Journal of Economic Integration
Vol. 39, No. 2, June 2024, 253-295
https://doi.org/10.11130/jei.2024016

pISSN: 1225-651X eISSN: 1976-5525

Do Institutions Trump Everything Else?: The Essay of Rodrik et al. (2004) Revisited

Kamel Ghaddab[1,2+]

[1]*Sousse High Institute of Management, Tunisia*
[2]*Magellan, IAE Lyon 3, France*

Abstract In this paper, we aim to estimate the independent effect of trade openness on the long-term growth performance of nations. The empirical methodology is based on two separate steps. In the first step, we will introduce an alternative methodology to estimate a more precise instrument for trade openness, using an extended hierarchical-longitudinal method, thereby enhancing the framework developed by Frankel and Romer [1999]. In the second step, we will estimate the growth effect of trade openness within an automatic selection framework for the income equation of nations, while accounting for the effects of institutions and geography.

Keywords: trade openness, economic growth, institutions, geography, hierarchical linear models, automatic selection

JEL Classifications: HB

Received 10 December 2023, Revised 2 April 2024, Accepted 18 April 2024

I. Introduction

This paper delves into the longstanding controversy in economics surrounding the role of trade openness in driving economic growth, a debate dating back to mercantilist doctrine. Amid rival theories and empirical findings, researchers often face the challenge of subjective specification choices. Building on Hendry's approach in 2000, I adopt an objective approach to address this issue, guided by his insight on the elusive nature of "science" and the quest for objectivity in economic analysis.

The fact that no nation has developed solely through autarky, or that development was not solely a consequence of openness to trade, is sufficient to undermine the existence of a simple relationship between an economy's external strategy and its long-term performance. The quest for the search of the deep origins of development has made it more challenging to isolate, study, and predict the independent effect of one of the determinants of long-run growth. This

+**Corresponding Author:** Kamel Ghaddab
Ph.D., Sousse High Institute of Management, Tunisia,
Magellan, IAE Lyon 3, France
E-mail: kamel.ghaddab@isgs.u-sousse.tn

complexity underlies the questioning posed by Rodrik and Rosenzweig in the introduction to the Handbook of Economic Development in 2009, as they ask about the utility of current research findings in terms of their relevance to policymakers for devising various strategies for economic development.

The causality between trade openness and economic development is akin to a mystery or statistical alchemy, reminiscent of Keynes [1939], rather than a central scientific focus dominating economic literature since Adam Smith's era. Influential works in this field suggest an institutional and/or geographical fundamentalism, which argues against the existence of an independent effect of trade openness on national development. Instead, economic development is primarily attributed to a lengthy process of institutional evolution or inherent geographical factors, shaping the fate of developing nations in the long term.

Our research primarily focuses on empirical inquiries, aiming to reassess the empirical impact of trade openness on long-term growth compared to rival determinants like institutions and geography. Existing literature indicates the interdependence of these factors, prompting the question: Does trade openness independently influence long-term growth, or is its impact solely mediated by institutional and/or geographical factors?

Two assumptions are crucial for establishing this relationship. Firstly, there's the endogeneity of institutions concerning their connection to income levels or openness. Empirical studies pose the question: Are advanced countries wealthy because of high-quality institutions, or vice versa? As for the second assumption, studies within the framework of the New Institutional Economy have faced challenges due to the endogeneity problem between openness and institutions. Better institutions are certainly a prerequisite for trade expansion (Anderson and Marcouiller, 1999), and more openness to trade leads to higher-quality institutions (Wei, 2001). These intertwined links and feedbacks render simple empirical analyses of these interrelations highly suspect (Rodrik, 2003). The question must either be: Can openness be an independent determinant of income, or does its effect go through institutions? The second assumption arises due to the endogeneity between the trade openness of economies and their development levels. The question of instrumentation seems important concerning these two assumptions.

Our research primarily focuses on the intricate relationship between trade openness and growth. We will concentrate on the variable of trade volume for instrumentation. For the institutional variable, we will rely on the analyses by Acemoglu et al. [2001], Mauro [1995], and Hall and Jones [1999], whose instruments have proven statistically robust in representing institutional development.

In our methodological approach to instrumenting for trade openness, we aim to refine the estimation of the Frankel and Romer [1999] instrument (hereafter F&R99). Our objective is to separate the geographical component from countries' trade, assuming the latter as an exogenous factor in relation to economic wealth, thus tackling the endogeneity concern.

The research framework of this paper enables a comparative analysis of key essays on long-run growth, including F&R99, Rodrik et al. [2004], and Sachs [2003]. Due to competing assumptions about economic growth determinants, the chosen approach relies on econometric reduction theory, with recent extensions by Doornik [2009] and Hendry and Doornik [2014], which is optimal for addressing conflicting or complementary assumptions within a theoretical framework.

This paper will explore three main axes. Firstly, we'll revisit the econometric approach of F&R99 for trade openness instrument construction, employing a structural method of bilateral trade. Secondly, we'll refine our method for approximating trade openness, using a hierarchical-longitudinal approach to the gravity equation. Thirdly, we'll conduct empirical analysis on causal elements determining income dispersion in the long run, employing automatic model selection to test various assumptions on the causality between trade openness and economic growth.

II. Indicators of Trade Openness: A Survey

The aim of this section is to offer a succinct and critical overview of several indicators introduced in empirical economic literature, which serve as proxies for measuring the degree of trade openness among countries. To mitigate the inherent conceptual ambiguity in analyzing these indicators, we adopt the classification criterion utilized by Dowrick and Golley [2004], particularly in terms of distinguishing between indicators of liberalization based on revealed openness and those based on political openness.

A. Indicators of revealed openness

Trade openness is commonly approximated by the intensity of trade, typically expressed as the ratio of total trade to GDP, a widely accepted method because this indicator relies on precise measurement, supported by robust statistical data collection and standardization processes that typically face few controversies. However, this approximation method faces theoretical and quantitative challenges. Geographic factors, which distinguish natural openness from economic openness, can bias the trade-to-GDP ratio. High trade-to-GDP ratios do not necessarily indicate liberal trade policies. The observed volume of international trade alone cannot accurately reflect an economy's international orientation, especially concerning trade policies and distortions in international prices. Pritchett (1996, p. 309) suggests adjusting the measurement of trade openness to account for non-political determinants of trade intensity. It becomes necessary to separate the political dimension of revealed openness (trade distortion and international prices) from structural determinants of openness. Due to the diversity and complexity of sources of distortion of international prices, inferring the first dimension can only be done as a residual component after estimating the effect of structural factors on the intensity of trade. Two approaches can be described. The first follows the tradition of Chenery and Syrquin [1975, 1989], where the correction is based on the structural determinants of an economy, and this by using economic size elements (such as GDP per capita, total population, and land area), as well as other factors like transportation costs and natural resources. A similar correction logic was employed by F&R99 and Wei [2001], taking a gravity perspective. Here, the dependent variable is the bilateral volume of trade per pair of economies, which is structurally determined by the interaction of economic size determinants, transport costs, and other factors related to trade familiarity inertia, such as language similarity, culture, and institutional structure. The second empirical research approach follows the traditional study of Leamer [1988], focusing on a factorial determination of economies' trade volume. It involves estimating the trade equation derived from the Heckscher-Ohlin-Vanek model. For the small country case, trade in certain products is heavily influenced by resource supply, international prices, technologies, tastes, natural barriers, and artificial barriers. Thus, to isolate this last component, it is necessary either to ensure the constancy of other determinants or to estimate their effect directly (Leamer [1988], pp. 1-2).

B. Indicators of political openness

The second set of indicators focuses on the political stance of countries regarding international trade, particularly examining the impact of trade policy barriers on trade volume and structure. This category includes indicators that measure direct trade policy instruments, such as tariffs and non-tariff barriers, as well as indicators analyzing the inward or outward orientation of trade policy.

Direct trade policy indicators encompass various policies governing international exchanges and provide reliable insights into an economy's trade regime. They include average tariffs and the proportion of total trade affected by non-tariff barriers like quotas and import licenses, especially in underdeveloped economies.

However, these indicators face empirical and theoretical challenges. Empirically, data on tariff barriers are often estimated[1], which can introduce biases, particularly for very high tariffs where trade volume approaches zero, rendering estimated tariffs null. Similarly, measuring trade openness is questioned, with concerns raised about relying solely on direct trade policy approximations (Thirlwall and Pacheco-López [2008]).

Theoretical limitations also exist, as nominal tariffs may not accurately reflect allocation effects, necessitating consideration of effective protection at the sectorial level. Another issue is that non-tariff barriers, prevalent in developing countries, lack a theoretical metric like tariff

1) Actual rates applied by governments are not directly collected from official sources.

barriers, and they vary widely, making it difficult to gauge their impact compared to tariffs.

Various essays have been developed to find indicators of trade openness (Krueger [1972 and 1997], Krueger and Tuncer [1984], and Bhagwati [1978]), aiming to estimate protectionism's impact on development and approximate trade regime evolution over time. These essays often focus on anti-export bias to demonstrate the benefits of export promotion over Import Substitution Industrialization. This helps distinguish trade liberalization from economic liberalism, where an "open" economy is characterized by a trade regime that does not bias exports, even with significant government interventionism.

C. Indicators of price distortions

In a somewhat partial independence from indicators of political liberalization of international trade, a distinct category of indicators focuses on the indirect effects of trade policies, particularly on the variation of domestic and international prices. This line of empirical research has developed around the essays of Leamer [1988], Dollar [1992], and in a similar vein, Krueger [1998]. These indicators differ from those related to revealed openness and political openness in that they define trade openness based on price differences rather than observed volume. They capture the indirect effects of trade instruments on commodity prices.

The methodology for constructing a price distortion indicator varies between Dollar [1992] and Krueger [1998]. Dollar's approach involves two stages. Initially, using an international comparison of price levels, Dollar calculates a relative price index for each economy relative to a benchmark economy (the United States)[2]. Trade openness is approximated by this index, attributing price differences between economies to trade distortions. However, in the presence of non-tradable goods, this representation becomes complicated since even in the case of free trade, prices will differ. The second stage focuses on correcting these indices for factors explaining price differences of non-tradable goods, such as differences in factor composition.

Krueger [1998] introduces a conceptual distinction between trade openness and trade extraversion. Trade regime bias is expressed indirectly through international comparison of price levels. The relative significance of this indicator was not quantified, but the discussion suggests that trade openness tends to narrow the deviation given by the bias, moving towards an extraversion regime[3].

2) For each economy, this entails calculating a relative price index, denoted as $PR_i = 100{*}e{*}(P_i/P_{USA})$, where P_i represents the price level for country i and P_{USA} represents the price level for the United States. e is exchange rate. In a scenario where all goods are tradable, this relative price index would theoretically equal 100, indicating a situation of free trade.

3) Thus, according to Krueger, a trade regime aiming to extend quotas or replace quantitative barriers with tariffs would not be interpreted as an extroverted trade regime. The bias is expressed not in terms of direct trade policy instruments, but rather indirectly through international comparisons of price levels. Formally, it is defined as "the relationship between the ratio of prices of import-substitute goods compared to international prices and the ratio

D. Composite indicators of trade openness

A fourth category of trade openness indicators has emerged in empirical literature. This category aims for comprehensiveness by encompassing various dimensions of openness beyond narrow aspects like trade policy instruments, as reflected in the title of the paper by Sachs and Warner [1995], "Economic Reform and the Process of Global Integration." This essay utilizes multiple criteria to assess the trade openness of countries, including: 1) an average tariff rate exceeding 40%; 2) a coverage rate of non-tariff barriers exceeding 40%; 3) a black market premium at least 20% lower than the official market; 4) an official monopoly on major exports, and 5) a socialist economic system.

Countries meeting at least one of these criteria are categorized as "closed." Tariff and non-tariff barriers represent direct restrictions, while a black market premium on the exchange rate can have similar effects by acting as a trade restriction. The inclusion of the state monopoly on exports and a socialist regime dummy variable also reflects trade restrictions imposed by centralized planning in economies (Wacziarg and Welch [2008]).

Other indicators have been proposed to quantify the trade openness of countries. For instance, the World Bank's 1987 report classified developing countries based on their trade orientation from strongly extraverted to strongly autarkic based on various criteria[4]. Additionally, the Heritage Foundation's Index of Economic Freedom evaluates ten criteria, including trade openness measured through tariff protection, non-tariff barriers, and the extent of customs authorities' corruption[5].

III. Approach by the structural determinants: The essay of Frankel and Romer [1999]

In the essay of F&R99, the form adopted for the gravity equation is given by:

$$Log\ \frac{T_{ij}}{Y_i} = \beta_0 + \beta_1\ Log\ D_{ij} + \beta_2\ Log\ S_i + \beta_3\ Log\ S_j + e_{ij} \tag{1}$$

---

of prices of exportable goods compared to international prices" (p. 1521).

4) The 1987 report classified 41 developing countries over 1963-1985 into: 1) Strongly extraverted, minimal anti-export bias; 2) Moderately extraverted, weak domestic market bias; 3) Moderately autarkic, biased towards domestic market with high protection and overvalued exchange; and 4) Strongly autarkic, biased towards domestic production with high protection and direct control measures.

