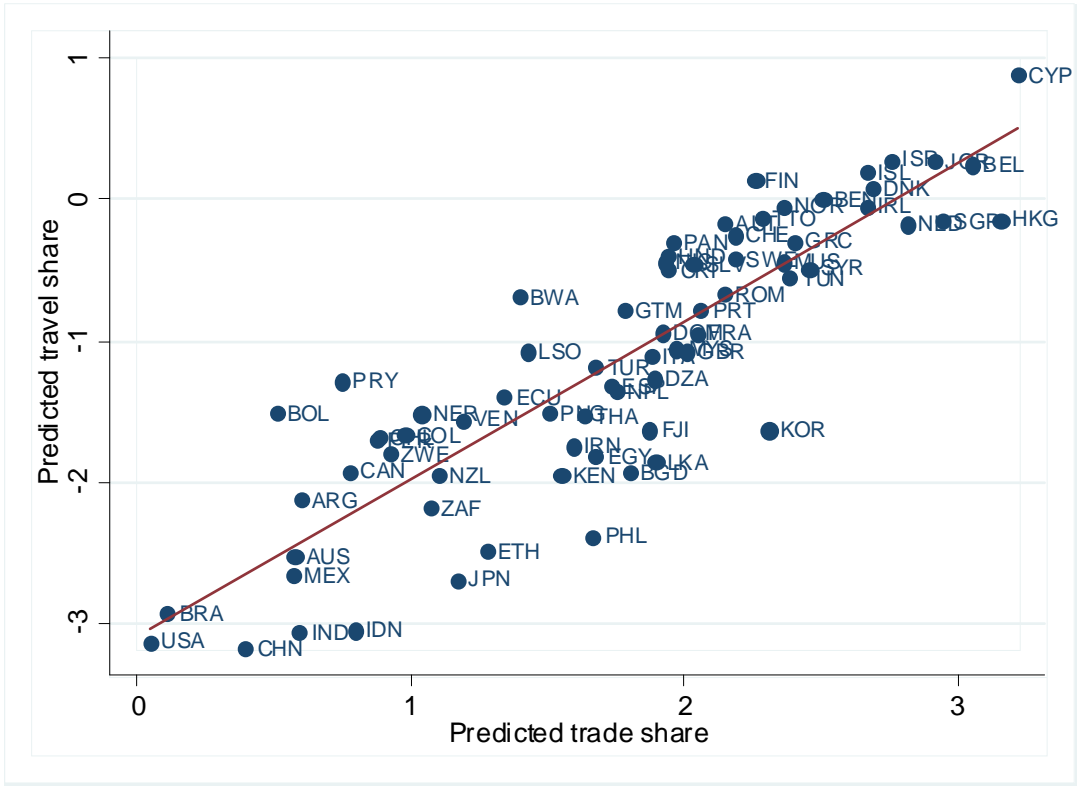


Figure 4. Predicted trade share versus predicted travel share.
 Notes: Number of observations is 72; correlation coefficient is 0.87.

Table 1: Summary Statistics for Selected Variables

Panel A							
Variable	Observations	Mean	Median	Std. Dev.	Min	Max	
TRAVEL	72	-0.72	-0.62	1.64	-4.9	1.83	
log(A)	94	8.48	8.72	0.75	6.01	9.54	
log(y)	94	9.38	9.56	1.14	6.45	10.96	
INSTITUTIONS	93	0.63	0.58	0.21	0.23	1	
LANDLOCK	89	0.18	0	0.38	0	1	
TROPICS	89	0.19	0.03	0.27	0	1	
LATITUDE	92	0.27	0.22	0.19	0	0.72	

Panel B							
Correlation Matrix							
TRAVEL	1						
log(A)	0.67	1					
log(y)	0.79	0.90	1				
INSTITUTIONS	0.65	0.56	0.76	1			
LANDLOCK	-0.11	-0.35	-0.33	-0.08	1		
TROPICS	-0.44	-0.42	-0.46	-0.39	-0.07	1	
LATITUDE	0.57	0.44	0.60	0.70	-0.05	-0.70	1

Notes: Correlations are calculated for the 72-country sample where TRAVEL is available. Panel A provides standard summary statistics, whereas Panel B provides the correlation matrix. All variables are explained in the main text.

