On Causality and Robustness:

Quantifying the causal effects between income, institutions and trade openness

Lindgren, Johan[1] and Senderovic, Alisa[2]
Master thesis in Economics, Stockholm School of Economics
Tutor: Mats Lundahl

May, 2011

Abstract

We investigate bi-directional causality between GDP per capita, the Kaufmann Rule of Law Index and trade as share of GDP using the Granger causality test. A larger data set than previously used in studies on causality is employed. We find that the various causal relationships depend on the countries under examination. Furthermore we show that the instrumental variable method is insufficient in quantifying the effects of the reverse feedback from income to institutions and trade openness. Moreover, changing the source of data for the very same measure leads to substantially different point estimates when employing the two-stage least squares method. Altogether, the results highlight the importance of adequate robustness tests.

Keywords: Causality, Granger test, instrumental variable, institutions, robustness
JEL Classifications: C1, C3, F1, O1

[1] 40039@student.hhs.se

[2] 40056@student.hhs.se

CONTENTS

1 INTRODUCTION

Development and economics are closely related, not only because there is a sub-field of the discipline called *development economics*, but because economics is *needed* in the development of countries. Economics is the science that teaches us how to deal with scarce resources and stopping there would most likely make an exhaustive explanation, but furthermore, economics provides us with the tools that make development understandable and quantifiable and hence also *affectable*. One splendid example on how science enlightens and scientists (and, as well as, politicians) transform this into a guide for presidents, dictators, prime ministers or whoever is in charge over any country in the world, is UNDP's Human Development Report[3]. Here one can find hands-on tips of the kind "do this, and that will be the result"; aiming for increased levels of development. It is of course *based* on science; however, there is a political dimension present which is not to be forgotten (a further discussion on this topic will not be held within this paper since it is far beyond the scope of it – this is science, and at least not *yet* politics).

On p. 157 in Rodrik et al (2004), the paper in focus, we can however read the following:

> *How much guidance do our results provide to policymakers who want to improve the performance of their economies? Not much at all. /…/ the operational guidance that our central result on the primacy of institutional quality yields is extremely meager.*

Basically, the study tells us that e.g. property rights are important, but not *how* important they are, nor in what way they could be strengthened.

The findings are however of great use for science. *Stand on the shoulder of giants* is a well-known saying, coined by Newton. This paper, however, makes a humble attempt to actually specify and quantify the effects found in Rodrik et al. (2004), but also, which perhaps is of more importance, to investigate how sensitive, or robust, the results are:

Performing robustness tests is often a matter of checking cross-correlations and running regressions stepwise in different combinations in order to present results that remain un-biased. But what does the word *robust* really mean? It must be that the results are stable and stay universally significant when changing samples, data sources, or methods. This is the view of the authors of this paper and hence a postulate for the study.

[3] See http://www.undp.org

Here is an example: If one projects a picture on a white screen, the audience can clearly see what the picture represents, say, an elephant. If the screen is removed, and the picture is projected on the wall behind, which might be of completely different quality than the screen, for instance with wallpaper, the audience can still see and understand the picture – the elephant does not change into a zebra. The same goes for an empirical study: If the same method and variables are used, but the *data* is changed, the results should not change; nor should they if the original data is used, but the method is changed – if the results are in fact accurate.

But still, we have to turn to the depths of our souls and ask again, just like Marcus Aurelius proposes: What is in its very nature – once again; why is this important? Perhaps one must be more specific:

In Rodrik et al. (2004) they conduct a number of robustness tests, however none of them involves changing the data source or the method. Hence, it is of most interest to investigate whether the results are still stable – significant or even the same – if the dependent variable GDP per capita is collected from another data source, or if the original dataset is used, but with another method estimating the causal relationships.

Most of the empirical studies making an attempt at examining the causal relationship between institutions and growth use the instrumental variable method, which is a purely cross section method. Such studies include Mauro (1995), Knack and Keefer (1995) as well as Acemoglu et al. (2001) on which the Rodrik et al. paper is based. The studies on causality using time-series techniques, concerning growth and other potentially influential variables usually cover individual countries or a set of countries. The results are mixed and the direction of causality seems to be dependent on the country and the time period under examination.

The purpose of this paper is twofold. Firstly, the purpose is to replicate a widely quoted study of the relation between institutions and growth, using another method than what is commonly used. The aim is to examine the causal relationship between institutions and growth using time series analysis in the form of the Granger causality method. Secondly, this paper attempts to highlight the importance of data quality and to show that most results presented in empirical studies are subject to data choice and method.

In this study, the authors use a method making it possible to show the very direction of impact of the different factors of the model. This, in turn, makes it possible for policy makers to decide which factors improving development to boost as well as what their effect on other factors will be, which is very difficult based on the Rodrik et al. study.

The hypothesis of this paper is that changing the data source will change the results more than what would have been the case if the results were truly robust. Furthermore, we argue that the IV method is inappropriate to prove causality. If it in fact would be appropriate, the results after changing the method will remain the same.

To our knowledge, this study uses a larger sample of countries than what has been used in previous studies. Furthermore, in this paper, the same source of data is used, but two different econometric methods are used. In addition, the same method as in the Rodrik et al. paper is used, but with another data source. Such a thorough investigation to check the stability of the results has not to our knowledge been seen. The results highlight the importance of testing for such obstacles.

First, the background on the subject is to be examined. Within this context the Rodrik et al. study will explicitly be explained and discussed. In section three, the method is outlined and in section four the analysis is carried out. Finally, the implications of our results are evaluated and a discussion of policy implications is provided.

1.1 DELIMITATIONS

There are several delimitations of the paper; some of which are unintentional, for instance due to data limitations. One of the most obvious limitations is that the study uses rather few number (10) of time observations. Generally, the fewer the observations, the more meager the information basis. This must however not be the case when using this specific method, since for each time observation, which is used as a variable, there is a *vector* of observations - countries. Hence, if only one explaining variable were used, considering the largest sample, one would still obtain 138 observations.

However, the sample and all variables might be considered as a small window of information in a wall, which represents reality. There is hence a lot of unavailable and inaccessible information. In order to deal with this problem in an as accurate way as possible, two tests are used: On one hand according to Granger's original set up, which results in a fewer number of variables; on the other a similar test including all variables (time lags).

In addition, since the study is based on Rodrik et al. (2004), which in turn uses instruments developed by Acemoglu et al. (2001) and Kaufmann (2009) and these instruments are unavailable for larger sample sizes than those used, this is also a delimitation that must be mentioned.

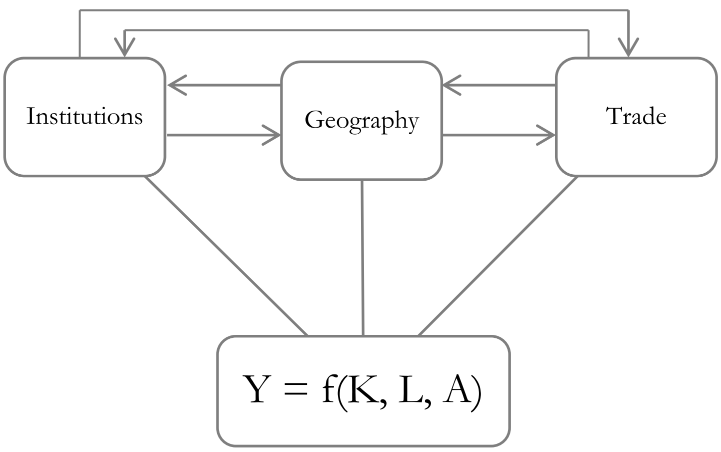
Studies of this kind always run the risk of being overwhelmingly extensive since there are numerous details involved that are of interest to investigate further; aiming to reveal new insights. Somewhere along the way, one must however decide on delimitations for practical reasons. In this particular case, it would have been possible to conduct a multivariate regression setup instead of the bivariate, which was decided on. A detailed discussion on why a bivariate analysis was deemed appropriate follows in part 3.4.

2 BACKGROUND ON GROWTH, INSTITUTIONS AND TRADE

2.1 THEORETICAL FRAMEWORK

According to standard growth theory, endogenous and exogenous, growth is a function of capital accumulation, human capital, and technological progress. This leads us to the more interesting question: W*hy* are some countries better at accumulating capital, educating their labour or innovating? The ideas provided in the literature may be summarized in the following way:

Figure 2.1. Possible causal determinants of growth



Geographical factors such as soil condition and access to markets could have an effect on income growth. Increased trade could also affect income by facilitating knowledge flows between countries. The idea that institutions affect growth stems from the idea that political and economic institutions create the “rules of the game in society” (North, 1990,

p. 2) in the form of property rights, contract enforcement and rule of law which all facilitate, if not are the source of, economic activity. But *how* would institutions in the form of e.g. property rights, lead to increased growth? Interactions based on "the contract" rely on the assumption that agents play by the rules. For this to happen, contracts must be enforceable. With no, or weak, institutions in place contracts may not be enforceable and there is a possibility of coercion. North (1981) argues that good institutions will support private contracts as well as provide protection against expropriation by the rulers or other elite groups. Acemoglu et al. (2005) mean that political institutions and distribution of resources in an initial period determine the economic and political institutions which in turn affect future economic performance and distribution of resources.

$$\begin{array}{ccc}
\begin{array}{c}\text{political}\\ \text{institutions}_t\end{array} & \Longrightarrow & \begin{array}{c}\text{de jure}\\ \text{political}\\ \text{power}_t\end{array}\\
 & & \&\\
\begin{array}{c}\text{distribution}\\ \text{of resources}_t\end{array} & \Longrightarrow & \begin{array}{c}\text{de facto}\\ \text{political}\\ \text{power}_t\end{array}
\end{array}
\left.\vphantom{\begin{array}{c}a\\a\\a\\a\\a\\a\\a\end{array}}\right\}
\begin{array}{ccc}
\Longrightarrow & \begin{array}{c}\text{economic}\\ \text{institutions}_t\end{array} & \Longrightarrow \left\{\begin{array}{c}\text{economic}\\ \text{performance}_t\\ \&\\ \text{distribution}\\ \text{of resources}_{t+1}\end{array}\right.\\
 & & \\
\Longrightarrow & \begin{array}{c}\text{political}\\ \text{institutions}_{t+1}\end{array} &
\end{array}$$

Source: "Institutions as the Fundamental Cause of Long-Run Economic Growth", Acemoglu, Johnson and Robinson (2005).

To get a better understanding of the relation between institutions and growth we must look at the different views on how institutions arise and what their impact on economic activity is.

According to the incidental institutions view, institutions develop as a result of incidental actions. An advocate of this view is Tilly who, in his book *Coercion, capital and European states,* AD 990-1992 (1992), seeks to explain the formation of states in Europe over a period of thousand years. He points to the fact that authoritarian rulers did not have a plan for forming states –they formed states to serve their subjective purposes. The formation of states was constrained by certain geographical conditions, e.g. it is more difficult to conduct military operations in a region with extreme climate. In other words, state formation was a consequence of other actions. Tilly argues that different combinations of coercion and capital created diverse types of states. This view, when applied to figure 2.1. would be in favor of different geographical situations giving rise to different institutions which in turn would affect the distribution of resources.

Those who are advocates of the efficient view on institutions argue that societies choose the institutions that maximize the total utility of the society, no matter what the distribution of this utility is. Given no negotiation costs, individuals can bargain to internalize negative and positive externalities, as according to the Coase Theorem [4]. From the efficient institutions point of view, the more constraints on the government the better since institutions that maximize society's surplus will be chosen by society itself. This is in line with North's definition of institutions as constraints on the government and the advocates of this view often argue for laissez-faire policies. Because this view relies on no bargaining costs, it is difficult to test empirically.

The social conflict view of institutions is similar to the efficient institutions view in the sense that it is individuals, or groups of individuals, that choose society's institutions. This does, according to the social conflict view on institutions, not always lead to efficient institutions. Instead, groups with political power create institutions that will benefit themselves. The reason why groups with political power seek to benefit themselves is due to the problem of commitment and contract enforcement. The state itself cannot credibly enforce private contracts if there is no outside agent who can enforce the contract between the state and its citizens and in turn sanction misbehavior by the state. This view is in line with Acemoglu et al[5]. (2001) as well as the empirical paper by Rodrik et al. (2004) who use a measure of the rule of law as a measure of institutions.

In summary, from theory we know by now that the way institutions can affect economic growth is i) by providing a commitment mechanism that will make the state protect the citizens from expropriation thus e.g. creating incentives for investment and ii) through constraints on the government and efficient institutions maximizing each society's surplus. The challenge lies in testing these empirically and because the former is, for reasons stated above, more easily tested empirically, it has gained wide acceptance among the empirical studies.

2.2 LITERATURE – EMPIRICAL STUDIES

In the literature on economic growth three main strands, and possibly a fourth, have evolved. The notion of geography as the main source of growth, increased trade as a

[4] See for instance Advanced Microeconomic Theory by Jehle and Reny, Addison-Wesley, 2001

[5] Acemoglu et al. also stress that the political power is a result of the initial distribution of resources which could be affected by geographical conditions.

fundamental driver of growth, and – perhaps empirically the most valid – institutions as a main source of growth. The fourth refers to different combinations of geography, trade and institutions as the main drivers of growth.

According to theory, growth is a result of capital accumulation and technological change. More interesting is the question of *why* some societies manage to accumulate capital and develop technological change. As mentioned preciously, geography, trade and institutions could provide an answer to this question.

In their paper from 1998, Gallup et al. note that landlocked countries and countries in tropical zones have lower incomes than countries in temperate climate that are not landlocked with a few exceptions which they explain by highly developed infrastructure or high degree of integration and thus low trade barriers for some European countries. They find, through an analysis of cross-section data, that once controlling for institutions and policy, geography matters for growth, as do institutions and policy. This however says nothing about the causal effect of geography and growth.

Jeffrey Sachs (2001) makes a similar conclusion as Gallup et al.:

> *Economies in tropical ecozones are nearly everywhere poor, while those in temperate ecozones are generally rich. And when temperate economies are not rich there is typically a straightforward explanation, such as decades under communism or extreme geographical isolation.*

The method used by Sachs (2001) is to regress income growth over a time interval on initial income, initial level of schooling, variables capturing institutions and policy and a variable measuring the share of population in temperate climate zones. He finds that temperate zone countries have incomes of 1.6 times those in non-temperate areas, and that climate variables "perform better" than simple geographical variables such as distance from the equator. Sachs performs a cross-section analysis. Thus the only conclusion that can be inferred from the results is that temperate zone countries happen to have higher growth rates than non-temperate ones.

There is extensive research on the indirect effect of geography, through institutions, on income. Engerman and Sokoloff (1994) make an attempt at explaining income differences between North and South America. They argue that because of geographical conditions such as favorable climate, high concentration of native population – and hence labor – and high concentration of natural resources, South American colonies were more suitable for large scale production and inequalities arose where native labor was used for extractive purposes. Consequently, in these colonies, inequalities gave rise to institutions that were unfavorable for growth. In contrast, geographical conditions in USA and Canada made small-scale production more favorable. There was thus a lower degree of inequality which favored democratic institutions and more equal distributions of wealth which was more favorable for long run growth.

Sala-i-Martin and Subramanian (2003) examine the Nigerian experience and find that some natural resources such as oil and minerals create opportunities for corruption and thus have a negative and nonlinear effect on growth through institutional quality. Easterly & Levine's (2003) study of cross country data supports the view that geography affects income only through the effect on institutions. They find that "tropics, germs and crops" affect income through institutional quality. Moreover, they find no evidence of these factors affecting income directly. The method used by Easterly & Levine is the instrumental variable technique which, at best, allows one to take into account the direct and indirect effect of endowments on income, but cannot estimate the size if the feedback from income on institutional quality and thus says very little about the actual causal effect of endowments on income.

Sachs and Warner (1995) show that the trade regime matters for income growth: Poor economies suffer from lack of convergence because they have been closed for a large part of the last decades. Dollar and Kraay (2004) note that more than half of the developing world lives in countries that have increased their trade openness. Moreover, their analysis shows that trade openness leads to more rapid growth and reduced poverty in developing countries. They employ time-series analysis to examine the relation between trade and income and use changes in trade volume from one decade to another as a proxy variable for changes in trade policy. They stress the advantage of using a time series technique since cross-country variation in trade volume often are a result of countries geographical conditions. Furthermore, using *changes* from one period to another helps to avoid the possibility that the results could be driven by some unobserved country specific variable that varies little over time, but drives both growth and trade (e.g. institutional quality).

