Economic complexity and growth*

Guzmán Ourens†

September, 2017

Abstract

An increasing literature highlights the importance of economic complexity in the development process. Recent contributions propose ways to measure economic complexity, by considering what economies export. We improve the proposed measures by constructing the indicators on data featuring greater product disaggregation and a larger number of countries. We apply this measure to evaluate the correlation between complexity and economies' current income and future growth. Complexity is robustly correlated to countries' income and can predict growth for a large list of countries. Since complexity indicators are fed with information of each country relative to the others, the inclusion of economies for which export figures hardly reflect true complexity creates a noise that reduces the indicators' correlation with growth for the entire sample. Finally, we show that this correlation remains significant when many of the usual determinants are included in growth regressions.

Keywords: Economic complexity; specialization; growth.
JEL Classification Numbers: O47, O33, F14.

*A previous version of the present work circulated in the past as "Can the Method of Reflections help predict future growth?". I gratefully acknowledge financial support from Fonds Nationale de la Recherche Scientifique, FRS-FNRS, grant FC 99040, Belgium. I would like to thank Raouf Boucekkine, David de la Croix, Julien Martin, Florian Mayneris, Sandra Poncet, Felipe Starosta de Waldemar and Hylke Vandenbussche, for their valuable comments. I would also like to thank participants at seminars in Univerist catholique de Louvain, Aix-Marseille School of Economics, Paris School of Economics and the 2015 RIEF Meeting in Orleans. I gratefully acknowledge financial support from FNRS and from the Belgian French-speaking Community (convention ARC 09/14-018 on Sustainability). Remaining errors are my own responsibility.

†FNRS and IRES-Université catholique de Louvain. Collège L. H. Dupriez, 3 Place Montesquieu B, office d119, 1348 Louvain-la-Neuve (Belgium). Tel:+32(0)10473984. e-mail: guzman.ourens@uclouvain.be

1 Introduction

The role of economic complexity to explain development differences across nations has been increasingly stressed in the literature.[1] Over their growth process, some economies acquire a diversity of productive capabilities that makes them more prone to innovation and allows them to adapt to changing market circumstances, enabling them to remain at the forefront of economic development. The concept of economic complexity is naturally related to technological diversification and connectedness between economic activities. Given that technological proximity between products differs for any pair, connectedness becomes important since producing highly connected goods provides greater spillovers enabling further diversification, and hence promotes complexity gains (see Hidalgo et al., 2007, Kali et al., 2013 or Aghion, 2014). Recent works have managed to introduce these concepts into economic theory by underlining uneven connections between sectors as important sources of spillover differences, leading to divergent development paths between economies. The intuitive idea is that highly connected sectors generate larger knowledge spillovers promoting diversification and innovation. Economies producing in highly connected industries where growth prospects are relatively large, are expected to present more stable development paths with higher growth rates (see Koren and Tenreyro, 2013 and Baldwin and Venables, 2015 respectively). In such context, complex economies, i.e. those endowed with a large set of productive capabilities, are more likely to reach the production of growth-promoting activities.

Measuring economic complexity is not straightforward given the large range of features that constitute productive capabilities, however, some proposals have emerged in recent years. Building on previous contributions[2] Hidalgo and Hausmann (2009) developed the *Method of Reflections* (MR), which provides an original approach to the measure of *product sophistication* and *economic complexity*, based only on product's international trade data. The main assumption is that the required capabilities for the production of one good can only be partially substituted by others. This implies that the set of capabilities an economy has determines what it can potentially produce. Sophisticated goods (i.e. those requiring a large set of capabilities) are produced only by complex economies (i.e. those having a large diversity of capabilities). Their proposal consists in evaluating the capabilities that a country has by analysing its current production. The MR uses export data to determine which products are being exported by which countries. An iterative procedure allows, not only to count how many products are being exported by a country, but also how many countries are exporting the same products, how many products are being exported by those countries, and so on. While the basic indicator provides a simple measure of diversification for country A, the first iteration adds information regarding how easy it is to export what A is exporting and the second includes information on how diversified are in average those economies exporting A's exports. HH09 explain that highly iterated indicators measure economic complexity since they are highly correlated to countries' current income. They also show that deviations from that relationship can predict growth in the short and long run. Moreover Hausmann and Hidalgo (2011) argue that the pay-off a country gets from a new capability depends on the amount of existing capabilities in that economy.

[1] Some of the earliest contributions on this can be found in Lall (2000), Sachs and Warner (2001) and Hausmann and Rodrik (2003).

[2] In a first attempt, Hausmann et al. (2007) proposed an indicator for the contribution that each product makes to growth by computing the average income of economies producing each of them. Having a value for each product, they also suggested that growth possibilities of a certain region could be approximated by the average value of goods produced in it.

A country with many existing capabilities can use a new one to produce a large variety of products and therefore the pay-off it gets from acquiring a new capability is larger than that of a country where existing capabilities are scarce. The resulting increasing returns to diversification help explain the divergence in growth across countries.

This paper improves the economic complexity measures presented in HH09 by constructing the MR indicators over a dataset that features a higher product disaggregation and a larger country sample than the one used in HH09. This allows for greater accuracy in the distinction of country capabilities and product requirements, providing a richer setting for the construction of the indicators. Moreover, a larger country sample enables the evaluation of the indicators' performance for a larger set of developing countries. This is of primary importance given that MR indicators are likely to be used, as their predecessors,[3] in development reports for poor and middle-income countries. Our paper evaluates the robustness of the results to changing the value of the main parameter, the revealed comparative advantage threshold, and modifying the country sample. Results show that MR complexity indicators are robustly correlated with per capita GDP. Their correlation with future growth in the long and short term is significant when the indicators are constructed for a large sample of countries, but not all. Countries with sudden jumps in export figures create a distortion in the construction of the indicators that greatly reduces their correlation with growth. The MR uses information for each country relative to all other countries, so keeping countries with volatile exports affects results for the entire sample. We show that the countries to exclude are not random but have certain characteristics. They are relatively small, poor, highly specialized and their degree of openness to the world economy is relatively low and volatile. This result should not be overlooked when constructing the indicators in policy-oriented reports that are particularly interested in the growth potential of economies with these characteristics.

As a further contribution, we evaluate how complexity performs along with other typical explanatory variables in growth regressions. Our results show that complexity is able to significantly explain future growth even within specifications including many of these controls. In fact, the complexity indicator outperforms most of the growth regressors evaluated here. However, when variables closely related to geographical location (e.g. regional dummies or variables measuring topographical characteristics of the economy) are included, complexity loses significance in growth regressions. The same happens when country fixed-effects are introduced, which could be interpreted as complexity being correlated with deep growth fundamentals.

A final section tackles some of the most common concerns regarding the use of export data for constructing the MR indicators. First, the increasing fragmentation of production across national boundaries due to firms' strategical decisions artificially inflates export figures and could be considered as a source of distortion between real economic complexity and what the MR measures. On the other hand ignoring trade in final goods implies a great informational loss that can harm greatly the indicators. We constructed the MR upon a database that excludes final products and show that the indicators' correlation with growth decreases in this case, providing evidence that trade in final goods provides valuable information to the MR's measure of complexity. Second, we account for the possibility of *fake diversification* coming from arbitrary decisions in the structure of a given product classification. My results show that

[3] A non-exhaustive list of papers using the previous batch of complexity indicators includes Hausmann and Klinger (2006) for the case of South Africa, Record and Nghardsaysone (2010) for the case of Lao, Abdon and Felipe (2011) for Sub-Saharian Africa or Jankowska et al. (2012) which compares development prospects in Asia and Latin America.

this can severely affect low-iterated indicators but is not important for highly-iterated complexity indicators. Finally, using international trade data to approach economic complexity could introduce a bias in the measure. Indeed, economic complexity could be underestimated for economies that do not export all products for which they have the requirement capabilities. We address this issue by tackling two potential sources of this problem. One could be that fast growing economies have not discovered all they can do with their recently developed resources. We evaluate how this affects the MR by constructing the indicators upon a database that excludes fast growing economies. Another source of potential underestimation of economic complexity could arise if, as shown in Cadot et al. (2011), the relationship between diversification and income is not linear, but instead is closer to an inverted-U. This suggests that after certain income threshold, economies might not be producing all the products for which they have the required capabilities. As an alternative exercise to deal with this issue, we construct the MR indicators excluding countries located at the right hand side of the hump. In both cases we find that complexity indicators computed on the databases with a restricted set of countries present a diminished correlation with growth. This indicates that the iterative procedure in the MR feeds from the non-monotonic relationship between diversification and income.

This work aims at contributing to the literature on the influence economic complexity has on growth. Moreover, it aims at providing useful information to policy-makers in charge of industrial policies by showing under which circumstances MR complexity indicators provide useful information for policy recommendations. The rest of the work is organized as follows. Section 2 presents the database used and justifies its choice. Section 3 present the MR indicators, their most important features and limitations. Results for the different exercises performed in this work are shown in sections 4, 5 and 6. Finally, Section 7 concludes.

2 Data

This work constructs the MR indicators using export data from the *Base pour l'Analyse du Commerce International* (International Trade Database at the Product Level, BACI from here on), as reported by Gaulier and Zignago (2010). BACI reports bilateral trade flows of products in the first version of the Harmonized Commodity Description and Coding System (HS) at a six-digit aggregation level for the period 1995-2007. This database matches reports from importers and exporters using records by UN-COMTRADE. GDP per capita at PPP comes from Penn World Tables 7.0 (PWT) as reported by Heston et al. (2011) and is available for the period 1950-2009. Therefore, this work is going to be able to compute growth rates for a maximum of 15 years after the initial year for which export data is available. Other auxiliary data is detailed in the Appendix.

UNCOMTRADE data does not include flows below 1,000 US dollars but accounts for more than 95% of total world trade in goods. In order to have complete information on GDP and exports for the same countries every year, it is necessary to drop some observations.[4] The final database used here includes 177 countries and 4948 products.

The selected database differs from that used in HH09 to test the correlation between complexity and future growth. The main source of information in that work is

[4]Countries being left-out are the Occidental Palestinian Territories, Serbia and Montenegro, San Marino, Timor-Leste and the Vatican City which represent less than 0.09 % of total trade in the sample in any year considered here.

Feenstra et al. (2005) which gathers UNCOMTRADE data at a four-digit aggregation level (775 products in SITCrev2 classification) and also matches export and import reports, covering the period 1962-2000. The authors complete their sample using raw UNCOMTRADE data for the years after 2000. However, they only report results for regressions in the time span 1985-2005.[5]

The choice of BACI as our database comprises a trade-off. On one hand, given that BACI reports information at a greater disaggregation level and a larger number of countries, it should feed the MR with richer information and thus enhance their properties. In fact, if it is true that we can evaluate capabilities by looking at who is exporting what, then the use of more detailed data should allow for a sharper distinction between the capabilities required for each product and possessed by each country. Having six-digit data allows for example to differentiate between product 847010 which is the code for electronic calculators and product 847050 which denotes cash registers. If there is some country producing one of these items and not the other, then that information, which is fundamental for the MR, is lost when four-digit data is used. Something similar can be said regarding the benefits of having a larger number of countries. Indeed, a larger pool of countries allows for a sharper distinction of the required capabilities to export a given product.

On the other hand, BACI offers information starting at the year 1995 which leaves a shorter time span for the analysis. This could constitute a shortcoming since complexity indicators are expected to perform better as predictors of long-term growth. Results in HH09 show that these indicators performing better as predictors of 20-year average growth rates than for shorter periods (10 or 5 years). Still, since significant results are also found for the short time frames, it should be possible to find significant results for the 15-years span covered here.

3 The method of reflections

The first step to compute the MR indicators is to build a matrix linking countries to the products they export with revealed comparative advantage. We use Balassa's index of comparative advantage (Balassa, 1964), $RCA_{c,p}$:

$$RCA_{c,p} = \frac{x_{c,p} / \sum_{i=1}^{N_p} x_{c,i}}{\sum_{j=1}^{N_c} x_{j,p} / \sum_{j=1}^{N_c} \sum_{i=1}^{N_p} x_{j,i}} \tag{1}$$

where $x_{c,p}$ is the value of exports of product p by country c, N_p is the total number of products considered (here $N_p = 4948$) and N_c the total number of countries used in the dataset (here $N_c = 177$). The $RCA_{c,p}$ gives the importance of a product p in country c's export basket relative to the importance that the same product has in worldwide trade. Let $M_{N_c \times N_p} = [m_{cp}]$ be the matrix where each entry m_{cp} takes value 1 if country c has $RCA_{c,p} > R^*$ in product p and value 0 otherwise. HH09 propose R^*=1 as a natural threshold to build the indicators so we use that benchmark in our

[5] The authors explain they checked the validity of their results with other databases. In particular they used UNCOMTRADE HS data at four-digit level (covering 1241 products and 103 countries) and North American Industry Classification System (NAICS) with data at six-digit aggregation level (318 products, 150 countries), but the results stemming from the use of these datasets are not presented. Although the NAICS has the same aggregation level than BACI the quantity of products contained in the former is much smaller.

computations. As robustness, we also construct the indicators for alternative values as is explained in the next section.

Using $M_{c,p}$ it is possible to build the MR's simpler indicators as follows:

$$k_{p,0} = \sum_{j=1}^{N_c} m_{j,p} \tag{2}$$

$$k_{c,0} = \sum_{i=1}^{N_p} m_{c,i} \tag{3}$$

According to (2), $k_{p,0}$ measures the number of countries exporting product p, so it is a measure of that product's ubiquity. Similarly, following equation (3) $k_{c,0}$ counts the number of products exported by country c, and so it measures country's diversification. But these are only rough approximations to the concepts of product sophistication and economic complexity. Suppose we have two different countries, both with similar diversification levels, but one is small and has achieved its diversification level by acquiring different capabilities, while the other one is large and the only reason why it has a diversified export basket is because of its size. A desirable property for an indicator of economic complexity is that it assigns a greater value to the first of these economies. Something similar can happen when evaluating product sophistication: two products can be exported by the same amount of countries but their technological requirements might be very different. This is why the MR proposes to build on the initial measures of diversification and ubiquity by further exploiting the information contained in $M_{N_c \times N_p}$. This is done by the iterative process described in the following equations:

$$k_{p,n} = \frac{1}{k_{p,0}} \sum_{j=1}^{N_c} m_{j,p}.k_{c,n-1} \tag{4}$$

$$k_{c,n} = \frac{1}{k_{c,0}} \sum_{i=1}^{N_p} m_{c,i}.k_{p,n-1} \tag{5}$$

where n is the number of iterations used to define indicators $k_{p,n}$ and $k_{c,n}$. The result of these iterations yields two vectors: $k_p = \{k_{p,0}, k_{p,1}, ..., k_{p,n}\}$ and $k_c = \{k_{c,0}, k_{c,1}, ..., k_{c,n}\}$. This work focuses on k_c since it gathers different indicators at the country level. By comparing $k_{c,i}$ results for very different countries it is possible to get a better idea of how the MR uses data to approach economic complexity. Let us take a look for example at the different trajectories that India and Japan follow as i grows for $k_{c,i}$, using data for the last year in the exports sample (2007). Results show that $k_{India,0} = 1693$ which ranks India in the 7^{th} place among the 177 countries, above Japan that has a $k_{Japan,0} = 1259$ and reaches the 17^{th} place. Since $k_{c,0}$ only counts the quantity of products country c is exporting with $RCA \geq 1$, it is probably influenced to a great extent by country size. To approach economic complexity more accurately further information is required. In particular, it is desirable to include information regarding how difficult it is to produce each of the products that country c is exporting.