Table 2: OLS Regressions

Dependent variable	(1) log(A)	(2) log(A)	(3) log(A)	(4) log(A)	(5) log(y)	(6) log(A)	(7) log(y)
TRAVEL	0.251*** (0.026)	0.190*** (0.039)	0.167*** (0.040)	0.165*** (0.041)	0.254*** (0.041)		
BTRAVEL						0.123*** (0.045)	0.180*** (0.041)
INSTITUTIONS		0.688** (0.319)	0.918* (0.463)	0.918* (0.471)	2.071*** (0.354)	1.063** (0.439)	2.417*** (0.346)
LANDLOCK			-0.547*** (0.200)	-0.459** (0.193)	-0.460*** (0.161)	-0.551* (0.313)	-0.603** (0.256)
TROPICS			-0.780*** (0.239)	-0.760*** (0.222)	-0.630** (0.273)	-0.314 (0.365)	-0.407 (0.296)
LATITUDE			-0.804 (0.501)	-0.687 (0.485)	-0.188 (0.522)	0.270 (0.730)	0.443 (0.650)
CONSTANT	8.868*** (0.049)	8.369*** (0.254)	8.648*** (0.243)	8.558*** (0.403)	8.651*** (0.307)	8.212*** (0.474)	8.500*** (0.367)
CONTINENTS	NO	NO	NO	YES	YES	YES	YES
Observations	72	71	68	68	68	62	62
R-squared	0.46	0.48	0.59	0.61	0.86	0.61	0.86

Notes: Ordinary Least Squares estimates. The dependent variable in Columns (1)-(4) and (6) is the log of productivity in 1996; in Columns (5) and (7) it is the log of real GDP (PPP) per worker in 1996. Variables are described in the main text. CONTINENTS refers to whether the model was estimated with continental dummies (Africa, America, Asia & Oceania and Europe) included. All standard errors (reported in parenthesis) are heteroskedasticity robust. Asterisks ***, **, * indicate significance at a 1%, 5% and 10% level, respectively.

Table 3: United States: Top 20, in- and outbound travel, 2005

Destination country for outbound travel	Number of visits	Country of origin for inbound travel	Number of visits
Mexico	20,325,274	Canada	14,862,096
Canada	14,390,696	Mexico	12,665,000
United Kingdom	3,436,000	United Kingdom	4,344,957
France	3,048,000	Japan	3,883,906
Puerto Rico	2,828,600	Germany	1,415,530
Italy	2,326,802	France	878,648
Germany	1,949,825	Korea, Rep.	705,093
China	1,555,450	Australia	581,773
Bahamas	1,380,083	Italy	545,546
Jamaica	1,059,640	Brazil	485,373
Dominican Republic	986,934	Netherlands	448,650
Spain	883,523	Spain	385,640
Hong Kong, China	877,900	Ireland	383,400
Ireland	854,000	India	344,926
Japan	822,033	Venezuela	340,315
Brazil	793,559	Colombia	325,398
Costa Rica	758,134	Sweden	290,530
Switzerland	657,060	Israel	284,310
India	611,165	China	270,272
Thailand	585,476	Switzerland	256,730

Source: World Tourism Organization.

Table 4: Nigeria: Top 20, in- and outbound travel, 2005

Destination country for outbound travel	Number of visits	Country of origin for inbound travel	Number of visits
United Kingdom	141,000	Niger	620,658
Saudi Arabia	96,671	Benin	393,215
Ghana	74,983	Liberia	107,401
United States	47,397	Cameroon	107,108
South Africa	28,995	Chad	85,077
Benin	20,500	Italy	65,593
China	19,916	Sudan	63,046
Israel	10,726	France	61,871
India	10,049	Germany	60,348
Ethiopia	5,524	Cote d'Ivoire	50,668
Canada	4,395	Mexico	41,744
Zimbabwe	3,367	United Kingdom	38,628
Burkina Faso	3,366	Ethiopia	37,117
Togo	3,356	Mali	36,986
Korea, Rep.	3,167	Algeria	36,577
Belgium	2,644	Togo	34,036
Bahrain	2,052	China	33,290
Brazil	1,819	Burkina Faso	31,842
Angola	1,799	Pakistan	27,701
Lebanon	1,592	India	25,946

Source: World Tourism Organization.