The most fundamental question regarding institutions and growth is perhaps whether it is growth that leads to improved institutions, if better institutions cause growth, or if the causality runs both ways. Glaeser et al. (2004) conduct a study of the relation between institutions and growth and provide an example of the idea that growth may create better institutions:

> *While on average, looking over the half century between 1950 and 2000, South Korea obviously had better institutions as measured by constraints on the executive, these institutions are the outcome of economic growth after 1950 rather than its cause.*

(Glaeser et al. 2004, pp. 273)

As always with individual cases, the question of generalizability arises, i.e. whether the case of South Korea can be used to explain the relation between institutions and growth elsewhere in the world and in time. The authors conclude through their analysis that institutions do not cause growth. However, because the method used by Glaeser et al. to conclude this is a simple cross section OLS regression – and this does not say anything about causality – their analysis neither disproves nor proves the hypothesis that institutions cause growth.

Moreover, Glaeser et al. find that the variables used in the literature to measure institutions are in fact not indicators of institutions but merely *outcomes* of institutions. According to the analysis made by the authors, the commonly used measures of institutions do not reflect permanent or durable constraints on the government or long run features of the political climate – as institutions were defined by North (1981). Measures such as government effectiveness, as compiled by Kaufmann et al., are ex-post outcomes. Singapore, which has a one-party rule – and should thus be considered to have "bad" institutions" – is one of the best ranked countries according to this index because its one-party government has *chosen* a policy that favors investor protection. Other measures such as the measure by Polity IV[6] attempts at measuring constraints on the government, but when looking at how the measure is compiled one notices that this too is an ex-post outcome measure. As Glaeser et al. note:

> */…/ Haiti gets the worst score of 1 under the dictatorship during 1960-1989, jumps up to 6 when Aristide is elected in 1990, goes back to 1 when he is ousted during 1991-1993. /…/ The data make it obvious that Polity IV provides a rapidly moving assessment of electoral outcomes over time, not a measure of actual political constraints on government, and certainly not a measure of anything permanent or durable.*

(Glaeser et al., Do Institutions Cause Growth?, 2004)

---

[6] See http://www.systemicpeace.org/polity/polity4.htm

Most of the empirical studies making an attempt at examining the causal relationship between institutions and growth use the instrumental variable method, which is a cross section method. Such studies include Mauro (1995), Knack and Keefer (1995) as well as Acemoglu et al. (2001) on which the Rodrik-paper is based. Mauro (1995), for instance, uses an ethnic division index as an instrument for institutions. However, this instrument may not resolve the reverse causality problem.

Some studies on causality concerning growth and other potentially influential variables using time-series techniques are present in the literature.

Ghartey (1993) examined the causal relation between trade and economic growth in USA, Japan and Taiwan and found that for the United States GDP causes exports, while in Taiwan the opposite causal relationship prevails. Chaudry et al. (2010) explore the causality relationship between trade liberalization, human capital and economic growth in Pakistan. They use quarterly data over the period of 1972-2007 and find that there is causality running from trade liberalization to economic growth. Ingianni (2010) uses a Granger-causality test and finds a unidirectional causal relationship from trade openness to GDP per capita, in Czech Republic, Poland, Latvia and Slovenia. The same direction of causality is found by Henriques and Sadorsky (1996) who find that for Canada, exports cause growth.

Dutt and Ghosh (1996) use a relatively large sample of countries, 26, to examine the causality between exports and growth. Their findings are mixed. In contrast to Henriques and Sadorsky (1996) they do not find any causal relationship between exports and growth for Canada, while in the United states, they found that causality runs from GDP to exports. Ahmad and Harnhirun (1996) find that trade in the form of exports cause growth, but not the other way around in Indonesia, Malaysia, the Philippines, Singapore and Thailand. This is further supported by Furuoka (2007) who came to the same conclusion for Malaysia.

Clearly, previous empirical studies are far from unanimous on the direction of causality between trade openness and income. The results seem to depend on the country and region under investigation as well as the time period examined .

There are a few studies on the causal effect of institutions on growth, and reverse causality of this relationship. Groenewold and Tang (2005) conduct a time-series analysis of Hong Kong to investigate whether democratic improvements of institutions have affected Hong Kong's growth rate of GDP from 1984 to 2003. They find that in the short run, democratic improvement causes GDP. Chong and Calderon (2000) use panel-data for 55-

countries over a time period and find that growth generates improvements in institutional quality. However, when excluding 20 of the most developed countries they find that causality runs the opposite way. By looking at a period over 1970-2007, Alfaro et al. (2008) find that institutional quality is the reason why capital does not flow to poor countries. Acemoglu et al. (2001) and Rodrik et al. (2004) show and argue that institutions are the fundamental source of growth. Alfaro et al. (2008) mean that it is through inflow of capital to poor countries that institutions matter for income levels.

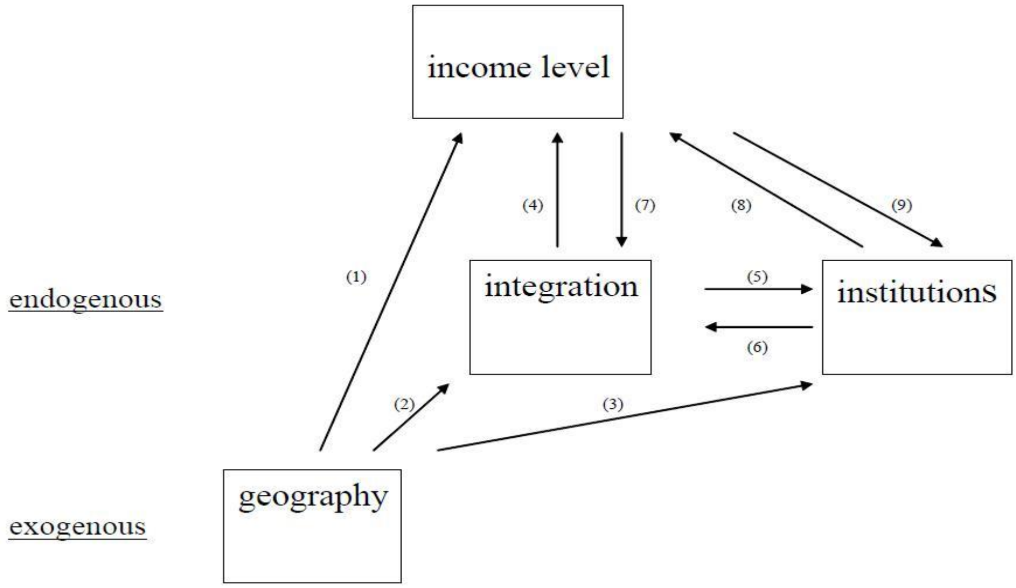
Anderson (2001) suggests that bad institutions inhibit trade due to high costs and risks of trade. This implies that there is a potential causal relationship between institutions and trade openness. Méon and Sekkat (2004) use both cross-section and time-series data to examine the relation between the quality of institutions and trade integration into the world economy of the MENA (Middle East North African) countries. They conclude that a worsening of the quality of institutions is associated with low performance of exports as well as foreign direct investment where the relation is stronger for exports. Winters (2004) argues that openness to trade can have an effect on institutional development. One reason for this is Wei's (2000) argument that more open countries suffer more from corruption (which is included in the rule of law measure) than less open countries since corruption is disproportionately associated with foreign trade.

2.3 RODRIK ET AL. STUDY

In this section, we describe the Rodrik et al. study, which is replicated and analyzed in this paper.

The study by Rodrik et al. aims at explaining the contributions of institutions, geography and trade in determining income levels. Figure 2.1.illustrates the effects of geography, integration and institutions on income that the authors are trying to capture with the help of the instrumental variable technique.

Figure 2.1 The "deep" determinants of income by Rodrik et al.



Source: Rodrik et al., "Institutions Rule: The primacy of Institutions over Geography and Integration in Economic Development", 2004

They find that once institutions are controlled for, geography and trade openness has only weak direct effects on income but stronger indirect effects through the quality of institutions.

Rodrik et al. try to address the problems of causality by utilizing the instrumental variable technique. The main regression to be estimated is the following:

$$\log y_i = \mu + \alpha R\hat{U}LE_i + \beta LC\hat{O}PEN_i + \gamma DISTEQ_i + \varepsilon_i \quad (2.1)$$

RULE, LCOPEN and DISTEQ are measures for institutions, trade integration and geography, respectively. It is plausible, as previously noted, that institutions and trade may affect income and that income may affect the quality of institutions and the level of trade. In order to isolate the direct effect of institutions and integration one needs to fins an instrument that correlates highly with the endogenous explanatory variables but is uncorrelated to the error term in regression 2.1.

Rodrik et al. use the Acemoglu, Johnson and Robinson (2001) measure of settler mortality and the Frankel-Romer constructed trade shares to instrument for the exogenous variation in institutions and trade, respectively. This is the first stage of the regressions.

$$RULE_i = \lambda + \delta LOGEM4_i + \phi LOGFRANKROM_i + \psi DISTEQ_i + \varepsilon_{RULEi} \quad (2.2)$$

$$LCOPEN_i = \theta + \sigma LOGEM4_i + \tau LOGFRANKROM_i + \omega DISTEQ_i + \varepsilon_{LCOPENi} \quad (2.3)$$

LOGEM4, LOGFRANKROM and DISTEQ represent measures for settler mortality, constructed trade shares by Frankel and Romer (1999), and geography as distance in degrees from the equator, respectively.

Next, they use the predicted values from these regressions for institutions and integration and use them as explanatory variables in regression 2.1.

Lastly, they regress the predicted variables of institutions and the measure of geography on integration and the predicted variables of integration and the measure of geography on institutions. This is done to isolate the direct effect of each and one on of the variables on income.

The variable "settler mortality" needs some further explanation: In order to capture the effect of institutions on today's level of income, Acemoglu et al. seek a factor – a variable – that carries information on institutions in earlier time periods and has actual effect on GDP per capita today. Acemoglu et al. claim that there were different types of colonization policies, resulting in different types of institutions. At one extreme, the Europeans set up what they call "extractive states". The Belgian colonization of Congo is one example of this. The purpose of this kind of colonization was to transfer resources from the colony to the colonizer and the institutions did not provide much protection for private property, nor did they provide balances against government expropriation.

At the other extreme, Europeans emigrated and the settlers replicated European institutions, which in fact provided the checks and balances against government expropriation and protection of private property that were lacking in the "extractive states". Examples of colonies in this latter extreme are Australia, New Zeeland, United States and Canada.

According to their theory, the colonization strategy was influenced by the feasibility of settlements. If the environment was not favorable and morality rates high, the formation of the extractive state was more likely. The institutions was inherited by the later independent state and persisted.

Acemoglu et al. use the mortality rates of soldiers, bishops and sailors stationed in the colonies. They are claimed to give "a good indication of the mortality rates of settlers" (p. 1370) and the measure is largely based on the work of the historian D. Curtin.

Now, returning to the Rodrik et al. paper. The estimation is done for three different samples: The original sample used by Acemoglu, Johnson and Robinson (2001) of 64 countries, an 80-country sample for which the instrument on settler mortality is available and a sample with 140 countries where the authors replace the settler mortality instrument with the fraction of population speaking English and western European languages.

For the 64-country sample, their results show that it is only institutions that have a significant effect on income, and this effect is positive (a coefficient of 1.78). The same is true for the 80-country sample as well as the 140-country sample, but with a stronger effect of institutions on income, a coefficient 0f 2.0, in the 80-country sample, and a smaller effect in the 140-country sample, a coefficient of 1.32.

The authors find strong effects of geography on institutions and trade integration. Since geography is statistically insignificant in the regression with income as the dependent variable, the authors conclude that geography has only indirect effects on income, through institutions and trade integration. Furthermore, institutions have a weak effect on integration and integration has a weak effect on institutions.

2.3.1 CRITIQUE

The study replicated is the result of a tremendous amount of work. It is also considered to be influential and hence accurate. They use the then latest instruments, also considered to be robust. Since the Acemoglu and Rodrik paper is widely cited, the instruments are considered to be at least "usable", in the sense of trust. This is what it is all about – trustworthy results. Here is a perhaps rather naïve, yet describing example:

A study, SA, investigates some relationship. The study builds on some dataset DSA, whose characteristics regarding quality to a great extent is unknown. The relationships within SA is thoroughly investigated, by using DSA. Every single regression and hypothesis test is well-documented.

Another study, SB, investigates some other relationship. This study builds on another dataset, DSB. This is a well-known dataset; widely used and well-recognized. The study is however performed in a rather imprecise way and is not particularly transparent.

A third study, SC, builds on some third dataset DSC, not well-known and not transparently constructed. This is however discussed and contingencies are corrected. The methods are transparent and all steps are shown in detail, as in SA.

Which one of the studies is preferable? Of course, SC. The SB type is sadly too common. Many authors seem to think that as soon as a well-known dataset is used, anything goes. This is perhaps the same disease, but in a lighter version, which SA is contaminated by: If a discussion on the quality of data is missing, the reader cannot decide whether the study is as precise as it appears to be, by the perhaps significant results.

Why is this? There are two different types of studies: qualitative and quantitative. In a qualitative study, numbers are seldom crucial; at least there is no inference drawn based on them. In quantitative studies, the very inference is the study. Here, data and method are at the center of attention, since inference and hence the results are completely dependent on them. This seem to be as obvious as forgotten. The discussion on data quality, or even a discussion on, not a reference to, the data is, in most economic papers, dead. Or at least missing. On the contrary, in basically any statistical paper, this discussion is central.

Every empirical study begins with an idea as well as with data. On data follows the very technical performance, inference and conclusions. The orthodox reader might think that every study must be perfect in the sense of accurate and precise results to be considered valid. This is achieved by advanced econometric methods and intelligent processing of the data, in order to present stunning inference.

The authors of this paper claim this to be too narrow a perspective: Every study is important, no doubt about it. Every inference drawn, even if it clearly contradicts common sense, is in some sense valid, important and worthwhile – if it is totally and unconditionally transparent. As long as the reader can follow every step, she can also understand what can be interpreted or done differently. Here, it is of most importance to stress that most studies performed by economists – or, rather, empirical studies carried out by econometricians by training – do not lack this. The paper in focus, by Rodrik et al. is a brilliant example of this. Every single regression is remarkably thoroughly described, as well as the inference and reasoning behind the conclusions. Very well; every link in the chain is well-described and analyzed, but not the very foundation – the data, nor the methods.

Furthermore, on page 136 in their 2004 paper, (Rodrik et al., 2004), it is stated that

> *Much of our paper is devoted to checking for the robustness of central results. In particular, we estimate our model for three different samples /…/ Finally, we compare and contrast our results to those in some recent papers that have undertaken exercises of a similar sort. Where there are differences in results, we identify and discuss the source of the differences and explain why we believe our approach is superior to on conceptual or empirical grounds.*

This is of course an honorable approach, but what is actually done is nothing more than checking for cross-correlations and running OLS regressions on the very same data in different combinations. One problem is of course that there is no standard for ”robustness tests” and no consensus between the disciplines statistics and econometrics concerning this.

The authors would like to state that the paper *Institutions Rule* is an interesting piece of work. It is easily understood that a tremendous amount of work and sharp-thinking is put into effort and the result is most interesting. The authors' intentions with this paper are not to re-do the study exactly, nor to integrate the new parts in the prime study. Instead, it is to show the effects in terms of sensitivity of significant results when basal conditions or prerequisites change; for instance, or rather, specifically, data and methods. This puts the study in a context and makes it important.

One obvious issue that comes to mind is the choice of instruments and variables. According to what is written on page 135 in Rodrik et al. (2004), the instruments used are those who give the best hope at that point of time of “unraveling the tangle of cause-and-effect relationships involved”. One motivation of the choice is that they have passed the American Economic Review-test. According to the webpage of American Economic Association (AEA)[7], the American Economic Review test is a transparency and accessibility test regarding the possibility to replicate the study, not any kind of robustness- or feasibility test. This is however not proposed by Rodrik et al., but should be mentioned here. The instruments are actually not motivated further.

The choice of independent variables is also an issue. When examining the other indices in Kaufmann’s World Governance Indicators, it is discovered that none of them can be instrumented on with the instruments that Rodrik et al. use. The coefficients of the instrument are insignificant which violates one of the requisites for a good instrument.

