Gravity, Market Potential, and Economic Development *

Keith Head[†] Thierry Mayer[‡]

October 26, 2010

Abstract

This paper provides evidence on the long-term impact of market potential on economic development. It derives from the New Economic Geography literature a structural estimation where the level of factors' income of a country is related to its proximity to large markets, referred to as "market potential." The empirical part evaluates this market potential for all countries in the world with available trade data over the 1965–2003 period and relates it to income per capita. Overall results show that market potential is a powerful driver of increases in income per capita.

1 Introduction

Krugman (1991) demonstrated that the same modeling "tricks" that he had used to explain international trade patterns could also contribute to explaining the tendency of economic activity to agglomerate. In particular the combination of monopolistic competition à la Dixit and Stiglitz with transport costs à la Samuelson yields a minimal model capable of analyzing endogenous regional concentration. A year later, Krugman suggested in his Ohlin lectures (published in book form in 1995) that these same tools could provide microeconomic underpinnings for a set of empirically useful, but theoretically fuzzy, relationships that Krugman referred to as "social physics." One of these relationships was the gravity equation for bilateral trade. The second was market potential. In a NBER working paper that year (1992), Krugman showed how to derive a relationship between wages and a construct that closely resembled the geographers' formulation of market potential. This market potential relationship was then successfully brought to data by Gordon Hanson (2005), while Redding and Venables (2004) used explicitly the structural link between gravity and market potential to guide estimation.

Our paper builds on this line of work and provides evidence on the long-term impact of market potential on economic development. Providing explanations for cross-country differences in development levels is perhaps one of the most important questions in economics. A large number of alternative frameworks have been proposed, and the literature has recently focused on whether

*We would like to thank Rodrigo Paillacar for his help and Souleymane Coulibaly for very fruitful discussions during the drafting of this paper.

[†]University of British Columbia and CEPR

[‡]Sciences-Po, CEPII, and CEPR.

physical geography, culture or institutions matter most in the long term economic performance of countries.[1] We focus here on a different explanation, where Krugman-inspired *economic* geography, synthesized and measured though the market potential index is key in economic development. The paper derives from the New Economic Geography literature a structural estimation where the level of factors' income of a country is related to the trade costs (for which distance is a key proxy) it faces to reach large markets. We also show that those predictions are actually more general than the original Dixit-Stiglitz-Krugman framework they originated from, and extend to many alternative trade models. The empirical part evaluates this market potential for all countries in the world with available trade data over the 1965–2003 period and relates it to income per capita. We also make the different constructions of market potential available to the profession on CEPII's website. Overall results show that market potential is a powerful driver of increases in income per capita.

This paper extends our knowledge on how market potential affects development in several dimensions. First, we show that the cross-sectional striking success of economic geography in predicting income per capita in Redding and Venables (2004) holds when considering panel data. This reinforces their finding strongly, and confirms other recent panel data results, mostly done on an intra-national basis. Second, the results are robust to an instrumentation strategy intended to capture omitted variable bias that would survive the introduction of country-level fixed effects. Third, we allow for a larger set of trade costs variables, notably border effects, colonial preferences and regional agreements.[2] All of them have a time-varying effect in our specification. In addition to these empirical contributions, by linking wage equations to gravity equations in a more general way, we build confidence in the broader applicability of the wage equation.

A limitation of the paper is that it does not consider any full-fledged alternative model of the determination of income per capita. The implicit null hypothesis is that incomes are determined by productivity, proxied here by schooling and a country fixed effect. Fingleton (2008) conducts a statistical comparison of market potential versus the neoclassical factor accumulation model. The key regression adds a term for population growth adjusted for investment rates. This variable is insignificant when controlling for market potential.

The remainder of the paper proceeds as follows: Section 2 lays out the theoretical underpinnings of the empirical wage equation, treating the Dixit-Stiglitz-Krugman assumptions used by Redding and Venables (2004) as a special case. Section 3 describes the data used, while Sections 4 and 5 present respectively econometric results for the gravity estimates that help build the market potential and the economic development regressions themselves. An online appendix shows how many different formulations of the gravity equation can be thought of as special cases of a general formulation.

2 Theory

Redding and Venables (2004) and Hanson (2005) were the first contributions to apply empirically the implications of the Krugman-type economic geography model in terms of wage differentials across US counties for Hanson, and across income per capita levels in the world for Redding and Venables. The relationship uncovered explains the level of factor incomes in a country i (wages

[1] Acemoglu et al. 2005 and Rodrik et al. 2004 provide nice summaries of the different theories in competition, arguing strongly in favor of the institutions' view.