Table 5: Global travel patterns, 2005 (regional breakdown)

	East Asia & Pacific	Europe & Central Asia	Latin America & Caribbean	Middle East & North Africa	North America	South Asia	Sub-Saharan Africa	Total outbound
East Asia & Pacific	171,000,000	10,100,000	472,658	1,101,891	7,383,103	917,408	508,647	191,483,707
Europe & Central Asia	15,100,000	322,000,000	7,564,503	10,700,000	12,700,000	2,429,692	4,139,770	374,633,965
Latin America & Caribbean	361,354	2,085,582	14,300,000	164,787	16,600,000	40,862	152,488	33,705,073
Middle East & North Africa	948,448	4,463,870	126,821	19,200,000	595,807	214,708	219,156	25,768,810
North America	8,123,513	19,500,000	34,100,000	1,231,948	29,500,000	1,011,592	637,365	94,104,418
South Asia	2,701,082	698,950	18,701	1,602,559	505,248	1,372,341	167,729	7,066,610
Sub-Saharan Africa	4,722,024	876,828	78,110	641,166	352,366	202,042	11,500,000	18,372,536
Total inbound	202,956,421	359,725,230	56,660,793	34,642,351	67,636,524	6,188,645	17,325,155	745,135,119
GDP per capita (PPP)	6671	19476	8901	8161	41797	2134	1733	
Total Population (millions)	2122	875	553	349	328	1478	762	
Countries	21	29	31	20	3	6	43	

Sources: Travel data from World Tourism Organization (authors calculations); GDP and population from World Development Indicators.

Table 6: Travel gravity equation (constrained linear regression)

Dependent variable	$\log(\text{TRAVEL}_{ij})$
$\log(\text{DISTANCE}_{ij})$	-1.582*** (0.067)
$\log(\text{POPULATION}_i)$	-0.308** (0.124)
$\log(\text{POPULATION}_j)$	0.706*** (0.026)
$\log(\text{AREA}_i)$	-0.112 (0.096)
$\log(\text{AREA}_j)$	-0.120*** (0.016)
CONTIGUITY_{ij}	1.719 (2.319)
LANDLOCK_i	-0.975*** (0.167)
LANDLOCK_j	-0.975*** (0.167)
$\text{CONTIGUITY}_{ij} \times \log(\text{DISTANCE}_{ij})$	1.287*** (0.270)
$\text{CONTIGUITY}_{ij} \times \log(\text{POPULATION}_i)$	-0.190 (0.171)
$\text{CONTIGUITY}_{ij} \times \log(\text{POPULATION}_j)$	-0.167 (0.136)
$\text{CONTIGUITY}_{ij} \times \log(\text{AREA}_i)$	-0.110 (0.175)
$\text{CONTIGUITY}_{ij} \times \log(\text{AREA}_j)$	-0.125 (0.136)
$\text{CONTIGUITY}_{ij} \times \text{LANDLOCK}_i$	0.359 (0.220)
$\text{CONTIGUITY}_{ij} \times \text{LANDLOCK}_j$	0.360 (0.220)
CONSTANT	1.777 (1.208)
Number of countries	177
Number of observations	14111
R-squared	0.398

Notes: The dependent variable is the (log of) the sum of in- and outbound travel between countries i and j divided by the size of the labor force in country i . Variables are described in the main text. Asterisk *, **, *** denote significance at 1, 5, 10 % respectively. All standard errors (reported in parenthesis) are cluster robust. Following Frankel and Romer (1999), the coefficients on LANDLOCK_i and LANDLOCK_j and the coefficients on their interaction with CONTIGUITY_{ij} have been constrained to be equal.