[7] See http://www.aeaweb.org/aer/data.php

Therefore, Rodrik et al. have chosen Rule of Law, perhaps because it can be instrumented on using available and widely used instruments. This however does not serve as an argument as to why Rule of Law would be an appropriate measure of institutions.

The results might hint of a somewhat reversed engineering approach, in the sense of independent variables have being chosen on basis of the instruments – not the other way round, which is according to norm. This discussion is lacking in the Rodrik et al. paper.

Another interesting issue is raised in note 12 in the paper:

> *Note that these calculations omit the feedback effect from income to trade and institutions since we are unable to estimate these. Our numbers can hence be viewed as impact effects, taking both direct and indirect channels into account, but ignoring the feedback from income. (pp. 143)*

This is somewhat a – even if it is a rather strong expression – failure. The backward causality problem is of most importance to estimate and account for, in order to avoid spurious results. This is not done, leading to inexhaustive results and sheer doubt about them. In an alternative set up, presented in this paper, this is actually tested.

3 METHODOLOGY

3.1 A DISCUSSION ON CAUSALITY

The very aim of the paper by Rodrik et al. is to investigate whether the new instruments, claimed to be the very best, innovative and highly accurate, together produce significant results regarding (history of) domestic institutions in colonies on today's income. The method used by Rodrik et al. is two stage least squares regression, 2SLS. Intuitively, the method works like this:

One would like to estimate the effect of one variable on another, or on some set of variables. However, running an ordinary least squares regression only reports correlation and direct effect of the explaining variables. Omitted variable bias, confounding and backward causality are only a few problems that might arise. In order to correct these problems, the 2SLS method accounts for variation in the error term that is due to some exogenous source, affecting the dependent variable only through the independent variables.

Since the variation in the error term of the equation of interest is captured by the first stage of the method, the general standard is to claim a causal relationship between the instrumented variables and the dependent variable if the regression in the two stages are run properly and are proved significant. The method is a superb way of describing an equilibrium, or rather a set of equilibria which for instance would be described by the demand curve sliding on a fixed supply curve. The different observed equilibria of supply matching demand would otherwise be nothing but a cluster if plotted. But does the method really describe, or prove, causality?

A statistician would most likely say no, since causality within this setting ultimately is proven by arguments. There are of course several potential problems with the 2SLS method, like weak instruments. These are however technical problems. The 2SLS method not pointing out causality is however a methodological problem. Here is an example:

A farmer owns one acre of arable soil. In order to improve the results of her work, the farmer pours artificial fertilizer on the acre. The more fertilizer, the better the growth of the crops. If the relationship is 1:1, we will see an entirely linear trend if plotted. Here, proving causality is a matter of arguing. There are however other factors that would fit in the equation as variables, like weather or number of sun hours.

One important factor of cultivation is however for the soil to lie fallow. Suppose we have historical data on the frequency with which the farmer has let the soil to lie fallow and we construct a variable that describes the ratio of fallow to operational years and that the soil corresponds positively on fertilizer when having been fallow. Now the variable would generally be considered to carry a lot of information, which in the setting described above, would be in the error term. In the first course in regression analysis, one is taught that the independent variables must not be correlated with one another, but correlated with the dependent variable. In this particular case, suppose the fallow variable is not correlated directly with the dependent variable; the quantity of the harvest, but affects the dependent variable only through the other independent variables. If so, it would be a perfect scenario on which to apply the instrumental variable method.

Now, all time periods are in some sense concentrated to one single time period, which we are using. Suppose now that the first stage regression is fertilizer on the fallow variable and the second is quantity of crop on the in the first stage (hopefully) improved fertilizer variable. This is the full 2SLS, instrumental, method.

Well now, why so much text and besides, about some farming issues? The example has clear parallels to what is investigated in the Rodrik et al. paper. There are time series available, but not used. Instead of directly investigating whether artificial fertilizer or letting the soil lie fallow in fact affects the quantity of the harvest, one tries to catch some variability in the error term, which might originate from some exogenously given variable.

Consider the case when time series are used: On the same initial lecture on regression analysis, we learned that an equation with $n + m$ variables, where $m \geq 1$, always explains at least as much, or more, than an equation with n variables. The same should go for time observations. Several observations in time on a variable must per definition carry as least as much information as one observation.

The general opinion among economists of today, is that the 2SLS method provides a tool to test for exogenous variation and causality (Heckman, 2008; Wooldridge, 2009). However, causality must still be argued for in text. On causality, the 2SLS method adds no more information than does OLS. If it mathematically-technically and empirically is proved, however, that one variable historically drives another, then one can conclude that one variable causes another. In order to prove this, historically, one might would like to use time series. That is the fundamental idea behind testing for Granger causality. Here, one explicitly investigates whether one variable causes another. Of course, a whole lot of interesting things might be found in the error term. Causality might be just one of them.

The purpose of this passage is, which has to be stressed, not to falsify a generally accepted method, widely used and accepted by economists That would in some sense be an intellectual suicide. Robust and significant results are per definition considered as valid, but must in the cumulative spirit of Popper and Kuhn (Popper 1945; Kuhn 1962) be reconsidered, questioned and replicated.

Rodrik et al. used cross section data in their study. But after limited amount of work, time series were found from the very same sources. The obvious challenge is to investigate whether additional observations in time add more information. Replicating the work of other scientists and analyzing the results is the very foundation upon which much of the modern research relies.

The hypothesis is that the use of another method, still applied within the same framework, variables and basically the same data, will affect the results. That the method captures and presents significant results is undoubtedly true, but to prove causality, it is most likely not enough. Hence, when data is available, it is possible to test with another method, which in some sense would be more accurate and that is exactly what is undertaken within the framework of the authors' study.

In his 1969 article *Testing for causality and feedback* (Econometrica, 1969), Nobel laureate Clive Granger argues that if historical data of a variable Y adds significant information to a regression with a variable X on historical data of X, then one can claim that Y causes X. This is an indeed useful and hence widely used method, suitable for large datasets. Furthermore, and this is what really makes it of most interest, it is a method of *technically* showing causality (ibid); not only to argue by text. Based on this, the authors of this paper would like to make a humble attempt to employ the method within the below specified framework, in order to test the above described hypotheses.

3.2 TECHNICAL SPECIFICATION OF THE GRANGER CAUSALITY METHOD

The technical procedure of the Granger-method is as follows:

The method builds on using two separate, stationary time series: X_t and Z_t with expected value 0 (zero). First, a regression is run on X with values from earlier years of X (regression R1). Thereafter, another regression on X is run but now with lagged values of X in addition to earlier values of Z as explaining variables (see equation R2). In order to investigate whether earlier values of Z adds some explanatory power of explaining X, one then performs a standard F-test.

$$X_{i,t} = \gamma_1 X_{i,t-1} + \gamma_2 X_{i,t-2} + \ldots + \gamma_j X_{i,t-j} + \varepsilon_i \qquad \text{(R1)}$$

$$X_{i,t} = \gamma_1 X_{i,t-1} + \gamma_2 X_{i,t-2} + \ldots + \gamma_j X_{i,t-j} + \mu_1 Z_{i,t-1} + \mu_2 Z_{i,t-2} + \ldots + \mu_j Z_{i,t-j} + \varepsilon_i, \qquad \text{(R2)}$$

where $X_{i,t}$ and $Z_{i,t}$ in the models tested in this paper are the differences of values between years at a certain point in time. $X_{i,t}$ is thus the variable for country i at time t, where t is the (first) difference between the years. This is done to make the series stationary.

The Granger causality method makes it possible to test also for backward causality. Hence, the step above is repeated, but the other way round: Regressions of Z on earlier values of Z is run and thereafter regressions on Z with earlier values of Z and X. In this way the method is impartial in the sense that it does not presuppose a specific relationship – other than the *thought* represented by the *hypothesis* – nor does it add some weight in one way or another to the specification. Testing for backward causality lies within the foundation of the method.

The assumption that historical values of one variable contain predictive information about later observations, implies that the model is *autoregressive*. The number of lags to be included in (R1) is determined by first estimating an autoregressive model for X whereby only significant lagged values of X are retained to perform (R1). The number of Z lagged values to be included in (R2) is decided upon by augmenting (R1) only with lagged values of Z that are individually significant.

The null hypothesis, H_0, is that the additional variables (Z) do *not* help explaining X:

$$H_0 = \mu_1 = \mu_2 = \ldots = \mu_j = 0$$

$$H_1 = \text{at least one of } \mu_{1,} \mu_{2,\ldots,} \mu_j \neq 0$$

The null hypothesis is rejected if i) at least one significant lagged value of Z is retained in regression two, *and* ii) if the lagged values of Z provide additional explanatory power to the regression, i.e. if they are jointly significant in regression two.

The latter is tested by an F-test regarding the additional variables:

$$F_{obs} = \frac{RSS_1 - RSS_2 \,/(p_2 - p_1)}{\mathrm{RSS}_2 \,/(n - p_2)}, \text{ where}$$

RSS_1 is the residual sum of squares from regression one (R1),

RSS_2 the residual sum of squares from regression two (R2),

p_2 and p_1 is the number of explaining number of explaining variables in regression two and one respectively,

and

n is the number of observations.

If $F_{obs} > F_{p_2 - p_1, n - p_2}$,

then the null hypothesis is rejected and one can, based on the model and the tests, argue that Z Granger causes X.

Many economists claim that the method is a matter of prediction. But if we stand at $t = 0$ and throw a ball to a person at $t = 1$, we can predict that the ball will land by the person in $t = 1$ and the person in $t = 1$ will of course realize that the cause of the ball hitting her on the head was us throwing it. Causality and prediction are in this sense hence the two sides of a coin.

3.3 Data sources and variables

Following variables are used in our analysis:

Variable	Measuring	Explanation	Source	Collected
LCGDP	Income	Logarithm of GDP per capita	Penn World Tables/IMF/World Bank	20110404
DISTEQ	Geography	Distance from the equator in degrees	Rodrik et al.	20112503
RULE	Institutions	Rule of Law index	Kaufmann	20112503
LCOPEN	Integration	Openness as a trade-to-GDP ratio	Penn World Tables	20112503
LOGEM4	Institutions	Settler mortality (see 2.3)	Acemoglu et al	20112503
LOGFRANKROM	Constructed Openness	Openness as a trade measure, weighted for the distance to markets	Frankel and Romer	20112503
ENGFRAC	Fraction of population speaking English		Rodrik et al.	20112503
EURFRAC	Fraction of population speaking other Western European languages		Rodrik et al.	20112503

As seen from the table, several sources of GDP per capita have been used since this is one of the basic ideas of the paper. It is clearly specified in each case what source is used. Below is a table with descriptive statistics of the variables used.

Table 3.1. Descriptive statistics for the variables used in IV-regressions

Variable	Obs	Mean	Std. Dev.	Min	Max
LCGDP_PW	139	8,407	1,141	5,770	10,450
LCGDP_WB	139	8,213	1,302	5,656	10,570
LCGDP_IMF	139	8,258	1,313	5,584	10,609
RULE	139	0,000	1,000	-2,323	1,926
LCOPEN	139	4,022	0,566	2,550	5,780
DISTEQ	139	0,000	1,000	-1,448	2,469
LOGFRANKROM	139	2,929	0,801	0,830	5,640
LOGEM4	79	4,671	1,218	2,150	7,990
ENGFRAC	139	0,078	0,243	0,000	1,000
EURFRAC	139	0,242	0,387	0,000	1,000

3.4 EMPIRICAL PERFORMANCE

The analysis performed in this paper is begun by replicating Rodrik et al.'s original results using the original data used in their paper. Three different samples are used in this setting; i) the original 64-country sample used by Acemoglu et al. (2001) ii) a larger 79-country sample for which the instrument on settler mortality is available and iii) a 138-country sample which includes countries that were never colonized where the fraction of population speaking English and other Western European languages is used as an instrument for institutions instead of settler mortality. Zimbabwe is excluded from this sample since it was found to be an outlier.

In the first stage of the two-stage-least-squares regression we regress rule of law and log of openness on the instruments settler mortality, constructed trade shares by Frankel-Romer and distance from the equator in the 64-country sample as well as the 79-country sample. For the 138-country sample, we regress rule of law and log of openness on the instruments fraction of people speaking English and other Western European languages, constructed trade shares by Frankel-Romer (1999) and distance from the equator.

The predicted values of rule of law and log of openness from the first stage are used in the second stage regression. Here, log of GDP per capita is regressed on the predicted values of rule of law and log of openness as well as distance from the equator.

The next step in our analysis involves exchanging the original data of GDP per capita from Penn World Tables to exactly the same measure from the World Bank and IMF, respectively. A comparison is made of the coefficients produced from data from Penn World Tables vis-à-vis data from the World Bank and IMF.

The next step in the analysis is to use the Granger causality method to investigate the causal relationship between income, institutions and openness to trade. The different relationships are investigated for the same three samples used in the two-stage least squares method. The Granger-causality method makes it possible to check for reverse feedback from income to openness and/or trade. The instrumental variable technique however used by Rodrik et al. to establish for causality ignores this effect. Our hypothesis is that the Granger-causality method will yield substantially different results compared to the instrumental variable method since plausibly – and as revealed by previous time-series studies – there is a feedback effect from income to institutions as well as to openness to trade.

First, we present results from the original Granger-causality test where we include only the lagged values that are significant. This specification is the most accurate and it is upon these results that we base our analysis. However, since there are only a few lagged variables that are significant, information is lost when excluding some lagged variables. Therefore, we present an alternative specification where we include all lagged values. These results should be interpreted with caution since it is not certain that all the lagged variables add any additional explanatory information to the regressions. Hence, the degrees of freedom change substantially. Instead, the results are presented to highlight how much results may differ if one does not choose appropriate lags to include in the regressions.

3.5 POTENTIAL PROBLEMS AND REMEDIES

Something on potential problems with the method should be mentioned: X_i and Y_i are assumed to be stationary time series with expected value zero (0). One might suspect that the time series do not own this latter property, but this is in fact not a problem, since the regression coefficients would be the same, given a model with an intercept, for centered values of X and Y. The former condition, however, might imply potential problems: Times series of GDP are, as commonly known, usually not stationary, nor heteroskedastic. Here, two-dimensional time series are used with many observations among countries, but not that many observations in time. Hence, this implies that testing for stationarity will be yield incorrect results, if not being impossible. This problem has, however, one nice property: Based on earlier studies (Gerdham and Löthgren, 2000), it is assumed that the time series are integrated of 1st order I(1). This implies that they will likely be stationary by taking first differences. This is hence done and the regressions are run on the corresponding values and will hence not be biased or leave spurious results.

Furthermore, it must be mentioned that there is always a way to further expand and broaden the scope of a study, but limits must due to scarce resources be decided on. One example of this stems from a methodological problem: Consider the two time series represented by the variables X and Y. In this case, say, X causes Y and this is empirically proven according to the method used. There is however a possibility that a variable Z, perhaps also with different lags, actually drives the relation. This would be a problem of exogenous variation, which is a common problem within statistics. The variable Z would, if so, be e.g. a confounding factor or an omitted variable, implying the relationship between X and Y to be spurious. If the causal effect is estimated by e.g. the Granger causality method, the final regression is X on lagged values of X and Y with the variable Z not present. Hence, there would be a confounding problem, since the factor Z would be correlated with both the dependent and independent variables.

The observant reader will soon understand that the method used to estimate causality, the Granger causality method, is only used in a bivariate set up; not multivariate. However, a nice property of the general set up used in this paper, is that the studies done before this one, are based on two-stage least squares regression (2SLS). One of the main features of the 2SLS method, without describing it in detail, is that it corrects for the exogenous variation in the error term of the main specification. Hence, the omitted variable bias is eliminated. Since the confounding problem is related to correlation between some variable in the error term and both the dependent and independent variables, the 2SLS method would *also* eliminate the problems within a framework where a regression of X on lagged values of itself and Y is run – a confounding problem within this specific framework; not only the omitted variable bias.