5) The index, ranging from 1 to 100, measures freedom across criteria like government size, corruption, business and trade freedom, taxes, monetary system, investment environment, financial activity, property rights, and employment. Trade openness within it is gauged by tariff protection, non-tariff barriers coverage, and customs corruption.

T_{ij}, is the observed volume of bilateral trade between countries, i and j. Y_i, is the aggregate production of country i. D_{ij}, is the distance between countries, i and j. S_i and S_j are the respective components relative to the size of countries, i and j. e_{ij}, is an error term. The two components of size, S_i and S_j, are given according to variables relative to each country. For economy i, for example, these variables are the total population, N_i, geographical area, A_i and the geographical landlookness L_i.

The substitution of the components S_i and S_j in the gravity equation enables us to derive the extended quantitative form of the gravity equation of trade used by the authors to estimate the exogenous geographical component of trade of countries:

$$\begin{aligned} Log\frac{T_{ij}}{Y_i} = \beta_0 + \beta_1 Log \sum D_{ij} + \beta_2 Log\ N_i + \beta_3 Log\ A_i + \beta_4 Log\ N_j + \beta_5 Log\ A_j \\ + \beta_6 Log\ (L_i + L_j) + \beta_7 B_{ij} + \beta_8 B_{ij} Log\ D_{ij} + \beta_9 B_{ij} Log\ N_i + \beta_{10} B_{ij} Log\ A_i \\ + \beta_{11} B_{ij} Log\ N_j + \beta_{12} B_{ij} Log\ A_j + \beta_{13} B_{ij}(L_i + L_j) + e_{ij} \end{aligned} \quad (2)$$

The estimation of the quantitative equation of trade will permit to derive the geographical component of trade for country i:

$$\hat{\tau}_i = \sum_{j \neq i} e^{\hat{\beta} X_{ij}} \quad (3)$$

$\hat{\beta}$, is the vector of the estimated coefficients. X_{ij} is the matrix of all variables used. τ_{ij} is the share of the bilateral trade of country i with country j in GDP of country i. B_{ij} is a dummy variable that takes 1 if countries i and j share a common border and 0 otherwise.

In the following sections, we will demonstrate how estimating the gravity equation using ordinary least squares, as done in the F&R99 paper, leads to an inaccurate estimation of the geographical component of trade. This estimation problem arises from two biases: the alpha inflation and the heterogeneity of parameters, also known as the shrinkage phenomenon. To illustrate these biases, we will conduct a comparative analysis of two forms of the gravity equation. The first corresponds to the form used in the paper under critique, while the second involves the mixed linear form proposed in our study.

In order to simplify the comparative analysis, we transform the general form of the gravity equation of F&R99 as follows:

$$Log\ \tau_{tp} = \beta_0 + \beta_1 x_{tp} + \gamma_p + \epsilon_{tp} \quad (4)$$

$$\xi_{tp} = \gamma_p + \epsilon_{tp} \tag{5}$$

This general model can be specified in two ways, depending on the assumption regarding the individual component of variance, γ_p denoted as Ψ. In the standard form of the gravity equation, this variance is assumed to be null, whereas in the mixed linear form of the equation, it is considered non-null.

The first bias of the standard form of the equation relates to the estimation of standard errors and the probabilities associated with individual coefficient tests. Suppose bilateral trade data between countries are observed at two levels. The first level corresponds to time units (t), during which the size variables are observed. The second level corresponds to pairs of countries (p), which serve as the units for the second level of estimation. For simplicity, let's assume a balanced case where we have the same number of level one units for every level two unit.

Let's denote $\widehat{se}(\hat{\beta}_{1,ols})$ as the estimated standard error of the coefficient β_1 for the standard form of the gravity equation, estimated by ordinary least squares in a standard regression model. Similarly, let's denote $\widehat{se}(\hat{\beta}_{1,ml})$ as the estimated standard error of the coefficient in the case of a specification with a random constant using the maximum likelihood method.

For the scenario of a pure between-cluster covariate where the within-cluster variance is zero, the estimated error of the Maximum Likelihood (ML) estimator for the random intercept form of the model is given by $\widehat{se}_B(\hat{\beta}_{1,ml})$. Rabe-Hesketh and Skrondal [2007, 2022] and Bryk and Raudenbush (2002) have demonstrated that in this case, the OLS standard error is underestimated $\left(\widehat{se}_B\left(\hat{\beta}_{1,ml}\right) \rangle\ \widehat{se}(\hat{\beta}_{1,ols})\right)$, leading to narrower confidence intervals and smaller p_values.[6)]

Conversely, for the scenario of a pure within-cluster covariate where the between-cluster variance is zero, the estimated error of the Maximum Likelihood (ML) estimator for the random intercept form of the model is given by $\widehat{se}_W(\hat{\beta}_{1,ml})$. In this case, the OLS standard error is overestimated $\left(\widehat{se}_W\left(\hat{\beta}_{1,ml}\right) \langle \widehat{se}(\hat{\beta}_{1,ols})\right)$[7)], resulting in wider confidence intervals and larger p-values[8)].

The intuitive conclusion drawn from this discussion about the consequences of misspecification problems (using standard regression modeling for clustered data) is that robust standard errors

6) In survey analysis, the design effect, or Kish factor, assesses the impact of clustering on the precision of survey data estimates, correcting for variance inflation due to the grouped sampling units. It's given by $DE = 1 + (n-1).\rho$, where n is the group size and ρ is the intra-class correlation. As ρ increases, variance inflates, reducing estimate precision.

7) See Rabe-Hesketh and Skrondal (2007, pp. 167-168) for formal details.

8) Refer to Problem 3-6 in Rabe-Hesketh and Skrondal (2022, pp. 291-292) for a detailed technical discussion on the implications of considering the hierarchical nature of data when estimating coefficients in panel data analysis.

for clustered data should be employed when using standard regression models for such data[9]. Employing a standard regression model when the random intercept model is true can result in incorrect standard errors and p-values, depending on whether the covariate is within-cluster or between-cluster.

Therefore, we assert that gravitational equation data display a natural hierarchical or clustered structure, with a longitudinal structure at level one and a clustered structure at a second level of observation. This assertion depends on the relevance of the mixed linear form specification for the gravitational equation (with constant and possibly random slopes) and the assumption on variance for the heterogeneity effect (γ_p). The intra-class correlation coefficient is used for this type of application.

The second bias is manifested by the heterogeneity of the estimated coefficients. Referring to the discussion by Aiken et al. [2003], three alternative approaches can be distinguished for testing the heterogeneity of the groups in an ordinary estimation form. The first approach is described as "disaggregated", as it completely disregards the hierarchical structure of the population under estimation. Formally, the effect of an independent variable on the dependent variable is represented by a single coefficient, assumed to be homogeneous for all observed units, and is referred to as the total regression coefficient by Aiken et al. [2003, p. 539]. This type of approach is susceptible to the alpha inflation bias discussed above[10].

The second approach considers the hierarchical structure of the population by analyzing the data at the group level. This method involves dividing the total population into distinct groups, with each group serving as the unit of analysis. Referred to as an aggregate analysis, this approach entails conducting separate estimations for each group, resulting in average data for each group. However, this approach has been criticized for its implicit assumption of generalization, which may lead to what Robinson (1950) described as the ecological fallacy. Additionally, the estimated coefficient in this approach measures the between-group effect, and in certain cases, there may be a sign reversal (Kreft and De Leeuw, 1998).

The third approach, often referred to in econometric literature as Least Squares Dummy Variables (LSDV) analysis, involves incorporating dummy variables for groups alongside the independent variables of the model during estimation. This approach allows for two specific forms of estimation.

The first form assumes that the slope (the effect of the explanatory variable on the dependent variable) is common across all groups, while the groups themselves exhibit heterogeneity in terms of the constant, i.e., the mean value of the dependent variable. In the presence of dummy variables, the estimated coefficient represents the coefficient of intra-class correlation.

9) Refer to the VCE (Variance Component Matrix for Estimates) option in Stata software for calculating robust standard errors in panel data analysis. For a detailed technical discussion, consult Rabe-Hesketh and Skrondal (2022).
10) See Cohen et al. (2003) for the problem of alpha inflation and overestimation of significance.

The second form introduces differentiation at the level of the slopes, estimating the interaction between group membership and the explanatory variable[11].

The hierarchical longitudinal approach to the gravity equation of trade appears to be more flexible and effective in handling the variability of trade volume. It accommodates both the longitudinal impact of size variables (such as gross domestic product and/or population) and the heterogeneous specificities of trade partners that remain constant over time. This approach helps address the issue of coefficient heterogeneity by treating these coefficients as random variables.

IV. An Hierarchical Longitudinal Approach of the Frankel and Romer [1999] Instrument

A. The modified gravity equation of trade

The gravity equation at level 1 of the estimation which takes the longitudinal nature of the data is given by:

$$L\tau_{tij} = \beta_{0ij} + \underbrace{\beta_{1ij}\left(LPOP_{tij} - \overline{LPOP}_{.ij}\right)}_{Size\ component\ of\ bilateral\ trade\ per\ pair} + \epsilon_{tij} \tag{6}$$

Where $\tau_{tij} \equiv T_{tij}/Gdp_{ti}$ represents the average share of bilateral trade for the pair of countries i and j, relative to the gross domestic product of country i in year t. $LPOP_{tij}$ is the logarithm of the product of the populations of countries i and j, while $\overline{LPOP}_{.ij}$ represents the average per-pair logarithm of the product of the total populations of the countries. ϵ_{tij}, denotes the error term[12].

$\hat{\beta}_{0ij}$, measures, in the case of per-pair centering of the variables at the first level, the volume of trade for the pairs when the size variable is equal to its respective average per pair. This component can be interpreted as the volume of bilateral trade for the pairs, adjusted for the effect of size. It's important to note that the parameter $\hat{\beta}_{1ij}$ should be interpreted cautiously. This is because, in the case of centering strategy for first-level variables, this parameter measures the effect of a variation in this variable compared to the average per pair on the estimated variation of bilateral trade volume.

11) Refer to Rabe-Hesketh and Skrondal (2007, pp. 114-124) for a detailed technical discussion on the implications of considering the hierarchical nature of the data on the estimation of the coefficients.

12) Data on bilateral trade are from the Direction of Trade: https://comtrade.un.org/data/ (date of access, 20/03/2022).

By assuming a random constant, the average trade volume per pair becomes the dependent variable at the second level of the regression. The unit of analysis in this case is the trade pairs defined from all countries in the sample. At this level, geographical variables are integrated to characterize the heterogeneity of pairs based on factors such as language similarity, shared colonial heritage, geographical proximity (including bilateral distance as a natural trade barrier), common borders, and geographical landlockedness.

The equation of the second level of estimation, related to the random intercept, is given by:

$$\beta_{0ij} = \underbrace{\gamma_{00} + \gamma_{01} Ldist_{ij} + \gamma_{02} Border_{ij} + \gamma_{03} Colony_{ij} + \gamma_{04} Comlang_{ij} + \gamma_{05} Landl_{ij} + \gamma_{06} Island_{ij}}_{\text{Geographic component of trade per pair}} + \mu_{00} \quad (7)$$

The variables,,,,, andrespectively represent the geographical distance between countries i and j, and dummy variables for common borders (1 or 0), colonial linkage (1 or 0), common language (1 or 0), landlockedness (2, 1, or 0), and a dummy variable indicating if one or both countries are islands (2, 1, or 0). The data for these variables are sourced from Rose [2017].

$$\hat{\beta}_{0ij} = \hat{\gamma}_{00} + \Sigma_k \hat{\gamma}_{0k} X_{kij} \quad (8)$$

The estimated value of the constant in the preceding equation measures the average volume of bilateral trade for a pair, after accounting for the specific characteristics of trade partners and adjusting for size effects. X_{ij} is a matrix of all explanatory variables for the average volume of trade for pair ij.

A second assumption is made for the second level of estimating the gravity equation. We assume that the pairs exhibit heterogeneity in terms of the coefficient associated with the size effect, represented by a random slope for the population variable. An additional equation is included in the second-level estimation, which corresponds to the slope:

$$\beta_{1ij} = \gamma_{10} + \mu_{10} \quad (9)$$

γ_{10}, is the slope of the total population (all pairs). μ_{10}, is the error term which corresponds to the stochastic form of the slope.

The absence of explanatory variables for the slopes means that the cross-level effect, i.e., the interaction between the variables of the first and second levels of the hierarchical structure of the gravity equation, is manifested only through the random component of the slope, μ_{10}

$\left(LPOP_{tij} - \overline{LPOP}_{.ij}\right)$.

The residual stochastic component of the gravity equation is given by:

$$\epsilon_{tij} + \mu_{00} + \mu_{10}\left(LPOP_{tij} - \overline{LPOP}_{.ij}\right) \tag{10}$$

ϵ_{tij} measures the traditional error term. In this hierarchical longitudinal approach, with random constants and slopes, two random components are added: μ_{00} , which is related to the difference between the constant specific to a pair of countries and the fixed constant of the total population of the pairs, and μ_{10}, which represents the difference between the slope of the variable (population) specific to a pair ij and that of the total sample population.