Table 7: 2SLS Regressions

Dependent variable	(1) log(A)	(2) log(A)	(3) log(A)	(4) log(A)	(5) log(y)	(6) log(A)	(7) log(y)
TRAVEL	0.187*** (0.051)	0.173*** (0.057)	0.167*** (0.046)	0.171*** (0.053)	0.183*** (0.057)		
BTRAVEL						0.191** (0.077)	0.178*** (0.064)
INSTITUTIONS		1.165* (0.616)	1.656** (0.667)	1.375** (0.574)	2.336*** (0.499)	0.502 (0.678)	2.012*** (0.619)
LANDLOCK			-0.533*** (0.192)	-0.454** (0.186)	-0.492*** (0.164)	-0.612** (0.286)	-0.625*** (0.241)
TROPICS			-0.897*** (0.189)	-0.795*** (0.198)	-0.753*** (0.275)	-0.170 (0.378)	-0.354 (0.321)
LATITUDE			-1.446** (0.597)	-0.985** (0.503)	-0.278 (0.498)	0.634 (0.913)	0.774 (0.779)
CONSTANT	8.822*** (0.061)	8.036*** (0.459)	8.367*** (0.349)	8.312*** (0.433)	8.537*** (0.366)	8.585*** (0.448)	8.658*** (0.411)
CONTINENTS	NO	NO	NO	YES	YES	YES	YES
OID (p-value)		0.04	0.16	0.25	0.43	0.25	0.21
Kleibergen-Paap statistic	30.60	4.53	4.80	7.75	7.75	8.55	8.55
Observations	72	71	68	68	68	62	62

Notes: Two-Stage Least Squares estimates. The dependent variable in Columns (1)-(4) and (6) is the log of productivity in 1996; in Columns (5) and (7) it is the log of real GDP (PPP) per worker in 1996. Excluded instruments are the fitted travel share and the two European language variables. The Kleibergen-Paap statistic can be used to test the strength of instruments; critical values are supplied by Stock and Yogo (2004). All variables are described in the main text. CONTINENTS refers to whether the model was estimated with continental dummies (Africa, America, Asia & Oceania and Europe) included. All standard errors (reported in parenthesis) are heteroskedasticity robust. Asterisks ***, **, * indicate significance at a 1%, 5% and 10% level, respectively.

Table 8: LIML Regressions

Dependent variable	(1) log(A)	(2) log(A)	(3) log(A)	(4) log(y)	(5) log(A)	(6) log(y)
TRAVEL	0.156** (0.075)	0.166*** (0.048)	0.171*** (0.053)	0.182*** (0.058)		
BTRAVEL					0.194** (0.079)	0.177*** (0.066)
INSTITUTIONS	1.497 (1.022)	1.773** (0.779)	1.396** (0.596)	2.338*** (0.506)	0.472 (0.711)	1.971*** (0.660)
LANDLOCK		-0.531*** (0.193)	-0.454** (0.186)	-0.492*** (0.165)	-0.615** (0.286)	-0.627*** (0.241)
TROPICS		-0.916*** (0.197)	-0.796*** (0.198)	-0.754*** (0.276)	-0.163 (0.381)	-0.348 (0.326)
LATITUDE		-1.543** (0.677)	-0.999* (0.512)	-0.279 (0.500)	0.654 (0.931)	0.808 (0.804)
CONSTANT	7.801*** (0.747)	8.320*** (0.392)	8.301*** (0.444)	8.536*** (0.368)	8.604*** (0.458)	8.674*** (0.423)
CONTINENTS	NO	NO	YES	YES	YES	YES
OID (p-value)	0.05	0.16	0.25	0.43	0.25	0.21
Kleibergen-Paap statistic	4.35	4.80	7.75	7.75	8.55	8.55
Observations	71	68	68	68	62	62

Notes: Limited Information Maximum Likelihood estimates. The dependent variable in Columns (1)-(3) and (5) is the log of productivity in 1996; in Columns (4) and (6) it is the log of real GDP (PPP) per worker in 1996. Excluded instruments are the fitted travel share and the two European language variables. The Kleibergen-Paap statistic can be used to test the strength of instruments; critical values are supplied by Stock and Yogo (2004). All variables are described in the main text. CONTINENTS refers to whether the model was estimated with continental dummies (Africa, America, Asia & Oceania and Europe) included. All standard errors (reported in parenthesis) are heteroskedasticity robust. Asterisks ***, **, * indicate significance at a 1%, 5% and 10% level, respectively.