This study cumulatively builds upon great research conducted by both Rodrik et al. and Acemoglu et al., which both have dug deeply into the matter of relevant and significant variables and decided on only a few variables and instruments, *which in turn are used within this study*. If Rodrik et al. and Acemoglu et al. would have gotten results that indicate that the relation is spurious and insignificant, then there would have been a reason to actively have searched for this third factor Z (see also Wooldridge, 2009). The results presented in both Rodrik et al. (2004) and Acemoglu et al. (2001) are however significant and correct according to their set up and methods and hence *in that sense* trustworthy. This is motivation enough to keep working within this framework with the variables in question and not to actively start searching for a third driving factor Z. This factor might of course however be present, but apparently not at a disturbing rate (the search of such a variable is certainly a splendid topic for the further research within the field). This is an extremely important passage, motivating the very empirical set up of the project.

The very observant reader will also soon understand that this third driving factor, which according to the passage above is eliminated, in fact might not be some random exogenous variable, lurking in the shades. It might well be one of the variables included and discussed in the set up. If so, a multivariate analysis would be appropriate to sort the hidden relationships out. However, including the variable would not reveal much information about the very strength of the relationship, which instead stratification would do: If the relationship studied changes a lot between the strata, one could conclude that the variable stratified by is the confounding variable.

In this particular case, there are three samples as described above. They are not explicitly stratified by for instance income, but it is reasonable to think that this would be an appropriate variable to stratify by. In part 4.2, one will see that the relationship regarding income changes significantly between the samples. Hence, the problem of bi- versus multivariate analysis is in this respect dealt with and a bivariate analysis is hence more appropriate and in accordance with Occam's razor.

4 EMPIRICAL FINDINGS

4.1 RODRIK ET AL. METHOD – TWO STAGE LEAST SQUARES

First, a pure replication of Rodrik et al.'s main specification is performed. This is done in order to check that the data treatment and the method used in the regressions with other data sources is the same as the authors of the replicated paper are using. The results are presented in table 4.1.

Table 4.1. Replication of Rodrik et al.'s original results

	1	2	3	4	5	6	7	8	9
Second stage results	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP
Geography (DISTEQ)	0,726	-0,319	-0,549	0,789	-0,243	-0,644	0,736	0,009	-0,138
	(-4,49)*	(-1,21)	-0,19	(5,25)*	(-1,01)	(-1,40)	(9,87)*	(-0,08)	(-0,85)
Institutions (RULE)		1,536	1,780		1,500	1,955		1,112	1,328
		(5,03)*	(3,87)*		(5,34)*	(3,70)*		(7,24)*	(5,92)*
Integration (LCOPEN)			-0,302			-0,484			-0,296
			(-0,42)			(-1,17)			(-1,50)
Number of observations	*64*	*64*	*64*	*79*	*79*	*79*	*139*	*139*	*139*
R-squared	*0,25*	*0,55*	*0,56*	*0,26*	*0,51*	*0,53*	*0,42*	*0,55*	*0,55*
Adj R-squared	*0,23*	*0,54*	*0,54*	*0,25*	*0,50*	*0,51*	*0,41*	*0,54*	*0,54*
First stage results		RULE	LCOPEN		RULE	LCOPEN		RULE	LCOPEN
Geography (DISTEQ)		0,458	-0,146		0,505	-0,119		0,623	-0,027
		(3,21)*	(-2,00)**		(3,69)*	(-1,60)		(10,88)*	(-0,81)
Settler morality (LOGEM4)		-0,333	-0,146		-0,294	-0,144			
		(-4,10)*	(-3,50)*		(-3,93)*	(-3,53)*			
Constructed openness (LOGFRANKROM)		0,246	0,662		0,227	0,580		0,310	0,508
		(1,95)***	(10,29)*		(2,10)**	(9,79)*		(4,30)*	(12,35)*
Population speaking English (ENGFRAC)								0,752	0,428
								(2,70)*	(2,69)*
Population speaking other Western European languages (EURFRAC)								0,419	-0,177
								(2,38)**	(-1,75)***
F-statistic		*17,23*	*41,51*		*18,46*	*38,68*		*43,44*	*41,9*

*,**,*** =significant at 1,5 and 10% level, respectively

The coefficients in each of the regressions are in almost all cases identical to the coefficients presented in the paper by Rodrik et al. The same is true for the t-statistics. Occasional differences in coefficients are extremely small – at most they differ with two decimal places – and are likely due to the exclusion of Myanmar from our samples since GDP data was unavailable in the data set published on the website. For reasons illustrated next, it was important to use the original data used by Rodrik et al.

Since it is our belief that changing the data *source*, not measure, for a variable could be sufficient to change the results, we used Rodrik et al.'s original data to replicate the results above. Since the GDP measure is complied as an estimate, changing the data source should in our belief provide different coefficients. As a comparison to the PPP-adjusted GDP per capita data in US dollars from Penn World Tables used by Rodrik et al. we use data on PPP-adjusted GDP per capita in US dollars from the World Bank. The results of the second stage regressions are presented in table 4.2. The results from the first stage regressions are the same as in table 4.1 since GDP per capita does not enter these regressions.

Table 4.2 Replication of Rodrik's analysis changing the data source to World Bank

	1	2	3	4	5	6	7	8	9
Second stage results	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP
Geography (DISTEQ)	0,717	-0,439	-0,728	0,791	-0,366	-0,865	0,795	-0,079	-0,257
	(3,78)*	(-1,40)	(-1,47)	(4,50)*	(-1,23)	(-1,51)	(9,03)*	(-0,52)	(-1,19)
Institutions (RULE)		1,698	2,007		1,682	2,248		1,337	1,599
		(4,66)*	(3,62)*		(4,85)*	(3,42)*		(6,57)*	(5,35)*
Integration (LCOPEN)			-0,382			-0,602			-0,359
			(-0,85)			(-1,17)			(-1,37)
Number of observations	64	64	64	79	79	79	139	139	139
R-squared	*0,19*	*0,46*	*0,56*	*0,21*	*0,45*	*0,48*	*0,37*	*0,52*	*0,53*
Adj R-squared	*0,17*	*0,44*	*0,54*	*0,20*	*0,44*	*0,46*	*0,37*	*0,51*	*0,52*

*,**,*** =significant at 1,5 and 10% level, respectively

Using World Bank data for PPP-adjusted GDP per capita in US dollars instead of the same data from Penn World Tables does in fact change the size of the point estimates. It also changes the confidence intervals, however not enough to change the significance level established in Rodrik's original IV-regressions. The adjusted R-squared as well as R-squared change when using the World Bank data set. It is of importance to note that the measure for GDP in Penn World tables and the GDP measure in the data from World Bank does in fact claim to measure the very same variable, PPP-adjusted GDP per capita in US dollars measure, yet the size of the coefficients change. This highlights that when one interprets the point estimates, one should be careful when interpreting the size of the coefficients.

To ensure that the differing results are not merely a matter of coincidence due to characteristics specific for the World Bank GDP data we perform the same regressions using data on PPP-adjusted GDP per capita in US dollars from the IMF. The results from the second stage regressions are presented in table 4.3 below.

Table 4.3 Replication of Rodrik's analysis changing the data source to IMF

	1	2	3	4	5	6	7	8	9
Second stage results	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP
Geography (DISTEQ)	0,676	-0,484	-0,719	0,773	-0,342	-0,775	0,812	-0,066	-0,245
	(3,62)*	(-1,57)	(-1,51)	(4,48)*	(-1,21)	(-1,46)	(9,21)*	(-0,43)	-0,253
Institutions (RULE)		1,704	1,955		1,621	2,111		1,342	1,605
		(4,77)*	(3,68)*		(4,92)*	(3,47)*		(6,68)*	(5,45)*
Integration (LCOPEN)			-0,310			-0,521			-0,360
			(-0,72)			(-1,09)			(-1,39)
Number of observations	64	64	64	79	79	79	139	139	139
R-squared	*0,17*	*0,48*	*0,49*	*0,21*	*0,45*	*0,46*	*0,38*	*0,53*	*0,53*
Adj R-squared	*0,16*	*0,47*	*0,47*	*0,20*	*0,43*	*0,44*	*0,38*	*0,52*	*0,52*

*,**,*** =significant at 1,5 and 10% level, respectively

Using IMF data for PPP-adjusted GDP per capita data provides point estimates different in size as compared to the point estimates using data from Penn World tables. Hence, the same conclusion can be made regarding the importance of being careful when interpreting the point estimates. Next, we show how much the significant point estimates change when using World Bank data as compared to using Penn World tables and IMF data as compared to using Penn World tables, respectively. The differences are summarized in table 4.4 and 4.5.

Table 4.4 Differences in point estimates using GDP data from World Bank vs. Penn World tables

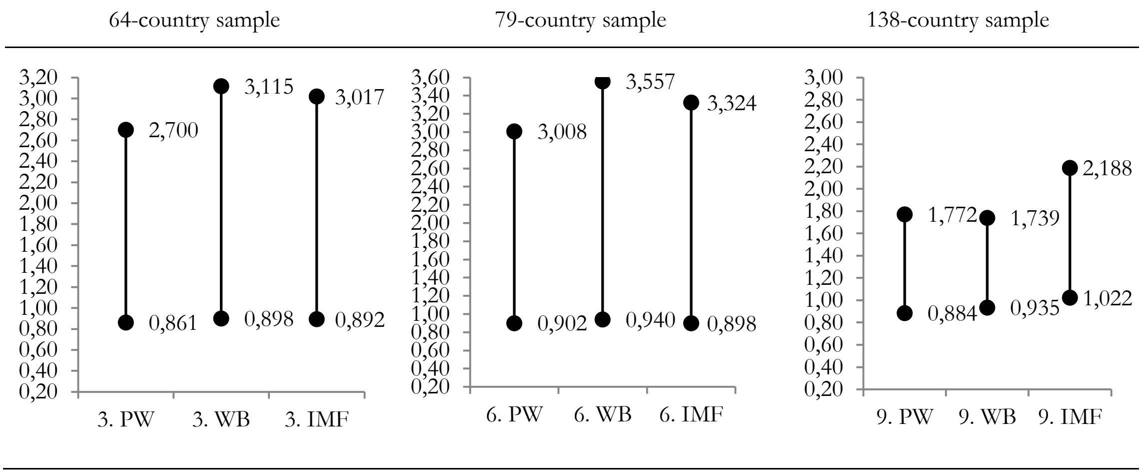
	1	2	3	4	5	6	7	8	9
	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP
Geography (DISTEQ)	-1,25%			0,22%			8,09%		
Institutions (RULE)		10,54%	12,70%		12,16%	15,01%		20,29%	20,42%
Integration (LCOPEN)									
Number of observations	64	64	64	79	79	79	139	139	139

Table 4.5 Differences in point estimates using GDP data from IMF vs Penn World tables

	1	2	3	4	5	6	7	8	9
	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP	LCGDP
Geography (DISTEQ)	-6,88%			-2,06%			10,34%		
Institutions (RULE)		10,97%	9,79%		8,07%	8,00%		20,75%	20,86%
Integration (LCOPEN)									
Number of observations	64	64	64	79	79	79	139	139	139

As seen from the table above, the point estimates differ a lot; up to almost 21%. This is undeniably a large difference although the point estimates of the beta coefficients from table 4.1-4.3 do not differ significantly. However, a closer look at the 95% confidence intervals when they are compared between the PW, WB and IMF datasets respectively, the end points differ very much indeed. The 95% confidence intervals for the rule of law coefficient corresponding to regression three, six and nine in table 4.1-4.3 is presented below.

Table 4.6 95% confidence intervals for the coefficient for rule of law



Hence, the impact of for instance the RULE index, which represents the impact of historical institutions, might be under- or over estimated by 85% or 18,3% respectively.

4.2 GRANGER CAUSALITY ANALYSIS

To perform the Granger-causality analysis we have focused on the un-instrumented measures of income, institutions, trade and geography. These measures are the natural log of per capita GDP, rule of law index, natural log of imports and exports share of GDP and distance from the equator in degrees, respectively.

Each lag of log of per capita GDP represents a difference in income from one year to another, for a vector of countries, i.e. $LCGDP_{i,t}$ represents the difference in income between the latest available year for our dataset and the year before. The most recent difference in log of GDP per capita is regressed on previous differences in income between years.

Next, log of GDP per capita is regressed on previous differences in income between years as well as previous differences in the rule of law index between years. The choice of how many lags to include is based on the number of lags that are significant.

$$LCGDP_{i,t} = \gamma_1 LCGDP_{i,t-1} + \gamma_2 LCGDP_{i,t-2} + \ldots + \gamma_j LCGDP_{i,t-j} + \varepsilon_i$$

$$LCGDP_{i,t} = \gamma_1 LCGDP_{i,t-1} + \gamma_2 LCGDP_{i,t-2} + \ldots + \gamma_j LCGDP_{i,t-j} + \mu_1 RULE_{i,t-1}$$
$$+ \mu_2 RULE_{i,t-2} + \ldots + \mu_j RULE_{i,t-j} + \varepsilon_i$$

One obtains the residual sum of squares from regression one and two to calculate the observed F-value as according to the technical specification above.

If the observed F-value is larger than some critical value, one can *reject* the hypothesis that the additional variables, in this case rule of law, do not help explain log of GDP per capita. Hence, if the observed F-value is larger than the critical value, one can conclude that rule of law Granger causes log of GDP per capita.

To check if the causality runs the opposite way, one first regresses the latest difference between years in the rule of law index on previous differences between years in the rule of law index. The same regression but including previous differences between years in log of GDP per capita, is then run. If the observed F-value is larger than the critical value, one rejects the hypothesis which states that GDP does not help explain rule of law, i.e. one can conclude that GDP per capita Granger causes rule of law.

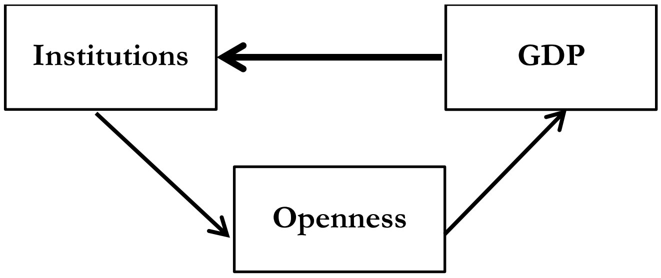
Next, we report our results together with the calculated observed F-values. F-values in bold represent a relationship on at least 10% level of significance.

Table 4.1. Granger causality relationships

64-country sample		79-country sample		138-country sample	
Fobs | LCGDP causes RULE	**6,351**	Fobs | LCGDP causes RULE	**3,123**	Fobs | LCGDP causes RULE	0,000
Fobs | RULE causes LCGDP	1,277	Fobs | RULE causes LCGDP	0,000	Fobs | RULE causes LCGDP	**3,910**
Fobs | LCGDP causes LCOPEN	2,324	Fobs | LCGDP causes LCOPEN	**5,060**	Fobs | LCGDP causes LCOPEN	**5,195**
Fobs | LCOPEN causes LCGDP	**5,333**	Fobs | LCOPEN causes LCGDP	**5,978**	Fobs | LCOPEN causes LCGDP	0,000
Fobs | RULE causes LCOPEN	**5,892**	Fobs | RULE causes LCOPEN	0,000	Fobs | RULE causes LCOPEN	0,000
Fobs | LCOPEN causes RULE	1,424	Fobs | LCOPEN causes RULE	0,000	Fobs | LCOPEN causes RULE	**21,325**

Next, we summarize the relationships and the respective significance levels, sample by sample. The beta coefficients referred to in the text can be found in the regression output from regression two in Appendix.

Figure 4.1. Granger-causality relationships in the 64-country sample



As graphically illustrated above, institutions Granger cause openness at 5% significance level. A possible interpretation of the beta coefficient from the regression that implies this relationship, is that an improvement (deterioration) of the rule of law index with one unit (1) will drive an improvement (deterioration) in the openness measure two years ahead with 0,48. Since the rule of law index stretches from -2,5 to 2,5, an improvement of one unit is rather unlikely, however, the relationship also holds for smaller changes than one unit, which could yield large changes in the openness measure.

A closer look at the countries in the 64-country sample provides a possible explanation of the relationship: The sample contains predominantly developing countries where an improvement of the rule of law matters to increase openness. This is in line with Anderson's (2001) argument that bad institutions affect trade negatively through high costs and risks and Méon and Sekkat's (2004) findings that deterioration in the quality of institutions in the form of rule of law, among others, is associated with lower performance of exports of the MENA countries.