B. Quality of estimation of the modified instrument

In this paragraph, our aim is to consider the longitudinal hierarchical form of the gravity equation, which will enable us to derive the modified form of the instrument used by F&R99. The focus of our current research is less on the estimation of the coefficients and the fixed and random components of the equation, and more on the post-estimation assessment of error quality, the correlation between the observed and predicted components of bilateral trade, and the construction of the instrument[13]. This construction involves isolating the exogenous geographical component from the trade of the countries that constitute the pairs in the gravity equation.

The estimation of the hierarchical-longitudinal form of the gravity equation of bilateral trade involves two stages corresponding to the two levels of regression. The first level deals with the temporal component of the variability of bilateral trade among pairs, while the second level estimates the effect of invariant geographical variables on trade volume. This estimation allows for the consideration of specific constants and slopes for each trade pair, overcoming the bias of assuming coefficient homogeneity.

At the first level, Equation (6) is estimated, while at the second level, Equations (7) and (9), which respectively represent the random constant and the random slope of the size effect, are integrated. To conduct a comparative analysis with the approach used in F&R99, we incorporate a component of inter-class effect that explains the random slope using pair-specific variables. Equation (9) will take the following modified form:

$$\beta_{1ij} = \gamma_{10} + \gamma_{11}Ldist_{ij} + \gamma_{12}Border_{ij} + \gamma_{13}Colony_{ij} + \gamma_{14}Comlang_{ij} + \gamma_{15}Landl_{ij} + \gamma_{16}Island_{ij} + \mu_{10} \tag{11}$$

13) For space constraints, the estimation results for the first and second levels of the trade equation were not given in the text. These results are available and can be provided upon request.

Substituting equations (7) and (11) into the gravity equation (6) yields the two components, fixed and random, of the total variation of bilateral trade. Therefore, the hierarchical-longitudinal form is as follows:

$$\begin{aligned} L\tau_{tij} = \; & \gamma_{00} + \gamma_{01}Ldist_{ij} + \gamma_{02}Border_{ij} + \gamma_{03}Colony_{ij} + \gamma_{04}Comlang_{ij} + \gamma_{05}Landl_{ij} \\ & + \gamma_{06}Island_{ij} + \gamma_{10}\left(LPOP_{tij} - \overline{LPOP}_{.ij}\right) + \gamma_{11}Ldist_{ij}\left(LPOP_{tij} - \overline{LPOP}_{.ij}\right) \\ & + \gamma_{12}Border_{ij}\left(LPOP_{tij} - \overline{LPOP}_{.ij}\right) + \gamma_{13}Colony_{ij}\left(LPOP_{tij} - \overline{LPOP}_{.ij}\right) \\ & + \gamma_{14}Comlang_{ij}\left(LPOP_{tij} - \overline{LPOP}_{.ij}\right) + \gamma_{15}Landl_{ij}\left(LPOP_{tij} - \overline{LPOP}_{.ij}\right) \\ & + \gamma_{16}Island_{ij}\left(LPOP_{tij} - \overline{LPOP}_{.ij}\right) + \underbrace{\mu_{00} + \mu_{10}\left(LPOP_{tij} - \overline{LPOP}_{.ij}\right) + \epsilon_{tij}}_{Random\ component\ of\ bilateral\ trade} \end{aligned} \quad (12)$$

Our methodology will adopt two specifications. The first specification assumes only one random component, represented by the constant. The second specification assumes that all coefficients of the gravity equation are random, including the effect of size. These two specifications are considered nested, meaning the second is derived from the first.

Table 1 allows us to summarize various reliability criteria for the random effects considered in estimating the gravity equation. The variance component derived from the empty model (where the trade share is explained only by a random constant, without any independent variable) enables us to calculate the intra-class correlation. In this case, the intra-class correlation represents the proportion of variance in bilateral trade share between pairs and is estimated by substituting the estimated variance components for their respective parameters, resulting in $\hat{\rho} = \hat{\Psi}/(\hat{\Psi} + \hat{\sigma}^2) = 0.84$. Here, trade pairs vary in their intercepts with a variance component $\hat{\Psi}$, while within the pairs, the estimated residual standard deviation around the pair-specific regression lines is represented by σ. The value of ρ indicates that approximately 84% of the variance in bilateral trade share is explained by pair-specific characteristics.

In terms of reliability, as discussed in Bryk and Raudenbush (2002), the reliability of a random effect measures the ratio of the true score or parameter variance to the observed score or total variance of the sample mean. This reliability tends to approach 1 when the group means vary substantially across level-2 units (holding constant the sample size per group), and when the sample size is large.

We note that the two studied random components (constant and coefficient) are statistically reliable, with reliability values of 0.98 and 0.74, respectively[14]. This result suggests that trade

14) For example, in the simple group-mean model $Log\, \bar{\tau}_{.p} = \beta_{0p} + \bar{r}_{.p}$, $with\ \bar{r}_{.p} \sim N(0, V_p), V_p = \sigma^2/n_p$:. The level 2 equation is $\beta_{0p} = \gamma_{00} + u_{0p}$, $u_{0p} \sim N(0, \tau_{00})$. The reliability of the paramter β_{0p} is given by the indicator $\lambda_j = Var(\beta_{0p})/Var(Log\ \bar{\tau}_{.p})$, λ_j is the ratio of parameter variance to the sum of parameter variance and error variance. It is often multiplied by 100 to obtain a percentage. This coefficient cannot exceed 1. In the case of

partners are characterized by significant differentiation, which is dependent on specific variables that are time-independent. Additionally, the effect of size appears to vary between trade pairs, as evidenced by a heterogeneous correlation between the rate of bilateral trade and the population variable. This statistical finding is corroborated by the Chi-squared test associated with each component.

Table 1. *Statistical Reliability of the Random Effects of the Trade Equation*

	Empty Model	Random Intercept	Random Coefficients
	Random Reliability		
Constant	0.981	0.975	0.987
Slope	-	-	0.740
	Estimated s.d.		
Constant	3.472	2.926	2.938
Slope	-	-	3.604
Chi-square	Stat (p value)	Stat (p value)	Stat (p value)
Constant	4323861.05(0.00)	3130279.96(0.00)	4225433.71(0.00)
Slope	-	-	321025.51(0.00)
s.d. of level 1	1.493	2.483	1.277
	Quality of estimation		
Deviation(-2ll)	3169874.51	3135579.93	2961645.74

Number of observations: 854443. Number of trade pairs: 32692

Regarding the specifications of the gravity equation, the comparative analysis favors those that assume all coefficients as random variables. This conclusion is based on the comparison of the deviation criterion associated with the likelihood function's value (-2LL in the lower block in Table 1).

We can infer from these findings that the quality of the instrument derived by F&R99 is significantly compromised, primarily for two important reasons. Firstly, the authors of that work assumed that the slope related to size is fixed according to the pairs (and implicitly countries)[15]. Secondly, the limitations of the data used in this study, particularly those concerning trade volume, along with the assumption of coefficient homogeneity, led them to use the estimated parameters of countries selected for the construction of the predicted volume of trade for the rest of the sample.

It is important to highlight two additional points regarding the analysis of the quality of

a hierarchical model, specific coefficients are assigned to level 2 units. The HLM algorithm informs us about the mean of $\lambda_{\hat{\beta}_{qp}}$, p for level 2 units and q for subscript for coefficient.

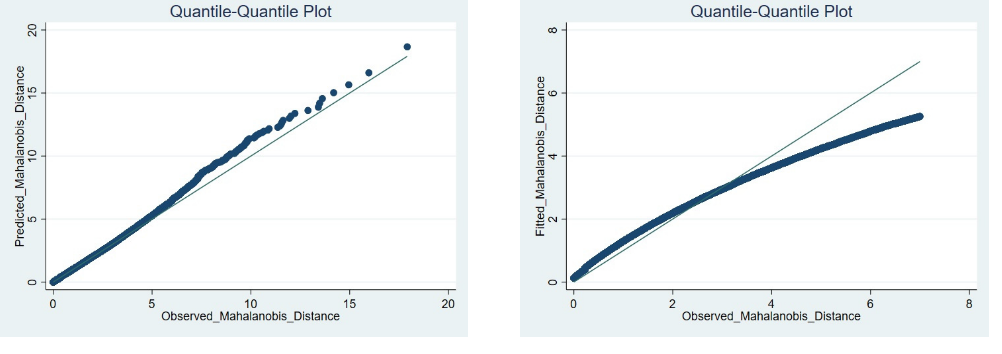
15) The use of the HLM algorithm permits the derivation and visualization of the large dispersion of slopes between pairs. We observe a great diversity in the amplitude of these slopes, and even an inversion of the sign for the slope of the size variable in some cases.

the instrument proposed in our work. Firstly, the normality of the errors derived from the estimation of the gravity equation must be carefully considered, especially in hierarchical models where the residual component of the regression can have multiple origins. The errors at the first level, ϵ_{tij}, represent deviations from the actual values predicted by the first-level variables. Regarding the errors at the second level, they consist of a complex stochastic component $\mu_{00} + \mu_{10}\left(LPOP_{tij} - \overline{LPOP}_{.ij}\right)$. However, analyzing the normality of errors at the higher level (pairs) is more complex. It involves testing the multivariate normality of all the random effects included in the mixed linear specification. This type of analysis follows the approach developed by Mahalanobis [1936] for assessing the robustness of multivariate models.

As mixed linear models can serve as alternatives to multivariate models of normality, the analysis tools developed by Mahalanobis [1936] can be applied (Heritier et al. [2009], pp. 110-111). The basic idea is to identify outlier points whose values significantly differ from the rest of the data. The Mahalanobis distance represents the squared distance of a unit from the center of a distribution with dimension v (the number of random effects studied). This distance is then compared to the estimated value of the distance, which is the difference between the empirical Bayes estimates, $\hat{\beta}_{qij}^{EB}$ [16], and its estimated value, $\gamma_{q0} + \Sigma_s \gamma_{qs} X_{sij}$.[17].

A QQ plot allows us to directly examine the dispersion of the observed and predicted distances, enabling the study of the multivariate normality of the errors at the second level of the regression. Specifically, the convergence of this curve towards the 45° line indicates normality.

Figure 1. Normality of levels 2 residuals (Random Intercept/Coefficient Models)



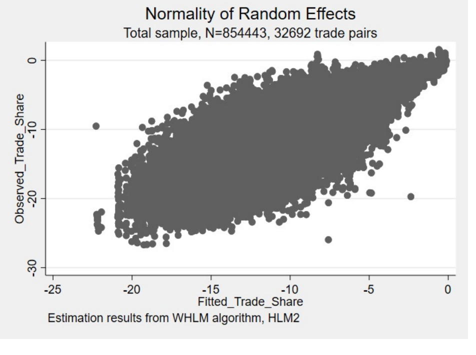
16) For the simple group-mean model $Log\ \bar{\tau}_{.p} = \beta_{0p} + \bar{r}_{.p}$, $Log\ \bar{\tau}_{.p}$ is an unbiased estimator of β_{0p} with variance V_j, γ_{00} could also be considered as a common estimator for each β_{0p}. A Bayes estimator is given by β^{*}_{0p}. It's an optimal combination $\lambda_p . Log\ \bar{\tau}_{.p} + (1-\lambda_p) . \gamma_{00}$, qualified as empirical Bayes estimates (See Bryk and Raudenbush (2002) for discussion).

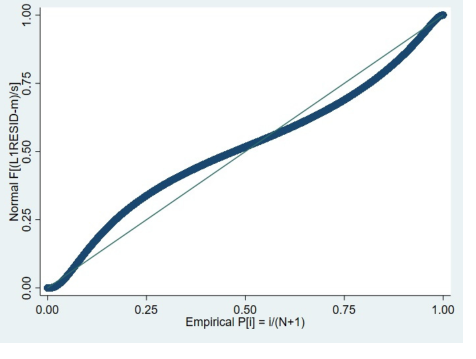
17) q and s are subscripts for coefficients and level 2 predictors. See Bryk and Raudenbush (2002, pp. 46-47), and Raudenbush et al. (2004, Chapter 1) for more conceptual and formal details.

The Figure 1 enables us to assess this type of normality in the context of the hierarchical specification of the estimated gravity equation. It serves as a test for the normality of the errors at the second level of the regression. We observe that the same graph also facilitates the detection of outlier points (pairs of countries), the precise identification of which poses significant technical challenges in this type of application. Furthermore, as discussed in Heritier et al. [2009], the reliability of this tool depends on the size of each unit (i.e., the number of observations per pair). Achieving a high level of reliability in our case of an unbalanced panel is not easily achievable.

In the same vein, Figure 2 presents a diagnostic analysis of level 1 residuals. On the left side of the graph, we observe a strong correlation between the predicted and actual components of the trade ratio. This correlation indicates that the trade equation was estimated with high precision. On the right side, the graph demonstrates that the normality of residuals at level 1 for the trade equation cannot be infirmed, allowing us to interpret and utilize these estimates accurately.