[2] Bosker and Garretsen (2010) also use an augmented gravity specification for sample of sub-Saharan countries.

if labor is the only factor) by a weighted sum of expenditures of all countries in the world. The weights are bilateral trade costs from i to each of the destination countries for i's exports. The resulting term is labeled Market Access (MA) by Redding and Venables (2004), Market Potential by Hanson (2005) and Real Market Potential (RMP) by Head and Mayer (2004), the "real" aspect being explained below. Here we use the market potential terminology to avoid confusion with the WTO definition of market access as the "tariff and non-tariff measures, agreed by members for the entry of specific goods into their markets."[3]

The relationship between factor incomes and market potential is referred to as the *wage equation* by Fujita et al. (1999). The founding contributions use the Dixit-Stiglitz type of monopolistic competition combined with iceberg trade costs. One might argue that this is not the most relevant framework for developing economies, especially for their resource-oriented sectors. We show here that the wage equation prediction arises under more general conditions.

2.1 Gravity and the wage equation

The best-known derivation of the wage equation is based on the zero-profit condition for symmetric, monopolistically competitive firms.[4] Here we propose a new derivation of the wage equation based upon the gravity equation for bilateral trade flows. Gravity involves two important constraints: budget allocation for the importer and market-clearing for the exporter. Consider an exporter country i and an importer country j. Budget allocation divides total expenditure, X_j, of the importer j across the exporting countries with Π_{ij} denoting the proportion of income allocated to country i.[5] By definition, bilateral exports, X_{ij}, are given by

$$X_{ij} = \Pi_{ij} X_j, \tag{1}$$

where $\sum_i \Pi_{ij} = 1$ and $\sum_i X_{ij} = X_j$.

The important step to derive a gravity equation from (1) is to show that Π_{ij} can be expressed in the following multiplicatively separable form:

$$\Pi_{ij} = \frac{A_i \phi_{ij}}{\Phi_j}. \tag{2}$$

Loosely speaking, A_i represents "capabilities" of exporter i, $0 \leq \phi_{ij} \leq 1$ represents the ease of access of market j to exporters in i, and Φ_j measures the set of opportunities of consumers in j or, equivalently, the degree of competition in that market.

A wide range of different micro-foundations imply equation (2). These include Dixit-Stiglitz monopolistic competition, Anderson and van Wincoop's (2003) model based on national product differentiation, and comparative advantage models such as Eaton and Kortum (2002). Recent models incorporating firm heterogeneity such as Chaney (2008) also imply similar multiplicative relationships. All those models have their budget allocation rule imply a gravity equation for bilateral trade which takes a simple multiplicative form:

$$X_{ij} = A_i \times \phi_{ij} \times X_j / \Phi_j, \tag{3}$$

[3] http://www.wto.org/english/tratop_e/markacc_e/markacc_e.htm

[4] See Fujita et al. (1999, pp. 52–54).

[5] For readability, we suppress time subscripts until we reach the regression specification.

where $\Phi_j = \sum_h \phi_{hj} A_h$, with different definitions of A_i and ϕ_{ij} depending naturally on the specific structure of the model. In the online appendix we show the A_i and Φ_j corresponding to each of these derivations.

As a second accounting identity, referred to as market clearing, holds that the sum of i's shipments to all destinations—including itself—equals the total value of i's production, noted Q_i.

$$Q_i = \sum_j X_{ij} = A_i \sum_j \frac{\phi_{ij} X_j}{\Phi_j}. \tag{4}$$

If B_i is country i's trade balance, we have $Q_i \equiv X_i + B_i$. At the world level, $\sum_j B_j = 0$, and therefore production must be equal to expenditure, $Q = X$.

If we have data on both expenditures X_j and production, Q_i, then the market-clearing condition tells us something about the unobserved attribute of the exporter, A_i. To see this define $s_j^X = X_j/X = X_j/Q$ as country j's share of world expenditure. Next, define the following term:

$$\Phi_i^* = \sum_h \frac{\phi_{ih} s_h^X}{\Phi_h}. \tag{5}$$

This term is central in what follows. It is an index of market potential.[6] Relative access to individual markets is measured as ϕ_{ih}/Φ_h. Hence, Φ_i^* is an expenditure-weighted average of relative access.

Hence, using (5) and (4), market-clearing conditions yields a very simple relationship between the exporter's capabilities A_i, its share of production $s_i^Q \equiv Q_i/Q$ and its market potential index Φ_i^*:

$$A_i = s_i^Q (\Phi_i^*)^{-1}. \tag{6}$$

This relationship is very general since it relies only on the gravity assumptions, namely the multiplicative budget allocation rule and market clearing.