Table 9: OLS Regressions

Dependent variable	(1) log(A)	(2) log(A)	(3) log(A)	(4) log(A)	(5) log(Y)	(6) log(A)	(7) log(A)	(8) log(A)	(9) log(A)	(10) log(Y)
TRAVEL	0.241*** (0.061)	0.211*** (0.063)	0.164** (0.067)	0.155** (0.066)	0.266*** (0.063)					
BTRAVEL						0.250*** (0.074)	0.186*** (0.069)	0.179** (0.084)	0.176* (0.092)	0.173* (0.088)
REAL OPENNESS	0.023 (0.118)	-0.064 (0.118)	0.008 (0.116)	0.024 (0.118)	-0.028 (0.106)	0.003 (0.124)	-0.144 (0.122)	-0.116 (0.143)	-0.134 (0.165)	0.017 (0.156)
INSTITUTIONS		0.766** (0.314)	0.903* (0.472)	0.878* (0.459)	2.118*** (0.393)		1.637*** (0.419)	1.123** (0.522)	1.226** (0.467)	2.396*** (0.348)
LANDLOCK			-0.548*** (0.196)	-0.465** (0.187)	-0.453*** (0.154)			-0.798*** (0.250)	-0.557* (0.309)	-0.602** (0.259)
TROPICS			-0.784*** (0.267)	-0.773*** (0.236)	-0.615** (0.286)			-0.270 (0.376)	-0.327 (0.361)	-0.405 (0.293)
LATITUDE			-0.797 (0.501)	-0.663 (0.507)	-0.216 (0.544)			-0.019 (0.687)	0.137 (0.702)	0.460 (0.648)
CONSTANT	8.786*** (0.434)	8.543*** (0.457)	8.628*** (0.416)	8.495*** (0.603)	8.725*** (0.449)	9.360*** (0.595)	8.602*** (0.618)	8.992*** (0.666)	8.726*** (0.790)	8.434*** (0.744)
CONTINENTS	NO	NO	NO	YES	YES	NO	NO	NO	YES	YES
Observations	72	71	68	68	68	67	66	62	62	62
R-squared	0.46	0.48	0.59	0.61	0.86	0.36	0.45	0.58	0.62	0.86

Notes: Ordinary Least Squares estimates. The dependent variable in Columns (1)-(4), (6)-(9) is the log of productivity in 1996; in Columns (5) and (10) it is the log of real GDP (PPP) per worker in 1996. Variables are described in the main text. CONTINENTS refers to whether the model was estimated with continental dummies (Africa, America, Asia & Oceania and Europe) included. All standard errors (reported in parenthesis) are heteroskedasticity robust. Asterisks ***, **, * indicate significance at a 1%, 5% and 10% level, respectively.

Table 10: Trade gravity equation (constrained linear regression)

Dependent variable	(2) log(REAL OPENNESS _{ij})
log(DISTANCE _{ij})	-1.132*** -0.057
log(POPULATION _i)	-0.247*** -0.062
log(POPULATION _j)	0.869*** (0.023)
log(AREA _i)	-0.084 (0.057)
log(AREA _j)	-0.190*** (0.018)
CONTIGUITY _{ij}	0.369 (1.418)
LANDLOCK _i	-0.756*** (0.118)
LANDLOCK _j	-0.756*** (0.118)
CONTIGUITY _{ij} × log(DISTANCE _{ij})	0.551* (0.297)
CONTIGUITY _{ij} × log(POPULATION _i)	-0.011 (0.142)
CONTIGUITY _{ij} × log(POPULATION _j)	-0.094 (0.125)
CONTIGUITY _{ij} × log(AREA _i)	-0.149 (0.143)
CONTIGUITY _{ij} × log(AREA _j)	-0.099 (0.136)
CONTIGUITY _{ij} × LANDLOCK _i	0.277 (0.208)
CONTIGUITY _{ij} × LANDLOCK _j	0.277 (0.208)
CONSTANT	-7.651*** (0.516)
Number of countries	133
Number of observations	9426
R-squared	0.368

Notes: Gravity equation underlying the Alcalá and Ciccone (2004) study. The dependent variable is the (log of) the sum of import and export between countries *i* and *j* divided by real GDP in country *i*. Variables are described in the main text. Asterisk *, **, *** denote significance at 1, 5, 10 % respectively. All standard errors (reported in parenthesis) are cluster robust. Following Frankel and Romer (1999), the coefficients on LANDLOCK_i and LANDLOCK_j and the coefficients on their interaction with CONTIGUITY_{ij} have been constrained to be equal.