Figure 2. Analysis of level 1 residuals (32692 pairs)





In Table 2 presented below, we summarize the estimation results of the observed trade ratio using the modified F&R99 instrument and two size variables, population and geographical area, as explanatory variables. To conduct a comparative analysis with the estimates provided by the authors, three specifications were estimated. These correspond to the explanatory power of the instrument taken separately, the size variables alone, and the combination of all variables. From these results, several observations can be made. First, similar to the analysis at the level of trade pairs, examining the instrument using aggregate country-level data indicates a statistically significant explanatory power, accounting for 52% of the observed trade ratio dispersion in our sample countries. Secondly, the size variables (population and land area) show expected correlations and statistical precision in their impact on the observed trade share. Lastly, when all three variables are used simultaneously, the correlations and estimation precision of these variables remain consistent in explaining the observed trade ratio dispersion. In the next section, we will employ this modified estimate of the F&R99 instrument in the income equation of the benchmark empirical study by Rodrik et al. [2004].

Table 2. *Correlation between Observed and Constructed Trade Shares*

Specification	(1)		(2)		(3)	
Dep. Var.: Trade share	*Coef.*	*se*	*Coef.*	*se*	*Coef.*	*se*
Constructed trade share	-	-	0.678*	0.050	0.538*	0.042
Size var.: Population	-0.122*	0.026	-	-	-0.082*	0.019
Size var.: Land area	-0.045*	0.021	-	-	-0.034**	0.015
Intercept	2.156	0.273	0.234	0.050	1.834	0.196
Adj. R^2	0.40		0.52		0.70	
N	166		166		166	

*And** denote statistical precision at 1% and 5% risk threshold.

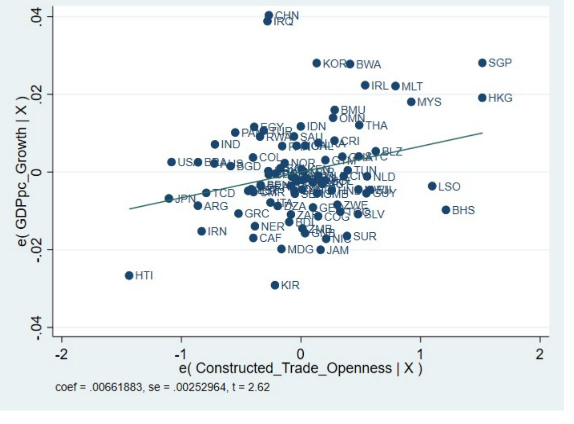
V. Trade Openness and Economic growth: In Search of Deep Determinants

In this section, we aim to address the central question of our empirical research, which concerns the examination of the causal relationship between trade openness and long-term economic growth. Our methodology involves revisiting the empirical literature that delves into the impact of trade openness on long-term income. We utilize our modified version of the instrument for estimating the observed trade volume of countries, as derived in the previous section. Our objective is to conduct a comparative analysis of our estimation results with those of the studies by F&R99 and Rodrik et al. [2004].

A. Trade and economic growth: The frankel and romer approach revisited

In this paragraph, we will assess the validity of the instrument constructed earlier by examining its correlation with predicted trade share (derived from the gravity equation) and long-term income. This analysis will be conducted in comparison to the estimation results obtained by F&R99. Initially, we will evaluate the explanatory power of the predicted trade ratio within a quantitative equation of nations' growth. The variables included in this estimation are the initial level of gross domestic product per capita, the investment rate, and population growth. After accounting for the traditional convergence effect, population growth, and the investment rate, Figure 3 illustrates a statistically significant association between the constructed trade ratio and average growth over the period 1970-2020. Although this correlation is estimated with 1% accuracy, its magnitude is relatively modest. As extensively discussed in Rodrik [2003], this modesty can be attributed to the simultaneous presence of "proximate" and "deep" determinants of economic growth. Proximate variables are endogenous as they both determine and are fully determined by economic growth, such as physical and human capital accumulation and total productivity. On the other hand, deep determinants are either partially endogenous or entirely exogenous. Among these, geography stands out as the only exogenous factor in this categorization. Trade and institutions are evidently endogenous and co-evolve with economic performance. Nevertheless, it is beneficial to conceptualize them as deep causal factors to the extent that they are not solely determined by incomes per se (Rodrik, 2003). Trade, for instance, is substantially influenced by a country's deliberate economic policies, while institutional development is at least partly shaped by a nation's political choices.

Figure 3. Correlation between constructed trade share and long-run growth



The same empirical modeling approach was undertaken, following the methodology employed by F&R99 and Rodrik et al. [2004], utilizing the real per capita GDP level in 2020 as a proxy for long-term national performance. The explanatory variables included population and geographical area. Three specifications of the income equation were estimated for a sample of 176 countries, with observations averaged over the period 1970-2020.

The benchmark specification considered the international income dispersion in the long run, utilizing the observed share of trade in GDP along with the two respective size variables, population, and geographical area. The two alternative specifications involved using the constructed share of trade as an instrument for the observed trade ratio. The instruments employed were those of F&R99 and a modified version derived from our hierarchical longitudinal approach.

Table 3 summarizes the estimation results for these three specifications. Consistent with the reference works, the included variables exhibit correlations of theoretically predicted signs and are statistically significant. Compared to the results of ordinary least squares (OLS) estimation, the signs of the respective correlations with the variables used remain unchanged, except for the physical area variable, which changes in the third specification. The correlation of this variable with income level is not precisely estimated in the specifications using instrumental variables.

Table 3. *Estimation of the Income Equation (OLS Vs. Instrumental Variables)*

Model	Base Model		F&R99		$F\&R99_M$	
Method	OLS		IV		IV	
Dep. Var.: $Lrgdppc_{2020}$	*Coef.*	*se*	*Coef.*	*se*	*Coef.*	*se*
Trade share in GDP	0.690*	0.157	0.892	0.692	1.349*	0.291
Size var.: Population	0.172*	0.068	0.261*	0.110	0.281*	0.077
Size var.: Land area	-0.129*	0.052	-0.093	0.076	-0.075	0.057
Intercept	7.345	0.993	6.654	1.737	6.273	0.936
N	176		130		176	
Fisher stat (*p_value*)	9.11(0.00)		-		-	
Wald stat (*p_value*)	-		1.41(0.237)		21.76(0.00)	
Minimum eigenvalue	-		17.197		173.282	

*denotes statistical precision of the coefficient at 1% risk threshold.

Regarding the explanatory power of the trade openness indicator, it is estimated with statistical precision, indicating that a 1% increase in observed trade generates nearly a 0.70% increase in long-run income per capita. Contrary to traditional expectations, the endogeneity bias of the trade ratio variable leads to an overestimation of its correlation with per capita income. The amplitude of the instrument used for this variable is significantly revealed to be more important. This increase in amplitude can be interpreted as the attenuation of the bias caused by measurement error on this variable overcompensating for the opposite bias of causality, and thus endogeneity, likely resulting in an overestimation of the coefficient in the OLS specification compared to that using instrumental variables (Rodrik et al., 2004).

The estimation results using instrumental variables indicate that after accounting for the effect of the two size variables, the impact of openness appears to be significant. Specifically, it is estimated that a 1% increase in the constructed trade ratio leads to a respective increase of 0.89% and 1.35% in long-run per capita income, using the instruments of F&R99 and its modified version derived from the hierarchical approach. These estimates are statistically significant, particularly for the modified version of the instrument.

For the explanatory power of the estimation specifications, the Fisher and Wald statistics indicate significant statistical reliability for our modified instrument. The last row of Table 3 contains the minimum eigenvalue for each of the two instruments used. This value allows us to test, following the method of Stock and Yogo [2005], the reliability of the instruments used for a variable considered theoretically endogenous[18]. Formally, the minimum eigenvalue corresponds to the Fisher statistic in the case of only one endogenous variable. In comparative terms, we can conclude that the modified version of the F&R99 instrument proved to be more reliable than the original estimates constructed by the authors.

A comparative analysis of the statistical reliability of the two instruments is presented in Table 4. The first row of this table provides the values of the partial coefficient of determination for these two instruments. This coefficient measures the correlation between the observed trade ratio, considered as the endogenous variable, and the instrument after adjusting for the effects of the exogenous variables (population and geographical area). It is calculated from the OLS estimation of the residuals of the endogenous variable regressed on the residuals of the exogenous variable on the instrument. It is approximately 12% for the $F\&R_{99}$ instrument and 52% for the modified F&R99M instrument. The difference indicates the advantage of the latter instrument's estimates.

Table 4. *Comparative Analysis of the Instruments*

	F&R99		$F\&R99_M$	
Criterion				
Shea's Partial R^2	0.120		0.516	
Test for instrument: H_0: instrument=0				
	Stat.	*p_value*	*Stat.*	*p_value*
Robust F(1, 164)	17.197	0.00	173.282	0.00
2SLS size of nominal 5%	16.38		16.38	
Test for redundant instrument				
Chi-square	12.364	0.00	71.863	0.00

In the same table, the results of the weak instrument test are provided. The test is based on the Fisher statistic, which exceeds the threshold value of 10, as recommended by Staiger and Stock (1997) for a single endogenous variable, for both instruments. Another test is conducted specifically for the modified version of the instrument. This test examines whether the modified instrument is redundant in the income equation. The chi-square statistic allows us to reject the null hypothesis that the modified instrument is redundant, indicating that it has explanatory power for the endogenous variable. The same test was conducted for the F&R99 version as an additional instrument alongside our modified estimates, F&R99M.

18) See Stock and Yogo (2005, p. 84) and Cameron and Trivedi (2009, pp. 190-192) for econometric details on the reliability of the instruments.

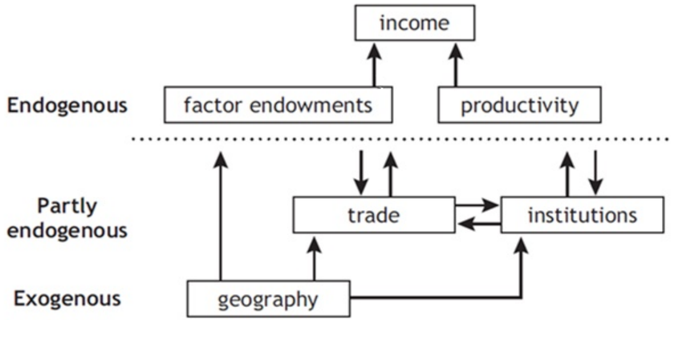
B. Trade openness, institutions and geography: Is it possible to overcome the causality tangle?

In this paragraph, we will assess the explanatory power of trade openness within a framework that considers deep determinants, including institutions and geography. Previously, we discussed Rodrik's (2003) distinction between "deep" and "proximate" determinants of a nation's economic growth, based on their endogenous or exogenous characteristics. Apart from variables related to the accumulation of productive factors, which are largely endogenous to growth, we identify three deep determinants of economic growth. These determinants are termed "deep" because they are partially endogenous (semi-endogenous), influencing and being influenced by economic growth. The search for reliable instruments for these variables in the literature is motivated by their nature. Only the geographical determinant possesses a completely exogenous nature, as its causality is unidirectional with respect to long-run growth. It's worth noting that a portion of this geographical dimension of development is semi-endogenous, referred to as "second nature geography" by Gallup et al. (1999). This concept relates to spatial interdependencies among economic agents that impact the state and accumulation rate of productive factors within an economy.

The significance of this empirical approach lies in testing the comparative explanatory power of each variable as a determinant of observed differences in the development of nations. Our focus is on the trade determinant, which has been the subject of both theoretical and empirical controversy. Therefore, we aim to determine whether trade openness has a statistically significant effect on long-run growth in the presence of institutional and geographical variables.

Our methodology will be based on Figure 4, reproduced from Rodrik [2003, p. 5], which synthesizes the theoretical causal pathways linking economic growth to its fundamental determinants. Using our modified version of the F&R99 instrument for trade openness and commonly used instruments for institutions, we will develop an empirical modeling strategy, comparing it to the approach taken by Rodrik et al. [2004].

Figure 4. All of growth economics in one page



On the econometric level, we will use three families of variables as statistical proxies for the three determinants of development: trade openness, institutions, and geography. Various specifications will be estimated based on the chosen approximations for these variables. Given the potential endogeneity of some variables, we will employ instrumental variables. This method involves referencing economic theoretical foundations to determine the exogenous versus endogenous status of variables. The remainder of the work will proceed in two stages. Firstly, we will estimate each endogenous variable using all exogenous variables. Then, in the second stage, we will focus on the independent effects of the variables used in the income equation for each of the studied specifications.

In this empirical modeling, the statistical variables and their approximations are described as follows. As the dependent variable, we will use the real level of per capita gross domestic product for the year 2020, in constant prices of 2015 ($Lrgdppc_{2020}$, Source: WDI 2022), as an approximation for long-term economic growth. In such empirical studies, we use the per capita income at the end of the period rather than the average growth rate as the dependent variable, allowing for analysis of economic performance over the very long term.