The first iteration includes that information since it counts the number of exports of c by the number of countries exporting each of them. With $i = 1$ results are $k_{India,1} = 13.9$ and $k_{Japan,1} = 11.4$. Notice that even though India exports more products, the average number of countries exporting what Japan is exporting is lower, which could imply that Japan is actually a more complex economy. This result is less related to country size but on the other hand it does not say much about the quantity of products being exported by each country. It could be the case that Japan's exports are scarce mainly because it is a country that has a very rare natural resource and it concentrates its exports among low sophisticated derivatives of that resource. So the first iteration brings new information but the outcome cannot be considered a proper index of complexity. A further iteration measures the average diversification of economies exporting what country c is exporting. According to our computations, when $i = 2$ results are $k_{India,2} = 1011.5$ and $k_{Japan,2} = 1126.8$ ranking India in the 22^{th} place while Japan reaches the 2^{nd} position. The average diversification of countries exporting what Japan exports is greater than the same figure for India. This gives the idea that the Japanese economy can be considered more complex than the Indian economy. By definition $k_{c,2}$ is less correlated with features like country size or specialization in extraction-type products than the less iterated components of the k_c vector. This is also shown empirically in the next section.

The example shows that an indicator with $i = 2$ is closer to the idea of economic complexity than other indicators where $i < 2$. The interpretation of the MR indicators gets more complicated as the number of iterations is increased, since every component in $k_{c,i}$ gathers information from the preceding ones. But this also means that elements coming from higher iterations have richer information and their correlation with economic complexity will be stronger. As the next section shows, when the number of iteration increases indicators converge to a mean and their variability decreases. Nevertheless we show that within a narrow range, highly iterated complexity indicators give a more stable ranking of countries, since the marginal informational contribution of iterations decreases approaching zero, which sets a limit to the iterative procedure.

3.1 Main features of k_c indicators

Table 1 shows how k_c indicators relate to different building blocks of complexity. The table shows pooled cross-section estimations with even k_c indicators as dependent variables and different indexes representing what could be considered as different components of economic complexity as explanatory variables. The logarithm of the gross rate of enrolment in tertiary education is used to approximate human capital, the inflation rate is used to approximate macroeconomic stability and the real interest rate is incorporated as a measure of the financial cost of engaging in a productive project. We included other regressors such as a measure of economic openness, the share of government consumption in GDP which measures the importance of the public sector in the economy, the share of industrial production in GDP and total rents from natural resources as a percentage of GDP. Finally, a control for the size of the economy was included which is measured by the logarithm of the economy's population. We used standardized variables in order to make coefficients comparable with each other.

According to Table 1 all complexity indicators are negatively related to higher inflation and interest rates while positively related to human capital. Nevertheless, the way that other of the features included in the regressions relate to MR indicators change as the level of iterations increases. Population size is highly correlated with $k_{c,0}$ but, as explained in the example of the previous section, its importance is reduced with the iterations. The degree of openness seems to gain importance with iterations indicating that having an open economy contributes to complexity more than it does to diversification. Similarly, government size and the weight of natural resources and manufactures in total production have a different effect on low or high iterated indicators. More importantly, it is remarkable how, although each regression is significant

Table 1: Pooled cross-section estimations of variables explaining k_c (even indicators).

Dep. var.:	skc0 (1)	skc2 (2)	skc4 (3)	skc6 (4)	skc8 (5)	skc10 (6)	skc12 (7)	skc14 (8)	skc16 (9)	skc18 (10)
Population (logs)	7.097***	3.917***	2.927***	1.905***	1.056***	0.779*	0.704*	0.684*	0.678*	0.677*
	(8.87)	(7.46)	(5.83)	(4.49)	(2.68)	(1.96)	(1.77)	(1.71)	(1.70)	(1.69)
Openness	0.061**	0.036*	0.035*	0.055***	0.064***	0.067***	0.068***	0.068***	0.068***	0.068***
	(2.27)	(1.66)	(1.80)	(3.43)	(3.42)	(3.29)	(3.26)	(3.25)	(3.25)	(3.25)
Inflation rate	-0.031	-0.025*	-0.028**	-0.031***	-0.028***	-0.027***	-0.026***	-0.026***	-0.026***	-0.026***
	(-1.27)	(-1.72)	(-2.24)	(-3.34)	(-4.27)	(-4.55)	(-4.61)	(-4.62)	(-4.63)	(-4.63)
Real interest rate	-0.196***	-0.151***	-0.182***	-0.233***	-0.226***	-0.218***	-0.215***	-0.215***	-0.214***	-0.214***
	(-2.92)	(-2.76)	(-2.98)	(-3.69)	(-3.92)	(-3.94)	(-3.94)	(-3.94)	(-3.94)	(-3.94)
Gov. expend. (%)	0.951***	0.713***	0.616***	0.178	-0.189	-0.312**	-0.348***	-0.358***	-0.361***	-0.362***
	(4.85)	(3.98)	(3.41)	(1.18)	(-1.50)	(-2.57)	(-2.89)	(-2.98)	(-3.00)	(-3.01)
Tertiary enroll. (logs)	5.108***	7.288***	7.601***	5.308***	2.658***	1.719***	1.451***	1.376**	1.354**	1.349**
	(5.30)	(9.07)	(9.04)	(7.42)	(4.57)	(3.14)	(2.68)	(2.55)	(2.52)	(2.51)
Industry GDP (%)	0.057	0.160	0.067	-0.181*	-0.307***	-0.336***	-0.342***	-0.344***	-0.344***	-0.345***
	(0.36)	(1.61)	(0.67)	(-1.89)	(-3.37)	(-3.74)	(-3.82)	(-3.84)	(-3.84)	(-3.84)
Primary GDP (%)	-0.590***	-0.611***	-0.396***	0.031	0.271***	0.332***	0.346***	0.350***	0.351***	0.351***
	(-4.20)	(-5.89)	(-3.70)	(0.34)	(3.52)	(4.40)	(4.60)	(4.65)	(4.66)	(4.66)
Obs.	1052	1052	1052	1052	1052	1052	1052	1052	1052	1052
Adj.-R^2	0.662	0.690	0.553	0.184	0.061	0.054	0.055	0.055	0.056	0.056
F-test	25.978	38.203	31.913	19.141	7.677	7.114	7.432	7.558	7.596	7.606
Prob> F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

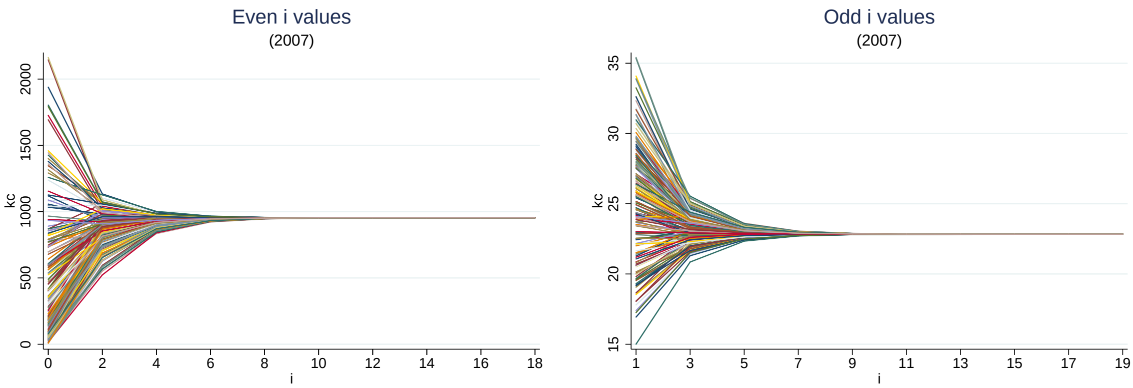
Notes: Pooled cross-section OLS estimations using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Data used covers the period 1995-2007. Dependent variable $sk_{c,i}$ is the standardized i-th component of the k_c vector divided by its standard deviation. *Population (logs)* is the logarithm of total population, *Openness* stands for economic openness (%) at 2005 constant prices, *Inflation rate* is the inflation rate, *Real interest rate* is the real interest rate (%), *Gov. expend. (%)* is the share of GDP destined to consumption, *Tertiary enroll. (logs)* is the logarithm of the gross rate of enrolment in tertiary education, *Industry GDP (%)* is the share of GDP coming from the secondary sector, *Primary GDP (%)* is the share of GDP coming from rents over natural resources. All explanatory variables were extracted from the WDI Database except for *Openness* which was taken from PWT. Some observations are lost since there are countries for which there is no information for the regressors. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

as a whole at 1%, the percentage of the k_c indicators explained by these variables is high for low iterated indicators, and low for high iterated indicators. The fact that the greatest part of the variation in highly-iterated indicators cannot be explained by our selected variables can be viewed as evidence supporting the hypothesis that highly iterated indicators are much richer and capture many more unobservable dimensions of economic complexity than less iterated indicators.

Figure 1 shows the value of each $k_{c,i}$ indicator for each country as i grows (for even and odd indicators separately). Only data from 2007 has been used to simplify the exposition but the same picture emerges for each year. As shown in both panels, when i grows the MR indicators converge to a mean, which is not surprising given that they are built as averages of other averages. The figure also shows that odd components inside a vector will converge to a certain value while even components tend to another. This is due to the way indicators are constructed: in building $k_{c,i}$ information from $k_{p,i-1}$ (and therefore $k_{c,i-2}$) is used but information from $k_{c,i-1}$ is not. Thus, odd components do not contribute with any information in the construction of even components within the same vector, and vice versa. The figure also shows that the final mean for even components (i.e. that of $k_{c,18}$) is 953.63 while that for odd components (corresponding to $k_{c,19}$) is 22.84. The difference in the units of measure of each of these families is given by the fact that even components count products while odd components count countries.

The convergence to a mean in measures of complexity implies that highly iterated indicators have a very narrow range (the largest standard deviation among all years for

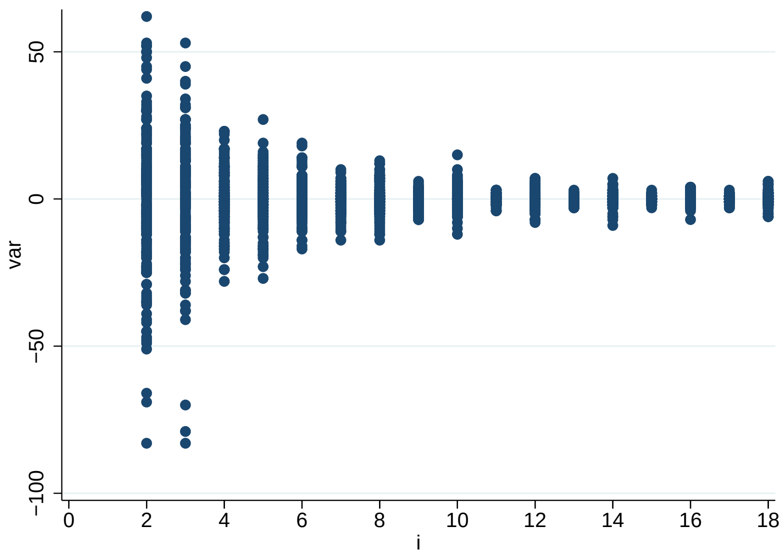
Figure 1: $k_{c,i}$ values for each country and different i level (2007).



Notes: The x-axis indicates the degree of iteration i of the $k_{c,i}$ indicator, while the y-axis measures the value of the corresponding indicator. Each line comprises values for one country. Even indicators (left) and odd indicators (right) are presented separately.

$k_{c,18}$ amounts to 0.007 in 1995 when the minimum value is 930.98 and the maximum is 931.02). Despite the narrow range, there is some distribution around the final mean which is fairly symmetric every year of our period (see histograms for $k_{c,18}$ and $k_{c,19}$ in Figure A.1). More importantly, highly iterated indicators yield more stable rankings of countries than less iterated indicators. Figure 2 shows the variation for each country in its $k_{c,i}$ ranking position as i increases. This figure also uses data for one year (2007) but the same conclusion holds for any other year in our database. Notice how the sorting tends to stabilize as i increases. This means that when i is low, information extracted from trade data in a further iteration has an important marginal contribution to sort countries according to their complexity, while when i is high the marginal informational contribution of an extra iteration is very low.

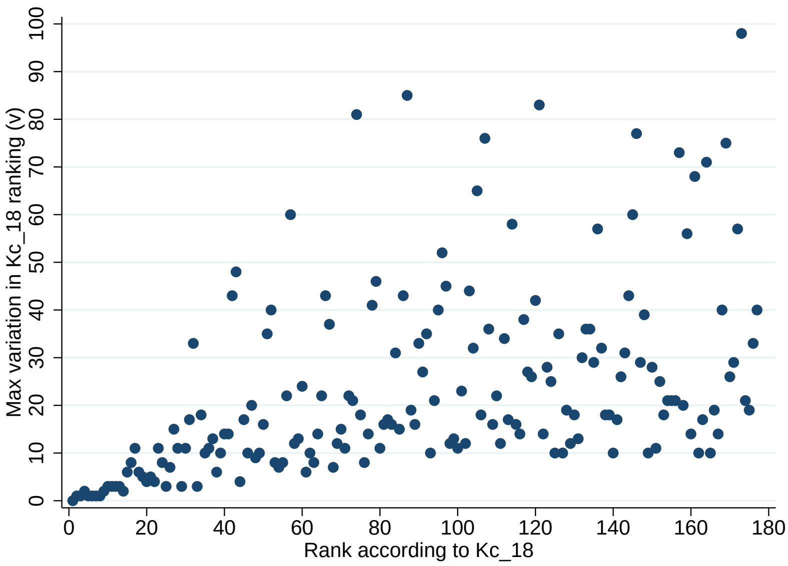
Figure 2: Ranking change between $k_{c,i-2}$ and $k_{c,i}$ for all countries as i grows (2007).



Notes: The x-axis indicates the degree of iteration i of the $k_{c,i}$ indicator, while the y-axis measures the difference $k_{c,i-2} - k_{c,i}$ for each country. Each dot shows that difference for a country and an iteration value i.

Although even and odd components nourish from different sources of information, highly iterated indicators from both families yield very similar rankings for countries and products. In fact the absolute value of the correlations between $k_{c,18}$ and either $k_{c,17}$ or $k_{c,19}$ is greater than 0.995 for every year in the sample used. Strong correlations can also be found between lower iterated indicators, e.g. the absolute value of the correlation between $k_{c,4}$ and $k_{c,5}$ is greater than 0.91 for every year.

Figure 3: Maximum changes in $k_{c,18}$ ranking for each country (v^m).