In the online supplemental materials we show that the A_i in all of the models that we use to derive gravity equations can be expressed as a power function of wages in the exporting country,

$$A_i = G_i w_i^{-\lambda}, \tag{7}$$

where G_i stands for how "good" country i is as a producer. It incorporates the number of firms (N_i) in CES monopolistic competition, the quality of products (b_i) in national product differentiation, and absolute advantages in all industries (T_i) in the Ricardian model. The λ is $\sigma - 1$, a demand parameter, in national production differentiation and CES monopolistic competition and θ, a distributional shape parameter, in the models featuring consumer or firm heterogeneity.

The next step is specify s_i^Q. To do so we impose balanced trade, $B_i = 0\ \forall i$, and follow Chaney (2008) in assuming that any profits in the economy go to a "global mutual fund" of which each worker holds w_j shares and the dividend per share is given by δ. This implies aggregate expenditure in country j of of $X_j = w_j L_j (1 + \delta)$. Combining these assumptions we have

$$s_i^Q = s_i^X = (X_j + 0)/X = w_i L_i (1 + \delta)/X. \tag{8}$$

[6] Redding and Venables (2004), Head and Mayer (2004) and Hanson (2005) develop very similar terms with one notable difference that Φ_i^* is defined in terms of expenditure shares rather that total expenditures.

Substituting equations 7 and 8 into 6 and solving for w_i we have the wage equation:

$$w_i = \left[\frac{G_i}{L_i(1+\delta)}\right]^{1/(\lambda+1)} \times (X\Phi_i^*)^{1/(\lambda+1)} \tag{9}$$

This equation implies a simple power function relationship between wages and market potential as long as G_i/L_i is determined by the parameters of the model. For empirical purposes we require that G_i/L_i not be affected by either wages or market potential in i. This is clearly the case in the CES monopolistic competition model, where G_i is proportional to N_i, as shown in the online supplement. Since all firms are the same size in that model, L_i is also proportional to N_i. Thus the term in square brackets is a constant. In other models G_i/L_i may differ across countries, suggesting the usefulness of panel data where country i fixed effects can be employed. Since several of the models make G_i a function of a country-wide productivity parameter, we think it makes sense to control for cross-country differences in productivity. We will do so by incorporating average years of education. We now turn to an empirical estimate of the term in parentheses, $X\Phi_i^*$, which we call real market potential.

2.2 Market Potential computation

Since $s_i^X = X_i/X$, we can re-express equation (5) as

$$\Phi_i^* = (1/X)\sum_h \phi_{ih}(X_h/\Phi_h),$$

Taking the log of the bilateral trade equation (3) yields

$$\ln X_{ij} = \ln A_i + \ln \phi_{ij} + \ln(X_j/\Phi_j) \tag{10}$$

Redding and Venables (2004) discovered that the last two terms in this equation are precisely what we need to calculate an estimate of market potential. The ϕ are estimated by specifying a vector of observed trade costs (distance, etc.) and the $\ln(X_j/\Phi_j)$ are estimated as fixed effects for each of the importing countries, denoted FE_j. Real market potential (RMP) can therefore be constructed as

$$\text{RMP}_i = X\widehat{\Phi_i^*} = \sum_h \widehat{\phi_{ih}}\exp(\widehat{\text{FE}_j}). \tag{11}$$

Taking logs of 9, adding time subscripts and an error term, and substituting in the expression for RMP, we have our estimating equation,

$$\ln w_{it} = \mu_i + \beta \ln \text{RMP}_{it} + \epsilon_{it}, \tag{12}$$

where μ_i is the country specific effect based on the bracket term in 9 and $\beta = 1/(\lambda+1) = 1/\sigma$ in CES monopolistic competition with symmetric firms and $\beta = 1/(\theta+1)$ in the models with heterogeneous firms or consumers.

We follow Redding and Venables (2004) in using income per capita in i as the proxy for the local price of immobile factors, w_i. RMP is therefore an element explaining income per capita of the country. An empirical issue with RMP is that it contains own income X_i, leading to endogeneity. This problem is all the more important given that internal trade costs are lower than international trade costs. A solution proposed by Redding and Venables (2004) is to replace RMP with "foreign market potential" (FMP), where

$$\mathrm{FMP}_i \equiv \sum_{h \neq i} \widehat{\phi_{ih}} \exp(\widehat{\mathrm{FE}_j})$$

This does not include own demand of country i. While removing the circularity inherent in regressing average incomes on RMP, neglecting internal demand violates the underlying theory. As discussed below, a better solution is to instrument for RMP with geographic centrality.

3 Data

The first stage of the empirical work, the fixed effect gravity estimation, requires bilateral trade flows over a long time period, obtained from IMF DOTS, and a vector of trade impediments, obtained from CEPII.[7] The second stage involves factor incomes on the left hand side, and productivity on the right hand side, combined with the first stage market potential estimate. Income per capita (the proxy for w) comes from the World Bank's World Development Indicators.[8] Average years of schooling (the proxy for productivity) come from Barro and Lee (2000).