Table 11: First Stage Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent variable	TRAVEL	REAL OPENNESS	INSTITU- TIONS	TRAVEL	REAL OPENNESS	INSTITU- TIONS	BTRAVEL	REAL OPENNESS	INSTITU- TIONS	BTRAVEL	REAL OPENNESS	INSTITU- TIONS
Fitted travel share	0.480 (0.370)	0.111 (0.222)	-0.062 (0.046)	0.558 (0.565)	0.196 (0.307)	0.001 (0.057)	0.918** (0.374)	0.090 (0.244)	-0.039 (0.048)	1.186** (0.576)	0.304 (0.323)	0.073 (0.053)
Fitted trade share	0.592 (0.520)	0.644** (0.294)	0.107* (0.059)	0.426 (0.845)	0.525 (0.445)	-0.004 (0.075)	0.110 (0.630)	0.672* (0.375)	0.066 (0.069)	-0.383 (0.966)	0.376 (0.514)	-0.104 (0.079)
ENGLISH LANGUAGE FRACTION	0.818 (0.540)	0.755** (0.296)	0.111* (0.062)	0.888 (0.588)	0.664** (0.288)	0.118* (0.071)	0.901 (0.548)	0.729** (0.327)	0.131* (0.072)	0.835 (0.526)	0.492 (0.308)	0.105 (0.070)
EUROPEAN LANGUAGE FRACTION	0.961** (0.366)	0.192 (0.203)	0.097** (0.044)	0.954*** (0.338)	0.366* (0.202)	0.144*** (0.048)	1.205** (0.565)	0.273 (0.347)	0.067 (0.068)	1.666*** (0.540)	0.894*** (0.327)	0.228*** (0.062)
LANDLOCK	0.134 (0.557)	0.351 (0.276)	0.038 (0.046)	0.067 (0.561)	0.445 (0.282)	0.009 (0.047)	0.990 (0.677)	0.471 (0.401)	0.006 (0.063)	1.051 (0.760)	0.675 (0.407)	-0.026 (0.063)
TROPICS	-0.754 (0.916)	0.455 (0.677)	0.205 (0.138)	-0.953 (1.083)	0.374 (0.706)	0.121 (0.144)	0.200 (1.275)	0.423 (0.847)	0.225 (0.147)	-0.282 (1.323)	0.212 (0.824)	0.159 (0.138)
LATITUDE	2.759*** (0.964)	1.139* (0.633)	0.855*** (0.129)	2.080 (1.763)	1.097 (1.022)	0.607*** (0.186)	3.092*** (1.136)	1.145 (0.800)	0.846*** (0.159)	1.581 (1.659)	0.927 (1.189)	0.674*** (0.188)
CONSTANT	-1.860 (1.949)	1.811 (1.117)	0.003 (0.230)	-0.919 (3.279)	2.135 (1.740)	0.489 (0.301)	-2.911 (2.236)	1.619 (1.482)	0.120 (0.274)	-0.576 (3.410)	2.388 (2.001)	0.700** (0.309)
CONTINENTS	NO	NO	NO	YES	YES	YES	NO	NO	NO	YES	YES	YES
F first stage	13.91	10.98	4.81				9.72	6.03	2.86			
Observations	68	68	68	68	68	68	62	62	62	62	62	62
R-squared	0.66	0.54	0.59	0.66	0.56	0.63	0.56	0.44	0.53	0.58	0.51	0.63

Notes: First Stage Regressions. Columns (1)-(3), (4)-(6), (7)-(9) and (10)-(12) are the first stage results associated with Columns (3), (4), (8) and (9) in Table 12, respectively. F first stage is the (robust) F-value associated with testing the null that the excluded instruments (fitted travel share, fitted trade share, English language fraction and European language fraction) are jointly equal to zero. All standard errors (reported in parenthesis) are heteroskedasticity robust. Asterisks ***, **, * indicate significance at a 1%, 5% and 10% level, respectively.