Notes: The x-axis shows the ranking position each country has according to $k_{c,18}$ in the year 2007. Lower values imply a higher positioning in the $k_{c,18}$ ranking. The y-axis measures the maximum variation each country shows in that ranking between any two consecutive years. Each dot represents a country.

When looking at how countries are ranked according to their $k_{c,18}$ throughout the years, there appears to be some cases of clear upward or downward trends that go along intuition. Countries like Malaysia, Thailand or Vietnam, which are well known cases of increasingly complex economies, exhibit a markedly upward trend in their positions over the years. Most countries however do not present such a clear trend. There are many for which their position is stable within a small range. For others their position in the world ranking is volatile. Variations from one year to the next (which we denote v) are more frequent and stronger among low complexity countries. Volatility is not a desirable property for indicators of economic complexity as it is expected that countries build their economic complexity through a long and costly process of acquiring (and destroying) capabilities. Strong and sudden changes can be a sign of undesirable noise. Figure 3 plots the maximum variation each country experiences in the $k_{c,18}$ ranking from any year to the next (called here v^m) against its $k_{c,18}$ level in 2007. The figure shows how the dispersion of v^m is increasing from the first positions until the 80^{th} place and remains stable after that, suggesting that more complex countries present a more stable position. This figure also shows that it is not usual to see changes in more than 40 positions of the $k_{c,18}$ ranking, and most countries never face a change of 20 positions.

The higher volatility for countries in lower positions of the $k_{c,18}$ ranking could be partially explained by misreports in export figures in those countries, but the fact that the dataset used here considers matched values from exporters and importers to reach the final value of a purchase reduces this possibility. Rather, it should be pointed out that, relying solely on export data, MR indicators are vulnerable to real sudden changes in export figures, which can happen in contexts of important changes

Table 2: Characterizing volatility in the $k_{c,18}$ ranking (1995-2007).

Dep. var.:	Changes in kc18 ranking from one year to the next (v)					
	(1)	(2)	(3)	(4)	(5)	(6)
GDPpc (logs)	-2.149***	-2.612***	-2.486***	-2.108***	-1.970***	-1.970***
	(-4.92)	(-7.00)	(-6.84)	(-6.17)	(-5.96)	(-5.94)
Population (logs)		-1.969***	-1.979***	-1.725***	-1.840***	-1.858***
		(-7.70)	(-7.78)	(-7.04)	(-6.89)	(-6.87)
Volatility of Openness			0.130**	0.106*	0.111*	0.123*
			(1.98)	(1.69)	(1.77)	(1.90)
HHc				12.878***	12.866***	12.753***
				(5.93)	(5.74)	(5.56)
Openness					-0.015*	-0.016*
					(-1.74)	(-1.78)
Year dummies						Yes
Obs.	2064	2064	2064	2064	2064	2064
Adj.-R^2	0.060	0.177	0.181	0.225	0.228	0.232
F-test	24.241	45.278	31.310	32.868	27.284	11.167
Prob> F	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Pooled OLS regression using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Each country-year represents one observation. *GDPpc (logs)* is the log of PPP converted GDP Per Capita (Laspeyres) at 2005 constant prices, *Population (logs)* is the log of population in thousands and *Openness* is the measure of openness reported as in PWT7.0. HH_c is the Hischmann-Herfindahl concentration index. *Volatility of Openness* is the number of the 20-quantile of the distribution of *Openness* to which each observation belongs. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

in trade or industrial policies, political or institutional environment, etc. This prevents the achievement of an accurate measure of economic complexity for countries where these kind of changes are sizeable and frequent. Table A.1 presents the list of all 177 countries sorted by v^m (computed under the benchmark case $R^* = 1$). The table also includes population (in thousands), per capita GDP at PPP ($GDPpc$) and the Hirschmann-Herfindahl index of export diversification (HH_c). This later index will be properly defined in equation (6) but it should be pointed that it ranges between 0 and 1 and that a value close to 1 indicates a very concentrated export basket. The table shows that those countries experiencing very large changes in the $k_{c,18}$ ranking (say v^m >40) are either very small, very concentrated or have a well known record of economic instability during the period considered here. Finally, in order to better characterize countries with high $k_{c,18}$ ranking volatility, Table 2 shows results of a cross-section estimation where v is the dependant variable and several explanatory variables are used. Results show that changes in $k_{c,18}$ are higher for relatively poor and small countries with concentrated exports and a degree of openness that is volatile and low.

3.2 Limitations of the MR

This section highlights some limitations of the MR and points at different strategies that this work uses to evaluate them. First, to measure economic complexity by looking at what countries export implies assuming that every country exports all products for which it has the required capabilities. But of course this might not be the case. In a context of uncertainty about the results of a given enterprise, some productions might have not being discovered yet, although required capabilities are available in the economy. In this context, the MR could underestimate economic complexity since some set of products that could be produced in a given economy are not accounted for in the MR, since no exports are reported. In Section 4 we show that the iterative process proposed in the MR is able to eliminate this potential problem, at least to some extent. By evaluating how MR indicators perform when fast growing countries are excluded from the sample, we show that the information these add enhance the correlation of MR complexity indicators and growth. A similar problem may arise for a different reason. One economy might have the capabilities to produce some product but decide to devote its resources to more productive activities. This last idea is reinforced by empirical studies. Cadot et al. (2011) show that the relationship between diversification and per capita income is not linear but hump-shaped, implying that as countries grow they diversify their production but, after certain income threshold, they start becoming specialized again in highly productive goods. In Section 6 we limit the country sample excluding countries that have surpassed the income threshold, to show that MR complexity indicators are not harmed but instead they nourish from the humped-shaped pattern.

Second, the fact that the MR considers as exported only those products with $RCA_{c,p} \geq R^*$ comprises another limitation since it implies ignoring productive processes that might add complexity to an economy even if they do not reach that threshold. In the next section we explore how robust are results when the RCA threshold is modified.

Third, it should be noticed that the original proposal for the construction of MR indicators, which uses only product exports, completely ignores services and the production of goods for the domestic market. This is a strong impediment when trying to get closer to an economy's technological capabilities, since both services and production for the domestic market are able to add technological knowledge into the productive structure of a country. This limitation is not easy to overcome since records on domestic market production and services are not available with the same level of comparability and with the same periodicity as product exports are. Hausmann and Hidalgo (2011) address this issue by constructing the MR indicators upon a database of total production from Chile. They conclude that results obtained from the MR indicators are not strongly influenced by the fact that it uses only product export data. More research should be done in order to explore whether this is the case for a larger number of countries.

Finally, the MR feeds from relative information so complexity measures for all countries are modified when the set of countries or products changes. This is due first to the fact that the $M_{c,p}$ matrix considers as exported only products for which a pair (c, p) fulfils $R_{c,p} > R^*$ and changing the set of countries or products used can change the value of $R_{c,p}$ for all pairs. Second, as the iterative procedure goes on, information regarding what other countries export is used over and over again in the evaluation of complexity for a given country. This means that, removing a set of countries with certain characteristics out of the sample can change the results for all the rest.

4 Testing the relationship between complexity, current income and future growth

A key feature of the MR stressed in HH09 is that complexity indicators are highly correlated with countries income and their future growth. Evidence for these results is provided in Tables S6-S10 in the Appendix of that work. In this section, we test the robustness of these conclusions replicating the exercises originally used, but constructing the MR on the more dissaggregated database provided by BACI. In particular we focus on evaluating whether these indicators are correlated with: i) countries' current income levels, ii) long-term growth, and iii) short-term growth. For each of these properties we proceed as follows. First, we present the results obtained by running the exact same specifications as in HH09. The only methodological departure in this first step is the use of our selected database.[6] If no significant results are found, the next step is to explore conditions under which significant coefficients could be found by changing the sample of countries used. The criterion for selecting the country sample is explained in detail. When significant results appear the last step is to test how much they depend on the choice of the R^* parameter.

4.1 Economic complexity and income

This section tests whether complexity indicators in the MR are positively correlated to countries income. First, we compute simple cross-section OLS regressions for a single year where the log per capita income at PPP is the dependent variable and a measure of complexity is the main regressor. We replicate the specification in Table S6 of HH09 by also including the Hirschmann-Herfindahl (HH_c) concentration index and (-1 times the) Theil's Entropy index (E_c). The HH_c index is a standard measure of market concentration but can be applied to a country's export basket to evaluate its concentration level, as is done here. The index is defined as follows:

$$HH_c = \sum_{i=1}^{N_p} s_{i,c}^2 \tag{6}$$

where $s_{i,c} = x_{c,i} / \left[\sum_{i=1}^{N_p} x_{c,i} \right]$ is the share of product i in country's c export basket. The index ranges between 0 and 1 and the higher the index the more concentrated the exports of country c are in fewer products.

The E_c index is a widely used measure of inequality and can be applied to measure the diversification of a country export basket when defined as follows:

$$E_c = -\sum_{i=1}^{N_p} s_{i,c}.log(s_{i,c}) \tag{7}$$

As defined here, E_c is always positive and higher values of the index imply that country c has a highly diversified export basket.

Table 3 shows the results for the year 2000, which amounts to replicating Table S6 of HH09, except the MR is computed upon our selected export database. All coefficient signs are as expected. The even components of the k_c vector are positively correlated with income, which goes with intuition given that they approximate economic complexity. Additionally, E_c has a positive sign which is as expected too since, according to the literature, diversification is positively related to countries' income levels (see for example Imbs and Wacziarg, 2003). This also explains why HH_c (which measures concentration) has a negative coefficient. The only odd component of k_c included in Table 3, i.e. $k_{c,1}$, has a negative coefficient since it measures the average ubiquity of products exported by each country. It can also be seen that, among even

[6] As a primary exercise we computed the MR using the same data as HH09. Besides checking that MR complexity indicators are correlated with growth for periods of 20, 10 and 5 years within the timespan 1985-2005, we expanded the time frame and verified the indicators perform well in similar regressions between 1965 and 2005.

components of the k_c vector, higher iteration levels yield greater coefficients. This is due to the fact that, as shown in Figure 1, the variability of the indicator decreases with its iteration level.

Table 3: Income regressions (2000).

Dep. var.:	Log per capita GDP									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ec	0.330***							0.062		0.374***
	(5.79)							(0.98)		(3.55)
HHc		-1.268**							0.554	2.881***
		(-2.52)							(1.25)	(3.77)
kc0			0.001***							
			(7.63)							
kc1				-0.138***						
				(-6.14)						
kc8					0.322***					
					(10.00)					
kc12						4.252***				
						(10.48)				
kc18							192.848***	177.641***	203.530***	156.785***
							(10.57)	(7.77)	(10.92)	(6.60)
Obs.	177	177	177	177	177	177	177	177	177	177
Adj.-R^2	0.155	0.026	0.212	0.210	0.344	0.341	0.333	0.332	0.334	0.373
F-test	33.572	6.340	58.274	37.659	100.019	109.875	111.787	55.250	61.751	47.969
Prob> F	0.000	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Cross-section OLS regressions using White's consistent estimator. Only data from the year 2000 was included. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the logarithm of per capita GDP. E_c is (-1 times) Theil's entropy index of diversification. HH_c is the Hischmann-Herfindahl concentration index. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Every component of the k_c vector is significant at 1% in each specification. Moreover columns 3-8 show how regressions using k_c indicators have greater adjusted-R^2 than regressions using only HH_c or E_c (columns 1 and 2). Given that the number of regressors and observations are the same in each of these specifications this means that the share of income explained by MR indicators is greater than that explained by diversification indexes. Notice also that the adjusted-R^2 grows when the specification uses higher iterated k_c components as regressors, reaching a level of 0.33 for $k_{c,18}$ in column 8. Finally, columns 9-11 show that there is no huge informational gain in introducing HH_c and E_c into the specification of column 8. This indicates that $k_{c,18}$ is a much richer indicator and diversification indexes do not have a large marginal contribution explaining income.

All previous conclusions are similar to those in HH09, and arise when performing this exercise for every year of the sample and also for a pooled regression of all years.[7] However, it is important to notice that the share of variability explained by the indicators when computed upon the new database is much lower than what HH09 find for the same year using a less aggregated dataset and a smaller set of countries. For instance they report an adjusted-R^2 of 0.535 for the same regression in column 8. This indicates that the positive correlation between of MR complexity indicators and income is lower when computed on our dataset.

4.1.1 Changing the R^* threshold in income regressions

This section checks how sensitive the former results are to changes in the R^* threshold. Increasing R^* would amount to be more restrictive regarding which cell of matrix $M_{c,p}$

[7] All omitted results are available upon request to the author.

takes value 1. In such a case, a smaller number of highly competitive exports would be considered and the matrix would be less dense in connections between countries and products. Reducing the value R^* would have the exact opposite effect.

Table 4: Pooled income regressions for different R^* threshold levels (1995-2009).

	Dependent variable: log per capita GDP									
$R^*=$	0.8		0,9		1		1,1		1,2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
kc12	5.550 ***		4.845 ***		4.365 ***		4.008 ***		3.780 ***	
	(12.82)		(13.10)		(13.26)		(13.55)		(13.48)	
kc18		306.432 ***		235.374 ***		187.807 ***		155.655 ***		134.297 ***
		(13.01)		(13.29)		(13.44)		(13.72)		(13.72)
Obs.	2,301	2,301	2,301	2,301	2,301	2,301	2,301	2,301	2,301	2,301
Adj.-R^2	0.360	0.339	0.364	0.345	0.365	0.345	0.366	0.345	0.364	0.344
F-test	23.42	42.94	24.31	29.42	24.89	47.26	24.91	94.09	25.95	37.98
Prob$>F$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Pooled cross-section OLS regressions using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the logarithm of GDP per capita. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant and year-dummies were also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4 shows results for pooled regressions using all years and including the same variables involved in columns 7 and 8 of Table 3. MR indicators used here were constructed with alternative values of the R^* threshold, namely 0.8, 0.9, 1.0, 1.1 and 1.2. Results show that MR complexity indicators still present highly significant coefficients in each case and the adjusted-R^2 in each specification is very similar between the alternative cases. When performing the same cross-section analysis as done in Table 3 for each individual year separately and each R^* threshold, the same conclusions always arise. Overall our results indicate that MR complexity indicators, especially the highly iterated versions, are positively correlated with countries' per capita income quite robustly.

4.2 Economic complexity and long-term growth

To test the performance of MR indicators as predictors of growth, the first step is to perform a similar analysis as done in Tables S7 and S8 of HH09. They present OLS regressions where complexity indicators are the main regressors explaining the average growth rate of the following 20 years (in Table S7) or 10 years (Table S8). Both tables exhibit significant coefficients at 1% for complexity indicators. Therefore, even though the longest average growth rate this work can compute is for 15 years, it should still be possible to find significant results. Nevertheless, as shown in Table 5, when computing the same regressions upon the disaggregated database and using the full sample of countries, results do not support the conclusion that any component of k_c can be used as predictors of 15-year growth. Sections 4.2.1 and 4.2.2 explore conditions under which significant coefficients could arise.