As a statistical approximation for the degree of trade openness, we will use the share of total trade in gross domestic product (Open, source: WDI2022). This variable is considered endogenous with respect to long-run growth. Therefore, it will be replaced by two instrumental variables: the Frankel & Romer instrument (labeled F&R99, Source: F&R99), and the modified version derived from our hierarchical-longitudinal form of the gravity equation ($F\&R99_M$).

The statistical approximations for the geographical determinant include: a) Geographical landlockedness (Landlock, Source: Gallup et al. [1999]); b) Proportion of the population living within 100km of the coast (POP_{100km}, Source: Tobler et al. [1995]); c) Proportion of geographical area located within 100km of the coast or a river ($Land_C$ and $Land_{CR}$. Source: Arc World Database Supplement and Arc Atlas Database 1996); d) Index of epidemiologic prevalence of malaria (Mal_{94}, Source: Gallup et al. [1999]). This index is calculated based on the geographical area affected by malaria and the rate of plasmodium transmission.; and e) Proportion of the area within the tropical zone (Tropic, Source: Arc World Supplement Database 1996).

We distinguish various dimensions of institutional determinants proposed in economic literature to address the endogeneity problem of institutions concerning income levels. The actual value of the institutional variable is approximated by the rule of law (Rule, Source: WGI 2022), which captures the formal dimension of institutions defined in New Institutional Economics literature. This dimension reflects perceptions of how well agents trust and adhere to societal rules, including contract enforcement, property rights, law enforcement, and the judicial system, as well as the likelihood of crime and violence (Kaufmann et al., 2013, pp. 4-5).

Another indicator, "Voice & Accountability" (Voice), introduced by Kaufmann et al. [2003], measures the extent to which citizens can participate in government selection and includes information on media independence, crucial for overseeing decision-makers' accountability.

The third indicator, Effectiveness, reflects bureaucratic quality, defined by public service and civilian independence from political pressure, as well as administrative credibility in policy design and control (Kaufmann et al. [2005], Governance Matters 4).

We also use components from the International Country Risk Group (ICRG), such as Bureaucracy, which assesses bureaucratic strength and expertise in maintaining policy continuity during government changes, thereby minimizing policy revisions and service interruptions.

Table 5. *Partial Correlation between Variables and Instruments*

Institutional variables			
Variable/instrument	Settler	English	Distance
Rule	-0.537	0.298	0.672
Voice	-0.400	0.370	0.561
Effectiveness	-0.610	0.296	0.672
Bureaucracy	-0.627	0.356	0.675
Trade variable			
Variable/instrument	F&R99		$F\&R99_M$
Open	0.689		0.730

All these approximations of the institutional determinant are considered to be endogenous in relation with income, and they must consequently be replaced by an instrument, which is correlated with these various dimensions of the institutional variable, and independent of the error. We will directly use the instruments developed in Acemoglu et al. [2001] and Hall and Jones [1999]. The first instrument is the settler mortality of the European colonists in a given country (labeled Settler, Source: Acemoglu et al. [2001]). In this case, the underlying assumption is that the expected rate of mortality of the colonists determines the colonial strategy regarding the construction of the institutional structures. The data are collected on the mortality of soldiers, bishops, and sailors in the colony in question, during the seventeenth and eighteenth centuries (Acemoglu et al. [2001], p. 1370). We assume that the colonial powers are instantaneously informed about the expected mortality rate. The second instrument used by Hall and Jones is related to the fraction of the total population speaking the English language (labeled English, Source: Hall and Jones [1999]). Another instrument often used in this context is related to the distance from the equator, labeled Distance (Source: Hall and Jones [1999]).

Table 5 above synthesizes the various partial correlations between each variable used and its instrument. For the institutional variables, we note that the instruments of distance to the equator and settler mortality of European colonists are strongly correlated with all the approximations used for the institutional determinant. The results for the instrument related to the fraction of the population speaking the English language inform us of the same degree of correlation. We deduce a clear advantage of the Hall and Jones [1999] instrument of distance to the equator as a presumably strong approximation for institutions, and it is consequently reliable to use these instruments to reflect the various institutional variables used.

In the same table, the partial correlation is calculated between the indicator of trade liberalization (share of total trade in product) and the two forms of the geographical instrument, F&R99 and its modified version by the longitudinal hierarchical approach. It appears that the modified form of the geographical instrument is strongly correlated with the observed share of trade. This is particularly due to the consideration of the random nature of the slope of the size variable and the treatment per pair of the geographical specificities of trade.

The following system summarizes the methodology of our estimation specification:

$$Lrgdppc_{i,2020} = \beta_0 + \beta_1 Open_i + \beta_2 Ins_i + \beta_3 Geo_i + \epsilon_i \tag{13}$$

$Lrgdppc_{i,2020}$, is the logarithmic level of the real GDP per capita of country i in 2020. $Open_i$, Ins_i and Geo_i the variables related to trade openness, institutions and geography. ε_i, is an error term.

Variables $Open_i$ and Ins_i are considered to be endogenous and partially determined by long-run economic growth. Consequently, these variables are replaced by their respective instruments. These instruments are expected to be correlated with the variables they replace and independent of the error.

In a first stage of the estimation, the endogenous variables will be estimated by using all the exogenous determinants as explanatory variables.

$$Open_i = \phi_0 + \phi_1 Instrument/Open_i + \phi_2 Instrument/Ins_i + \phi_3 Geo_i + \mu_i \tag{14}$$

$$Ins_i = \omega_0 + \omega_1 Instrument/Open_i + \omega_2 Instrument/Ins_i + \omega_3 Geo_i + \tau_i \tag{15}$$

In a second phase, the dependent variable is $Lrgdppc_{i,2020}$, and it is explained by the two instruments and all exogenous variables.

Tables 6, 7, 8, and 9 summarize the estimation results of the second stage of modeling the income equation using instrumental variables. Due to space constraints, the results of the first stage estimation are not presented in the text. In the first section of each table, the estimation results of eight benchmark specifications are provided. These specifications correspond to various approximations of the institutional variable using different instruments. In specifications 1 and 3, we utilize the institutional approximation used in Rodrik et al. [2004], which is related to the formal dimension of institutions, namely the rule of law. In the remaining specifications, we employ three other approximations related to government efficiency (specifications 2 and

4), voice & accountability (specifications 5, 6, and 7), and the bureaucracy index (specification 8). As for the instruments, we utilize the modified F&R99 instrument for trade openness, along with settler mortality of European colonists, the fraction of the population speaking English, and distance to the equator for the institutional variables.

The conclusion drawn from these results is that the explanatory power of the institutional variable, as a deep determinant of the international dispersion of income, is statistically significant and robust to nearly all the variations introduced in the institutional approximations and the instruments used for these variables.

The estimation results indicate that a 1% increase in institutional quality, approximated by indicators such as the rule of law, voice & accountability, government effectiveness, and bureaucracy, leads to a corresponding increase of slightly more than 1% in long-run per capita income. The only exception is for specifications that approximate the geographical variable by malaria transmission risk, where the magnitude is lower than 1%. The magnitude of this effect observed in our study appears to be smaller than that reported in Rodrik et al. [2004], possibly due to the use of different instruments for trade openness and institutional quality.

Table 6. *Trade Openness, Institutions and Geography*

$Lrgdppc_{2020}$	Specification 1				Specification 2			
	Settler		English		Settler		English	
	β	*se*	*β*	*se*	*β*	*se*	*β*	*se*
Open	.029	.217	.274	.204	-.039	.167	.234	.192
Rule	1.381*	.191	.858*	.161	-	-	-	-
Effectiveness	-	-	-	-	1.205*	.127	.874*	.150
Landlock	-.451***	.256	-.674*	.163	-.471*	.194	-.556*	.159
Diagnostic Tests : R^2_{Shea}								
Institution	.336		.136		.444		.138	
Open	.662		.365		.654		.347	
Wald test (*p*)	50.24 (0.00)		140.56 (0.00)		98.59 (0.00)		129.90 (0.00)	
N	70		109		70		109	

*et***denote statistical precision of the coefficient at the 1% and 10% risk thresholds

What significantly distinguishes our study from that of Rodrik et al. [2004] is the statistically significant effect of the geographical determinant, as approximated by variables such as the fraction of the population living within 100km of the coast, physical landlockedness, the proportion of land within 100km of the coast, the fraction of land in the tropical zone, and the proportion of area affected by malaria transmission. The effects of these variables are relatively substantial, ranging between 0.45% and 0.67% for landlockedness, 0.36% for POP_{100km}, between 0.36% and 0.92% for the tropical area, between 0.61% and 1.38% for malaria, and between 0.23% and 1.88% for distance from the coast. These magnitudes remain robust when changing the institutional instrument used.

Another distinction in our study compared to the seminal work of Rodrik et al. [2004] is that when we utilize institutional approximations such as voice & accountability and bureaucracy, alongside instruments such as settler mortality, fraction of English-speaking population, and distance to the equator for institutions, the openness variable exhibits a statistically significant correlation with long-run growth. This is a departure from previous studies employing similar modeling approaches.

This newfound emphasis on the geographical determinant, alongside that of institutional nature, can be attributed to several factors. Firstly, the choice of an approximation for institutional quality appears to significantly influence the statistical significance of the determinants, including geography and openness. As discussed in the New Institutional Economics literature, the institutional concept is multidimensional, making it challenging for a single indicator to encompass all dimensions of the institutional determinant.

Table 7. *The Income Equation with Other Approximations for Geography*

$Lrgdppc_{2020}$	Specification 3				Specification 4			
	Settler		Distance		Settler		English	
	β	*se*	*β*	*se*	*β*	*se*	*β*	*se*
Open	.504	.355	.125	.176	.156	.221	.214	.197
Rule	1.741*	.344	1.063*	.086	1.370*	.187	-	-
Effectiveness	-	-	-	-	-	-	.879*	.164
Pop_{100km}	-.297	.500	.367**	.189	-	-	-	-
Tropic	-	-	-	-	-.367	.321	-.656**	.288
Diagnostic Tests : R^2_{Shea}								
Institution	.201		.560		.368		.131	
Open	.478		.584		.646		.346	
Wald test (*p*)	16.31 (0.00)		45.30 (0.00)		17.68 (0.00)		48.81 (0.00)	
N	71		114		71		114	

*et**denote statistical precision of the coefficient at the 1% and 5% risk thresholds

Moreover, the selection of the geographical variable also impacts the statistical precision of the approximations used. Consistent with discussions in the New Economic Geography literature, the influence of geography on long-term economic prosperity varies depending on how we define and measure the geographical variable. The statistically significant correlation of geographical variables with long-term income levels suggests that these variables independently influence the growth trajectory of countries-particularly through proximate variables related to the accumulation of productive factors-after accounting for their effects on other endogenous variables in the model, such as the quality of institutions and trade openness of economies.

In Rodrik et al. [2004], the insignificance of the geographical variable was largely attributed to the choice of using distance to the equator as the approximation. However, as discussed in Sachs [2003], this choice has been criticized for being too weak to adequately test the relevance of geographical determinants (p. 4). Thus, a more thorough examination of the influence of geography on the international dispersion of income is warranted.

In our study, we utilize the distance to the equator variable as an instrument for institutions, as also employed by Hall and Jones [1999]. This instrument has demonstrated robustness, exhibiting strong correlations with various institutional approximations while maintaining its geographical nature and independence from errors.

Table 8. *The Income Equation with the Market Access Approximation for Geography*

$Lrgdppc_{2020}$	Specification 5				Specification 6			
	Settler		English		Settler		English	
	β	*se*	*β*	*se*	*β*	*se*	*β*	*se*
Open	.885*	.285	.578*	.220	1.476*	.557	.572*	.231
Voice	1.820*	.334	.954*	.251	2.497*	.649	1.014*	.279
Tropic	-.536	.439	-.920*	.362	-	-	-	-
$Land_C$	-	-	-	-	-1.880***	.990	.232	.327
Diagnostic Tests : R^2_{Shea}								
Institution	.224		.119		.130		.131	
Open	.700		.529		.471		.675	
Wald test (*p*)	15.33 (0.00)		37.02 (0.00)		11.30 (0.00)		26.02 (0.00)	
N	70		109		70		109	

*et***denote statistical precision of the coefficient at the 1% and 10% risk thresholds

Our research aims to introduce alternative approximations for the geographical variable. We find that distance to the equator is not necessarily the most suitable approximation for capturing geographical aspects of development. Unlike distance to the equator, we find that the proportion of the population living within 100km of a maritime or river coast explains the international dispersion of income independently of its effect on institutional quality. The mechanisms driving this effect were elucidated in Gallup et al. [1999], particularly concerning its impact on specialization structures and the natural level of integration among countries.

Similarly, our analysis suggests that physical landlockedness significantly influences trade costs for countries, thereby affecting their comparative advantage structures. These findings contribute to the geographical development thesis put forth by Gallup et al. [1999], Radelet and Sachs [1998], and Sachs [2003].