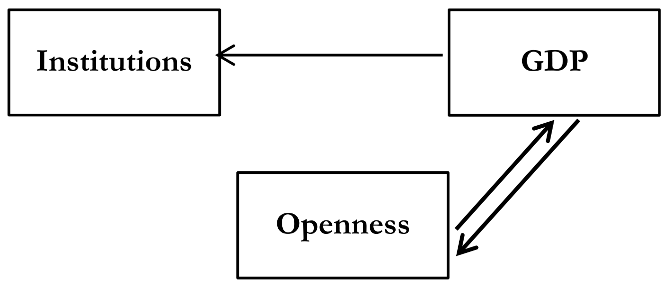
The relationship between openness and GDP is significant at 5% and suggests that if openness is increased with one (1) unit, the year-to-year change in (log) GDP increases by 0,21 percentage points. Due to scarce resources and usually lack of physical capital, trade can serve as a basis for access to physical capital in developing countries. Evidence for export-led growth has previously been found for Indonesia, Malaysia, Pakistan, Taiwan and Thailand (Ahmad and Harnhirun, 1996; Furuoka, 2007; Ghartey, 1993; and Chaudry et al., 2010) which are included in our 64-country sample.

GDP per capita Granger causes institutions at 1% significance level. One would expect the reverse relationship, especially since this sample contains countries that have low level of institutional quality. In other words, starting at a low level of rule of law, an improvement of the index should plausibly lead to a more favorable investment climate, increased foreign direct investment and hence eventually increases in GDP per capita (as according to e.g. Chong and Calderon, 2000). However, if only institutional improvements which are *unrelated* to GDP have occurred over the time period we have investigated, a relationship where institutional improvements affect GDP will not be present in the data. This is in fact likely since the rule of law index contains measures that might be unrelated to GDP. This might be the explanation for why the causality does not run from institutions to GDP.

The beta coefficients for the two lagged values of GDP that have a significant effect of institutions show ambiguous results. The changes in GDP values from 1998 to 2000 affect institutions negatively (the effect is as small as -0,07) while the change in GDP from 2007 to 2008 have a positive effect on institutions, with an effect of 0,26.

Next we analyze the causal relationships adding 15 countries to the 64-country sample.

Figure 4.1. Granger causality relationships in the 79-country sample

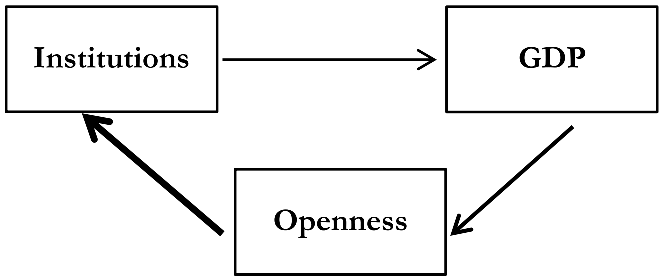


In the 79-country sample, the causal relationship between institutions and openness disappears. Furthermore, the causality runs both ways between openness and GDP at 5% significance level while the causality running from GDP to institutions is significant at 10% compared to 1% level of significance in the previous sample. The countries added to this sample include Benin, Burundi, Central African Republic, Chad, Fiji, Laos and Rwanda, among others. These countries have in common the fact that they are all underdeveloped - for instance, Central African Republic ranked 159th of 169 countries on the Human Development Index (Human Development Report, 2010) and Chad ranks as the 7th least developed country in the world according to the same source. Furthermore, their economies are based on agriculture with small industrial sectors - in Laos, 80% of the population live on subsistence farming - and therefore heavily trade dependent and integrated with their neighbours (The World Factbook, CIA). In the 79-country sample, openness to trade still causes GDP but because heavily trade dependent countries are added to the sample, the causality also runs in the reverse direction.

The significance level of the causal relationship from GDP to institutions decreases when we add more countries that are politically unstable. Fiji, for instance has suffered from several military coups[8] and Chad was declared the most corrupt country in the world by Transparency International in 2005 (Corruption Perceptions Index, 2005). In most of these additional countries, corruption has been present for a long time, and GDP has not increased enough to have an effect on institutions (see also next part).

The causal relationship between openness and institutions disappears when adding 15 politically unstable and trade dependent countries. The reason for this is probably that since they are in fact heavily trade dependent, their institutional improvements, or more likely deterioration in the quality of institutions, are unrelated to their need to trade.

Figure 4.1. Granger causality relationships in the 138-country sample



The largest sample adds countries that have experienced improvements in institutions that have driven increases in GDP per capita over the time period that we have examined. These countries include Albania, China, Russia, Turkey and a large sample of EU countries. When adding these countries, the causality between institutions and GDP changes direction. This implies that policies affecting GDP must have been carried out during the time period in question.

The beta coefficient corresponding to the differences in log of GDP per capita on rule of law is 0,18. The interpretation of this would be that an increase in the rule of law index by one (1) unit drives an increase in the differences between years in log of GDP per capita.

[8] See for example http://news.bbc.co.uk/2/hi/asia-pacific/country_profiles/1300477.stm

This means that implementing policies that lead to a change in the rule of law index by one unit leads to a future average rise in the change of log of GDP per capita by 0,18 – irrespective of a country's growth rate. A time-series study of the relation between institutional quality and GDP growth in Hong-Kong (Groenewold and Tang, 2005) supports these results.

The relation between openness and institutions is significant at 1 % level and shows that causality runs from openness to institutions. Since these results did not prevail in the previous samples, we can safely conclude that it is the additional countries that drive these results. Out of the 59 additional countries added to this sample, 21 were or became EU members during the period under examination. The more countries trade with each other over time, the more likely it is that its economic institutions will be integrated. In EU, which started with common economic institutions, further increases in trade and integration inevitably led to common political institutions. The findings in this sample are in line with Winter's argument that openness to trade can have an effect on institutional development. The reason is provided by Wei (2000); more open countries suffer more from corruption – which is part of the rule of law measure-because corruption is disproportionately connected to foreign trade. The same should be true for the heavily trade dependent countries in the 79-country sample. However, once again, if corruption has not changed over the period under examination such a relationship will not be shown in the data. Regarding EU, as well as the other additional countries in this sample, decreases in corruption as well as improvements in rule of law has in fact taken place – they are one of the prerequisites

In this sample, GDP causes openness at 5% level of significance. The feedback from openness to GDP disappears when including the additional countries. The reason is that these countries, as the additional countries added to the 79-country sample, are heavily trade dependent. With high level of economic integration comes specialization where some countries produce certain goods and services in which they have a comparative advantage. Since certain goods and services can be acquired more easily in other countries, when GDP per capita increases domestic demand is met by importing from foreign countries. These findings are in line with Ghartey (1993) and Henriques and Sadorsky (1996) who found that GDP causes export growth in USA, which has similar characteristics to the countries added to the 139-country sample.

4.3 Granger causality, alternative specification

Since lagged values of explanatory variables contain predictive information on the dependent variable, one loses information when excluding these. It should be noted that while one should only include lagged values that are individually and jointly significant. This is only possible when one has access to data sets that include many time observations and preferably with high frequency, say quarterly or monthly data.

Within our framework, we have only had access to ten time observations back in time. With annual data, this usually results in having one or two lagged values of explanatory variables that are significant. It is possible that if one had access to data further back in time, one would likely find more lagged values that affect the dependent variable significantly. For this reason, we choose to present results including *all* lagged values that were available.

Table 4.1. Granger-causality relationships, alternative specifications

64-country sample		79-country sample		138-country sample	
Fobs | LCGDP causes RULE	0,467	Fobs | LCGDP causes RULE	0,330	Fobs | LCGDP causes RULE	0,440
Fobs | RULE causes LCGDP	1,551	Fobs | RULE causes LCGDP	1,491	Fobs | RULE causes LCGDP	**2,149**
Fobs | LCGDP causes LCOPEN	0,596	Fobs | LCGDP causes LCOPEN	0,598	Fobs | LCGDP causes LCOPEN	0,908
Fobs | LCOPEN causes LCGDP	**1,783**	Fobs | LCOPEN causes LCGDP	**2,135**	Fobs | LCOPEN causes LCGDP	0,550
Fobs | RULE causes LCOPEN	1,224	Fobs | RULE causes LCOPEN	1,149	Fobs | RULE causes LCOPEN	0,941
Fobs | LCOPEN causes RULE	1,383	Fobs | LCOPEN causes RULE	0,858	Fobs | LCOPEN causes RULE	0,319

The results indicate that one should be careful when choosing what lags to include in the regressions. The results presented in our alternative specification of the Granger-causality regressions do not contradict any of the findings found in our original specification. However, they do not reveal some of the causal relationships found when including only significant lagged values of each variable.

The only causal relationship we can establish within the 64-country sample is the one running from openness to income, and this is only at 10% level of significance. As can be seen from the results, it is found that openness Granger causes GDP per capita also in the 79-country sample, this time at 5% level of significance. The only causal relationship found in the largest sample used is the one running from institutions to GDP per capita. This indicates that institutions, as measured by the rule of law index Granger cause GDP per capita. This relationship is valid on 5% level of significance.

5 CONCLUSION

The purpose of this study was to highlight the importance of performing thorough robustness tests for the validity of empirical results. Furthermore, this study aimed at, using what is in our opinion a more appropriate method, determining the causality between income, institutions and openness to trade – variables that are usually examined in a cross-section setting.

As shown by the analysis in the previous section, robustness tests should not be limited to testing one's results for differences among samples of countries. Instead, awareness is needed of the fact that only changing the data *source* could result in different point estimates. In the case of the Rodrik-study, these point estimates differ up to 21%.

By the same argument, robustness tests should moreover not be limited to only one method. This study employed time-series techniques in the form of Granger causality analysis to investigate the causality between income, institutions and trade openness. The method is more appropriate than the two-stage least squares method since it allows one to quantify the reverse feedback effect from income to institutions and trade openness. Rodrik et al. state that they cannot estimate this reverse feedback effect. The Granger causality analysis reveals that there *is* in fact such a reverse effect and that it is, in most instances, substantially large.

In contrast to Rodrik et al., we do not find that institutions affect income for all the samples under examination. In fact, we find that causality runs from income to institutions in the two smaller samples that include mainly developing countries. Our finding of a significant causal relationship between trade openness and income is in contrast to the one found by Rodrik et al. who find that the openness variable is insignificant once controlling for institutions. Since our method is a bivariate analysis, the relationship between openness and growth does not take into account the possibility that institutions may be driving the relationship. Despite this, since the different samples can be considered to represent different strata on institutional quality, the pair-wise causality tests for the different samples account for the effect of institutional quality on the relation between openness and growth.

Furthermore, the finding of *no* causal relationship between openness and institutions in the 79-country sample indicates that even if one had controlled for institutions in the relation between openness and growth.

In conclusion, our findings indicate that causal relationships between income, institutions and trade openness rely heavily on the countries under examination which further highlights the importance of robustness tests.

6 DISCUSSION AND POLICY IMPLICATIONS

The existence of every study needs to be motivated. Of course, every study is per definition contributive to what might be called the general knowledge of the world, irrespective of the results and perhaps also the quality of the work, but still, its existence needs to be motivated. This is a direct result of Kuhn's (1962) and Popper's (1945) thoughts on science: As long as we believe in the current paradigm, we do not even have thoughts on another one. Only for the, sometimes painful, exodus from one paradigm to another, we need to present new, overwhelming facts. In addition, we also consider a hypothesis to be falsifiable and not undoubtedly true. This is closely related to the simple question "why is this interesting?"

One of the purposes of this paper was to highlight the importance of thorough robustness tests. As shown, the results from the Rodrik et al. study differ by changing method and become unreliable when still employing the original method and set up, but changing the data source. The conclusion must be that the results, although the methods and data have undergone and passed robustness tests of the correlation- and regression type, are in fact not robust. This suggests several implications: Economists must take data more seriously. If some certain relationship is proven significant and – ceteris paribus – the method is changed, the results should not change. In this case, the instrumental variable method, which is considered a method solving the errors-in-variables problems, and the Granger causality method are at stake with each other, presenting different, but still significant results. Since the Granger causality method relies on more data, here from the same source, it is reasonable to consider its results as reliable. And even if the in the know reader distrusts the results, she should at least find the differences interesting.

As stated, such an easy maneuver as changing the data source of *one single variable* puts whole a study at stake. GDP is generally considered a stable and at least after a few years highly reliable variable. The results are hence remarkable.

Rodrik et al. claim that their study gives little guidance in policy decisions. This paper has on the contrary clearly shown the causal directions of the same variables, which from the policy perspective certainly is helpful: For instance, the impact of general integration among the EU countries is revealed and demonstrated. If GDP is increased, trade will apparently increase by a quantified factor. This in turn has effect on the rule of law variable, which has impact on GDP. An appealing feature of the method used, is that it provides a possibility to check for backward causality. In some cases, this backward running cause of events is highly present. This also puts gravity into the matter of correct checking for robustness.

The impact of the factors investigated on each other is clearly a result of a purposeful policy, which on average would have the same effect also in other regions. However, the results are not entirely universal. How the causes of effects differ among the samples are however shown.

Altogether, it must be stressed that causality analyses on policy effects are possible to perform. *How* the policies should be *implemented* is however a question that is not answered by the set up presented in this paper. This must be investigated and answered on a somewhat other, lower, scale.

Finally, some word on future research: Clearly, more case studies among countries need to be performed. Only then it is possible to show the exact relations and impacts among the factors – variables – studied. Furthermore, a standard for thorough robustness test must be developed. Otherwise, scientists will keep presenting results that are at stake with each other, when they perhaps should not be.

REFERENCES

Acemoglu, Daron; Johnson, Simon; Robinson, James A. (2001), The Colonial Origins of Comparative Development: An Empirical Investigation. *The American Economic Review*, Vol. 91, No. 5 (Dec., 2001), pp. 1369-1401

Acemoglu, Daron; Johnson, Simon; Robinson, James A. (2005), Institutions as the Fundamental Cause of Long-Run Growth, *Handbook of Economic Growth*, North Pole: 2005

Ahmad, J. and Harnhirun, S., (1996), Cointegration and Causality between Exports and Economic Growth: Evidence from the ASEAN Countries. The Canadian Journal of Economics, 29 (2), pp 406-413

Alfaro, Laura; Kalemli-Ozcan, Sebnem; Volosovych, Vadym (2008), Why Doesn't Capital Flow From Rich Countries To Poor Countries? An Empirical Investigation. *The Review of Economics and Statistics.* 90(2), 347–368

Anderson, J. E. (2001), Trade and Informal Institutions. https://www2.bc.edu/~anderson/TradeInstitutions2.pdf>, collected 20111505

Chaudhry, Imran Sharif; Malik, Ali; Faridi, Muhammad Zahir (2010), Exploring the causality relationship between trade liberalization, human capital and economic growth: Empirical evidence from Pakistan. *Journal of Economics and International Finance*, 2(8), pp. 175-182

Chong, Alberto; Calderón, César (2000), Causality and Feedback between Institutional Measures and Economic Growth. *Economics and Politics*, vol. 12, pp 69-81

Dutt, Swama D.; Ghosh, Dipak (1996), Export Growth-Economic Growth Nexus: A Causality Analysis. *The Journal of Developing Areas,* 30, pp. 167-82

Easterly, W. and R. Levine (2003) "Tropics, germs, and crops: how endowments influence economic development", *Journal of Monetary Economics*, 50:3-39

Engerman, Stanley L. and Kenneth L. Sokoloff. 1998. "Factor Endowments, Insitutitons , and Differential Paths of Growth among New World Economies: A View from Economic Historians of the United States?". In *How Latin American Fell Behind: Essays on the Economic Histories of Brazil and Mexico*, 1800-1914., Stephen Haber, Stanford, CA: Stanford University Press

Henriques, I.; Sadorsky, P. (1996), Export-led Growth or Growth-driven Exports? The Canadian Case. Canadian Journal of Economics, 29, issue 3, pp 540-556

Ghartey. E. E. (1993), Casual Relationship between Exports and Economic Growth: Some Empirical Evidence in Taiwan, Japan and the US, *Applied-Economics.* (9), pp. 1145-1152

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

Furuoka, Fumitaka, (2007), Do Exports Act as "Engine" of Growth? Evidence From Malaysia. *Economics Bulletin,* 6, No. 37, pp. 1-14

Gallup, John Luke, Jeffrey D. Sachs, and Andrew D.Mellinger. (1998), Geography and Economic Development, *NBER Working Paper 6849*, http://www.nber.org/papers/w6849