4.2.1 Filtering-out countries

As mentioned in section 3.1, the extreme changes some countries exhibit in their $k_{c,18}$ ranking position through the years point at some existing noise between the informational basis of the indicators and what they intent to measure. This noise can

Table 5: Regressions for long-term growth (no filter).

Dep. var.:	Average growth rate of GDP per capita (1995-2009)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDPpc	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(-1.23)	(-1.34)	(-1.42)	(-1.45)	(-1.44)	(-1.47)	(-1.58)	(-1.60)
Ec						-0.016		0.007
						(-0.74)		(0.19)
HHc							0.164	0.205
							(1.09)	(0.80)
kc0	0.000							
	(0.63)							
kc1	0.007							
	(0.50)							
kc4		0.000						
		(0.06)						
kc5		-0.012						
		(-0.04)						
kc12			0.172					
			(0.08)					
kc13			11.240					
			(0.05)					
kc18				2.529	144.597	197.896	198.049	187.910
				(0.34)	(1.16)	(1.10)	(1.09)	(1.03)
kc19					13982.428	18980.230	19044.553	18106.138
					(1.12)	(1.07)	(1.08)	(1.00)
Obs.	177	177	177	177	177	177	177	177
Adj.-R^2	0.004	0.001	0.002	0.008	0.007	0.004	0.006	0.001
F-test	1.034	1.146	1.205	1.463	1.739	1.237	1.600	1.274
Prob> F	0.379	0.332	0.309	0.234	0.179	0.298	0.191	0.282

Notes: Cross-section OLS regressions using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is 15-year average growth rate and values from the initial year (1995) are used for the explanatory variables. $GDPpc$ is PPP converted GDP Per Capita (Laspeyres) at 2005 constant prices as reported in PWT7.0. E_c is (-1 times) Theil's entropy index of diversification. HH_c is the Hischmann-Herfindahl concentration index. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

greatly harm regression analysis and so this work proposes to use a filter of countries based on the maximum $k_{c,18}$ ranking change from one year to the next (previously denoted v^m). That is, countries that have a ranking change between any two years greater than a certain threshold (v^*) are being dropped-out from the analysis. This choice for the filter provides us with a flexible criterion that shows how sensitive results are to changes in the sample used (changing v^*).[8] The filter can also help to put forward the conditions that the sample must fulfil in order for the MR indicators to be useful predictors of growth.

Significant results are found when filtering for observations with $v^m < v^* = 57$, that is, dropping countries for which the maximum variation in the $k_{c,18}$ ranking is greater than 57 positions between any two years. The reader can check in Table A.1 which are the countries being left aside, which in this case rise up to 16 (out of a total of 177). Table 6 shows the results obtained applying such a filter on the same specifica-

[8]The filtering can be done differently. One alternative is to base the criterion upon the quantity of times a country changes from one quantile of the $k_{c,18}$ distribution to another, but that would imply a greater probability of deletion of countries close to the quantile limit. The filtering choice made here avoids this problem and allows flexibility, without losing simplicity and transparency. Another possibility could be to drop countries in the higher part of the distribution of $k_{c,18}$ ranking's standard deviation. MR indicators stemming from this option have less significant correlation with growth than those stemming from the database that uses the filter proposed here. This indicates that the correlation of $k_{c,18}$ with future growth suffers more from the existence of large variations from one year to the next than from general volatility.

Table 6: Regressions for long-term growth (filter used $v^* = 57$).

	(1)	(2)	(3)	(4)	(5)	(6)
GDPpc	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(-0.84)	(-1.22)	(-1.35)	(-1.43)	(-1.34)	(-1.52)
Ec						-0.038
						(-1.03)
HHc						0.082
						(0.31)
kc0	-0.000					
	(-0.05)					
kc1	-0.001					
	(-0.09)					
kc4		0.002				
		(0.73)				
kc5		0.053				
		(0.18)				
kc12			1.835			
			(1.31)			
kc13			162.731			
			(1.15)			
kc18				8.651	286.965**	449.051***
				(1.42)	(2.35)	(2.88)
kc19					27439.275**	42699.509***
					(2.26)	(2.77)
Obs.	161	161	161	161	161	161
Adj.-R^2	-0.004	0.016	0.030	0.026	0.051	0.071
F-test	0.910	0.829	1.395	1.079	7.951	4.942
Prob> F	0.437	0.480	0.246	0.342	0.001	0.001

Notes: Cross-section OLS regressions using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is 15-year average growth rate and values from the initial year (1995) are used for the explanatory variables. *GDPpc* is PPP converted GDP Per Capita (Laspeyres) at 2005 constant prices as reported in PWT7.0. E_c is (-1 times) Theil's entropy index of diversification. HH_c is the Hischmann-Herfindahl concentration index. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

tions used in Table 5 (which itself replicated the specifications presented in HH09). It is possible to see how $k_{c,18}$ and $k_{c,19}$ are together significantly correlated to growth, and the level of significance increases when diversification indexes are included. Evidence presented here implies that MR complexity indicators contain valuable information to predict long-term growth when built for as many as 161 countries.

The quality of the MR can be further improved by keeping an even more selective list of countries. In Table 7 we show how results change by applying a more restrictive and a looser filter ($v^* = 47$ and $v^* = 67$ respectively) than those in the previous table. Keeping a larger number of countries into the sample yields non-significant results while using a more restrictive sample implies obtaining significant results for a greater variety of specifications (i.e. using $k_{c,18}$ as only predictor). If we set an even more restrictive filter (results omitted here), which implies filtering-out a larger group of countries, it is possible to find significant results for less iterated indicators (although only at a 10% level). Both the significance and magnitude of the effect MR indicators have on growth increases as v^* is reduced. This is also the case for the general fit of different specifications as measured by the adjusted-R^2 and the overall significance of the regression. In sum, the evidence suggests then that MR indicators are correlated to long-term growth for a large sample of economies as long as it does not include countries for which variations in export figures between any two years are too strong.

Table 7: Regressions for long-term growth. Sensitivity to different filter levels.

v* =	47				67			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDPpc	-0.000***	-0.000***	-0.000***	-0.000***	-0.000	-0.000	-0.000	-0.000
	(-3.75)	(-3.88)	(-3.67)	(-3.75)	(-1.00)	(-1.02)	(-0.95)	(-1.00)
Ec				-0.035				-0.029
				(-0.98)				(-0.79)
HHc				-0.043				0.035
				(-0.16)				(0.13)
kc12	2.048				0.104			
	(1.46)				(0.05)			
kc13	170.467				10.144			
	(1.21)				(0.05)			
kc18		13.871***	283.534**	394.983***		0.370	203.855	319.087
		(3.09)	(2.24)	(2.61)		(0.04)	(1.19)	(1.49)
kc19			26597.189**	37111.054**			20065.867	30926.415
			(2.17)	(2.47)			(1.20)	(1.49)
Obs.	158	158	158	158	167	167	167	167
Adj.-R^2	0.091	0.084	0.109	0.112	-0.002	0.004	0.008	0.006
F-test	6.688	7.524	8.935	7.365	1.227	1.176	1.290	1.009
Prob> F	0.000	0.001	0.000	0.000	0.302	0.311	0.278	0.405

Notes: Cross-section OLS regressions using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is 15-year average growth rate and values from the initial year (1995) are used for the explanatory variables. $GDPpc$ is PPP converted GDP Per Capita (Laspeyres) at 2005 constant prices as reported in PWT7.0. E_c is (-1 times) Theil's entropy index of diversification. HH_c is the Hischmann-Herfindahl concentration index. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

4.2.2 Removing fast growing and decreasing countries

As explained in Section 3.2, the MR could in theory, underestimate economic complexity in countries that are experiencing rapid structural changes. Given that countries do not necessarily learn to produce right away all products for which they have gathered the required capabilities, exports of fast growing countries (or countries facing important depressions) could under-represent (over-represent) complexity. If this is the case and the bias is important, the quality of the MR can be harmed for the entire sample of countries as a whole given the iterative nature of the method. This justifies the introduction of an alternative filter: one that excludes countries going through extraordinarily rapid processes of economic growth (either positive of negative). Table 8 presents results for regressions with $k_{c,18}$ and $k_{c,19}$ as main regressors and having dropped observations belonging to either the top or bottom 5% of the distribution of the 15-year average growth rate. Comparing these results from those in Tables 6 and 7 it is possible to see that the significance of complexity indicators is not improved by the exclusion of these countries.

4.2.3 Changing the R^* threshold

The significance of the correlation between MR indicators of complexity and income proved to be fairly robust to changes in R^* even with the full sample of countries, but this is not the case for regressions evaluating growth. Table A.2 shows results for alternative R^* values when applying the same country filter used in Table 6 ($v^* = 57$). Significant results disappear when the threshold R^* is changed. However, it is possible to find a v^* for which significant results are obtained with all alternative R^* values. In Table 9 results are presented for each of the alternative R^* thresholds explored here and with $v^* = 30$ (which implies keeping over 130 countries). The table shows

Table 8: Regressions for long-term growth. Cleaning for growth outliers.

Dep. var.:	Average growth rate of GDP per capita (1995-2009)								
v* =		47			57			67	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GDPpc	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(-3.93)	(-3.78)	(-3.89)	(-3.99)	(-3.86)	(-3.97)	(-3.92)	(-3.79)	(-3.88)
Ec			-0.024			-0.017			-0.014
			(-0.84)			(-0.60)			(-0.50)
HHc			0.028			0.118			0.117
			(0.12)			(0.50)			(0.52)
kc18	8.494**	195.535*	287.691**	8.735**	177.723	278.836**	8.126**	177.952*	267.850***
	(2.21)	(1.96)	(2.41)	(2.28)	(1.58)	(2.24)	(2.14)	(1.72)	(2.94)
kc19		18427.547*	27117.679**		16653.031	26225.240**		16732.098	25262.196***
		(1.88)	(2.35)		(1.47)	(2.20)		(1.62)	(2.86)
Obs.	147	147	147	149	149	149	153	153	153
Adj.-R^2	0.093	0.112	0.119	0.092	0.106	0.119	0.084	0.098	0.106
F-test	8.627	9.318	6.615	8.793	12.545	6.651	8.607	12.901	8.637
Prob> F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Cross-section OLS regressions using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is 15-year average growth rate and values from the initial year (1995) are used for the explanatory variables. *GDPpc* is PPP converted GDP Per Capita (Laspeyres) at 2005 constant prices as reported in PWT7.0. E_c is (-1 times) Theil's entropy index of diversification. HH_c is the Hischmann-Herfindahl concentration index. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

significant results for $k_{c,18}$ as predictor of growth in every case at a 5% significance level.

Table 9: Regressions for long-term growth ($v^* = 30$ and different R^* values).

	Dep. var.: avg. gr. rate of p/c GDP (1995-2009)				
R^*=	0.8	0.9	1.0	1.1	1.2
	(1)	(2)	(3)	(4)	(5)
GDPpc	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(-3.390)	(-3.416)	(-3.279)	(-3.457)	(-3.571)
kc18	18.254**	16.188***	10.961**	9.438**	7.675**
	(2.363)	(2.632)	(2.277)	(2.329)	(2.239)
Obs.	130	129	127	130	133
Adj.-R^2	0.069	0.074	0.065	0.071	0.074
F-test	5.840	5.836	5.465	6.121	6.768
Prob> F	0.004	0.004	0.005	0.003	0.002

Notes: Cross-section OLS regressions using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is 15-year average growth rate and values from the initial year (1995) are used for the explanatory variables. *GDPpc* is PPP converted GDP Per Capita (Laspeyres) at 2005 constant prices as reported in PWT7.0. $k_{c,18}$ is the 18-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

4.3 Economic complexity and short-term growth

Table S9 in HH09 presents regressions with the 5-year average growth as dependent variable and the initial value of MR indicators as main regressors. That table uses observations from years 1985, 1990, 1995 and 2000 and their results show highly significant correlation between MR indicators and short-term growth. Due to data constraints, Table 10 is built using years 1995, 2000 and 2005 with the purpose of comparing results using the same specifications on our database (even when we have only three observations per country instead of four). Our results show that, as in the case of regressions of long-run growth, significant coefficients for MR indicators are not to be found unless the country sample excludes some of the countries with volatile export figures. Notice that even with a country filter of $v^* = 47$ (i.e. excluding 19 countries) the coefficient for $k_{c,18}$ is very small. Moreover, controlling for diversification yields non-significant coefficients for the MR complexity indicators (column 3).

Table 10: Regressions for short-term growth. Different values of v^*.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GDPpc	-0.000***	-0.000***	-0.000***	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(-4.14)	(-3.91)	(-4.52)	(-1.37)	(-1.22)	(-1.30)	(-1.62)	(-1.51)	(-1.64)
Ec			0.015**			0.011			0.012
			(2.00)			(1.13)			(1.20)
HHc			0.065			0.093			0.110
			(1.03)			(1.08)			(1.27)
kc12	0.007*			0.007*			0.004		
	(1.79)			(1.74)			(0.90)		
kc13	-0.064			-0.070			-0.038		
	(-1.35)			(-1.43)			(-0.73)		
kc18		0.002***	0.005		0.001*	0.006		0.001	0.003
		(2.71)	(1.37)		(1.83)	(1.46)		(0.96)	(0.83)
kc19			-0.045			-0.057			-0.033
			(-0.93)			(-1.15)			(-0.64)
Obs.	474	474	474	483	483	483	531	531	531
Adj.-R^2	0.035	0.033	0.038	0.010	0.008	0.008	0.003	0.004	0.003
F-test	8.687	11.482	5.961	2.594	2.533	1.682	1.608	1.913	1.218
Prob> F	0.000	0.000	0.000	0.052	0.080	0.137	0.186	0.149	0.299

Notes: Pooled cross-section OLS estimations using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the average growth rate for 5-year periods starting in 1995, 2000 and 2005. Explanatory variables take their initial value of each 5-year period. $GDPpc$ is PPP converted GDP Per Capita (Laspeyres) at 2005 constant prices as reported in PWT7.0. E_c is (-1 times) Theil's entropy index of diversification. HH_c is the Hischmann-Herfindahl concentration index. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

These results are strongly influenced by the last of the 5-year periods considered here. Indeed, reaching some of the years of the latest global depression proves to be detrimental to the results presented here as many of the more complex economies experienced unusually low growth rates during this period. Table 11 shows results for the same specifications in Table 10 but considering only 5-year growth for the years 1995 and 2000. The coefficients obtained are much larger and significant even with diversification controls, although the conclusion that non-significant results appear when the full country-sample is used remains. The argument could be made that it is justified to exclude the crisis period from the analysis, as the last world depression had little to do with economic complexity as understood here. Still, the above result could also be interpreted as pointing at a fundamental limitation of economic complexity indexes when used as growth predictors. The fact that economic depressions are often caused by factors that are orthogonal to domestic productive capabilities, implies that indicators measuring those capabilities should probably not be used as the main single tool when evaluating growth prospects in the short term or business cycles in general. Instead these tools seem more suitable to evaluate long long-term growth trends.

As shown for regressions with long-term growth, results are not robust to changes in R^*. Table A.3 presents results for alternative values of that parameter and shows that, even with a filtered country sample ($v^* = 47$) and excluding the latest 5-year period, non-significant coefficients appear.