RMP calculations require measurement of ϕ_{ii}, the freeness of a country's trade with itself. In addition to having shorter distance, self-trade has a preferential dimension, that has been widely documented in the border effect literature (see Anderson and van Wincoop, 2004 for a survey of the evidence). Redding and Venables (2004) deal with this by adjusting the distance coefficient, which they divide by two for self-trade in their preferred specification. Head and Mayer (2004) adopt a different approach by estimating those border effects in the first step. This method involves measuring self trade for all countries in the world over the period. At the industry level, this is fairly easy, one just has to take global production of an industry, and subtract total exports to obtain "exports" to self. For aggregate trade, this is a little bit more subtle, since one needs to subtract total exports from the value of production that is actually tradable in the country. We follow Wei's (1996) method here and consider the non-service part of a country's GDP to be its tradable part. In what follows, we present results using Head and Mayer's (2004) method handling border effects, but results using the Redding and Venables (2004) method are quite similar (the dataset provided online includes both methods).

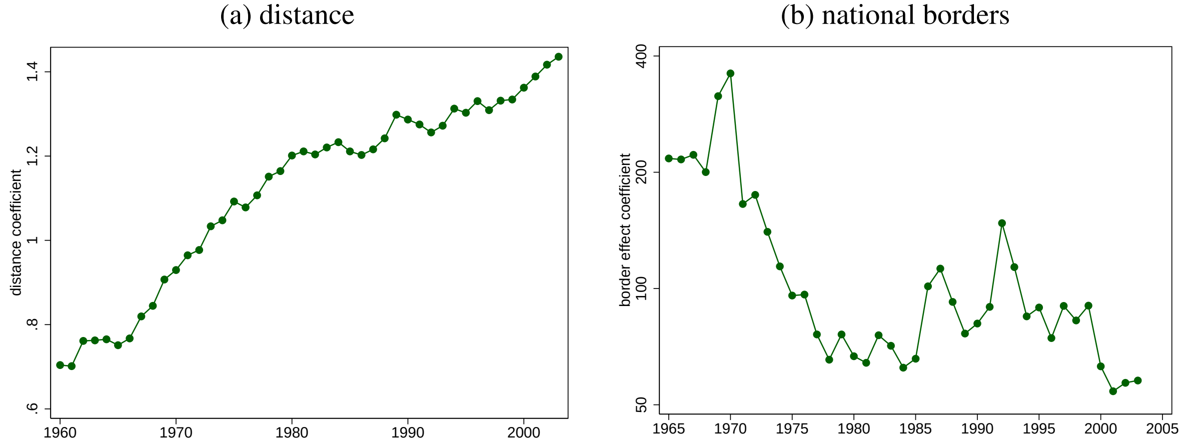
4 Gravity results

The first step estimates a gravity-type relationship where bilateral trade is regressed each year on a set of importer and exporter dummies and on a vector of trade impediments that is larger than the one used by Redding and Venables (2004), who focus on distance and contiguity only. The components of ϕ_{ij} include distance and contiguity, but also common language, colonial links, dummies for common membership of a regional trade agreement (RTA), a currency union and

[7] `http://www.cepii.fr/anglaisgraph/bdd/distances.htm`

[8] It is possible to go into deeper industry level detail, where the LHS variable becomes average wage in the industry. Head and Mayer (2006) do this for a European sample, Paillacar (2008) provide data and analysis for a much larger sample of countries and years.

Figure 1: The effects of distance and national borders on trade



WTO membership. Summarizing results from the estimation, the average fit is 0.73, with an average number of observations around 13,000. The average coefficients on trade costs are very much in line with existing findings. The average coefficient for distance is very close to -1 and common language, RTA and WTO membership have comparable mean effects around 0.4.

We present figures of the resulting coefficients over time. The most interesting and puzzling result is the increasing coefficient of distance on trade flows over time in panel (a) of figure 1. This trend is not isolated in the literature. Disdier and Head (2008) report such an evolution in their meta-analysis of distance coefficients in gravity equations. In what is perhaps the most comparable set of results in terms of estimation method, Redding and Schott (2003) show in their Table 1, that the coefficient on distance starts at -1.18 in 1970 and rises gradually to end at -1.49 in 1995 (they only include contiguity in the regression as a control for trade costs, which might explain the slightly lower impact of distance in their case in all years).