Gerdham, G. and Löthgren, M. (2000), On Stationarity and Cointegration of International Health Expenditure and GDP. *Journal of Health Economics*, 19, 461-475

Glaeser, Edward L., La Porta, Rafael Lopez-de-Silanes, Florencio and Shleifer, Andrei, (2004) "Do Institutions Cause Growth?, *Journal of Economic Growth*, 2004, vol. 9, pp. 271-303

Granger, Clive (1969), Testing for Causality and Feedback. Econometrica. 37, 424-438

Groenewold, N., and S.H.K. Tang, (2007), Killing the Goose that Lays the Golden Egg: Institutional Change and Economic Growth in Hong Kong, *Economic Inquiry* 45, no. 4: 787–799

Heckman, James J. (2008), Econometric Causality. IZA Discussion Paper No. 3525, pp. 4

Ingianni, Andrea (2010), Intra-European Union Trade Openness and New Members' Output Convergence: a Time-series Analysis. *In: 9th Annual Meeting of the EEFS International Conference;* 3 - 6 June 2010, Athens, Greece. (Unpublished)

Jehle, Geoffrey A and Philip J Reny, Advanced Microeconomic Theory, 2nd ed, Addison-Wesley, 2001, pp. 543

Kaufmann, Daniel; Kraay, Art; Mastruzzi, Massimo (2009), Governance Matters VIII: Aggregate and Individual Governance Indicators, 1996-2008. *World Bank Policy Research Working Paper No. 4978*

Knack, P. and S. Keefer (1995), Institutions and Economic Performance: Cross-country Tests Using Alternative Institutional Measures. *Economics and Politics*, 7, 3, 207–27

Kuhn, Thomas (1962), The Structure of Scientific Revolutions. *The University of Chicago Press*. pp. 24–25

Mauro, P. (1995), Corruption and Growth. *Quarterly Journal of Economics*, **110**, 3, 681–712

Méon, Pierre-Guillaume; Sekkat, Khalid (2004), Does the Quality of Institutions Limit the MENA's Integration in the World Economy?. *The World Economy*, vol. 27, issue 9, pp. 1475-1498

North, Douglas (1990), Institutions, Institutional Change and Economic Performance, *Cambridge: Cambridge University Press*

North, Douglass C. (1981), Structure and Change in Economic History, *Norton*: Norton

Popper, Karl (1945), *The Open Society and Its Enemies.* Routledge

Rodrik, Dani; Subramanian, Arvind; Trebbi, Frencesco (2004), Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development. *Journal of Economic Growth,* 9, pp 131-165

Sachs, J.D., (2001), Tropical Underdevelopment, *NBER Working Paper 8119*., http://www.nber.org/papers/w8119

Sachs, J.D., & Warner, A., (1995), Economic reform and the process of global integration, *Brookings Papers on Economic Activity*, 1-118

Sala-i-Martin, Xavier and Subramanaian, A., (2003), Addressing the Natural Resource Curse: An Illustration from Nigeria, *NBER Working Paper 9804*, Cambridge/Mass

Tilly, Charles (1990), Coercion, Capital, and European States, AD 990-1990, *Oxford: Basil Blackwell*

United Nations Development Programme (2010), The Real Wealth of Nations: Pathways to Human Development, http://hdr.undp.org/en/reports/global/hdr2010/

Wei, S-J. (2000), Natural openness and good government. *National Bureau of Economic Research,* Working Paper no. 7765

Winters, Alan, (2004), Trade Liberalization and Economic Performance: An Overview. *The Economic Journal,* vol 114, 493, pp F4–F21

Wooldridge, Jeffrey M. (2009), *Introductory Econometrics – A Modern Approach*. 4th edition, South-Western Cengage Learning, pp. 506

Human Development Report, United Nations Development Programme, 2010, http://hdr.undp.org/en/reports/global/hdr2010/, collected 20110505

The world Factbook, CIA, https://www.cia.gov/library/publications/the-world-factbook/geos/la.html, collected 20110502

Corruption Perceptions Index, Transparency International, 2005 http://www.transparency.org/news_room/in_focus/2005/cpi_2005#cpi

APPENDIX 1: LIST OF COUNTRIES IN EACH SAMPLE

64-country sample	79-country sample	139-country sample	
ALGERIA	ALGERIA	ALBANIA	MALAYSIA
ANGOLA	ANGOLA	ALGERIA	MALI
ARGENTINA	ARGENTINA	ANGOLA	MALTA
AUSTRALIA	AUSTRALIA	ARGENTINA	MAURITANIA
BAHAMAS	BAHAMAS	AUSTRALIA	MAURITIUS
BANGLADESH	BANGLADESH	AUSTRIA	MEXICO
BOLIVIA	BARBADOS	BAHAMAS	MONGOLIA
BRAZIL	BELIZE	BAHRAIN	MOROCCO
BURKINA FASO	BENIN	BANGLADESH	MOZAMBIQUE
CAMEROON	BOLIVIA	BARBADOS	NAMIBIA
CANADA	BRAZIL	BELGIUM	NEPAL
CHILE	BURKINA FASO	BELIZE	NETHERLANDS
COLOMBIA	BURUNDI	BENIN	NEW ZEALAND
CONGO	CAMEROON	BOLIVIA	NICARAGUA
CONGO, DEM. REP.	CANADA	BOTSWANA	NIGER
COSTA RICA	CENTRAL AFRICAN REPUBLIC	BRAZIL	NIGERIA
DOMINICAN REPUBLIC	CHAD	BULGARIA	NORWAY
ECUADOR	CHILE	BURKINA FASO	OMAN
EGYPT	COLOMBIA	BURUNDI	PAKISTAN
EL SALVADOR	CONGO	CAMBODIA	PANAMA
ETHIOPIA	CONGO, DEM. REP.	CAMEROON	PAPUA NEW GUINEA
GABON	COSTA RICA	CANADA	PARAGUAY
GAMBIA	DJIBOUTI	CAPE VERDE	PERU
GHANA	DOMINICAN REPUBLIC	CENTRAL AFRICAN REPUBLIC	PHILIPPINES
GUATEMALA	ECUADOR	CHAD	POLAND
GUINEA	EGYPT	CHILE	PORTUGAL
GUYANA	EL SALVADOR	CHINA	QATAR
HAITI	ETHIOPIA	COLOMBIA	ROMANIA
HONDURAS	FIJI	COMOROS	RUSSIA
HONG KONG	GABON	CONGO	RWANDA
INDIA	GAMBIA	CONGO, DEM. REP.	SAUDI ARABIA
INDONESIA	GHANA	COSTA RICA	SENEGAL
IVORY COAST	GUATEMALA	CUBA	SEYCHELLES
JAMAICA	GUINEA	CYPRUS	SIERRA LEONE
KENYA	GUINEA-BISSAU	CZECH REPUBLIC	SINGAPORE
MADAGASCAR	GUYANA	DENMARK	SOUTH AFRICA
MALAYSIA	HAITI	DJIBOUTI	SPAIN
MALI	HONDURAS	DOMINICAN REPUBLIC	SRI LANKA
MALTA	HONG KONG	ECUADOR	SUDAN
MEXICO	INDIA	EGYPT	SURINAME
MOROCCO	INDONESIA	EL SALVADOR	SWAZILAND
NEW ZEALAND	IVORY COAST	EQUATORIAL GUINEA	SWEDEN

NICARAGUA	JAMAICA	ETHIOPIA	SWITZERLAND
NIGER	KENYA	FIJI	SYRIA
NIGERIA	LAOS	FINLAND	TAIWAN, CHINA
PAKISTAN	MADAGASCAR	FRANCE	TANZANIA
PANAMA	MALAYSIA	GABON	THAILAND
PARAGUAY	MALI	GAMBIA	TOGO
PERU	MALTA	GERMANY	TRINIDAD AND TOBAGO
SENEGAL	MAURITANIA	GHANA	TUNISIA
SIERRA LEONE	MAURITIUS	GREECE	TURKEY
SINGAPORE	MEXICO	GUATEMALA	UGANDA
SOUTH AFRICA	MOROCCO	GUINEA	UNITED KINGDOM
SRI LANKA	NEW ZEALAND	GUINEA-BISSAU	UNITED STATES
SUDAN	NICARAGUA	GUYANA	URUGUAY
TANZANIA	NIGER	HAITI	VENEZUELA
TOGO	NIGERIA	HONDURAS	VIETNAM
TRINIDAD AND TOBAGO	PAKISTAN	HONG KONG SAR, CHINA	YEMEN
TUNISIA	PANAMA	HUNGARY	ZAMBIA
UGANDA	PAPUA NEW GUINEA	ICELAND	ZIMBABWE
UNITED STATES	PARAGUAY	INDIA	
URUGUAY	PERU	INDONESIA	
VENEZUELA	RWANDA	IRAN	
VIETNAM	SENEGAL	IRELAND	
	SIERRA LEONE	ISRAEL	
	SINGAPORE	ITALY	
	SOUTH AFRICA	IVORY COAST	
	SRI LANKA	JAMAICA	
	SUDAN	JAPAN	
	SURINAME	JORDAN	
	TANZANIA	KENYA	
	TOGO	KOREA, SOUTH	
	TRINIDAD AND TOBAGO	KUWAIT	
	TUNISIA	LAOS	
	UGANDA	LEBANON	
	UNITED STATES	LESOTHO	
	URUGUAY	LUXEMBOURG	
	VENEZUELA	MADAGASCAR	
	VIETNAM	MALAWI	

APPENDIX 2: LIST OF ADDITIONAL COUNTRIES ADDED TO EACH SAMPLE

Additional countries in the 79-sample

BARBADOS
BELIZE
BENIN
BURUNDI
CENTRAL AFRICAN REPUBLIC
CHAD
DJIBOUTI
FIJI
GUINEA-BISSAU
LAOS
MAURITANIA
MAURITIUS
PAPUA NEW GUINEA
RWANDA
SURINAME

Additional countries in the 138-sample

ALBANIA
AUSTRIA
BAHRAIN
BELGIUM
BOTSWAN
BULGARIA
CAMBODIA
CAPE VERDE
CHINA
COMOROS
CUBA
CYPRUS
CZECH REP.
DENMARK
EQUATORIAL GUINEA
FINLAND
FRANCE
GERMANY
GREECE
HUNGARY
ICELAND
IRAN
IRELAND
ISRAEL
ITALY
JAPAN
JORDAN
KOREA
KUWAIT
LEBANON
LESOTHO
LUXEMBOURG
MALAWI
MONGOLIA
MOZAMBIQUE
NAMIBIA
NEPAL
NETHERLANDS
NORWAY
OMAN
PHILIPINES
POLAND
PORTUGAL
QATAR
ROMANIA
RUSSIA
SAUDI ARABIA
SEYCHELLES
SPAIN
SWAZILAND
SWEDEN
SWITZERLAND
SYRIA
TAIWAN
THAILAND
TURKEY
UNITE
YEMEN
ZAMBIA

APPENDIX 3: GRANGER OUTPUT, 64-COUNTRY SAMPLE

RULE causes LCGDP

Regression Statistics 1	
Multiple R	0,95
R Square	0,89
Adjusted R Square	0,87
Standard Error	0,11
Observations	64

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	12	5,52	0,46	36,12	0,00
Residual	51	0,65	0,01		
Total	63	6,17			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	0,00	0,02	0,15	0,88	-0,03	0,03	-0,03	0,03
GDPDIFF2	0,45	0,16	2,76	0,01	0,12	0,77	0,12	0,77
GDPDIFF1	-1,00	0,05	-19,51	0,00	-1,10	-0,89	-1,10	-0,89
RULEDIFF9	-0,09	0,07	-1,37	0,18	-0,23	0,04	-0,23	0,04
RULEDIFF8	0,02	0,12	0,15	0,88	-0,22	0,26	-0,22	0,26
RULEDIFF7	-0,14	0,08	-1,83	0,07	-0,29	0,01	-0,29	0,01
RULEDIFF6	0,06	0,13	0,46	0,65	-0,19	0,31	-0,19	0,31
RULEDIFF5	0,21	0,13	1,61	0,11	-0,05	0,46	-0,05	0,46
RULEDIFF4	0,21	0,15	1,37	0,18	-0,10	0,51	-0,10	0,51
RULEDIFF3	0,24	0,11	2,16	0,04	0,02	0,46	0,02	0,46
RULEDIFF2	0,45	0,25	1,77	0,08	-0,06	0,96	-0,06	0,96
RULEDIFF1	-0,28	0,24	-1,20	0,24	-0,76	0,19	-0,76	0,19

Regression Statistics 2	
Multiple R	0,93
R Square	0,87
Adjusted R Square	0,86
Standard Error	0,12
Observations	64

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	5,39	1,08	79,39	0,00
Residual	58	0,79	0,01		
Total	63	6,17			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	0,00	0,02	-0,07	0,94	-0,03	0,03	-0,03	0,03
GDPDIFF2	0,33	0,16	2,07	0,04	0,01	0,65	0,01	0,65
GDPDIFF1	-0,98	0,05	-19,18	0,00	-1,08	-0,88	-1,08	-0,88
RULEDIFF7	-0,18	0,07	-2,45	0,02	-0,33	-0,03	-0,33	-0,03
RULEDIFF3	0,25	0,11	2,20	0,03	0,02	0,47	0,02	0,47
RULEDIFF2	0,23	0,24	0,97	0,34	-0,25	0,71	-0,25	0,71

LCGDP causes RULE

Regression Statistics 1	
Multiple R	0,72
R Square	0,52
Adjusted R Square	0,40
Standard Error	0,08
Observations	64

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	13	0,34	0,03	4,20	0,00
Residual	50	0,31	0,01		
Total	63	0,65			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	0,02	0,02	1,08	0,29	-0,02	0,06	-0,02	0,06
RULEDIFF9	-0,13	0,04	-2,90	0,01	-0,22	-0,04	-0,22	-0,04
RULEDIFF2	-0,45	0,18	-2,53	0,01	-0,80	-0,09	-0,80	-0,09
RULEDIFF1	0,54	0,16	3,45	0,00	0,23	0,86	0,23	0,86
GDPDIFF9	-0,03	0,07	-0,50	0,62	-0,17	0,10	-0,17	0,10
GDPDIFF8	-0,11	0,04	-2,72	0,01	-0,19	-0,03	-0,19	-0,03
GDPDIFF7	-0,07	0,06	-1,24	0,22	-0,19	0,04	-0,19	0,04
GDPDIFF6	-0,08	0,10	-0,84	0,41	-0,29	0,12	-0,29	0,12
GDPDIFF5	-0,12	0,09	-1,33	0,19	-0,31	0,06	-0,31	0,06
GDPDIFF4	-0,09	0,09	-0,99	0,33	-0,28	0,09	-0,28	0,09
GDPDIFF3	0,00	0,13	-0,03	0,97	-0,27	0,26	-0,27	0,26
GDPDIFF2	0,22	0,17	1,29	0,20	-0,12	0,57	-0,12	0,57
GDPDIFF1	0,27	0,11	2,41	0,02	0,04	0,50	0,04	0,50
GDPDIFF	-0,07	0,09	-0,81	0,42	-0,24	0,10	-0,24	0,10

Regression Statistics 2	
Multiple R	0,67
R Square	0,45
Adjusted R Square	0,41
Standard Error	0,08
Observations	64

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	0,29	0,06	9,66	0,00
Residual	58	0,35	0,01		
Total	63	0,65			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	0,00	0,01	0,00	1,00	-0,02	0,02	-0,02	0,02
RULEDIFF9	-0,11	0,04	-2,69	0,01	-0,19	-0,03	-0,19	-0,03
RULEDIFF2	-0,52	0,16	-3,24	0,00	-0,84	-0,20	-0,84	-0,20
RULEDIFF1	0,55	0,14	3,82	0,00	0,26	0,83	0,26	0,83
GDPDIFF8	-0,08	0,03	-2,21	0,03	-0,14	-0,01	-0,14	-0,01
GDPDIFF1	0,26	0,09	2,92	0,00	0,08	0,44	0,08	0,44

LCGDP causes OPEN

Regression Statistics 1	
Multiple R	0,43
R Square	0,19
Adjusted R Square	0,03
Standard Error	0,10
Observations	64