Table 11: Regressions for short-term growth excluding the last 5-year period. Different values of v^*.

Dep. var.:	avg. growth (5-yr period starting in 1995, 2000 and 2005)								
v* =	47			57			No filter		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GDPpc	-0.000***	-0.000	-0.000***	-0.000*	-0.000	-0.000*	-0.000	-0.000	-0.000*
	(-3.26)	(-1.11)	(-3.16)	(-1.80)	(-0.59)	(-1.91)	(-1.49)	(-1.20)	(-1.72)
Ec			0.008			0.013			0.017
			(0.84)			(1.25)			(1.50)
HHc			0.025			0.122			0.141
			(0.33)			(1.12)			(1.18)
kc12	0.161***			0.114**			0.046		
	(3.63)			(2.50)			(0.64)		
kc13	-1.811***			-1.286**			-0.517		
	(-3.60)			(-2.49)			(-0.63)		
kc18		0.001*	5.350***		0.001	3.946**		0.001	0.969
		(1.93)	(2.84)		(1.10)	(2.06)		(0.61)	(0.32)
kc19			-60.738***			-44.800**			-11.003
			(-2.83)			(-2.05)			(-0.32)
Obs.	316	316	316	322	322	322	354	354	354
Adj.-R^2	0.052	0.008	0.045	0.018	-0.002	0.016	-0.002	-0.002	-0.002
F-test	5.622	2.491	3.130	2.223	0.787	2.103	0.924	1.085	0.991
Prob> F	0.001	0.084	0.009	0.085	0.456	0.080	0.429	0.339	0.423

Notes: Pooled cross-section OLS estimations using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the average growth rate for 5-year periods starting in 1995 and 2000. Explanatory variables take their initial value of each 5-year period. $GDPpc$ is PPP converted GDP Per Capita (Laspeyres) at 2005 constant prices as reported in PWT7.0. E_c is (-1 times) Theil's entropy index of diversification. HH_c is the Hischmann-Herfindahl concentration index. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Finally, we introduce country fixed-effects as done in Table S10 of HH09. Significant coefficients could not be found when controlling for time-invariant country characteristics even filtering out countries and dropping the last 5-year period as shown in Table A.4.[9]

5 Economic complexity in growth regressions

The previous section has established that, under certain conditions, complexity indicators contain valuable information that is correlated with future economic growth in the long-term. In evaluating that, we have kept specifications identical to those presented in HH09, changing only the database used, and considering different alternatives for the country sample and the value of the parameter R^*. We now evaluate how significant is that correlation, when controlling for other explanatory variables, often believed to be related to growth. For this task we depart from the specifications in HH09 and instead introduce complexity indicators into typical growth regressions. This implies using the average growth rate of each country in the very long run (period 1962-2007) as dependent variable, the initial value of the (standardized) main indicator of economic complexity ($sk_{c,18}$) as main explanatory variable and controlling for growth determinants suggested by economic theory.

Growth regressions have been subject to large scrutiny. The main drawbacks

[9] We computed the same regressions using moving averages of the 5-year growth rate to increase the number of observations per country but that does not improve the results with country fixed-effects. Cleaning-out for outliers from the distribution of the growth rate does not improve the significance of the coefficients either.

usually signalled are: a) too many variables have been identified by the literature as significant growth determinants, b) given the limited amount of countries reporting the required information, not all of these determinants can be included together in growth regressions, c) the list of robust growth regressors is sensitive to the construction of the dependant variable. Given points a) and b), evidence in favour of a variable measuring economic complexity should be welcomed, as it resumes many different aspects of the development of capabilities that other variables capture separately (remember our interpretation of the results in Table 1). Put simply, evidence showing that complexity is significantly related with growth even controlling for usual growth determinants, could indicate that this sort of resume measures can replace some of these controls and help overcome one of the most important criticism made to growth regressions. To overcome point c) we follow Ciccone and Jarocinski (2010) and evaluate our regressions using alternative measures of growth in the right-hand side, i.e. including different measures of income (from PWT6.3, PWT7.0, PWT8.1, WDI and data from the Maddison project as reported by Bolt and van Zanden, 2014) to compute growth rates and also computing labour productivity growth (using data from PWT8).

Given the long list of potential controls we could use in this exercise, we first include all variables identified in Sala-i Martin et al. (2004) as robust growth regressors. These controls constitute a wide range of measures of basic growth fundamentals (initial wealth, investment costs, human capital, etc.), as well as indexes of institutional quality, regional, cultural and geographical characteristics. We also include the initial level of investment, as this is one of the main drivers of growth in theoretical models and is also often used in growth regressions.

We complete our list of controls by including important growth regressors identified in the literature after the publication of Sala-i Martin et al. (2004). This includes deep-rooted determinants of institutional characteristics like genetic diversity or historical plough use which may determine different steady states between economies (see Ashraf and Galor, 2013 and Alesina et al., 2013 respectively) and macroeconomic variables like fiscal policy volatility and public debt as a fraction of GDP, for which new evidence was reported over the last years (Fatás and Mihov, 2013 and Eberhardt and Presbitero, 2015 cover these in depth). In using cross-section OLS regressions we aim at reducing potential biases from measurement errors (see Hauk and Wacziarg, 2009). The reader can check in the Appendix (Table A.5) the list of variables included, a precise definition of what they measure and the source were the information was extracted from.

In order to compute our main indicator of economic complexity for the year 1962 we go back to the use of 4-digit level export flows as reported in Feenstra et al. (2005), since 6-digit export data for that period is not available. We introduce controls one by one, in a specification in which the initial level of per capita income, investment, education and economic complexity are also included. Table 12 shows that most of the robust growth determinants can be included without making the coefficient for economic complexity lose its significance. This result is robust to the use of different measures of growth rates used as dependant variable. Table A.6 reports results using Maddison's income data and shows that an even larger set of controls can be included without turning the complexity indicator insignificant, and the level of significance for that indicator is larger in most regressions. A similar conclusion is reached using the other alternatives we propose here for the dependent variable.

Table 12: Cross-section growth regressions, robustness of significance for $sk_{c,18}$

Dep. var.:	Avg. growth rate 1965-2007.																	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
GDPpc (logs)	-0.207	-0.210	-0.199	-0.153	-0.170	-0.159	-0.157	-0.101	-0.075	-0.072	-0.125	-0.107	-0.101	-0.103	-0.025	0.003	0.001	0.021
	(-1.24)	(-1.22)	(-1.11)	(-0.84)	(-0.93)	(-0.88)	(-0.87)	(-0.61)	(-0.43)	(-0.47)	(-0.83)	(-0.74)	(-0.76)	(-0.80)	(-0.16)	(0.02)	(0.01)	(0.13)
Primary enroll.	0.008	0.009	0.009	0.009	0.008	0.008	0.009	0.008	0.012*	0.011*	0.010*	0.009	0.008	0.008	0.002	0.002	0.002	0.003
	(1.37)	(1.33)	(1.33)	(1.34)	(1.20)	(1.25)	(1.44)	(1.42)	(1.93)	(1.95)	(1.80)	(1.65)	(1.25)	(1.28)	(0.32)	(0.29)	(0.27)	(0.46)
Investment (%GDP)	0.007	0.005	0.005	0.004	0.006	0.002	0.001	-0.001	-0.002	-0.003	-0.003	-0.003	-0.002	-0.002	-0.000	-0.000	-0.000	0.000
	(1.06)	(0.71)	(0.75)	(0.56)	(0.94)	(0.30)	(0.29)	(-0.15)	(-0.47)	(-0.66)	(-0.69)	(-0.66)	(-0.33)	(-0.32)	(-0.06)	(-0.04)	(-0.07)	(0.08)
skc18	0.245**	0.235**	0.239**	0.197*	0.245**	0.206*	0.226**	0.226**	0.275**	0.267**	0.222*	0.246*	0.244*	0.243*	0.243*	0.361**	0.371**	0.323*
	(2.39)	(2.33)	(2.28)	(1.93)	(2.33)	(1.97)	(2.16)	(2.11)	(2.08)	(2.33)	(1.81)	(2.01)	(1.97)	(1.95)	(1.85)	(2.33)	(2.36)	(1.90)
Investment price PPP		-0.001	-0.001	-0.002	-0.002	-0.002	-0.001	-0.001	-0.001	0.000	0.001	0.001	0.002	0.002	-0.000	0.000	0.000	-0.000
		(-1.08)	(-1.07)	(-1.39)	(-1.31)	(-1.32)	(-1.00)	(-0.52)	(-0.47)	(0.15)	(0.45)	(0.48)	(0.85)	(0.83)	(-0.10)	(0.06)	(0.04)	(-0.00)
Malaria prevalence			0.052	0.107	0.155	0.102	0.126	0.148	0.213	0.392	0.331	0.378	0.359	0.354	0.244	0.323	0.320	0.294
			(0.20)	(0.40)	(0.58)	(0.39)	(0.47)	(0.57)	(0.77)	(1.41)	(1.12)	(1.26)	(1.18)	(1.19)	(0.66)	(0.85)	(0.83)	(0.76)
LATAM				-0.243*	0.251	0.200	0.252	0.341	0.419	0.433	0.562*	0.611*	0.507	0.511	0.528	0.341	0.349	0.429
				(-1.67)	(0.78)	(0.62)	(0.76)	(1.04)	(1.12)	(1.55)	(1.69)	(1.79)	(1.36)	(1.32)	(1.06)	(0.47)	(0.48)	(0.58)
Frm Spanish colony					-0.585*	-0.508	-0.513	-0.504	-0.571	-0.656**	-0.716**	-0.798**	-0.718*	-0.729*	-0.846*	-0.859	-0.875	-0.890*
					(-1.86)	(-1.61)	(-1.62)	(-1.62)	(-1.50)	(-2.35)	(-2.19)	(-2.31)	(-1.94)	(-1.82)	(-1.70)	(-1.60)	(-1.69)	(-1.71)
Coastal pop.						0.000***	0.000***	0.000***	0.000*	0.000**	0.000**	0.003**	0.003**	0.003*	0.004**	0.005**	0.004**	0.004*
						(3.21)	(3.05)	(2.74)	(1.70)	(2.37)	(2.24)	(2.04)	(2.05)	(2.02)	(2.61)	(2.71)	(2.11)	(2.06)
Muslim pop.							0.148	0.270	0.514**	0.475*	0.530**	0.466*	0.469*	0.471*	0.523*	0.605**	0.608**	0.668**
							(0.58)	(1.04)	(2.13)	(1.89)	(2.10)	(1.89)	(1.90)	(1.85)	(1.88)	(2.24)	(2.22)	(2.42)
Buddhist pop.								0.970**	1.042**	0.779**	0.636**	0.650**	0.602*	0.604*	0.685*	1.268***	1.294***	1.248***
								(2.21)	(2.03)	(2.10)	(2.12)	(2.15)	(1.90)	(1.90)	(1.96)	(3.41)	(3.41)	(3.26)
Gov. expenditure									-0.011	-0.004	-0.003	-0.006	-0.010	-0.012	-0.020	-0.033	-0.033	-0.029
									(-0.46)	(-0.20)	(-0.15)	(-0.28)	(-0.45)	(-0.45)	(-0.71)	(-1.08)	(-1.03)	(-0.87)
RER distortions										-0.006***	-0.005**	-0.005**	-0.005**	-0.005**	-0.003	-0.002	-0.002	-0.002
										(-2.76)	(-2.46)	(-2.49)	(-2.25)	(-2.16)	(-1.19)	(-0.57)	(-0.58)	(-0.50)
Years open											0.419*	0.430*	0.432*	0.425*	0.283	0.152	0.147	0.108
											(1.88)	(1.90)	(1.88)	(1.78)	(0.87)	(0.43)	(0.40)	(0.27)
Pop. density												-0.003*	-0.003*	-0.003*	-0.004**	-0.005**	-0.005**	-0.005**
												(-1.93)	(-1.96)	(-1.92)	(-2.47)	(-2.52)	(-2.60)	(-2.42)
Tropic land													0.221	0.220	0.178	0.325	0.349	0.284
													(0.95)	(0.96)	(0.86)	(1.42)	(1.26)	(0.99)
Fiscal policy volatility														-0.040	-0.123	-0.084	-0.076	-0.059
														(-0.18)	(-0.49)	(-0.32)	(-0.26)	(-0.20)
Linguistic diffs.															0.129	0.176	0.184	0.234
															(0.31)	(0.43)	(0.45)	(0.56)
Genetic homogeneity																-21.893	-22.812	-14.896
																(-0.75)	(-0.75)	(-0.43)
Gen. homogeneity (sq)																35.957	37.406	24.892
																(0.77)	(0.76)	(0.45)
East-Asia																-1.042**	-1.044**	-1.021**
																(-2.59)	(-2.54)	(-2.50)
Confucian pop.																	2.724	1.835
																	(0.20)	(0.14)
Mining (%GDP)																		-1.406
																		(-0.80)
Obs.	67	67	67	67	67	67	67	67	59	59	59	58	58	58	49	49	49	49
Adj.-R^2	0.150	0.149	0.136	0.156	0.187	0.208	0.200	0.291	0.355	0.455	0.473	0.485	0.487	0.475	0.402	0.400	0.378	0.367
F-test	5.774	4.788	3.920	3.206	4.645	7.653	7.102	7.241	14.911	20.707	17.500	15.578	17.390	16.048	24.589	36.068	19.079	17.991
Prob$>F$	0.001	0.001	0.002	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Pooled cross-section OLS estimations using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the average growth rate for the period 1965-2007. Income data corresponds to PWT7.0. Detail for growth determinants can be found in the Appendix. $sk_{c,i}$ is the standardized i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

On the other hand, the introduction of some other controls has an importantly negative effect on the significance of the complexity index. In Table 13, we show our most basic growth regressions including only one of each of these controls and, as a result, the coefficient of $k_{c,18}$ turns non-significant. In the case above, as in the other cases omitted here, controls turning complexity non-significant are related to geographical location (Sub-Saharian Africa dummy) or deep-rooted institutional characteristics (reflected in labor allocation across genders and picked up by historical plough use or in life expectancy). Notice that, although column (2) shows a significant coefficient for $k_{c,18}$ when a dummy variable signalling Sub-Saharian african countries is included, that result is not robust to the inclusion of other controls in the specification as the next column reveals.