Panel (b) of figure 1 shows a more expected result, namely that the impact of national borders decreases over time. Note, however, that the estimated negative impact of crossing a national border on trade flows remains considerable in 2003, with a dividing factor around 50. This figure naturally aggregates very different situations, and is probably driven by developing countries that are usually estimated to have much larger border effects. In the online supplemental materials we also depict and discuss the evolution of the coefficients on the colony and RTA variables.

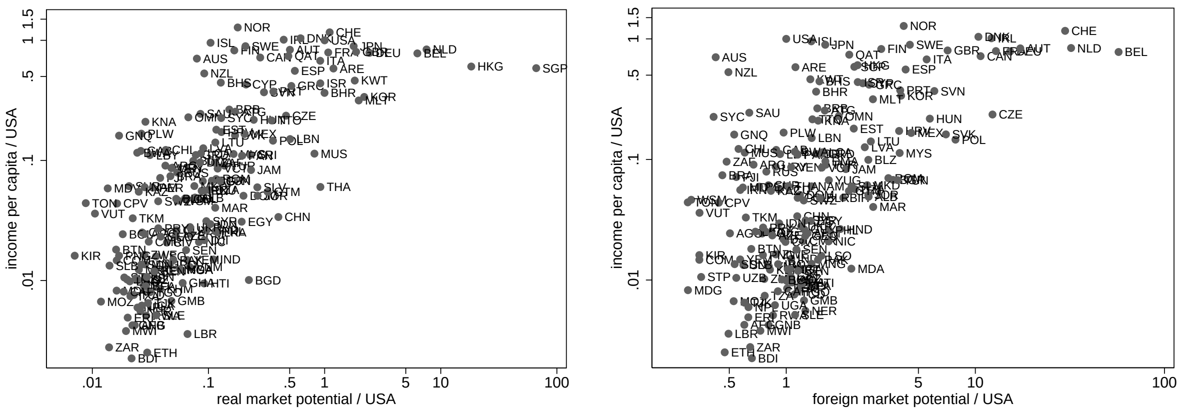
5 Market potential results

The above summarized gravity equations enable a computation of market potential indices, RMP_i and FMP_i, as described in section 2.2, for all countries with available trade data over the 1965–2003 period. This permits replication of the relationship between income per capita and market potential uncovered by Redding and Venables (2004). We start by reproducing their figure depicting GDP per capita in i graphed against RMP_i and FMP_i. We express both relative to the USA in 2003, in order to ease the reading of the axes on figure 2. There is a strong positive relationship between

Figure 2: Market Potential and development in 2003

(a) Real Market Potential

(b) Foreign Market Potential



market potential and income per capita.[9] Larger and/or more centrally located countries are much richer than countries characterized by a small local market and few or small neighbors. The cases of Belgium (BEL) and Netherlands (NLD) are interesting: with the exception of the very small territories of Hong Kong (HKG) and Singapore (SGP), Belgium and the Netherlands are the two top countries in terms of RMP. Panel (b) shows that this arises mainly from their advantageous location, as for Switzerland (CHE). Opposed to the case of those countries are the United States (USA) and Japan (JPN). Both are among the top RMP economies, but that comes almost entirely from their internal demand, since in terms of FMP, panel (b) shows a quite weak position. China (CHN) and Thailand (THA) are similar cases for the developing world. Both have a quite high RMP (which should warrant higher average wages, according to panel (a), but a fairly average FMP.

5.1 Baseline regression results

Table 1 presents our benchmark regression results, in which we use the Head and Mayer (2004) method of estimating market potential (which introduces border effects directly, rather than through a differential effect of internal distance used by Redding and Venables).[10] Column (1) replicates the benchmark Redding and Venables (2004) econometric specification, providing results for a cross section of 180 countries in 1995 (they use 101 countries in 1996, but the skills data we later use is only available every five year, including 1995). The coefficient is 0.80, roughly doubling the 0.395 they obtain. This difference in coefficients stems mainly from the different construction of the market potential variable, since we also obtain larger coefficients than Redding and Venables

[9] The correlation between log RMP and log average income is 0.72 in 2003. The correlation with log FMP is lower as would be expected since the own income term is removed but it is still substantial: 0.60.

[10] In unreported results (available in the working paper version) we run the exact same set of regressions of this section, using the Redding and Venables (2004) method. Results are very comparable, with a slightly lower fit in general, and smaller coefficients for market potential variables.

even when estimating the wage equation on the same set of countries.[11]

Column (2) pools over the whole set of years available for our countries, and column (3) presents results with country fixed effects. These are the first *within* estimates of the relationship between income and RMP.[12] The within results are particularly interesting. Market potential can potentially be correlated with a vast number of other variables relevant to the level and growth of income per capita. This is the rationale behind Table 2 of Redding and Venables (2004), that includes a large number of controls draw from the development literature. Those include primary resource endowments, other features of physical geography, but also measures of property rights protection, and a dummy if the country was under socialist rule between 1960 and 1985. Those controls offer variance that is predominantly or even exclusively cross-sectional. The use of panel data with country fixed effects permits to control for those and all other factors that are constant over time, and which can affect the level of income per capita. As expected, the coefficient on market potential drops but stays very significant and within a range comparable to the literature on this type of estimates.