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	10	0,13	0,01	1,22	0,30
Residual	53	0,55	0,01		
Total	63	0,68			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,12	0,02	-5,15	0,00	-0,17	-0,08	-0,17	-0,08
OPENDIFF3	-0,30	0,17	-1,75	0,09	-0,65	0,05	-0,65	0,05
GDPDIFF9	0,13	0,08	1,63	0,11	-0,03	0,29	-0,03	0,29
GDPDIFF8	0,08	0,05	1,51	0,14	-0,03	0,18	-0,03	0,18
GDPDIFF7	0,02	0,07	0,31	0,76	-0,12	0,17	-0,12	0,17
GDPDIFF6	0,09	0,13	0,68	0,50	-0,17	0,34	-0,17	0,34
GDPDIFF5	-0,05	0,12	-0,41	0,68	-0,30	0,20	-0,30	0,20
GDPDIFF4	-0,03	0,12	-0,30	0,77	-0,27	0,20	-0,27	0,20
GDPDIFF3	-0,07	0,17	-0,41	0,68	-0,41	0,27	-0,41	0,27
GDPDIFF2	0,04	0,21	0,20	0,84	-0,37	0,46	-0,37	0,46
GDPDIFF1	-0,01	0,13	-0,07	0,95	-0,27	0,26	-0,27	0,26

No lag is individually nor jointly significant, therefore no second regression is performed, and only the first regression is presented

LCOPEN causes LCGDP

Regression Statistics 1	
Multiple R	0,63
R Square	0,40
Adjusted R Square	0,26
Standard Error	0,12
Observations	64

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	12	0,53	0,04	2,85	0,00
Residual	51	0,79	0,02		
Total	63	1,31			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,02	0,03	-0,83	0,41	-0,08	0,03	-0,08	0,03
LCGDPDIFF2	0,79	0,22	3,59	0,00	0,35	1,23	0,35	1,23
LCGDPDIFF1	-0,46	0,15	-2,98	0,00	-0,77	-0,15	-0,77	-0,15
OPENDIFF9	0,12	0,12	1,00	0,32	-0,12	0,35	-0,12	0,35
OPENDIFF8	0,24	0,10	2,32	0,02	0,03	0,45	0,03	0,45
OPENDIFF7	-0,09	0,14	-0,60	0,55	-0,38	0,20	-0,38	0,20
OPENDIFF6	0,00	0,20	-0,02	0,98	-0,40	0,39	-0,40	0,39
OPENDIFF5	-0,13	0,23	-0,54	0,59	-0,60	0,34	-0,60	0,34
OPENDIFF4	-0,09	0,16	-0,55	0,58	-0,41	0,23	-0,41	0,23
OPENDIFF3	0,25	0,22	1,14	0,26	-0,19	0,69	-0,19	0,69
OPENDIFF2	-0,05	0,17	-0,30	0,76	-0,40	0,29	-0,40	0,29
OPENDIFF1	0,06	0,24	0,24	0,81	-0,42	0,53	-0,42	0,53

Regression Statistics 2	
Multiple R	0,58
R Square	0,34
Adjusted R Square	0,31
Standard Error	0,12
Observations	64

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0,44	0,15	10,27	0,00
Residual	60	0,87	0,01		
Total	63	1,31			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,04	0,02	-2,48	0,02	-0,08	-0,01	-0,08	-0,01
LCGDPDIFF2	0,72	0,20	3,62	0,00	0,32	1,12	0,32	1,12
LCGDPDIFF1	-0,40	0,13	-2,99	0,00	-0,67	-0,13	-0,67	-0,13
OPENDIFF8	0,21	0,09	2,31	0,02	0,03	0,39	0,03	0,39

LCOPEN causes RULE

Regression Statistics 1	
Multiple R	0,66
R Square	0,44
Adjusted R Square	0,29
Standard Error	0,09
Observations	64

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	13	0,28	0,02	2,97	0,00
Residual	50	0,37	0,01		
Total	63	0,65			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	0,00	0,02	-0,15	0,88	-0,04	0,03	-0,04	0,03
RULEDIFF9	-0,15	0,05	-3,16	0,00	-0,24	-0,05	-0,24	-0,05
RULEDIFF2	-0,62	0,20	-3,14	0,00	-1,01	-0,22	-1,01	-0,22
RULEDIFF1	0,47	0,17	2,78	0,01	0,13	0,81	0,13	0,81
OPENDIFF9	0,10	0,08	1,20	0,24	-0,07	0,26	-0,07	0,26
OPENDIFF8	0,12	0,07	1,64	0,11	-0,03	0,26	-0,03	0,26
OPENDIFF7	0,15	0,10	1,47	0,15	-0,05	0,35	-0,05	0,35
OPENDIFF6	0,02	0,14	0,13	0,89	-0,26	0,29	-0,26	0,29
OPENDIFF5	-0,19	0,16	-1,22	0,23	-0,51	0,13	-0,51	0,13
OPENDIFF4	0,02	0,12	0,15	0,88	-0,22	0,25	-0,22	0,25
OPENDIFF3	0,27	0,15	1,87	0,07	-0,02	0,57	-0,02	0,57
OPENDIFF2	0,07	0,12	0,61	0,55	-0,16	0,31	-0,16	0,31
OPENDIFF1	-0,04	0,17	-0,22	0,83	-0,37	0,30	-0,37	0,30

Regression Statistics 2	
Multiple R	0,59
R Square	0,35
Adjusted R Square	0,31
Standard Error	0,08
Observations	64

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	0,23	0,06	7,97	0,00
Residual	59	0,42	0,01		
Total	63	0,65			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,01	0,01	-0,67	0,50	-0,03	0,02	-0,03	0,02
RULEDIFF9	-0,14	0,04	-3,28	0,00	-0,23	-0,05	-0,23	-0,05
RULEDIFF2	-0,56	0,17	-3,22	0,00	-0,91	-0,21	-0,91	-0,21
RULEDIFF1	0,47	0,15	3,10	0,00	0,17	0,78	0,17	0,78
OPENDIFF3	0,16	0,13	1,20	0,23	-0,10	0,42	-0,10	0,42

RULE causes LCOPEN

Regression Statistics 1	
Multiple R	0,42
R Square	0,18
Adjusted R Square	0,01
Standard Error	0,10
Observations	64

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	11	0,12	0,01	1,04	0,43
Residual	52	0,56	0,01		
Total	63	0,68			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,10	0,01	-6,55	0,00	-0,13	-0,07	-0,13	-0,07
OPENDIFF3	-0,42	0,18	-2,42	0,02	-0,78	-0,07	-0,78	-0,07
RULEDIFF9	0,03	0,06	0,50	0,62	-0,09	0,15	-0,09	0,15
RULEDIFF8	0,04	0,11	0,36	0,72	-0,18	0,26	-0,18	0,26
RULEDIFF7	-0,02	0,07	-0,33	0,74	-0,16	0,12	-0,16	0,12
RULEDIFF6	0,04	0,12	0,37	0,71	-0,19	0,28	-0,19	0,28
RULEDIFF5	-0,11	0,11	-0,97	0,34	-0,34	0,12	-0,34	0,12
RULEDIFF4	-0,01	0,14	-0,11	0,91	-0,29	0,26	-0,29	0,26
RULEDIFF3	0,00	0,10	-0,03	0,97	-0,21	0,20	-0,21	0,20
RULEDIFF2	0,50	0,23	2,14	0,04	0,03	0,97	0,03	0,97
RULEDIFF1	-0,11	0,22	-0,52	0,61	-0,55	0,32	-0,55	0,32

Regression Statistics 2	
Multiple R	0,38
R Square	0,15
Adjusted R Square	0,12
Standard Error	0,10
Observations	64

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	0,10	0,05	5,29	0,01
Residual	61	0,58	0,01		
Total	63	0,68			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,10	0,01	-7,34	0,00	-0,12	-0,07	-0,12	-0,07
OPENDIFF3	-0,37	0,15	-2,51	0,01	-0,67	-0,08	-0,67	-0,08
RULEDIFF2	0,48	0,20	2,43	0,02	0,08	0,88	0,08	0,88

APPENDIX 4: GRANGER OUTPUT, 79-COUNTRY SAMPLE

RULE causes LCGDP

Regression Statistics 1	
Multiple R	0,91
R Square	0,84
Adjusted R Square	0,80
Standard Error	0,13
Observations	79

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	14	5,45	0,39	23,22	0,00
Residual	64	1,07	0,02		
Total	78	6,53			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,02	0,02	-1,19	0,24	-0,07	0,02	-0,07	0,02
GDPDIFF9	0,21	0,08	2,51	0,01	0,04	0,38	0,04	0,38
GDPDIFF4	-0,19	0,14	-1,31	0,19	-0,47	0,10	-0,47	0,10
GDPDIFF2	0,37	0,19	1,92	0,06	-0,01	0,76	-0,01	0,76
GDPDIFF1	-0,91	0,07	-12,35	0,00	-1,06	-0,76	-1,06	-0,76
RULEDIFF9	-0,03	0,04	-0,62	0,54	-0,11	0,06	-0,11	0,06
RULEDIFF8	0,03	0,12	0,27	0,79	-0,21	0,28	-0,21	0,28
RULEDIFF7	0,06	0,08	0,80	0,43	-0,10	0,23	-0,10	0,23
RULEDIFF6	0,11	0,11	1,03	0,31	-0,11	0,33	-0,11	0,33
RULEDIFF5	-0,03	0,13	-0,26	0,79	-0,30	0,23	-0,30	0,23
RULEDIFF4	0,18	0,14	1,27	0,21	-0,10	0,46	-0,10	0,46
RULEDIFF3	0,22	0,13	1,71	0,09	-0,04	0,47	-0,04	0,47
RULEDIFF2	-0,02	0,21	-0,11	0,91	-0,45	0,40	-0,45	0,40
RULEDIFF1	-0,24	0,18	-1,31	0,19	-0,60	0,12	-0,60	0,12

No lag is significant, therefore no second regression is performed, and only the first regression is presented

LCGDP causes RULE

Regression Statistics 1	
Multiple R	0,49
R Square	0,24
Adjusted R Square	0,10
Standard Error	0,11
Observations	79

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	12	0,23	0,02	1,73	0,08
Residual	66	0,73	0,01		
Total	78	0,96			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	0,01	0,02	0,65	0,52	-0,03	0,06	-0,03	0,06
RULEDIFF9	-0,07	0,03	-2,04	0,05	-0,14	0,00	-0,14	0,00
RULEDIFF6	-0,24	0,09	-2,71	0,01	-0,41	-0,06	-0,41	-0,06
GDPDIFF9	-0,07	0,07	-1,13	0,26	-0,20	0,06	-0,20	0,06
GDPDIFF8	-0,10	0,05	-1,98	0,05	-0,19	0,00	-0,19	0,00
GDPDIFF7	0,03	0,07	0,43	0,67	-0,11	0,18	-0,11	0,18
GDPDIFF6	-0,05	0,11	-0,41	0,68	-0,27	0,18	-0,27	0,18
GDPDIFF5	-0,07	0,11	-0,71	0,48	-0,29	0,14	-0,29	0,14
GDPDIFF4	-0,03	0,11	-0,27	0,79	-0,25	0,19	-0,25	0,19
GDPDIFF3	-0,08	0,16	-0,54	0,59	-0,40	0,23	-0,40	0,23
GDPDIFF2	-0,16	0,16	-1,01	0,32	-0,47	0,15	-0,47	0,15
GDPDIFF1	0,11	0,10	1,09	0,28	-0,09	0,32	-0,09	0,32

Regression Statistics 2	
Multiple R	0,43
R Square	0,19
Adjusted R Square	0,15
Standard Error	0,10
Observations	79

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0,18	0,06	5,73	0,00
Residual	75	0,78	0,01		
Total	78	0,9635793			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	0,00	0,01	0,02	0,98	-0,03	0,03	-0,03	0,03
RULEDIFF9	-0,08	0,03	-2,71	0,01	-0,15	-0,02	-0,15	-0,02
RULEDIFF6	-0,21	0,08	-2,74	0,01	-0,36	-0,06	-0,36	-0,06
GDPDIFF8	-0,08	0,04	-1,77	0,08	-0,16	0,01	-0,16	0,01

LCGDP causes LCOPEN

Regression Statistics 1	
Multiple R	0,36
R Square	0,13
Adjusted R Square	-0,01
Standard Error	0,10
Observations	79

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	11	0,10	0,01	0,91	0,54
Residual	67	0,68	0,01		
Total	78	0,79			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,11	0,02	-5,24	0,00	-0,16	-0,07	-0,16	-0,07
OPENDIFF3	-0,23	0,14	-1,71	0,09	-0,51	0,04	-0,51	0,04
GDPDIFF9	0,06	0,06	1,04	0,30	-0,06	0,19	-0,06	0,19
GDPDIFF8	0,08	0,05	1,65	0,10	-0,02	0,17	-0,02	0,17
GDPDIFF7	0,01	0,06	0,09	0,93	-0,12	0,14	-0,12	0,14
GDPDIFF6	0,08	0,11	0,70	0,48	-0,14	0,30	-0,14	0,30
GDPDIFF5	-0,07	0,10	-0,73	0,47	-0,28	0,13	-0,28	0,13
GDPDIFF4	0,02	0,10	0,17	0,87	-0,19	0,22	-0,19	0,22
GDPDIFF3	-0,05	0,15	-0,35	0,73	-0,35	0,25	-0,35	0,25
GDPDIFF2	0,02	0,15	0,16	0,87	-0,27	0,32	-0,27	0,32
GDPDIFF1	0,07	0,10	0,68	0,50	-0,12	0,26	-0,12	0,26

Regression Statistics 2	
Multiple R	0,30
R Square	0,09
Adjusted R Square	0,07
Standard Error	0,10
Observations	79

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	0,07	0,04	3,83	0,03
Residual	76	0,71	0,01		
Total	78	0,79			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,11	0,01	-8,18	0,00	-0,14	-0,08	-0,14	-0,08
OPENDIFF3	-0,18	0,13	-1,43	0,16	-0,44	0,07	-0,44	0,07
GDPDIFF8	0,09	0,04	2,25	0,03	0,01	0,17	0,01	0,17

LCOPEN causes LCGDP

Regression Statistics 1	
Multiple R	0,92
R Square	0,84
Adjusted R Square	0,80
Standard Error	0,13
Observations	79

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	14	5,47	0,39	23,65	0,00
Residual	64	1,06	0,02		
Total	78	6,53			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,02	0,03	-0,75	0,46	-0,08	0,04	-0,08	0,04
GDPDIFF9	0,20	0,08	2,43	0,02	0,03	0,36	0,03	0,36
GDPDIFF4	-0,26	0,13	-1,92	0,06	-0,52	0,01	-0,52	0,01
GDPDIFF2	0,45	0,18	2,53	0,01	0,09	0,81	0,09	0,81
GDPDIFF1	-0,84	0,08	-10,94	0,00	-0,99	-0,68	-0,99	-0,68
OPENDIFF9	-0,01	0,12	-0,07	0,95	-0,24	0,23	-0,24	0,23
OPENDIFF8	0,20	0,10	2,00	0,05	0,00	0,40	0,00	0,40
OPENDIFF7	-0,17	0,13	-1,29	0,20	-0,43	0,09	-0,43	0,09
OPENDIFF6	-0,05	0,18	-0,28	0,78	-0,41	0,31	-0,41	0,31
OPENDIFF5	-0,09	0,19	-0,45	0,66	-0,48	0,30	-0,48	0,30
OPENDIFF4	-0,14	0,16	-0,88	0,38	-0,45	0,18	-0,45	0,18
OPENDIFF3	0,02	0,19	0,13	0,90	-0,35	0,40	-0,35	0,40
OPENDIFF2	0,04	0,17	0,26	0,79	-0,29	0,38	-0,29	0,38
OPENDIFF1	-0,05	0,23	-0,23	0,82	-0,51	0,41	-0,51	0,41

Regression Statistics 2	
Multiple R	0,91
R Square	0,83
Adjusted R Square	0,82
Standard Error	0,12
Observations	79