Table 12: LIML Regressions

Dependent variable	(1) log(A)	(2) log(A)	(3) log(A)	(4) log(A)	(5) log(y)	(6) log(A)	(7) log(A)	(8) log(A)	(9) log(A)	(10) log(y)
TRAVEL	-0.022 (0.390)	0.495** (0.231)	0.325** (0.142)	0.442 (0.618)	-0.354 (1.988)					
BTRAVEL						0.320* (0.188)	0.436*** (0.154)	0.368** (0.158)	0.653* (0.391)	0.513 (0.361)
REAL OPENNESS	0.388 (0.688)	-0.649 (0.420)	-0.278 (0.214)	-0.413 (0.934)	0.830 (3.066)	-0.087 (0.346)	-0.367 (0.307)	-0.413 (0.255)	-0.822 (0.650)	-0.593 (0.620)
INSTITUTIONS		1.493* (0.874)	1.397** (0.701)	1.101 (0.931)	2.902 (2.329)		0.981 (0.923)	-0.043 (1.135)	0.526 (0.978)	1.961** (0.863)
LANDLOCK			-0.479*** (0.176)	-0.320 (0.384)	-0.757 (1.074)			-0.868*** (0.208)	-0.761** (0.310)	-0.737*** (0.264)
TROPICS			-0.574* (0.333)	-0.352 (1.035)	-1.633 (3.271)			0.068 (0.537)	0.018 (0.476)	-0.207 (0.419)
LATITUDE			-1.252** (0.630)	-0.895 (0.602)	-0.478 (1.332)			0.738 (1.240)	0.421 (0.951)	0.674 (0.888)
CONSTANT	7.402*** (2.537)	10.174*** (1.637)	9.450*** (0.874)	9.868*** (3.488)	5.401 (11.544)	9.877*** (1.674)	10.545*** (1.399)	11.032*** (1.245)	12.441*** (3.139)	11.474*** (2.996)
CONTINENTS	NO	NO	NO	YES	YES	NO	NO	NO	YES	YES
OID (p-value)		0.26	0.76	0.43	0.47		0.33	0.44	0.57	0.44
Observations	72	71	68	68	68	67	66	62	62	62

Notes: Limited Information Maximum Likelihood estimates. The dependent variable in Columns (1)-(4) and (6)-(9) is the log of productivity in 1996; in Columns (5) and (10) it is the log of real GDP (PPP) per worker in 1996. The set of excluded instruments is the fitted travel share, the fitted trade share and the two European language variables. All variables are described in the main text. CONTINENTS refers to whether the model was estimated with continental dummies (Africa, America, Asia & Oceania and Europe) included. All standard errors (reported in parenthesis) are heteroskedasticity robust. Asterisks ***, **, * indicate significance at a 1%, 5% and 10% level, respectively.

Table 13: Dynamic Panel Regressions

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					log(y)			
log(y), lagged	0.749*** (0.049)	0.655*** (0.058)	0.667*** (0.062)	0.593*** (0.084)	0.728*** (0.046)	0.633*** (0.057)	0.649*** (0.067)	0.609*** (0.086)
TRAVEL	0.070*** (0.018)	0.074*** (0.024)	0.074*** (0.027)	0.086*** (0.029)	0.060*** (0.016)	0.073*** (0.025)	0.064** (0.028)	0.065** (0.032)
REAL OPENNESS					0.042* (0.023)	-0.004 (0.040)	-0.042 (0.051)	-0.053 (0.048)
Time dummies	YES	YES	YES	YES	YES	YES	YES	YES
Long-run elasticity (TRAVEL)	0.28	0.21	0.22	0.21	0.22	0.2	0.18	0.17
Confidence interval, 95% (TRAVEL)	[0.13, 0.43]	[0.07, 0.43]	[0.07, 0.38]	[0.07, 0.35]	[0.11, 0.33]	[0.08, 0.32]	[0.04, 0.32]	[0.03, 0.31]
Wald test (long-run elasticity = 0.2, p-value)	0.29	0.85	0.79	0.87	0.70	1.00	0.81	0.65
Number of instruments	224	103	80	55	329	142	107	69
Hansen J test (p-value)	1.00	0.32	0.15	0.14	1.00	0.69	0.21	0.11
Test for residual autocorrelation								
First order (p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Second order (p-value)	0.95	0.89	0.91	0.83	0.77	0.81	0.90	0.88
Observations	1195	1195	1195	1195	1118	1118	1118	1118
Countries	120	120	120	120	112	112	112	112

Notes: Arellano-Bond (DIF-GMM) estimates in all columns. The dependent variable is the log of real GDP (PPP) per worker. TRAVEL and, when included, REAL OPENNESS are treated as endogenous in all columns, i.e. levels lagged at least twice are valid instruments. Column (1) employs all available instruments, whereas Columns (2)-(4) sequentially lower the maximum number of lags to be used as instruments from 4 (column (2)) to 2 (column (4)). All standard errors (reported in parenthesis) are robust. Asterisks ***, **, * indicate significance at a 1%, 5% and 10% level, respectively.