The last three columns report coefficients using foreign market potential in place of the full market potential construction. Recall that this is a way to alleviate the endogeneity problem, but not a perfect one. Theory predicts own market size to affect the level of factors' income of a country, since those sales to domestic consumers often represent a large part of overall sales. Coefficients are larger for FMP than for the complete market potential, as was also the case in Redding and Venables (2004). Once again, the use of panel data reduces the estimated impact of market potential, but leaves it strongly positive and significant. With country fixed effects the elasticity with respect to RMP is about the same as for FMP.

The main lesson from Table 1 is therefore that the impact of market potential is robust to panel data estimation. This is the first important and comforting finding of our paper. The impact of economic geography (market access) on income per capita is not driven by some fixed omitted variable in the cross-sectional regression. The within impact is smaller than in the cross-sectional one, as expected, but remains economically large in magnitude.[13] Pushing further the inspection of the impact of market potential, one can naturally be worried that some *time varying factor* might be omitted from the regression. The first such factor of concern is of course the evolution of average skills in the population. Theory and dozens of empirical paper tells us that education should enter this equation, and might possibly have a relationship with market potential, for instance if the incentives to accumulate human capital are larger in large/central markets. Note that Redding and Venables (2004) do not control for skills although Redding and Schott (2003) show that the level of skills in a country is related to its market access. More recently, some related geography papers have included education levels as controls: Head and Mayer (2006) on a regional level basis, Fally et al. (2010) and Hering and Poncet (2010) at the individual level for Brazilian and Chinese workers respectively.[14]

[11] Our gravity equations have more countries and a larger set of trade cost controls.

[12] Fingleton (2008) estimates a wage equation with fixed effects using the Harris formulation of market potential (the sum of incomes divided by distance).

[13] Indeed, the inverse of the estimated coefficient implies a value of σ that exceeds one but lies below the normal range of estimates. The IV estimates presented in the final set of results yield smaller coefficients and therefore larger implied σs.

[14] Duranton and Monastiriotis (2002) for the UK, Combes, Duranton and Gobillon (2008) for France, and Mion and Naticchioni (2005) for Italy, had all already shown (in specifications less grounded in economic geography theory)

Table 2 includes the Barro and Lee (2000) measure of average years of schooling (among the 25+ population). The cross-sectional and pooled results of market potential in columns 1, 2, 4 and 5 are lowered as expected. The preferred within specification maintains a very significant and high coefficient on both the complete and foreign measures of market potential. Note also the very high coefficients on the skills variable in the non-within specifications. The lower values obtained when using the country fixed effects reinforce the attractiveness of those specifications: the estimates averaging around 0.10 are now more comparable to what has been found in the above quoted literature (Head and Mayer, 2006, Fally et al., 2010, and Hering and Poncet, 2010).

Table 1: Market Potential and GDP per capita

	Dependent Variable: ln GDP/cap					
	(1)	(2)	(3)	(4)	(5)	(6)
ln RMP	0.80^a	0.70^a	0.59^a			
	(0.06)	(0.05)	(0.02)			
ln FMP				0.88^a	0.88^a	0.58^a
				(0.11)	(0.10)	(0.02)
Time Frame	1995	1965–2003	1965–2003	1995	1965–2003	1965–2003
Country FE	No	No	Yes	No	No	Yes
Observations	180	6245	6245	180	6245	6245
R^2	0.521	0.547	0.748	0.280	0.318	0.711

Note: Standard errors in parentheses. Levels of statistical significance are signaled by a (1%), b (5%), and c (10%). Heteroskedasticity-robust standard errors clustered by country in columns (2), (3), (5) and (6). Those columns also include a full set of year dummies. Within R^2 reported in columns (3) and (6).