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	5,41	1,08	70,71	0,00
Residual	73	1,12	0,02		
Total	78	6,53			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,04	0,02	-1,89	0,06	-0,08	0,00	-0,08	0,00
GDPDIFF9	0,21	0,07	3,00	0,00	0,07	0,34	0,07	0,34
GDPDIFF4	-0,26	0,12	-2,12	0,04	-0,50	-0,02	-0,50	-0,02
GDPDIFF2	0,45	0,17	2,69	0,01	0,12	0,78	0,12	0,78
GDPDIFF1	-0,86	0,06	-14,62	0,00	-0,98	-0,74	-0,98	-0,74
OPENDIFF8	0,22	0,09	2,44	0,02	0,04	0,39	0,04	0,39

LCOPEN causes RULE

Regression Statistics 1	
Multiple R	0,45
R Square	0,20
Adjusted R Square	0,06
Standard Error	0,11
Observations	79

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	12	0,20	0,02	1,40	0,19
Residual	66	0,77	0,01		
Total	78	0,96			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,01	0,02	-0,61	0,55	-0,05	0,03	-0,05	0,03
RULEDIFF9	-0,09	0,03	-2,57	0,01	-0,16	-0,02	-0,16	-0,02
RULEDIFF6	-0,19	0,09	-2,21	0,03	-0,36	-0,02	-0,36	-0,02
OPENDIFF9	-0,10	0,09	-1,17	0,25	-0,28	0,07	-0,28	0,07
OPENDIFF8	0,05	0,08	0,63	0,53	-0,11	0,22	-0,11	0,22
OPENDIFF7	0,05	0,11	0,41	0,68	-0,18	0,27	-0,18	0,27
OPENDIFF6	0,05	0,16	0,35	0,73	-0,26	0,36	-0,26	0,36
OPENDIFF5	-0,17	0,16	-1,10	0,27	-0,48	0,14	-0,48	0,14
OPENDIFF4	0,00	0,12	0,03	0,98	-0,24	0,25	-0,24	0,25
OPENDIFF3	0,06	0,15	0,37	0,71	-0,25	0,36	-0,25	0,36
OPENDIFF2	0,10	0,14	0,71	0,48	-0,17	0,37	-0,17	0,37
OPENDIFF1	0,01	0,19	0,08	0,94	-0,36	0,39	-0,36	0,39

No lag is significant, therefore no second regression is performed, and only the first regression is presented

RULE causes LCOPEN

Regression Statistics 1	
Multiple R	0,35
R Square	0,13
Adjusted R Square	-0,02
Standard Error	0,10
Observations	79

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	11	0,10	0,01	0,88	0,57
Residual	67	0,69	0,01		
Total	78	0,79			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,10	0,01	-7,23	0,00	-0,12	-0,07	-0,12	-0,07
OPENDIFF3	-0,21	0,14	-1,54	0,13	-0,49	0,06	-0,49	0,06
RULEDIFF9	0,04	0,03	1,05	0,30	-0,03	0,10	-0,03	0,10
RULEDIFF8	0,02	0,08	0,19	0,85	-0,15	0,18	-0,15	0,18
RULEDIFF7	0,02	0,06	0,36	0,72	-0,10	0,15	-0,10	0,15
RULEDIFF6	-0,01	0,08	-0,18	0,86	-0,18	0,15	-0,18	0,15
RULEDIFF5	-0,12	0,10	-1,18	0,24	-0,32	0,08	-0,32	0,08
RULEDIFF4	0,09	0,11	0,80	0,42	-0,13	0,30	-0,13	0,30
RULEDIFF3	0,05	0,09	0,59	0,55	-0,13	0,23	-0,13	0,23
RULEDIFF2	0,18	0,15	1,22	0,23	-0,12	0,49	-0,12	0,49
RULEDIFF1	-0,04	0,14	-0,25	0,80	-0,32	0,25	-0,32	0,25

No lag is significant, therefore no second regression is performed, and only the first regression is presented

APPENDIX 5: GRANGER OUTPUT, 138-COUNTRY SAMPLE

RULE causes LCGDP

Regression Statistics 1	
Multiple R	0,84
R Square	0,71
Adjusted R Square	0,67
Standard Error	0,14
Observations	138

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	15	5,73	0,38	19,91	0,00
Residual	122	2,34	0,02		
Total	137	8,07			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,04	0,02	-2,09	0,04	-0,07	0,00	-0,07	0,00
GDPDIFF9	0,15	0,06	2,27	0,03	0,02	0,27	0,02	0,27
GDPDIFF8	0,11	0,05	2,07	0,04	0,00	0,21	0,00	0,21
GDPDIFF4	-0,09	0,09	-0,99	0,32	-0,27	0,09	-0,27	0,09
GDPDIFF2	0,31	0,15	2,06	0,04	0,01	0,62	0,01	0,62
GDPDIFF1	-0,82	0,06	-13,72	0,00	-0,94	-0,70	-0,94	-0,70
RULEDIFF9	0,00	0,04	0,08	0,94	-0,07	0,08	-0,07	0,08
RULEDIFF8	0,07	0,10	0,77	0,44	-0,12	0,26	-0,12	0,26
RULEDIFF7	-0,01	0,07	-0,09	0,93	-0,15	0,14	-0,15	0,14
RULEDIFF6	0,15	0,10	1,53	0,13	-0,04	0,34	-0,04	0,34
RULEDIFF5	0,00	0,12	-0,02	0,98	-0,23	0,23	-0,23	0,23
RULEDIFF4	0,12	0,12	1,03	0,30	-0,11	0,35	-0,11	0,35
RULEDIFF3	0,20	0,10	2,09	**0,04**	0,01	0,39	0,01	0,39
RULEDIFF2	-0,05	0,17	-0,30	0,77	-0,39	0,29	-0,39	0,29
RULEDIFF1	-0,07	0,15	-0,45	0,66	-0,36	0,23	-0,36	0,23

Regression Statistics 2	
Multiple R	0,84
R Square	0,70
Adjusted R Square	0,68
Standard Error	0,14
Observations	138

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	6	5,63	0,94	50,40	0,00
Residual	131	2,44	0,02		
Total	137	8,07			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,04	0,02	-2,51	0,01	-0,08	-0,01	-0,08	-0,01
GDPDIFF9	0,14	0,06	2,38	0,02	0,02	0,26	0,02	0,26
GDPDIFF8	0,11	0,05	2,35	0,02	0,02	0,21	0,02	0,21
GDPDIFF4	-0,12	0,09	-1,41	0,16	-0,29	0,05	-0,29	0,05
GDPDIFF2	0,25	0,14	1,77	0,08	-0,03	0,53	-0,03	0,53
GDPDIFF1	-0,81	0,06	-14,60	0,00	-0,92	-0,70	-0,92	-0,70
RULEDIFF3	0,18	0,09	1,95	0,05	0,00	0,36	0,00	0,36

LCGDP causes RULE

Regression Statistics 1	
Multiple R	0,31
R Square	0,09
Adjusted R Square	0,01
Standard Error	0,10
Observations	138

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	12	0,12	0,01	1,09	0,38
Residual	125	1,18	0,01		
Total	137	1,31			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	0,02	0,02	1,04	0,30	-0,01	0,05	-0,01	0,05
RULEDIFF9	-0,06	0,02	-2,58	0,01	-0,11	-0,01	-0,11	-0,01
RULEDIFF4	-0,03	0,08	-0,40	0,69	-0,19	0,13	-0,19	0,13
GDPDIFF9	-0,02	0,05	-0,52	0,61	-0,11	0,07	-0,11	0,07
GDPDIFF8	-0,04	0,04	-1,03	0,30	-0,11	0,04	-0,11	0,04
GDPDIFF7	-0,03	0,05	-0,76	0,45	-0,12	0,05	-0,12	0,05
GDPDIFF6	0,01	0,07	0,17	0,86	-0,13	0,15	-0,13	0,15
GDPDIFF5	-0,02	0,08	-0,31	0,76	-0,17	0,13	-0,17	0,13
GDPDIFF4	-0,07	0,06	-1,09	0,28	-0,20	0,06	-0,20	0,06
GDPDIFF3	-0,08	0,10	-0,77	0,44	-0,29	0,13	-0,29	0,13
GDPDIFF2	-0,05	0,11	-0,48	0,63	-0,26	0,16	-0,26	0,16
GDPDIFF1	0,07	0,06	1,06	0,29	-0,06	0,19	-0,06	0,19

No lag is significant, therefore no second regression is performed, and only the first regression is presented

LCGDP causes LCOPEN

Regression Statistics 1	
Multiple R	0,30
R Square	0,09
Adjusted R Square	0,01
Standard Error	0,11
Observations	138

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	11	0,14	0,01	1,14	0,34
Residual	126	1,45	0,01		
Total	137	1,59			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,09	0,02	-5,24	0,00	-0,13	-0,06	-0,13	-0,06
OPENDIFF8	-0,12	0,06	-1,82	0,07	-0,24	0,01	-0,24	0,01
GDPDIFF9	-0,01	0,05	-0,15	0,88	-0,11	0,09	-0,11	0,09
GDPDIFF8	0,06	0,04	1,46	0,15	-0,02	0,14	-0,02	0,14
GDPDIFF7	-0,09	0,05	-1,83	**0,07**	-0,20	0,01	-0,20	0,01
GDPDIFF6	-0,01	0,08	-0,08	0,94	-0,16	0,15	-0,16	0,15
GDPDIFF5	-0,10	0,08	-1,18	0,24	-0,26	0,07	-0,26	0,07
GDPDIFF4	-0,05	0,07	-0,65	0,51	-0,19	0,09	-0,19	0,09
GDPDIFF3	-0,04	0,12	-0,32	0,75	-0,26	0,19	-0,26	0,19
GDPDIFF2	-0,01	0,12	-0,08	0,94	-0,24	0,22	-0,24	0,22
GDPDIFF1	0,03	0,07	0,40	0,69	-0,11	0,17	-0,11	0,17

Regression Statistics 2	
Multiple R	0,221
R Square	0,049
Adjusted R Square	0,035
Standard Error	0,106
Observations	138

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	0,08	0,04	3,45	0,03
Residual	135	1,52	0,01		
Total	137	1,59			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,090	0,012	-7,542	0,000	-0,113	-0,066	-0,113	-0,066
OPENDIFF8	-0,114	0,060	-1,884	0,062	-0,233	0,006	-0,233	0,006
GDPDIFF7	-0,107	0,047	-2,279	0,024	-0,200	-0,014	-0,200	-0,014

LCOPEN causes LCGDP

Regression Statistics 1	
Multiple R	0,84
R Square	0,70
Adjusted R Square	0,66
Standard Error	0,14
Observations	138,00

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	15	5,66	0,38	19,11	0,00
Residual	122	2,41	0,02		
Total	137	8,07			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,05	0,02	-2,22	0,03	-0,10	-0,01	-0,10	-0,01
GDPDIFF9	0,13	0,06	2,17	0,03	0,01	0,26	0,01	0,26
GDPDIFF8	0,11	0,05	2,09	0,04	0,01	0,22	0,01	0,22
GDPDIFF4	-0,17	0,09	-1,90	0,06	-0,34	0,01	-0,34	0,01
GDPDIFF2	0,35	0,15	2,41	0,02	0,06	0,64	0,06	0,64
GDPDIFF1	-0,77	0,06	-12,67	0,00	-0,89	-0,65	-0,89	-0,65
OPENDIFF9	0,05	0,08	0,56	0,58	-0,12	0,21	-0,12	0,21
OPENDIFF8	0,13	0,09	1,54	0,13	-0,04	0,31	-0,04	0,31
OPENDIFF7	-0,06	0,10	-0,57	0,57	-0,27	0,15	-0,27	0,15
OPENDIFF6	0,10	0,13	0,75	0,45	-0,16	0,35	-0,16	0,35
OPENDIFF5	0,02	0,13	0,17	0,87	-0,23	0,27	-0,23	0,27
OPENDIFF4	0,00	0,10	0,01	0,99	-0,21	0,21	-0,21	0,21
OPENDIFF3	-0,05	0,14	-0,34	0,74	-0,31	0,22	-0,31	0,22
OPENDIFF2	0,06	0,12	0,48	0,63	-0,18	0,29	-0,18	0,29
OPENDIFF1	-0,06	0,15	-0,38	0,70	-0,36	0,25	-0,36	0,25

No lag is significant, therefore no second regression is performed, and only the first regression is presented

LCOPEN causes RULE

Regression Statistics 1	
Multiple R	0,34
R Square	0,11
Adjusted R Square	0,03
Standard Error	0,10
Observations	138

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	12	0,15	0,01	1,32	0,21
Residual	125	1,16	0,01		
Total	137	1,31			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,01	0,01	-0,74	0,46	-0,04	0,02	-0,04	0,02
RULEDIFF9	-0,06	0,02	-2,64	0,01	-0,11	-0,02	-0,11	-0,02
RULEDIFF4	-0,05	0,08	-0,69	0,49	-0,21	0,10	-0,21	0,10
OPENDIFF9	-0,10	0,06	-1,76	0,08	-0,21	0,01	-0,21	0,01
OPENDIFF8	0,01	0,06	0,09	0,92	-0,11	0,12	-0,11	0,12
OPENDIFF7	-0,03	0,07	-0,48	0,63	-0,17	0,11	-0,17	0,11
OPENDIFF6	-0,01	0,09	-0,14	0,89	-0,18	0,16	-0,18	0,16
OPENDIFF5	-0,12	0,08	-1,41	0,16	-0,28	0,05	-0,28	0,05
OPENDIFF4	0,05	0,07	0,65	0,52	-0,09	0,18	-0,09	0,18
OPENDIFF3	-0,05	0,09	-0,51	0,61	-0,23	0,13	-0,23	0,13
OPENDIFF2	-0,03	0,08	-0,32	0,75	-0,18	0,13	-0,18	0,13
OPENDIFF1	-0,07	0,10	-0,64	0,53	-0,27	0,14	-0,27	0,14

Regression Statistics 2	
Multiple R	0,27
R Square	0,07
Adjusted R Square	0,05
Standard Error	0,09
Observations	138

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0,10	0,03	3,62	0,01
Residual	134	1,21	0,01		
Total	137	1,31			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,01	0,01	-0,69	0,49	-0,02	0,01	-0,02	0,01
RULEDIFF9	-0,07	0,02	-2,96	0,00	-0,12	-0,02	-0,12	-0,02
RULEDIFF4	-0,03	0,07	-0,45	0,65	-0,18	0,11	-0,18	0,11
OPENDIFF9	-0,08	0,05	-1,58	0,12	-0,18	0,02	-0,18	0,02

RULE causes LCOPEN

Regression Statistics 1	
Multiple R	0,27
R Square	0,07
Adjusted R Square	-0,01
Standard Error	0,11
Observations	138

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	11	0,12	0,01	0,93	0,52
Residual	126	1,47	0,01		
Total	137	1,59			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	-0,11	0,01	-10,03	0,00	-0,13	-0,09	-0,13	-0,09
LCOPENDIFF8	-0,06	0,06	-0,99	0,32	-0,18	0,06	-0,18	0,06
RULEDIFF9	0,02	0,03	0,54	0,59	-0,04	0,07	-0,04	0,07
RULEDIFF8	-0,01	0,07	-0,20	0,84	-0,15	0,12	-0,15	0,12
RULEDIFF7	0,05	0,06	0,89	0,38	-0,06	0,16	-0,06	0,16
RULEDIFF6	-0,09	0,07	-1,25	0,21	-0,24	0,05	-0,24	0,05
RULEDIFF5	-0,12	0,09	-1,37	0,17	-0,30	0,05	-0,30	0,05
RULEDIFF4	0,04	0,09	0,50	0,62	-0,13	0,21	-0,13	0,21
RULEDIFF3	0,08	0,07	1,15	0,25	-0,06	0,23	-0,06	0,23
RULEDIFF2	0,12	0,13	0,93	0,35	-0,13	0,37	-0,13	0,37
RULEDIFF1	-0,08	0,12	-0,72	0,47	-0,31	0,15	-0,31	0,15

No lag is significant, therefore no second regression is performed, and only the first regression is presented

APPENDIX 6: CORRELATION TABLE ON VARIOUS SOURCES OF DATA FOR GDP PER CAPITA

	64-country sample		79-country sample		139-country sample	
	LCGDP_WB	LCGDP_PW	LCGDP_WB	LCGDP_PW	LCGDP_WB	LCGDP_PW
LCGDP_WB	1					
LCGDP_PW	0,9636	1				
LCGDP_WB			1			
LCGDP_PW			0,9553	1		
LCGDP_WB					1	
LCGDP_PW					0,9562	1