Table 13: Cross-section growth regressions, non-significant results for $sk_{c,18}$

Dep. var.:	Avg. growth rate 1965-2007.				
	(1)	(2)	(3)	(4)	(5)
GDPpc (logs)	-0.321*	-0.248*	-0.143	-0.088	-0.145
	(-1.88)	(-1.82)	(-1.15)	(-0.57)	(-1.07)
Primary enroll.	0.003	0.006	0.005	0.005	0.007
	(0.39)	(0.98)	(0.83)	(0.78)	(1.41)
Investment (%GDP)	0.006	0.004	0.002	0.009	-0.001
	(1.03)	(0.67)	(0.37)	(1.44)	(-0.17)
skc18	0.123	0.185*	0.088	0.134	0.045
	(0.96)	(1.80)	(0.87)	(1.41)	(0.43)
Life expectancy	0.030**				
	(2.60)				
S-S Africa		-0.542**	-0.822***		
		(-2.60)	(-2.81)		
Debt/GDP ratio in 1970				-0.001	
				(-0.25)	
Historical plough use					0.611***
					(4.17)
Investment price PPP			-0.000		
			(-0.15)		
Malaria prevalence			0.255		
			(0.88)		
LATAM			-0.484***		
			(-3.18)		
Obs.	67	67	67	55	66
Adj.-R^2	0.217	0.252	0.339	0.041	0.338
F-test	6.909	8.204	6.359	3.300	11.503
Prob> F	0.000	0.000	0.000	0.012	0.000

Notes: Pooled cross-section OLS estimations using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the average growth rate for the period 1965-2007. Income data corresponds to PWT7.0. Detail for growth determinants can be found in the Appendix. $sk_{c,i}$ is the standardized i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Finally, we also performed panel growth regressions with country-fixed effects, where the dependent variable is the average growth rate in 10-year periods starting in 1965, 1975, 1985, and 1995. Explanatory variables include MR complexity indicators together with primary education enrolment and investment level as controls. Results are presented in Table A.8 in the Appendix and do not yield significant coefficients for $sk_{c,18}$ in any of the specifications explored.[10] These results suggest that, at least for the period covered here, complexity indicators do not add much above country fixed effects

[10]In specifications including also $sk_{c,19}$, we find significant coefficients of this variable but the sign is at odds with intuition.

in the explanation of future growth. Altogether our results suggest that economic complexity could be determined by geographical conditions, natural endowments, deeply rooted institutional characteristics or some other form of relatively invariant growth determinants. Certainly, the study of what determines economic complexity, a matter that escapes the scope of the present work, should be subject of future work.

6 Potential issues arising from the database chosen

The use of export data to measure economic complexity could be contested. This section deals with some of the issues that are usually related to the construction of the MR using export data in general and six-digit export data in particular.

6.1 Increasing fragmentation in productive processes

The increasing fragmentation of production across national boundaries due to firms' strategical decisions, could be considered as a source of distortion to the ability of MR indicators to capture economic complexity. When firms decide to relocate a part of their production across national borders, international trade data may not reflect capabilities accurately. For instance, if a firm decides to design and build the most important pieces of a complex product but sends those pieces to a low wage country for assembly, when the final product is finally shipped to its consuming market it would count as exported by the second economy, even though that country does not necessarily has all capabilities required in manufacturing it.

To evaluate the effect that fragmentation has on the MR, we constructed the indicators upon a database that excludes final products with the aim of eliminating that source of distortion. The logic of this exercise is that complex economies should be able to export not only complex products but also its corresponding intermediates, while less complex economies are more prone to import intermediates and export final products. However, to eliminate trade in final consumption products from the analysis can also yield a great informational loss. We compare results stemming under this new setting with those previously obtained in order to evaluate how important final product trade information is to MR indicators. To filter out final products we use the upstreamness index (u_p) presented in Antràs et al. (2012). The index measures the average distance each product has to the final consumer market. MR indicators are constructed upon a database keeping only products with $u_p > 1.4362$, i.e. dropping those in the first quartile in the distribution of the index when considering only manufacturing industries. Products erased are therefore only those which are closer to final consumption.

The resulting MR indicators exhibit even greater jumps from one year to the next and their correlation to future growth is importantly harmed. So much so, that even imposing very restrictive filter levels for the country sample, significant coefficient for the MR are hardly found in regressions measuring the relationship of complexity indicators and long-term growth. Table 14 shows results following similar specifications as those in Table 7. Columns 5-8 in Table 14 correspond to a very restrictive filter level ($v^* = 40$) which leaves 135 observations. As can be seen, MR indicators do not present significant coefficients under such setting. In columns 1-4 an even more restrictive filter is applied ($v^* = 20$) leaving only 83 countries, and even in that context only one regression exhibits a significant coefficient for kc_{18} at 5% confidence level. Comparing these results with those in Table 7, it is clear that the removal of final consumption goods from the construction of the MR implies an informational loss and this loss is

Table 14: Long term regressions filtered for upstreamness

Dep. var.:	Average growth rate for the period 1995-2009							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDPpc	-0.000***	-0.000***	-0.000***	-0.000***	-0.000	-0.000	-0.000	-0.000
	(-3.14)	(-3.41)	(-3.28)	(-3.49)	(-0.64)	(-0.70)	(-0.63)	(-0.67)
Ec	0.049		0.060		-0.020		-0.018	
	(0.78)		(0.94)		(-0.41)		(-0.35)	
HHc	0.231		0.273		0.041		0.065	
	(0.58)		(0.70)		(0.15)		(0.24)	
kc18	9.469	16.284**	-248.337	-174.385	4.914	-1.573	-204.521	-201.185
	(0.83)	(2.12)	(-1.08)	(-0.76)	(0.23)	(-0.10)	(-0.83)	(-0.71)
kc19			-24692.997	-18396.422			-20185.803	-19244.160
			(-1.06)	(-0.86)			(-1.01)	(-0.71)
Obs.	83	83	83	83	135	135	135	135
Adj.-R^2	0.119	0.128	0.122	0.125	-0.015	-0.005	-0.019	-0.009

Notes: Pooled cross-section OLS estimations using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the average growth rate for the period 1995-2009. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant was also included in all specifications. Columns 1-4 correspond to $v^* = 20$, while columns 5-8 apply $v^* = 40$. * significant at 10%; ** significant at 5%; *** significant at 1%.

too large for the MR to keep its properties, which suggests that the indicators are actually nourished from information of trade in final goods.

6.2 Impact of classification used

It should be noticed that the six-digit HS classification in which BACI is presented, is quite uneven between different production lines. Having information at six-digits, instead of four-digits, can be very informative in some of these lines (where there can be up to 25 different products) while some other four-digit lines are not further disaggregated at six-digits. The heterogeneity in the classification is not strictly based in technological aspects but it mainly pursues legal/tax considerations. As a consequence of this, countries producing in product lines where disaggregation is higher will appear as more diversified and that could artificially bias the complexity measures of MR indicators. The next exercise proposed here shows that this is not the case since the iterative process of MR indicators gets rid of this distortive effect too.

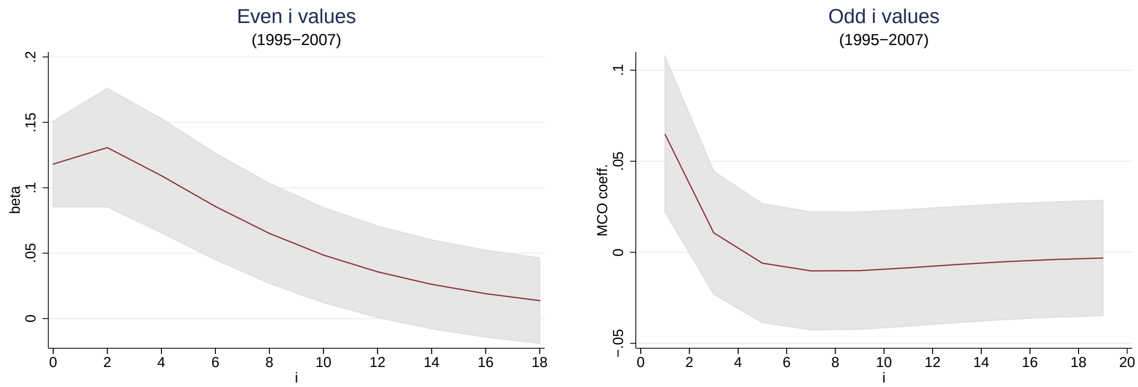
In a first step we construct an index that assigns a greater value to countries exporting products belonging to highly disaggregated product lines. Our proposal, sums all products exported by a country, weighting each by the number of different products listed in the six-digit classification inside each four-digit category (denoted n_p). The proposed index is therefore:

$$EXPn_c = \sum_{j=1}^{N_p} \frac{x_{c,i}}{\sum_{i=1}^{N_c} x_{j,p}} n_p \tag{8}$$

This index, a version of the $EXPY$ introduced in Hausmann et al. (2007), yields the average n_p of a country's export basket and therefore will take larger values for countries that export many of the products being largely disaggregated by the classification used here.

A high correlation between this indicator and MR complexity indicators would imply that economies evaluated as complex achieved that status partly because their diversification level was artificially enhanced by the classification used.

Figure 4: Correlation between the importance of diversification driven by 6-digit classification and $k_{c,i}$ as i grows. All countries (1995-2007).



Notes: The x-axis indicates the degree of iteration i of the $k_{c,i}$ indicator, while the y-axis measures the coefficient value of the standardized $EXPn_c$ as sole explanatory regressor of standardized $k_{c,i}$ in cross section regressions using White's consistent estimator. 95% confidence interval is represented by the shadowed area. All countries and years are included. An omitted constant was also included in all specifications. Even indicators (left) and odd indicators (right) are presented separately.

The second step is then to evaluate whether this diversification index is significantly correlated to different complexity measures. We do this by computing simple cross section regressions where $EXPn_c$ is the main regressor of each $k_{c,i}$ indicator. Figure 4 plots the main coefficient of these regressions for each value of i along with their 95% confidence interval. As expected, differences in the classification can be of some importance for low iterated MR indicators, since they are directly affected by the number of products being exported by a country. However, as the iterative process goes further the potential distortion loses significance.

6.3 Inverted-U relationship between income and diversification

Using six-digit product data Cadot et al. (2011) show that the relationship between diversification and per capita income is not linear but hump-shaped. This suggests that after surpassing certain development level, complex economies may find it optimal to specialize in a subset of high productivity goods, among all of those plausible of being produced according to their capabilities. Then, economies do not necessarily produce all products for which they have the required capabilities, a fact that could harm the MR. Still, given the way the MR is constructed, we expect the indicators to pick-up this pattern: if very rich but specialized countries export different products than other countries with the same degree of specialization but lower income levels, this should feed MR indicators with valuable information.

We evaluate the effect that this pattern could have by constructing the MR indicators upon a database that excluded countries to the right hand side of the hump (GDP per capita greater than \$24,000 in 2009, the final year of our sample). We find that doing so implies a more volatile complexity index kc_{18}. Also, our exercise does not increase but, if anything, reduces the correlation of the indicators with future growth. Table 15 report results for regressions evaluating the correlation between complexity indicators and long term growth, using similar specifications and applying the same country filter as in Table 7. Columns 1-4 correspond to a filter of $v^* = 47$ and leaves 149 observations, while columns 5-8 apply $v^* = 67$ leaving 161 countries in our sam-

Table 15: Long term regressions filtered for rich countries

Dep. var.:	Average growth rate for the period 1995-2009							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDPpc	-0.000***	-0.000***	-0.000***	-0.000***	-0.000	-0.000	-0.000	-0.000
	(-4.14)	(-3.99)	(-4.15)	(-3.96)	(-0.98)	(-0.97)	(-1.02)	(-1.03)
Ec	0.009		0.005		-0.003		0.009	
	(0.26)		(0.13)		(-0.07)		(0.25)	
HHc	0.210		0.211		0.029		0.031	
	(0.81)		(0.80)		(0.10)		(0.11)	
kc18	15.858**	14.347***	95.772	34.574	-0.191	-1.113	-211.686	-190.667
	(2.38)	(2.82)	(0.76)	(0.32)	(-0.02)	(-0.10)	(-0.62)	(-0.75)
kc19			7520.295	1922.676			-19936.380	-18051.464
			(0.65)	(0.19)			(-0.63)	(-0.78)
Obs.	149	149	149	149	161	161	161	161
Adj.-R^2	0.090	0.094	0.086	0.088	-0.006	0.007	-0.004	0.008

Notes: Pooled cross-section OLS estimations using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the average growth rate for the period 1995-2009. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant was also included in all specifications. Columns 1-4 correspond to $v^* = 47$, while columns 5-8 apply $v^* = 67$. * significant at 10%; ** significant at 5%; *** significant at 1%.

ple. Comparing these results to those in Table 7 shows that the significance of the relevant coefficient is not increased by our exercise. If the country filter applied is too loose then complexity indicators appear as non-significant in both tables. Applying a more restrictive filter is able to yield significant results in some cases but not others under the current exercise, which contrasts with that in Table 7 in which all specifications threw significant coefficients for complexity indicators (even with a larger country sample). The fact that excluding countries in the right hand side of the hump does not improve the indicators properties suggests that specialized rich countries also contribute with valuable information to the MR indicators and therefore they are not harmed but instead they nourish from the hump-shaped pattern.

7 Conclusions

Economic complexity as measured by the MR relies on information for one economy relative to all others. This means that results for each country are sensitive on the country sample used to build the indicators. There are plenty of sources of noise that affect the quality of the indicators. Abrupt commercial or productive policy changes, as well as episodes of internal instability, can greatly affect export figures, and therefore prevent MR indicators to accurately measure economic complexity. Our evidence suggests that the positive correlation between complexity and income per capita is strong enough, and the result is present even with our full sample of countries. However, the noise can prove too much when the correlation with future growth is evaluated. Indeed, we find that the correlation between complexity and future growth is non-significant when the country sample does not filter out cases for which $k_{c,18}$ varies greatly between any two years. Countries that need to be excluded so the index maintains this desirable property are not random. Rather they are characterized by having on average very low income, being small, highly specialised, closed to the world economy and exhibiting high volatility in their degree of openness. This should be kept in mind when using the MR to evaluate growth prospects of underdeveloped economies for which some of these characteristics may be present.

Our results indicate that complexity is significantly correlated with future growth also when controlling for many of the usual growth regressors, even outperforming most of them. Nevertheless, the indicator loses significance when variables related to geographical location (e.g. regional dummies or variables measuring topographical characteristics of the economy) are incorporated. The correlation also turns non-significant when including country fixed-effects. These results could be indicating that complexity is strongly correlated with deep growth fundamentals, a matter that could be explored in detail in future work.

References

Abdon, A. and Felipe, J. (2011). The Product Space: What Does It Say About the Opportunities for Growth and Structural Transformation of Sub-Saharian Africa? Technical Report 670, Levy Economic Institute of Bard College, New York.

Aghion, P. (2014). Afterword: Rethinking Industrial Policy. In Stiglitz, J. E. and Greenwald, B. C., editors, *Creating a Learning Society: A New Approach to Growth, Development and Social Progress*, pages 509–522. Columbia University Press, New York.

Alesina, A., Giuliano, P., and Nunn, N. (2013). On the Origins of Gender Roles: Women and the Plough. *The Quarterly Journal of Economics*, 128(May):469–530.

Antràs, P., Chor, D., Fally, T., and Hillberry, R. (2012). Measuring the Upstreamness of Production and Trade Flows. *American Economic Review Papers and Proceedings*, 102(3):412–416.

Ashraf, Q. and Galor, O. (2013). The Out of Africa Hypothesis, Human Genetic Diversity, and Comparative Economic Development. *American Economic Review*, 103(1):1–46.

Balassa, B. (1964). The Purchasing-Power Parity Doctrine: A Reappraisal. *Journal of Political Economy*, 72(6):584–596.