It is interesting to quantify a little bit more precisely those results, going further than statistical significance. Consider the following experiment: In 2003, take a country with a low RMP, say the Congo Democratic Republic, and one with a large RMP, say Thailand, which in 2003 has an RMP 66 times larger than CDR. Using the 0.37 estimate of column (2) in Table 2, raising the RMP of CDR to the one of Thailand is predicted to increase its GDP per capita by a factor of around 24, while the real ratio in 2003 is around 22. Part of this increase is in fact tautological since own GDP enters RMP as stated above. Another interesting experiment is to raise FMP of a country, which does not include own GDP. Still in 2003, we observe Brazil to be in the tenth percentile of the lowest FMP countries, while Mexico is ranked 18th in terms of FMP, among the top ten percent countries. Using the column (5) estimate, the model predicts that based on a 900% difference in FMP, Mexico should have a GDP per capita around five times higher than Brazil, the real factor being 2.24. Last, one wants to evaluate the size of the market potential impact based on within variance alone. Over the last ten years of our sample (1993–2003) the average growth of RMP is 111%, and the corresponding figure for FMP is 161%. Using estimates from columns (3) and (6), this corresponds to a predicted income per capita growth of 61% and 105% respectively. In addition to the very strong fit of the model, and the very high precision of market potential coefficients, the economic magnitude implied by the estimates is therefore quite large.

that geographic wage differentials are largely influenced by skill differences.

Table 2: Market Potential and GDP per capita, with skills control

	Dependent Variable: ln GDP/cap					
	(1)	(2)	(3)	(4)	(5)	(6)
Avg. years of schooling	0.37^a	0.29^a	0.08^b	0.42^a	0.36^a	0.12^a
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)
ln RMP	0.41^a	0.37^a	0.55^a			
	(0.06)	(0.05)	(0.04)			
ln FMP				0.42^a	0.39^a	0.65^a
				(0.09)	(0.06)	(0.06)
Time Frame	1995	1965–2003	1965–2003	1995	1965–2003	1965–2003
Country FE	No	No	Yes	No	No	Yes
Observations	108	866	866	108	866	866
R^2	0.809	0.791	0.804	0.773	0.747	0.792

Note: Standard errors in parentheses. Levels of statistical significance are signaled by a (1%), b (5%), and c (10%). Robust standard errors clustered by country in columns (2), (3), (5) and (6). Those columns also include a full set of year dummies. Within R^2 reported in columns (3) and (6).

5.2 Instrumented results

As stated above, substituting FMP to RMP helps to solve the endogeneity problem, since own income does not appear in the explanation of income per capita. However, it is a significant departure from the theory. Consider the case of the United States. While it has a much larger RMP than Canada and Mexico, it has a much lower FMP. If foreign demand was the only driver of factor incomes in the NEG model, Canada and Mexico should both be richer than the USA. On the contrary, the NEG model predicts that the United States should be richer than its two neighbors precisely because it has a large internal demand that makes it a more profitable location for firms. The same paradoxical prediction of FMP is very clearly appearing for Brazil. Hence FMP has nice features, but is clearly not ideal as a replacement or instrument for RMP. What is preferable is an instrument that does not use the income information altogether, but keeps the measures of trade costs, including trade costs to self. We seek an exogenous source of variance of RMP that would come from trade costs, in the cross-section and if possible in the time dimension as well.

Geographic centrality of i ($\sum_j d_{ij}^{-1}$) is a good candidate in terms of exogeneity since it depends only on the physical location of a country relative to the rest of the populated world. It was introduced by Head and Mayer (2006) and also used by Hering and Poncet (2010). Unfortunately, it does not vary over time. A related instrument that does vary over time is $\sum_j \phi_{ijt}$, that is the complete measure of trade costs, including time-varying memberships in free trade and common currency associations. It also varies over time because the first-step estimates of trade costs coefficients vary over time as we saw in Figure 1. The drawback of this variable is greater concern over endogeneity.

Table 3 reports results. The first stage F-test shows that the two proposed instruments are quite powerful determinants of RMP either in the cross-section or in the temporal dimensions. Column (4) is the most demanding, instrumenting while including the full sets of country and year dummies. The first stage regression exhibits an unreported coefficient of 0.74 on $\sum_j \phi_{ijt}$

explaining RMP in the pure within dimension, with a t-statistic of 7. The second stage result shows a smaller coefficients for log RMP and schooling. The implied estimated of σ is now three, bringing it closer to the consensus from the literature. Further the 0.1 return to schooling in column (4) is now in line with estimates based on individual data.

Table 3: Market Potential and GDP/cap with skills control and IV

	Dependent Variable: ln GDP/cap			
	(1)	(2)	(3)	(4)
ln RMP	0.40^a	0.40^a	0.30^b	0.35^a
	(0.12)	(0.10)	(0.12)	(0.10)
Average years of schooling	0.37^a	0.28^a	0.31^a	0.10^a
	(0.05)	(0.03)	(0.04)	(0.03)
Time Frame	1995	1965–2003	1965–2003	1965–2003
Country FE	No	No	No	Yes
IV	$\sum_j d_{ij}^{-1}$	$\sum_j d_{ij}^{-1}$	$\sum_j \phi_{ijt}$	$\sum_j \phi_{ijt}$
First stage F	31.83	23.21	12.65	50.83
Observations	108	866	866	855
R^2	0.809	0.791	0.789	0.797

Note: Standard errors in parentheses. Levels of statistical significance are signaled by a (1%), b (5%), and c (10%). Robust standard errors clustered by country in columns (2), (3) and (4). Those columns also include a full set of year dummies. Within R^2 reported in columns (4).