Another geographical approximation has been introduced, focusing on the growth rate of malaria (as seen in specifications 7 and 8 in Table 9). This aspect of geographical influence, termed "first nature", has been a point of convergence in various studies, notably Gallup et al. [1999] and Sachs [2003]. Due to potential endogeneity concerns regarding its relationship with income, we opted to use a variable related to malaria prevalence as an instrument for this approximation[19]. The variable representing malaria rates in 1994 demonstrates a direct and independent impact on the international dispersion of income. These findings align with previous studies by Sachs [2003], Gallup and Sachs [2000], Carstensen and Gundlach [2005], and Batten and Martina [2006].

The estimation results concerning the variable of trade openness partially align with those presented in Rodrik et al. [2004]. Despite the statistical reliability of the instrument used, it appears that the degree of trade openness does not exert an independent effect on the level of long-term income, except for its indirect impact on institutional quality (as seen in specifications 1 to 4). This finding echoes the thesis proposed by Wei [2001], suggesting that the "natural" level of integration among nations influences long-term income through the quality of institutions, often measured by "good governance". However, this relationship was not consistently observed across all institutional approximations, and our analysis indicates that trade openness can indeed serve as a deep determinant of long-run income dispersion when different institutional approximations and instruments are employed. These discrepancies suggest potential specification issues in the income equation, leading to instability in estimation results and sensitivity to the choice of approximations and instruments.

Table 9. *The Income Equation with Malaria Effect Approximation for Geography*

$Lrgdppc_{2020}$	Specification 7				Specification 8			
	Distance		English		Settler		Distance	
	β	*se*	*β*	*se*	*β*	*se*	*β*	*se*
Open	.498*	.164	.514*	.175	.343*	.165	.434*	.135
Voice	.708*	.222	.645**	.285	-	-	-	-
Mal_{94}	-1.136*	.311	-1.213*	.375	-.614*	.357	-1.381*	.224
Bureaucracy	-	-	-	-	.971*	.192	.426*	.136
Diagnostic Tests : R^2_{Shea}								
Institution	.155		.091		.244		.216	
Open	.629		. 537		.708		.613	
Wald test (*p*)	27.59 (0.00)		22.36 (0.00)		9.50 (0.00)		33.12 (0.00)	
N	117		117		63		101	

*et**denote statistical precision of the coefficient at the 1% and 5% risk thresholds

19) The prevalence rate is calculated as the frequency of the total population affected by malaria. For conceptual and measurement details, refer to Sachs [2003].

The results presented above are further supported by the statistical reliability of the instruments utilized for the two considered endogenous variables. As indicated by the tests summarized in the second part of Table 6, the Wald statistics on the joint null hypothesis of the two instruments, excluded in the structural form of the model, demonstrate that the instruments employed for openness and institutions exhibit statistical reliability across various specifications of the income equation. This test allows us to reject the null hypothesis of instrument invalidity, confirming the efficacy of the instruments, namely settler mortality, English fraction, and distance to the equator for institutions, and the modified F&R99 instrument for openness.

Another indicator of reliability, known as the partial coefficient of determination, can be employed, as qualified by Shea [1997]. This metric reveals the correlation between the instrument and the endogenous variable under consideration, after accounting for the influence of all other exogenous variables. Across all estimated specifications in our empirical modeling of the international income equation, this coefficient ranges between 34% and 70% for the openness instrument, and between 9% and 56% for the institutional instruments. These results obtained through instrumental variables analysis can be interpreted with a high degree of accuracy and precision.

Table 9 summarizes the estimation results for the remaining specifications of the income equation, incorporating two additional approximations for the institutional variable, namely voice & accountability and bureaucracy quality. The findings are qualitatively similar to those presented in the first set of specifications (Table 6), except for those including the variable of malaria rate as a geographical approximation. In this case, although the statistical precision remains consistent, the magnitude of the institutional variable decreases significantly, favoring the geographical variable. For instance, in specification (8), a 1% increase in the malaria rate leads to a decrease of over 1% in per capita income in the long run. Similar results were reported in Sachs's [2003] study. It's noteworthy that even in the specification incorporating the rule of law variable (Rodrik et al.'s preferred approximation [2004]), the magnitude of the geographical variable is important.

What can be inferred from the estimation results of the income equation is that the magnitude and statistical precision of each of the three deep determinants of economic growth vary depending on our approximation of these variables. When the effects of these variables are estimated with statistical precision, they range between 0.34 and 1.47 for openness, 0.42 and 2.49 for institutions, and 0.36 and 1.88 for geography.

A novel finding from our estimation is that the correlation between trade openness and the economic growth of nations is statistically significant. However, caution must be exercised in interpreting these estimation results because the sensitivity of the magnitude and statistical reliability to the specific approximation and instrument chosen tends to relativize the range of studied specifications by the deep determinants.

This result is likely because the deep determinants of development suffer from a measurement bias primarily due to their multidimensionality, and the fact that a single approximation cannot cover all dimensions of such a determinant, despite its precision. For example, regarding the institutional variable, the New Institutional Economics has enabled us to grasp the complexity of the relationship between institutions and development levels on one hand, and trade openness on the other.

Similarly, the relationship between geographical factors and economic development appears challenging to narrow down, whether for the direct effect of geography on long-term income or the effects on the spatial interdependencies of second nature geography. Redding and Venables [2004] and Radelet and Sachs [1998] attest to the complexity of such a relationship.

In the final section, we aim to address the challenge of disentangling the complex interplay of deep determinants of long-term development to isolate the specific impact of trade openness. This involves conducting an analysis using an automated selection process to identify the most relevant specifications for the income equation of nations. This approach has demonstrated its reliability as an econometric method for testing the relative significance of competing assumptions within a well-defined theoretical framework. Employing this approach does not necessarily imply a lack of suitable theoretical underpinnings for the causal relationships under study, but rather reflects a pragmatic methodology suited to navigating the complexities involved, as outlined by Rodrik [2003].

VI. An Automatic Selection Essay of the Alternative Hypothesis of Income Determination

In this section, our analysis will focus on the use of automatic methodology for selecting relevant variables in determining the observed dispersion of per capita income in the long run. The variable under controversy is trade openness, particularly in underdeveloped countries. Our aim is to study the relevance of this variable within the framework of deep determinants of economic growth.

As discussed in growth empirics, three main theses have been presented and studied in the empirical literature. The first, which we describe as integrationist following Sachs and Warner [1995], argues for the independent effect of trade openness as a deep determinant of long-run growth for nations.

The second thesis, described as institutional fundamentalism, builds on Rodrik et al. [2004], suggesting that the degree of openness cannot independently determine long-term growth; rather, its effects are mediated through institutional quality.

The third thesis, termed geographical fundamentalism, draws from studies by Diamond [1999], Gallup et al. [1999], and Sachs [2003], positing that only geographical determinants directly impact long-run growth, indirectly influencing trade openness (Wei, 2001), and institutional quality over time (Engerman and Sokoloff, 2002).

Given the intertwined characterization outlined in Rodrik [2003], and to distinguish between the direct and indirect effects of these deep determinants, our empirical methodology revolves around two elements: instrumenting the endogenous variables and automatically selecting alternative modeling specifications.

In the preceding section of this paper, we presented an approach to instrumenting trade openness. In this section, our empirical research focuses on estimating an income equation based on the deep determinants. Our primary objective is to ascertain whether trade openness, once instrumented, exerts an independent effect on income dispersion in the long run, as compared to the theses of institutional and geographical fundamentalism. Key reference studies include those by Rodrik et al. [2004] and Sachs [2003].

Our goal is to address the uncertainty that surrounds the debate regarding the role of trade openness as a determinant of development, particularly when compared to institutions and geography. In Rodrik et al.'s study [2004], the principal conclusion is that once the issue of endogeneity concerning institutions and trade openness is addressed, only institutions demonstrate a direct effect on the long-term growth of nations. The effects attributed to trade and geography either show an estimated effect contrary to theoretical predictions or an effect that we cannot statistically reject as null.

In the previous section, we attempted to demonstrate that the relevance of the findings in Rodrik et al. [2004] primarily stems from the choices made regarding institutional approximations and geographical variables. By using a modified version of the single instrument for trade openness (F&R99) and considering alternative institutional and geographical approximations, the estimation results reveal different insights into the relative importance of the deep determinants of national growth. While the differences in base periods and sample compositions exist, they likely result from inherent issues in specifying the income equation. This explanation is further justified by the absence of a directly estimable structural equation for the income equation and the complexity of the causal diagrams linking the determinants, which precludes certainty regarding the existence of a "general" model. This qualification is consistent with the theory of reduction, which delineates the transition from what we term the process of data generation to an unrestricted general model (GUM).

More precisely, while it may be feasible to isolate the effects of institutions and openness on development, the same cannot be said for the geographical determinant. This is because the correlation between this determinant and economic growth varies depending on our method of approximating the geographical variable. Certain approximations, such as the fraction of the population living within 100 kilometers of a coast (maritime or river), or the physical landlockedness of a country, represent geographical dimensions that influence the international level of economic integration and related aspects of international spatial interactions (second nature geography). On the other hand, other approximations, such as those related to the malaria risk population rate or the distance to the equator, represent geographical dimensions that directly affect the accumulation of productive factors (physical and human), and consequently, the proximate determinant of economic growth.

Below, we outline the general empirical strategy employed for modeling the long-term income equation. The first conceptual component is the Data Generation Process (DGP), which serves as the broad characterization of the wealth accumulation process, encapsulating the various interactions and complex economic decisions. This process is often represented by a general graph, as proposed by Rodrik [2003], categorizing income determinants into completely endogenous variables (productive factors), semi-endogenous variables (partly determined by income: institutions, openness, and geographical aspects), and completely exogenous variables (first nature geography).

The second element is the Local Data Generation Process (LDGP), which distinguishes between proximate and deep determinants of growth. It involves the generation of data specific to the variables under analysis. These variables include those related to trade openness, institutions (both formal and informal), and geographical factors influencing spatial interactions and other growth determinants.

The third element is the General Unrestricted Model (GUM), which depends on intrinsic data characteristics (availability, form, and transformation) and the nature of correlations considered (linear or non-linear) to formulate an unrestricted general model.

Finally, the last element involves determining the specific approximations of economic variables used, as well as their status as exogenous or endogenous variables, and selecting appropriate instruments.

In our methodology, two classes of variables are considered endogenous: the degree of openness of economies (measured by the observed share of trade in the gross domestic product) and the level of institutional development (formal and informal), as measured by approximations such as the rule of law, democracy accountability, and freedom index. Bureaucratic Quality (labeled Bureaucracy) reflects the institutional strength and quality of the bureaucracy, acting as a shock absorber to minimize policy revisions during governmental changes. High points are awarded to countries where the bureaucracy has the strength and expertise to govern without drastic policy changes or interruptions in government services.

The Freedom in the World Index (Freedom) assigns 0 to 4 points for each of 10 political rights indicators and 15 civil liberties indicators, representing the degree of freedom, with 0 indicating the least freedom and 4 the greatest[20].

The third class of variables pertains to the geographical dimension, considered exogenous in this unrestricted model. The commonly adopted approximations include physical landlockedness (Landlock), the fraction of the entire area located within 100km of a coast ($Land_C$), the fraction of the entire area located within 100km of a coast or river ($Land_{CR}$), the fraction of the entire area located in the tropical zone (Tropic), and the proportion of the population at risk of malaria (Mal_{94}).

The instruments used for trade openness and institutional development variables include the modified version of F&R99 ($F\&R99_M$, Source: Author), settler mortality of European colonists (Settler), and distance to the equator (Distance). The trade openness approximation is determined by the observed ratio of total trade to gross domestic product (Open).

Below, we will present the estimation results obtained when applying an automatic selection of the specifications for the income equation. Initially, we need to specify the strategy for the selective reduction step. Three strategies were introduced in the initial version of the selection algorithm, as discussed in the essays by Hoover and Perez [1999] and Hendry and Krolzig [2001]. The first strategy, described as "liberal", aims to minimize the deletion rate of significant variables. The second strategy, described as "conservative", aims to minimize the deletion risk of relevant variables, with a weaker emphasis on retaining irrelevant variables. The third strategy, described as "expert", allows the user to balance the deletion risk and the power of reduction[21]. For our purpose, we have chosen to adopt the 'liberal' strategy, particularly due to the controversy surrounding the rival nature of alternative theses of economic development.

Three blocks of results will be presented. The first relates to the estimation of the income equation with instrumental variables but without any reduction. The second pertains to the Student probabilities of the variables integrated into the initial unrestricted general model. The third will summarize the results after automatic selection, i.e., the outcome of the reduction and thus the final model to be retained after the battery of specification tests.

The first stage of automatic selection is the process of research and selection. We establish the unrestricted general model (GUM) from which the research begins to distinguish between relevant and irrelevant variables. After testing the adequacy of the general specification (GUM), we base the analysis on individual t-tests to classify the variables according to their statistical precision, using a fixed threshold for rejection. The general specification (GUM) defines the space of all potentially reducible models. From this total space, we define the space of possible paths for reduction. A path is defined for any insignificant variable[22].