Baldwin, R. E. and Venables, A. J. (2015). Trade policy and industrialisation when backward and forward linkages matter. *Research in Economics*, 69(2):123–131.

Bolt, J. and van Zanden, J. (2014). The Maddison Project: collaborative research on historical national accounts. *The Economic History Review*, 67(3):627–651.

Cadot, O., Carrère, C., and Strauss-Kahn, V. (2011). Export Diversification: What's behind the Hump? *The Review of Economics and Statistics*, 93(2):590–605.

Ciccone, B. A. and Jarocinski, M. (2010). Determinants of Economic Growth: Will Data Tell? *American Economic Journal: Macroeconomics*, 2(4):222–246.

Easterly, W. and Levine, R. (1997). Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics*, 112(4):1203–1250.

Eberhardt, M. and Presbitero, A. F. (2015). Public debt and growth: Heterogeneity and non-linearity. *Journal of International Economics*, 97(1):45–58.

Fatás, A. and Mihov, I. (2013). Policy volatility, institutions, and economic growth. *The Review of Economics and Statistics*, 95(May):362–376.

Feenstra, R. C., Lipsey, R. E., Deng, H., Ma, A. C., and Mo, H. (2005). World Trade Flows: 1962-2000. Technical report, NBER.

Gallup, J. L., Mellinger, A., and Sachs, J. D. (2001). Geography Datasets. Technical report.

Gaulier, G. and Zignago, S. (2010). BACI: International Trade Database at the Product-Level. The 1994-2007 Version. Technical report, CEPII.

Hauk, W. R. and Wacziarg, R. (2009). A Monte Carlo study of growth regressions. *Journal of Economic Growth*, 14(May):103–147.

Hausmann, R. and Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, 16:309–342.

Hausmann, R., Hwang, J., and Rodrik, D. (2007). What you export matters. *Journal of Economic Growth*, 12:1–25.

Hausmann, R. and Klinger, B. (2006). South Africa's Export Predicament. Technical Report 129, CID-Harvard University.

Hausmann, R. and Rodrik, D. (2003). Economic Development As Self-Discovery. *Journal of Development Economics*, 72(2):603–633.

Heston, A., Summers, R., and Aten, B. (2011). Penn World Table Version 7.0. Technical report, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26):10570–5.

Hidalgo, C. A., Klinger, B., Barabási, A.-L., and Hausmann, R. (2007). The product space conditions the development of nations. *Science*, 317(5837):482–7.

Imbs, J. and Wacziarg, R. (2003). Stages of diversification. *American Economic Review*, 93(1):64–86.

Jankowska, A., Nagengast, A., and Perea, J. R. (2012). The Product Space and the Middle-Income Trap: Comparing Asian and Latin American Experiences. Technical Report 311, OECD Development Centre.

Kali, R., Reyes, J., McGee, J., and Shirrell, S. (2013). Growth networks. *Journal of Development Economics*, 101:216–227.

Koren, M. and Tenreyro, S. (2013). Technological Diversification. *American Economic Review*, 103(1):378–414.

Lall, S. (2000). The Technological Structure and Performance of Developing Country Manufactured Exports, 1985-98. *Oxford Development Studies*, 28(3):337–369.

Record, R. and Nghardsaysone, N. (2010). Export Dynamics and Diversification in Lao PDR: An Analysis of the Product Space. Technical Report August 2009, Lao PDR Trade Development Facility, Vientiane, LAO PDR.

Sachs, J. D. and Warner, A. M. (2001). The curse of natural resources. *European Economic Review*, 45(4-6):827–838.

Sala-i Martin, X. (1997). I Just Ran Two Million Regressions. *American Economic Review*, 87(2):178–183.

Sala-i Martin, X., Doppelhofer, G., and Miller, R. I. (2004). Determinants of Long-Term Growth : A Bayesian Averaging of Classical Estimates (BACE) Approach. *American Economic Review*, 94(4):813–835.

Appendix

Table A.1: All countries sorted by v.

n	country	v^m	pop	$GDPpc$	HH_c
1	Chad	98	9886	1216.7001	0.89
2	Somalia	85	9238	455.28954	0.08
3	Eritrea	83	5358	668.07809	0.49
4	Micronesia	81	108	3459.2452	0.24
5	Palau	77	21	15042.989	0.41
6	Kiribati	76	108	3623.0613	0.11
7	Iraq	75	27500	3948.5904	0.98
8	Marshall Isds	73	61	7350.8817	0.30
9	Guinea-Bissau	71	1473	790.90988	0.79
10	Brunei Dar.	68	374	49063.004	0.52
11	Rwanda	65	10141	924.10574	0.13
12	Djibouti	60	694	2163.5826	0.16
13	Bahamas	60	302	29114.779	0.18
14	C. African R.	58	4544	614.92852	0.19
15	Eq. Guinea	57	600	18065.461	0.75
16	Maldives	57	365	4689.1134	0.16
17	Liberia	56	3270	403.54628	0.28
18	Burundi	52	8783	370.88008	0.16
19	Qatar	48	815	123306.83	0.27
20	Comoros	46	711	927.71454	0.13
21	Jamaica	45	2782	9232.0147	0.26
22	Cuba	44	11160	10955.648	0.20
23	Bhutan	43	673	4077.332	0.15
24	Mali	43	12769	937.23069	0.54
25	Sierra Leone	43	4918	830.02159	0.14
26	Seychelles	43	86	26659.864	0.24
27	Angola	42	12264	4353.2594	0.91
28	Cape Verde	41	486	3498.4351	0.19
29	Saint Lucia	40	159	12684.021	0.12
30	Solomon Isds	40	567	1598.0084	0.51
31	Niger	40	14215	532.4731	0.19
32	Gabon	40	1457	10300.196	0.52
33	Mozambique	39	20906	715.72619	0.25
34	Tonga	38	104	7585.6323	0.12
35	Ant. &Barb.	37	83.425	17449.365	0.28
36	Suriname	36	470.784	10169.251	0.29
37	Samoa	36	188	7000.8877	0.36
38	Libya	36	6037	19822.039	0.73
39	Senegal	35	11394	1494.6751	0.04
40	Bermuda	35	67	50487.918	0.26
41	Afghanistan	35	31890	736.4802	0.03
42	Algeria	34	33363	6062.3346	0.33
43	Congo	33	3802	1912.8881	0.65
44	Armenia	33	2972	5603.8337	0.06
45	St. V. & Gren.	33	105	6774.4992	0.22
46	Azerbaijan	32	8120	7969.982	0.61
47	Yemen	32	21591	2421.0603	0.56
48	Belize	31	295	9251.5921	0.09
49	Tajikistan	31	7077	1764.0826	0.43
50	Tanzania	30	39384	1039.0978	0.05
51	Mongolia	29	2952	3040.4315	0.20
52	Cameroon	29	18060	1826.4193	0.28
53	Burkina Faso	29	14797	904.82764	0.51
54	Gambia	28	1630	1325.6739	0.10
55	Cte d'Ivoire	28	19747	1299.9685	0.09
56	Macao SAR	27	526	50192.209	0.03
57	Zambia	27	12341	1794.781	0.32
58	Viet Nam	26	86519	2584.623	0.03
59	Guinea	26	9569	834.1302	0.24
60	Togo	26	6042	743.42531	0.11
61	Turkmenistan	25	4774	6195.1448	0.56
62	Fiji	25	852	4470.6179	0.05
63	Bahrain	24	1054	24388.667	0.46
64	Dominican R.	23	9426	9341.0853	0.04
65	Albania	22	2992	6017.5483	0.02
66	Georgia	22	4646	4708.7586	0.04
67	Kuwait	22	2376	51608.868	0.45
68	Grenada	22	106	12910.239	0.11
69	Nigeria	21	143312	1941.6279	0.74
70	Ghana	21	22981	1143.4527	0.16
71	Sudan	21	40526	2065.7254	0.79
72	Benin	21	8278	1254.2203	0.13
73	Kyrgyzstan	21	5284	2034.3604	0.08
74	Lebanon	21	3896	11102.844	0.02
75	Lao	20	6035	2338.7379	0.21
76	Bosnia Herz.	20	4552	6245.4561	0.01
77	UAE	19	4444	53847.789	0.28
78	Vanuatu	19	212	6042.9865	0.25
79	Uganda	19	30263	1063.8559	0.04
80	Papua N. G.	19	5806	2417.5555	0.15
81	Sao T. &P.	18	165	1522.3805	0.18
82	Jordan	18	5997	4412.8462	0.01
83	Madagascar	18	19449	776.60522	0.03
84	Bolivia	18	9426	3619.4953	0.19
85	Zimbabwe	18	11443	158.69639	0.07
86	Nicaragua	18	5408	2113.1228	0.04
87	Paraguay	18	6113	3785.1792	0.13
88	China	17	1310584	5889.7756	0.01
89	Syria	17	20488	3955.6206	0.15
90	Cyprus	17	1049	19463.163	0.03
91	Guyana	17	764	4028.2658	0.10
92	Mauritania	17	2982	1619.2027	0.17
93	Macedonia	17	2056	7330.3366	0.03
94	Panama	16	3258	9290.753	0.02
95	Peru	16	28050	6750.1931	0.07
96	Nepal	16	27828	1092.1796	0.03
97	Dominica	16	72	5585.7857	0.07
98	Iran	16	74093	10059.125	0.65
99	Moldova	16	4329	2376.1807	0.02
100	Belarus	15	9725	11177.406	0.11
101	Oman	15	2800	20761.767	0.40
102	Kazakhstan	15	15285	10845.011	0.26
103	D. R. Congo	14	64355	227.13513	0.14
104	Uzbekistan	14	27079	2081.9791	0.08
105	Uruguay	14	3279	9959.7523	0.03
106	Romania	14	22106	9523.5197	0.01
107	Malaysia	14	26896	11643.636	0.02
108	Malawi	14	14233	581.42325	0.18

n	country	v^m	pop	$GDPpc$	HH_c
109	T. & Tobago	14	1233	26286.33	0.15
110	Mauritius	14	1264	9001.029	0.06
111	Ecuador	13	14135	5972.9726	0.26
112	Malta	13	402	21891.878	0.10
113	Thailand	13	65110	7752.4983	0.01
114	Egypt	13	75677	4684.5714	0.05
115	Greece	12	10706	27963.833	0.02
116	Guatemala	12	12728	6043.275	0.02
117	Indonesia	12	234694	3626.2218	0.02
118	Hong Kong	12	6980	37366.706	0.01
119	Honduras	12	7516	3625.6007	0.04
120	Kenya	12	36914	1233.4624	0.04
121	Tunisia	11	10213	5938.2831	0.02
122	Philippines	11	94158	2963.9609	0.05
123	Ethiopia	11	79936	593.06326	0.10
124	Mexico	11	108701	12696.946	0.03
125	Costa Rica	11	4331	11215.63	0.07
126	Croatia	11	4493	15620.298	0.02
127	Estonia	11	1316	19047.44	0.02
128	Russia	11	141378	14495.691	0.15
129	Singapore	11	4553	48215.077	0.05
130	El Salvador	10	5982	6501.7185	0.02
131	Haiti	10	9500	1383.0115	0.18
132	Sri Lanka	10	20508	3738.7763	0.02
133	Barbados	10	282	24556.108	0.05
134	Morocco	10	30594	3327.0202	0.01
135	Lithuania	10	3575	15648.443	0.02
136	Cambodia	10	13719	1800.3898	0.06
137	Pakistan	10	175495	2291.7809	0.01
138	Latvia	10	2260	15486.282	0.04
139	Venezuela	10	26415	9545.4248	0.46
140	S. Arabia	10	24499	20449.167	0.62
141	Bangladesh	10	148894	1290.071	0.05
142	Iceland	9	302	43125.299	0.09
143	Chile	8	16304	12135.42	0.14
144	R. of Korea	8	48250	25061.674	0.02
145	Bulgaria	8	7323	10529.308	0.02
146	Portugal	8	10643	20700.2	0.01
147	Poland	8	38518	15249.073	0.01
148	Colombia	8	42597	7522.1774	0.06
149	Turkey	7	74768	10549.486	0.01
150	Ukraine	7	46300	7123.4969	0.01
151	Australia	7	20750	40290.085	0.03
152	N. Zealand	6	4132	28297	0.02
153	Slovakia	6	5448	19495.099	0.02
154	India	6	1124135	2999.8866	0.02
155	Hungary	6	9956	17486.939	0.02
156	Norway	5	4628	50959.308	0.17
157	Denmark	5	5468	36335.751	0.01
158	Argentina	4	40049	11344.785	0.03
159	Canada	4	32936	37703.047	0.02
160	Israel	4	6990	25473.586	0.06
161	Brazil	3	193919	9040.1881	0.01
162	Italy	3	59627	30199.592	0.01
163	Bel.-Lux.	3	10392	35573.98	0.01
164	France	3	63682	32015.431	0.01
165	Slovenia	3	2009	26593.094	0.01
166	Spain	3	45212	29133.475	0.01
167	Ireland	3	4420	39168.021	0.04
168	Czech R.	2	10229	23518.813	0.01
169	Switzerland	2	7555	39912.063	0.02
170	Netherlands	2	16571	40691.157	0.01
171	Germany	1	82237	33638.221	0.01
172	USA	1	301580	43697.458	0.01
173	Sweden	1	9031	37358.676	0.01
174	Finland	1	5239	34888.82	0.02
175	UK	1	61249	35649.013	0.01
176	Austria	1	8200	38232.841	0.00
177	Japan	0	127434	34223.768	0.01

Notes: v^m represents the maximum variation in the $k_{c,18}$ ranking that each country exhibited in the period 1995-2007. *pop* is the total population (in thousands) and $GDPpc$ is PPP converted GDP Per Capita (Laspeyres) at 2005 constant prices as reported in PWT (year 2007). HH_c is the Hischmann-Herfindahl concentration index.

Figure A.1: Histogram of highly-iterated complexity indicators (2007).

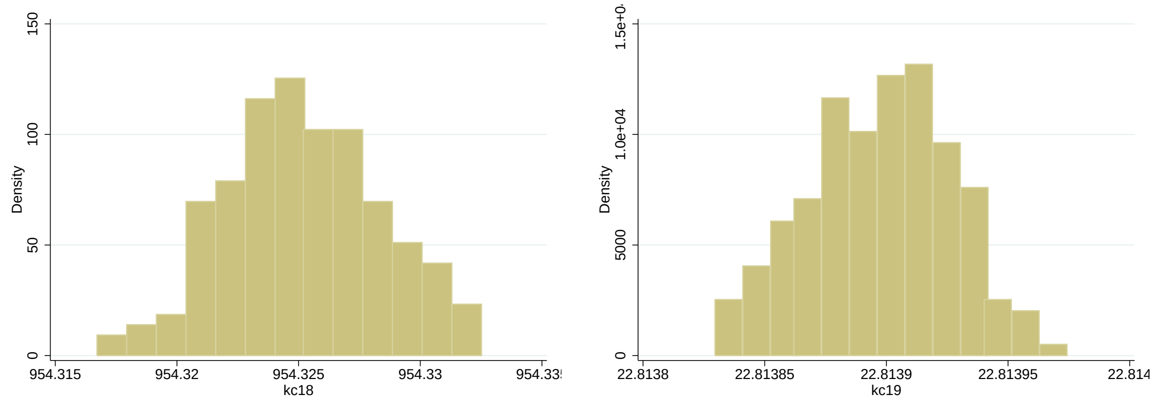


Table A.2: Regressions for long-term growth ($v^* = 57$ and different R^* values).