6 Conclusion

This paper provides evidence that access to markets, measured here as a theory-based index of market potential is an important factor in development. We generalize the theoretical motivation of Krugman (1992, 1995) and empirical implementation by Redding and Venables (2004) in many directions, and find very robust evidence that the economic geography of countries matter greatly in their income per capita trajectory. To illustrate, our results show that in 2003, bringing the market potential of the Congo Democratic Republic to the one of Thailand is predicted to increase its GDP per capita by a factor of around 24. The average growth of market potential due to neighbor countries between 1993 and 2003 in our sample is estimated to have raised income per capita by around 105%.

7 References

Acemoglu, D., S. Johnson, and J. Robinson, 2005, Institutions as the Fundamental Cause of Long-Run Growth, in Philippe Aghion and Stephen Durlauf, eds., *Handbook of Economic Growth* (North Holland).

Anderson, J.E. and E. van Wincoop, 2003, Gravity With Gravitas: A Solution to the Border Puzzle, *American Economic Review*, 93(1), 170–192.

Anderson, J.E. and E. van Wincoop, 2004, Trade Costs, *Journal of Economic Literature* Vol. XLII, 69–751.

Barro, R. J. and J.-W. Lee, 2000, "International Data on Educational Attainment: Updates and Implications," CID Working Paper No. 42.

Bosker, M. and H. Garretsen,2010, "Economic Geography and Economic Development in sub-Saharan Africa," mimeo.

Breinlich, H., 2006, The Spatial Income Structure in the European Union: What Role for Economic Geography?, *Journal of Economic Geography*, 6(5).

Chaney, T. 2008. Distorted Gravity: the Intensive and Extensive Margins of International Trade, American Economic Review 98(4): 1707-1721.

Combes P.-P., G. Duranton, and L. Gobillon, 2008, Spatial wage disparities: Sorting matters!, Journal of Urban Economics, forthcoming.

Disdier, Anne-Celia and Keith Head, 2008, The Puzzling Persistence of the Distance Effect on Bilateral Trade, *Review of Economics and Statistics*.

Duranton, G. and V. Monastiriotis, 2002, Mind the Gaps: The Evolution of Regional Earnings Inequalities in the U. K., 1982–1997, Journal of Regional Science 42(2), 219–256.

Eaton, J. and S. Kortum, 2002, Technology, Geography, and Trade, *Econometrica*, 70(5), 1741–1779.

Fally, T., R. Paillacar and C. Terra, 2010, Economic Geography and Wages in Brazil: Evidence from Micro-Data *Journal of Development Economics* 91(1), 155–168.

Fingleton, B., 2008, Competing models of global dynamics : evidence from panel models with spatially correlated error components, *Economic Modelling* 25, 542–558.

Fujita , M., P. Krugman and A. Venables, 1999, *The Spatial Economy: Cities, Regions, and International Trade*, (MIT Press, Cambridge).

Hanson, G., 2005, Market potential, increasing returns and geographic concentration *Journal of International Economics* 67(1): 1–24.

Head, K. and T. Mayer, 2004, Market Potential and the Location of Japanese Investment in the European Union, *Review of Economics and Statistics* 86(4), 959–972.

Head, K. and T. Mayer, 2006, Regional Wage and Employment Responses to Market Potential in the EU, *Regional Science and Urban Economics*, 36(5): 573–594

Hering, L. and S. Poncet, 2010, Market access impact on individual wages: evidence from China, *Review of Economics and Statistics*.

Krugman, 1991, Increasing returns and economic geography, *Journal of political economy* 99(3),

Krugman, 1992, A dynamic spatial model, NBER Working Paper #4219.

Krugman, 1995, *Development, Geography, and Economic Theory*. The MIT Press (Cambridge, MA).

Mion, G. and P. Naticchioni , 2005, Urbanization Externalities, Market Potential and Spatial Sorting of Skills and Firms, mimeo.

Paillacar, R., 2008, An empirical study of the changing world economic geography of manufacturing industries: 1980–2003, mimeo CEPII.

Redding, S. and P. Schott, 2003, Distance, Skill Deepening and Development: Will Peripheral Countries Ever Get Rich? Journal of Development Economics, 72(2), 515-41.

Redding, S. and A. Venables, 2004, Economic geography and International Inequality, Journal of International Economics 62(1), 53–82.

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(2): 131–165.