---

20) The data on bureaucracy quality is from the ICRG 2017, and the data on freedom in the world is from the Freedom House Database 2020.

21) Refer to Hendry and Krolzig [2001], and Granger and Hendry [2005] for an analysis of the simulation results comparing alternative reduction strategies. Generally, the differences observed are not significant and do not compromise the final reduction specifications.

22) For θ variables of this type, we have 2^{θ} models.

The reduction principles are pruning, branching, and chopping. Pruning involves systematically eliminating branches from the research for which rejection is not possible. Backtracking compared to the initial GUM still allows for the preservation of ineffective research. Branching involves proceeding with a logic of reduction not only by variable but also by branch[23]. Chopping involves definitively eliminating a variable or a "strongly" insignificant branch using a given deletion threshold.

The second stage involves backtracking. We note that the research process described above is iterative, allowing us to explore all possible paths. The logic of backtracking involves conducting a battery of tests for each reduction diagram obtained, comparing them to the initial general specification (GUM)[24]

The third stage involves the confrontation between the terminals and specification diagnosis. Initially, we define a Terminal Model as "a model which cannot be further reduced according to the criteria adopted for the selection" [Doornik, 2009, p.96]. Given the multiple paths of selection, there may be more than one final model. In such cases, we employ a logic of envelopment to select only one terminal, the final specification. If there are two models, we define their union and compare both terminals[25]. It's worth noting that at this final stage, the researcher may choose to retain the final specification that aligns with the theoretical framework. Alternatively, the algorithm can use an information criterion to make the decision[26].

Table 10 below presents an estimation of the income equation using the two-stage least squares method[27]. The endogenous variables of trade openness and institutions are replaced by their instruments. At this stage, no reduction process is employed. Institutions, geography, and openness are statistically significant, although this significance may change when different approximations and instruments for variables are chosen. Homoscedasticity can only be justified for the voice and democracy approximations.

23) We only need to conduct a joint Fisher test on the components of the branch. If the rejection is successful, we can reject all the variables simultaneously. If not, it is necessary to proceed with a reduction by variable.

24) For the battery of test statistics used in diagnostic phase, the following tests are runned: a) The Fisher test for specification; b) The normality test; c) The Chow test for parameter stability; d) The heteroscedasticity and The heteroscedasticity-X tests. See Owen [2002] and Doornik and Hendry [2013] for formal details on the use of these tests.

25) If one model dominates the other in their union, the latter will be selected as the final specification. If one of the models cannot be rejected, it will be considered final. If both models cannot be rejected, their union will be retained as the initial GUM, initiating another stage of the search process.

26) Refer to Hendry and Krolzig [2001] for a comprehensive analysis of the confrontation between terminals (Table 1, p. 837).

27) The numbers in brackets in Tables 10, 12 and 13 represent p_value for tests.

Table 10. *Estimation Results of the Income Equation by the IV*

Stage 0: The General Model without Reduction Process						
	Model I		Model II		Model III	
	(Voice)		(Democracy)		(Freedom)	
$Lrgdppc_{2020}$	β	*se*	β	*se*	β	*se*
Open/Y	0.431**	0.222	0.528*	0.208	0.335	0.222
Institutions/Y	0.769*	0.296	0.363	0.234	0.536***	0.295
$Land_C$/X	0.034	0.783	-0.214	0.715	-0.079	0.635
Pop_{100km}/X	0.159	0.633	0.576	0.567	0.615	0.479
Tropic/X	-0.833*	0.314	-1.148*	0.335	-0.887**	0.418
Landlock/X	-0.586**	0.249	-0.192	0.255	-0.010	0.218
$Land_{CR}$/X	-0.185	0.552	-0.052	0.511	0.067	0.399
Intercept	10.083	0.249	8.715	0.940	8.572	0.881
Diagnostic Tests						
Normality test: χ^2	7.955 [0.018]		7.196 [0.027]		1.152 [0.561]	
Hetero test: F	3.414 [0.000]		4.015 [0.000]		1.622 [0.096]	
Test b=0:χ^2	179.59 [0.000]		151.88 [0.000]		207.17 [0.000]	

*,**and***denote respectively the threshold of significance 1, 5 and 10%.

Additionally, a problem of normality persists in all three specifications when all geographical approximations are integrated simultaneously. These findings suggest the need to re-examine the statistical relevance of the specifications retained for the income equation, as well as the effect of the a priori selection of variables on the quality of the estimation.

In an approach to reduce the income equation, the first stage of the selection logic is based on analyzing the deletion probabilities of all variables modeled to explain income dispersion. Table 11 allows visualization of these probabilities for the three institutional approximations used. Additionally, information criteria provided for the three approximation models are given, along with the decision regarding the initial unrestricted general model (last line of the table). This decision is based on a Fisher test, which determines whether to accept the general model as a representation of the data generation process. For all institutional approximations, the decision is to accept this general model with a risk threshold of 1%. Consequently, the general model can undergo a reduction process.

From Table 11, we can deduce the tree of reduction necessary for this selection. As introduced in the general case by Doornik [2009], it defines the branches of selection composed of all the variables considered potentially non-significant, and therefore rejectable, as they are studied in Table 11. The principal branch in our case consists of the following variables: $Land_{CR}$, POP_{100km}, Tropic, and Landlock. These variables have been classified according to their increasing insignificance[28).

Table 11. *Individual Probabilities for Variables*

Student_prob. GUM_0 & Terminal Models			
	Model I	Model II	Model III
$Lrgdppc_{2020}$	(Voice)	(Democracy)	(Freedom)
Variable/Type	t_prob	t_prob	t_prob
$Land_{CR}$/Exogenous	0.656	-	-
$Land_{C}$/Exogenous	-	0.198	-
Pop_{100km}/Exogenous	0.608	-	-
Tropic/Exogenous	0.007	0.000	0.004
Landlock/Exogenous	0.015	-	0.011
LogLikelihood	-219.32	-228.58	-575.420
AIC	4.270	4.969	11.270
HQ	4.341	5.024	11.322
SC	4.446	5.105	11.398
$Fisher_{prob}$/Decision	0.00/kept	0.00/kept	0.00/kept

From this principal branch, we will define all the paths of selection, which consist of all the sub-branches reduced of the (most) insignificant variable. A model will be judged as a terminal candidate if no possibility of reduction is possible. It will be described as terminal if the backtracking test compared to the initial general model is accepted, then this process of reduction will converge towards the final specifications (of the general to specific) synthesized in Table 12[29].

A specification is known as final if all the necessary specification tests are checked. For the three institutional approximations used, the structure of the final specification for the income equation proves to be unchanged, i.e., made up of the three deep determinants of long-run growth. Therefore, our benchmark specification is where the quality of the institutions is approximated by the index of voice accountability, democracy accountability, and freedom index. In this case, all the specification tests, as those of the diagnosis, are justified. The

28) This tree is specific for Model I, where institutional quality is approximated by the variable, Voice.

29) In order to conduct a comparative analysis of the modified version of the $F\&R99_M$ instrument with the basic version of F&R99, it's noted that using the basic version of the instrument resulted in the following final specification as a 'terminal' of the reduction process: For the first benchmark institutional variable (Voice): $Lrgdppc_{2020} = 9.891 + 0.686.\,Voice^{*} + 0.037.Openness - 0.567..\ Landlock^{*} - 1.147.Mal94^{*}$ Specification test: Chi²(1) = 1.1378 [0.2861]. The instrument for institutions used is Distance to Equator. Normality cannot be rejected at 5%. The numbers in brackets represent p_value for tests. For the second benchmark institutional variable (Democracy): $Lrgdppc_{2020} = 8.581 + 0.376.Democracy^{*} + 0.107.Openness - 0.420.Landlock^{*} - 1.416.Mal94^{*}$. Specification test: Chi²(1) = 1.8866 [0.1696]. The instrument for institutions used is Distance to Equator. Normality cannot be rejected at 5%. For the third benchmark institutional variable (Freedom): $Lrgdppc_{2020} = 8.280 + 0.026.Freedom^{*} + 0.076.$ $Openness - 0.487.Landlock^{*} - 1.120.Mal94^{*}$.Specification test: Chi²(1) = 0.97008 [0.3247]. The instrument for institutions used is Distance to Equator. Normality cannot be rejected at 5%. All terminals confirm the absence of openness as a deep determinant of growth.* and ** denote respectively the risk thresholds of 1 and 5%. Detailed results are available upon request.

assumption of normality is respected; the problem of heteroscedasticity being eliminated for the case of the squared value of variables, as for the product of variables (heteroscedasticity-X); the Chow test indicates that, following a decomposition of the sample at a rate of 70% and 30%, the estimated parameters are stable.

Table 12. *Final Specification of the Income Equation (F&R99$_M$, Distance)*

Final specification of the income equation after reduction						
	Model I (Voice)		Model II (Democracy)		Model III (Freedom)	
Lrgdppc$_{2020}$	β	*se*	β	*se*	β	*se*
Openness	0.431**	0.222	0.472**	0.204	0.434***	0.262
Institutions	0.766*	0.274	0.351***	0.191	0.028*	0.008
Tropic	-0.834*	0.306	-1.188*	0.311	-0.817*	0.281
Landlock	-0.584**	0.237	-	-	-0.556*	0.216
Land$_C$	-	-	0.391	0.302	-	-
N	106		94		103	
Diagnostic Tests						
Normality test:χ^2	7.955* [0.01]		7.607** [0.02]		7.228** [0.02]	
Hetero test: F	3.847* [0.00]		5.852* [0.00]		6.587* [0.00]	
Test b=0:χ^2	181.63* [0.00]		150.22* [0.00]		134.75* [0.00]	

*,**and***denote respectively the threshold of significance 1, 5 and 10%.

An important deduction emerging from these results is in opposition to any fundamentalist thesis: the dispersion of income is explained by openness, institutions, and geography. These results are likely to challenge the institutionalism "Rule" of Rodrik et al. [2004] and the geographical determinism defended by Diamond [1999] and Sachs [2003]. More particularly, the results for openness as a deep income determinant permit to highlight the explanatory power, far from being marginal, for the observed income dispersion. Based on a series of statistically relevant specification tests, the difference in our study can be attributed to two important elements. On the one hand, the quality of the instrumentation of the variable related to openness, aggregated separately for each country, can explain the relevance of the variable under estimation. On the other hand, the logic of reduction adopted in our work is likely to mitigate the risk of deletion concerning the choice of approximation for the modeled variables.

Although the reduction approach itself serves as a relevant test of specification, surpassing traditional sensitivity and robustness analyses like Leamer's extreme bounds analysis (Ericsson, 2008)[30], we conducted robustness tests by incorporating regional dummy variables into the

30) The explanation is based on the fact that the logic of reduction relies on the principle of envelopment of all the terminals compared to the initial unrestricted general model. Ericsson [2008] demonstrated that the principle of envelopment surpasses the properties of the extreme values method.

initial unrestricted general model, particularly those for Sub-Saharan Africa (D_{SSA}), Latin America (D_{LATAM}), and South Asia (D_{SASIA}). Table 13 shows that these regional variables are included in the final specification, but the magnitude and statistical precision of the three deep determinants remain unchanged. This outcome suggests that while other factors specific to countries in these regions may influence income determination, their effects are not disrupted by the model specification that includes the three deep determinants discussed in the theoretical paradigm of growth modeling literature.

Two additional robustness tests were performed in our model, focusing on the use of another instrument commonly employed in such empirical studies-the settler mortality of Acemoglu et al. [2001] -and the inclusion of the malaria risk variable typically used in empirical studies. Both tests indicate that the final specification derived from the automatic reduction process appears stable in terms of its composition, the magnitude of the determinants, and the statistical precision of their effects.

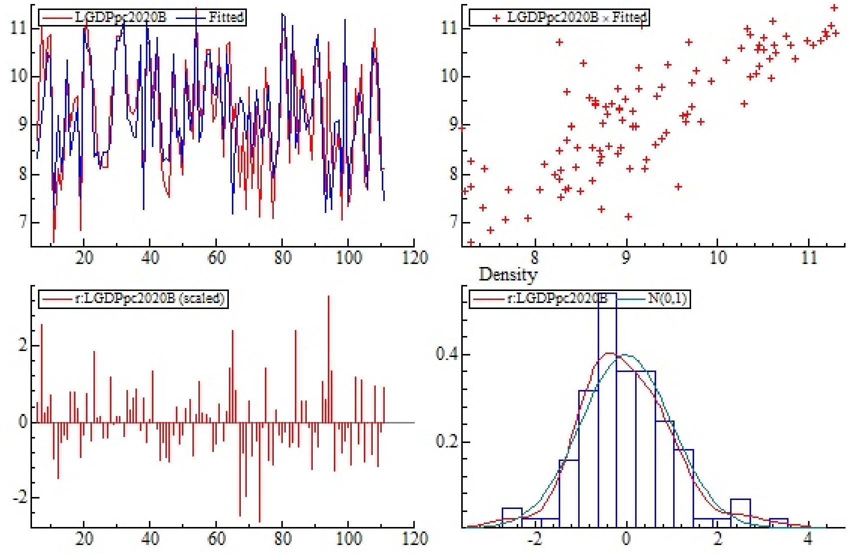
Table 13. *Robustness Analysis for Final Specification (F&R99$_M$, Distance)*

Final specification of the income equation after reduction						
	Model I (Voice)		Model II (Voice, Settler)		Model III (Voice, Regional dummies)	
$Lrgdppc_{2020}$	β	*se*	β	*se*	β	*se*
Openness	0.328***	0.198	0.533***	0.315	0.398**	0.169
Institutions	0.699*	0.196	1.499*	0.241	0.633*	0.117
Landlock	-0.594*	0.172	-	-	-	-
$Land_C$	-	-	-	-	-0.491**	0.115
Mal_{94}	-1.062*	0.290	-	-	-	-
$Land_{CR}$	-	-	-	-	-0.482*	0.221
D_{SSA}	-	-	-	-	-1.589*	0.163
D_{SASIA}	-	-	-	-	-0.868*	0.291
D_{LATAM}	-	-	-	-	-0.645*	0.159
N	106		66		106	
Diagnostic Tests						
Normality test:χ^2	5.342*** [0.06]		2.124 [0.34]		6.069**[0.04]	
Hetero test: F	3.653* [0.00]		4.212* [0.00]		2.418* [0.01]	
Test b=0:χ^2	223.88* [0.00]		39.018*[0.00]		369.56* [0.00]	

*,**and***denote respectively the threshold of significance 1, 5 and 10%.

Lastly, Figure 5 comprises a series of figures related to various elements of graphical analysis assessing the quality of the final specification chosen. We illustrate this with the example of the specification where institutional quality is proxied by voice & accountability, and the instruments used are the distance to the equator and the modified version of F&R99. The final specification appears relevant, both in terms of the comparative evolution between the actual and predicted values of per capita income in 2020, and in terms of the normality of regression errors. This graphical analysis is likely to support our conclusions regarding the statistical relevance of the reduction process, which allows us to identify the three deep determinants as significant explanations for the observed income disparities in the long run.

Figure 5. Visualization of the final specification of the income equation



VII. Conclusion

Overall, although our study offers new insights into the complex empirical question of comparative development, it raises more questions than it answers and prompts further research. While the anticipated outcome might have been to demonstrate that trade openness impacts long-term economic growth, this paper suggests reopening the empirical debate regarding the primacy of other determinants such as institutions and geography. The central aim was to reexamine the empirical literature concerning the causality between trade openness and the economic performance of nations, which has been marked by controversy between two empirical theses. One defends trade's role as a long-term growth determinant, as proposed by Sachs and Warner [1995]. The alternative thesis, situated within a deep determinants approach, asserts the supremacy of institutions and geography as the underlying factors in nations' comparative development experiences. Our study aligns with the latter perspective and refutes fundamentalist theses of long-term economic development. The robustness of these findings lies in the reduction approach, which minimizes bias in selecting approximations and instrumental variables. However, it's important to note that these results don't quantitatively justify openness as a dominant strategy for economic development. The direct effect on income determination is conditioned by institutional development and geographical factors. The conclusions of our empirical analysis are significantly influenced not only by the selected statistical approximations for variables but also by the instrumentation of variables and the chosen empirical approach to analyze complex causal relationships among fundamental determinants.

References

Acemoglu, D., Johnson, S., & Robinson, J. (2001). The colonial origins of comparative development: An Empirical Investigation. *The American Economic Review, 91*(5), 1369-1401.

Aiken, L., Cohen, J., Cohen, P., & West, S. (2003). *Applied Multiple Regression/Correlation Analysis for The Behavioral Sciences*. London: Lawrence Publishers.

Anderson, J. E., & Marcouiller, D. (1999). Insecurity and the Pattern of Trade: An Empirical Investigation. *Review of Economics and Statistics, 84*(2), 342-352.

Arc World Database Supplement and Arc Atlas Database (1996). ESRI Database.

Batten, A., & Martina, A. (2006, Avril). *Diseases Dominate. Working Paper N°470, School of Economics*. Australian National University, pp. 1-60.

Bhagwati, J. (1978). *Foreign Trade Regimes and Economic Development : Anatomy and Consequences of Exchange Control Regimes*. Cambridge, MA: Ballinger Press.

Bryk, A., & Raudenbush, S. W. (2002). *Hierarchical Linear Models for Social and Behavioral Research: Applications and Data Analysis Methods*. Newbury Park, CA: Sage.

Cameron, A. C., & Trivedi, P. K. (2009). *Microeconometrics Using Stata*. Stata Press. p. 692.

Carstensen, K., & Gundlach, E. (2005). *The Primacy of Institutions Reconsidered: The Effects of Malaria Prevalence in the Empirics of Development* (Working Paper N°1210). Kiel Institute for World Economics, pp. 1-33.

Chenery, H., & Syrquin, M. (1975). *Patterns of Development*. World Bank, Oxford: Oxford University Press.

Chenery, H., & Syrquin, M. (1989). *Patterns of Development: 1950-1985* (Discussion Paper N°41). World Bank.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Diamond, J. (1999). *Guns, Germs, and Steel: The Fates of Human Societies*. New York, Londres: W.W. Norton.

Dollar, D. (1992). Outward-Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976-1985. *Economic Development and Cultural Change, 40*(3), 523-544.

Doornik, J. (2009). Autometrics. In J. Castle, & N. Shaphard (Eds.), *The Methodology and Practice of*

Econometrics (pp. 88-121). New York: Oxford University Press.

Doornik, J., & Hendry, D. (2013). *Empirical Econometric Modelling -- PcGive 14* (Vol. I). Timberlake Consultants Ltd.

Dowrick, S., & Golley, J. (2004). Trade Openness and Growth: Who Benefits? *Oxford Review of Economic Policy, 20*(1), 38-56.

Engerman, S., & Sokoloff, K. (2002, October). *Factor Endowments, Inequality, and Paths of Development Among New World Economics* (NBER Working Paper N°9259) Cambridge, Massachusetts: National Bureau of Economic Research.

Ericsson, N. (2008). The Fragility of Sensitivity Analysis : An Encompassing Perspective. *Oxford Bulletin of Economics and Statistics, 70*(Supplement), 895-914.

Frankel, J., & Romer, D. (1999). Does Trade Cause Growth? *The American Economic Review, 89*(3), 379-399.

Gallup, J-L., Scahs, J., & Mellinger, A. (1999). Geography and Economic Development. *International Regional Science Review, 22*(2), 179-232.

Gallup, J-L, & Sachs, J. (2000). The Economic Burden of Malaria. *The American Journal of Tropical Medicine and Hygiene, 64*(1-2), 64-85.

Granger, C. W. J., & Hendry, D. F. (2005). A Dialogue Concerning a New Instrument for Econometric Modeling. *Econometric Theory, 21*(1), 278-297.

Hall, R., & Jones, C. (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others? *The Quarterly Journal of Economics, 114*(1), 83-116.

Hendry, D., & Krolzig, H. M. (2001). Computer Automation of General-to-Specific Model Selection Procedures. *Journal of Economic Dynamics & Control, 25*, 831-866.

Hendry, D. (2000). *Econometrics : Science or Alchemy*. Washington: Oxford University Press.

Hendry, D., & Doornik, J. A. (2014). *Empirical Model Discovery and Theory Evaluation*. MIT Press.

Heritage-Foundation. (2009, Mars). The Index of Economic Freedom. Consulted on www.heritage.org/Index/

Heritier, S., Cantoni, E., Copt, S., & Victoria-Feser, M-P. (2009). *Robust Methods in Biostatistics.* United Kingdom: A John Wiley and Sons, Ltd, Publication.

Hoover, K. D., & Perez, S. J. (1999). Data Mining Reconsidered: Encompassing and the General-to-Specific Approch to Specification Search. *Econometrics Journal, 2*, 167-191.

ICRG. (2017). ICRG 2017. Consulted Mars 2022, in https://www.statistic.lib.ou.edu/icrg (3B Research Database).

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2003). Governance Matters III: Governance Indicators for 1996-2002. *The World Bank Economic Review, 18*(2), 253-287.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2005). *Governance Matters IV : Governance Indicators for 1996-2004* (Policy Research Working Paper No. 3630). World Bank, Washington, DC. World Bank. https://openknowledge.worldbank.org/handle/10986/8221 License: CC BY 3.0 IGO.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2013). *The Worldwide Governance Indicators: Methodology and Analytical Issues* (Policy Research Working Papers). World Bank Group. https://doi.org/10.1596/1813-9450-5430.

Keynes, J. M. (1939). Official Papers. *The Economic Journal, 49*(195), 558-577.

Kreft, I., & De Leeuw, J. (1998). *Introducing Multivlevel Modeling*. Sage Publications, Inc.

Krueger, A. O. (1972). Evaluating Restrictionist Trade Regimes : Theory and Measurement. *The Journal of Political Economy, 8*, 48-62.

Krueger, A. O., & Tuncer, B. (1984). An Empirical Test of the Infant Industry Argument Reply. *The American Economic Review, 74*, 1112-1113.

Krueger, A. O. (1997). Trade Policy and Economic Development: How We Learn. *The American Economic Review, 87*, 1-22.

Krueger, A. O. (1998). Why Trade Liberalisation is Good for Growth. *The Economic Journal, 108*, 1513-1522.

Leamer, E. (1988). Measures of Openness. In *Dans NBER, Trade Policy Issues and Empirical Analysis* (pp. 145-204). NBER.

Mahalanobis, P. C. (1936). On the Generalized Distance in Statistics. *Proceedings of the National Institute of Science of India, 12*, 49-55.

Mauro, P. (1995). Corruption and Growth. *The Quarterly Journal of Economics, 110*, 681-712.

Owen, P. D. (2002). *General-to-Specific Modelling Using PcGets* (Discussion Papers No 0213). University of Otago Economics.

Rabe-Hesketh, & Skrondal, H. (2007). *Multilevel and Longitudinal Analysis Using Stata* (2nd ed.). Stata Press.

Rabe-Hesketh, & Skrondal, H. (2022). *Multilevel and Longitudinal Analysis Using Stata* (4th ed.). Stata Press.

Radelet, S., & Sachs, J. (1998). *Shipping Costs, Manufactured Exports and Economic Growth*. American Economic Association Meetings.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. T. (2004). *HLM 6: Hierarchical Linear and Nonlinear Modeling*. Scientific Software International, Inc. p. 310.

Redding, S., & Venables, A. (2004). Economic Geography and International Inequality. *Journal of International Economics, 6*, 55-84.

Robinson, W. S. (1950). Ecological Correlations and the Behavior of Individuals. *Sociological Review, 15*, 351-357.

Rodrik, D. (2003). *In Search of Prosperity: Analytic Narratives on Economic Growth*. Princeton University Press.

Rodrik, D., & Rosenzweig, M. R. (2009). Development Policy and Development Economics: An Introduction. In D. Rodrik, & M. Rosenzweig (Ed.), *Handbook of Development Economics* (Vol. vacancy, pp. 1-21). North Holland.

Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions Rule : The Primacy of Institutions Over Geography and Integration in Economic Development. *Journal of Economic Growth, 9*, 131-165.

Rose, A. (2017). Why do Estimates of the EMU Effect on Trade Vary So Much? *Open Economic Review, 28*, 1-18.

Pritchett, L. (1996). Measuring Outward Orientation in LDCs Can It Be Done. *Journal of Development Economics, 49*, 307-335.

Sachs, J. (2003). *Institutions Don't Rule: Direct Effects of Geography on Per Capita Income* (NBER

Working Paper N°9490) Cambridge, Massachusetts: National Bureau of Economic Research.

Sachs, J., & Warner, A. (1995, January). *Economic Reform and the Process of Global Integration.* Brookings Papers on Economic Activity.

Shea, J. (1997). Instrument Relevance in Multivariate Linear Models : A Simple Measure. *Review of Economics and Statistics, 79*, 348-352.

Staiger, D., & Stock, H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica, 65*, 557-586.

Stock, J. H., & Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In D. W. Andrews (Ed.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge: Cambridge University Press.

Thirlwall P., & Pacheco-López, P. (2008). *Trade Liberalization and The Poverty of Nations*. Northampton: Edward Elgar.

Tobler, W., Deichmann, U., Gottsegen, J., & Maloy, K. (1995). *The Global Demography Project* (Technical Report No. 95-6). NCGIA Geography Department, University of California. Santa Barbara, CA.

Wacziarg R., & Welch H. (2008). Trade Liberalization and Growth: New Evidence. *The World Bank Economic Review*, 187-231.

Wei, S.-J. (2001). *Good Governance and Natural Openness* (NBER Working Paper N°7765). Cambridge, Massachusetts: National Bureau of Economic Research.

World Bank. (1987). *Barriers to Adjustment and Growth in the World Economy: Industrialization and Foreign Trade*. Washington: The World Bank.

World Bank. (2022). *World Development Indicators CDROM.* Washington: The World Bank.

World Governance Indicators Project (WGI). (2022). Database available at http://info.worldbank.org/governance/wgi/. Washington: The World Bank.