	Dep. var.: avg. gr. rate of p/c GDP (1995-2009)							
R^*=	0.8				0.9			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDPpc	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(-1.227)	(-1.258)	(-1.258)	(-1.244)	(-1.177)	(-1.169)	(-1.209)	(-1.290)
Ec				-0.017				0.005
				(-0.473)				(0.135)
HHc				-0.103				0.106
				(-0.386)				(0.418)
kc12	0.991				-0.311			
	(0.231)				(-0.105)			
kc13	111.011				-39.791			
	(0.212)				(-0.125)			
kc18		6.104	22.331	32.884		1.845	-160.606	-155.881
		(0.434)	(0.094)	(0.126)		(0.170)	(-0.776)	(-0.794)
kc19			2,018.300	3,051.174			-18,135.561	-17,712.213
			(0.070)	(0.097)			(-0.835)	(-0.864)
Obs.	162	162	162	162	163	163	163	163
Adj.-R^2	0.000	0.006	-0.000	-0.012	0.001	0.007	0.006	-0.005
F-test	1.128	1.089	0.791	0.456	1.105	1.289	0.730	0.761
Prob> F	0.340	0.339	0.455	0.768	0.349	0.278	0.536	0.579

	Dep. var.: avg. gr. rate of p/c GDP (1995-2009)							
R^*=	1.1				1.2			
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
GDPpc	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(-1.066)	(-1.082)	(-1.016)	(-1.018)	(-0.913)	(-0.924)	(-0.921)	(-0.943)
Ec				-0.031				-0.013
				(-0.826)				(-0.323)
HHc				-0.056				-0.011
				(-0.208)				(-0.043)
kc12	0.228				0.187			
	(0.119)				(0.122)			
kc13	19.196				16.631			
	(0.118)				(0.144)			
kc18		0.742	118.316	198.055		-0.417	23.589	53.308
		(0.102)	(0.951)	(1.292)		(-0.065)	(0.181)	(0.284)
kc19			10,449.509	17,217.453			1,909.548	4,157.930
			(0.980)	(1.328)			(0.190)	(0.290)
Obs.	163	163	163	163	164	164	164	164
Adj.-R^2	-0.001	0.005	0.004	-0.003	-0.002	0.004	-0.002	-0.014
F-test	1.349	1.214	0.727	0.654	1.387	1.143	0.422	0.266
Prob> F	0.261	0.300	0.538	0.659	0.249	0.322	0.737	0.900

Notes: Cross-section OLS regressions using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is 15-year average growth rate and values from the initial year (1995) are used for the explanatory variables. *GDPpc* is PPP converted GDP Per Capita (Laspeyres) at 2005 constant prices as reported in PWT7.0. E_c is (-1 times) Theil's entropy index of diversification. HH_c is the Hischmann-Herfindahl concentration index. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.3: Regressions for short-term growth excluding the last 5-year period and $v^* = 57$. Different values of R^*.

	Dep. var.: avg. growth (5-year starting in 1995 and 2000)							
$R^* =$	.8		.9		1.1		1.2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDPpc	-0.000	-0.000	-0.000	-0.000	-0.000*	-0.000*	-0.000	-0.000
	(-1.471)	(-1.554)	(-1.369)	(-1.505)	(-1.756)	(-1.771)	(-0.839)	(-0.960)
Ec		0.010		0.013		0.006		0.011
		(0.854)		(1.113)		(0.664)		(1.016)
HHc		0.052		0.086		0.027		0.083
		(0.437)		(0.732)		(0.367)		(0.783)
kc18	3.103	1.712	2.120	1.007	3.734***	3.241**	0.336	-0.105
	(0.502)	(0.264)	(0.437)	(0.201)	(2.675)	(2.131)	(0.144)	(-0.043)
kc19	-24.103	-13.294	-20.524	-9.744	-42.495***	-36.882**	-3.819	1.212
	(-0.502)	(-0.264)	(-0.437)	(-0.201)	(-2.675)	(-2.130)	(-0.143)	(0.043)
Obs.	314	314	312	312	318	318	318	318
Adj.-R^2	0.002	-0.002	0.003	-0.001	0.027	0.022	-0.003	-0.007
F-test	1.972	1.272	1.864	1.072	2.537	1.568	0.667	0.532
Prob> F	0.141	0.281	0.157	0.376	0.0567	0.169	0.514	0.752

Notes: Pooled cross-section OLS estimations using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the average growth rate for 5-year periods starting in 1995 and 2000. Explanatory variables take their initial value of each 5-year period. *GDPpc* is PPP converted GDP Per Capita (Laspeyres) at 2005 constant prices as reported in PWT7.0. E_c is (-1 times) Theil's entropy index of diversification. HH_c is the Hischmann-Herfindahl concentration index. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.4: Regressions for short-term growth excluding the last 5-year period and including country fixed-effects.

Dep. var.:	avg. growth (5-yr period starting in 1995 and 2000)					
	(1)	(2)	(3)	(4)	(5)	(6)
GDPpc	-0.000***	-0.000***	-0.000**	-0.000***	-0.000***	-0.000**
	(-2.69)	(-2.94)	(-2.59)	(-2.60)	(-2.68)	(-2.57)
Ec				0.021		0.013
				(0.36)		(0.17)
HHc					-0.097	-0.047
					(-0.22)	(-0.07)
kc12	0.010					
	(0.06)					
kc13	-0.094					
	(-0.05)					
kc18		0.002**	-2.367	-3.022	-2.693	-2.936
		(2.18)	(-0.40)	(-0.47)	(-0.44)	(-0.47)
kc19			26.913	34.349	30.606	33.364
			(0.40)	(0.47)	(0.44)	(0.47)
Obs.	354	354	354	354	354	354
Countries	177.000	177.000	177.000	177.000	177.000	177.000
Adj.-R^2	0.090	0.093	0.091	0.090	0.090	0.088
F-test	3.335	4.594	2.955	2.367	2.306	2.419
Prob> F	0.038	0.033	0.034	0.073	0.078	0.050

Notes: Panel estimations with country fixed-effects. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the average growth rate for 5-year periods starting in 1995 and 2000. Explanatory variables take their initial value of each 5-year period. *GDPpc* is PPP converted GDP Per Capita (Laspeyres) at 2005 constant prices as reported in PWT7.0. E_c is (-1 times) Theil's entropy index of diversification. HH_c is the Hischmann-Herfindahl concentration index. $k_{c,i}$ is the i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.5: Controls used in growth regressions

var name	Description	Data source
East-Asia	Dummy for East-Asian countries.	Own construction following https://en.wikipedia.org/wiki/East_Asia
Primary enroll.	Enrolment rate in primary education (avg. 1962-1982).	Own construction using SE.PRM.TENR in WDI
Investment price PPP	Investment price level (avg. 1960-1964) PPP.	pi in PWT6.3 in Heston et al. (2011)
GDPpc (logs)	Log of GDP per capita in 1960.	rgdpl PWT6.3 in Heston et al. (2011)
Tropic land	Proportion of country's land area within geographical tropics.	lnd100km in geodata.dta in Gallup et al. (2001)
Coastal pop.	Coastal (within 100 km of coastline) population per coastal area in 1960's 1965.	dens65c in geodata.dta in Gallup et al. (2001)
Malaria prevalence	Index of malaria prevalence in 1966.	Mal66a in malaria.dta in Gallup et al. (2001)
Life Expectancy	Life expectancy in 1960.	X2 in Sala-i Martin (1997)
Confucian pop.	Fraction of population Confucian in 1960.	X53 in Sala-i Martin (1997)
S-S Africa	Dummy for Sub-Saharan African countries.	X4 in Sala-i Martin (1997)
LATAM	Dummy for Latin American countries.	X5 in Sala-i Martin (1997)
Mining (%GDP)	Fraction of GDP in mining.	X59 in Sala-i Martin (1997)
Frm Spanish colony	Dummy for former Spanish colonies.	X50 in Sala-i Martin (1997)
Years open	Number of years economy has been open between 1950 and 1994.	X23 in Sala-i Martin (1997)
Muslim pop.	Fraction of population Muslim in 1960.	X56 in Sala-i Martin (1997)
Buddhist pop.	Fraction of population Buddhist in 1960.	X51 in Sala-i Martin (1997)
Linguistic diffs.	Average of five different indices of ethnolinguistic fractionalization which is the probability of two random people in a country not speaking the same language.	muller in othervar.dta in Easterly and Levine (1997)
Gov. expenditure	Share of expenditures on government consumption to GDP in 1961.	NE.CON.GOVT.ZS in WDI
Pop. density	Population per area in 1960.	EN.POP.DNST in WDI
RER distortions	Real exchange rate distortions.	X41 in Sala-i Martin (1997).
Investment (%GDP)	Investment Share of Real GDP per capita (rgdpl).	ki in PWT6.3
Debt/GDP ratio in 1970	Debt to GDP ratio in 1970	Debt in data imf.dta in Eberhardt and Presbitero (2015)
Historical plough use	Historical usage of animal plough cultivation	plow in crosscountry dataset.dta in Alesina et al. (2013)
Genetic homogeneity	Predicted index of genetic diversity for contemporary national populations	phomo in country.dta Ashraf and Galor (2013)
Fiscal policy volatility	Fiscal policy volatility computed as in Fatás and Mihov (2013)	Data from PWT6.3 in Heston et al. (2011)

Table A.6: Cross-section growth regressions, robustness of significance for $sk_{c,18}$

Dep. var.:	Avg. growth rate 1965-2007.									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
GDPpc (logs)	-0.253	-0.221	-0.283*	-0.387***	-0.345**	-0.259*	-0.315*	-0.379***	-0.444***	-0.426***
	(-1.52)	(-1.37)	(-1.69)	(-2.98)	(-2.54)	(-1.80)	(-1.96)	(-2.90)	(-3.04)	(-2.90)
Primary enroll.	0.010*	0.009	0.010*	0.010*	0.012*	0.011*	0.016**	0.016***	0.015**	0.014**
	(1.74)	(1.63)	(1.68)	(1.74)	(1.94)	(1.80)	(2.43)	(2.73)	(2.64)	(2.52)
Investment (%GDP)	0.001	-0.002	0.002	0.001	0.000	-0.002	-0.004	-0.004	-0.005	-0.005
	(0.12)	(-0.21)	(0.25)	(0.10)	(0.03)	(-0.35)	(-0.63)	(-0.77)	(-0.85)	(-0.85)
skc18	0.262**	0.227**	0.275**	0.254**	0.276**	0.259**	0.357**	0.389***	0.302**	0.308**
	(2.31)	(2.06)	(2.45)	(2.23)	(2.33)	(2.10)	(2.47)	(3.07)	(2.21)	(2.22)
Investment price PPP		-0.002	-0.002	-0.001	-0.001	-0.001	-0.002	-0.000	0.000	0.000
		(-1.24)	(-1.11)	(-0.89)	(-0.76)	(-0.65)	(-1.12)	(-0.10)	(0.15)	(0.12)
Coastal pop.			-0.000*	-0.000*	-0.000*	-0.000**	-0.000**	-0.000***	-0.000***	0.001
			(-1.76)	(-1.93)	(-1.71)	(-2.13)	(-2.64)	(-2.88)	(-2.77)	(0.86)
Malaria prevalence				-0.334	-0.250	-0.214	-0.114	0.081	-0.014	0.010
				(-1.25)	(-0.89)	(-0.78)	(-0.40)	(0.32)	(-0.05)	(0.04)
Muslim pop.					0.292	0.368*	0.583**	0.582***	0.589***	0.561***
					(1.32)	(1.72)	(2.64)	(2.90)	(2.88)	(2.79)
Buddhist pop.						1.015**	1.000**	0.689*	0.480*	0.494*
						(2.37)	(2.04)	(1.87)	(1.86)	(1.95)
Gov. expenditure							-0.004	0.006	0.006	0.006
							(-0.16)	(0.29)	(0.26)	(0.24)
RER distortions								-0.007***	-0.007***	-0.007***
								(-3.34)	(-3.11)	(-3.07)
Years open									0.500*	0.502*
									(2.01)	(1.95)
Pop. density										-0.001
										(-1.10)
Obs.	66	66	66	66	66	66	58	58	58	57
Adj.-R^2	0.147	0.157	0.164	0.175	0.185	0.273	0.378	0.499	0.526	0.516
F-test	4.662	3.339	3.375	3.692	2.942	2.641	4.579	10.174	12.358	11.230
Prob> F	0.002	0.010	0.006	0.002	0.008	0.013	0.000	0.000	0.000	0.000

Notes: Pooled cross-section OLS estimations using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the average growth rate for the period 1965-2007. Income data from the Maddison Project. Detail for growth determinants can be found in the Appendix. $sk_{c,i}$ is the standardized i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.7: Cross-section growth regressions, robustness of significance for $sk_{c,18}$ (cont.)

Dep. var.:	Avg. growth rate 1965-2007.									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
GDPpc (logs)	-0.416***	-0.444***	-0.384***	-0.415***	-0.334*	-0.329	-0.300	-0.289	-0.243	-0.205
	(-2.81)	(-3.06)	(-2.95)	(-3.38)	(-1.76)	(-1.68)	(-1.41)	(-1.40)	(-1.16)	(-1.02)
Primary enroll.	0.013*	0.013**	0.009*	0.009*	0.007	0.007	0.006	0.006	0.007	0.005
	(1.97)	(2.10)	(1.72)	(1.75)	(1.10)	(1.11)	(0.82)	(0.84)	(0.89)	(0.68)
Investment (%GDP)	-0.003	-0.003	-0.000	0.003	0.002	0.002	0.001	0.002	0.002	0.001
	(-0.56)	(-0.56)	(-0.00)	(0.70)	(0.47)	(0.46)	(0.19)	(0.33)	(0.42)	(0.15)
skc18	0.318**	0.337**	0.243**	0.347***	0.387***	0.387**	0.432**	0.530***	0.474**	0.352*
	(2.30)	(2.45)	(2.13)	(3.26)	(2.79)	(2.74)	(2.59)	(2.99)	(2.31)	(1.72)
Investment price PPP	0.001	0.001	0.002	0.002	0.002	0.001	0.000	0.001	0.001	0.001
	(0.61)	(0.75)	(1.05)	(1.50)	(0.65)	(0.63)	(0.16)	(0.43)	(0.30)	(0.18)
Coastal pop.	0.001	0.001	0.001	0.002	0.003*	0.003	0.004	0.004*	0.004*	0.003
	(1.04)	(0.93)	(0.90)	(1.57)	(1.97)	(1.67)	(1.64)	(1.81)	(1.81)	(1.49)
Malaria prevalence	0.007	0.001	-0.199	-0.061	0.039	0.044	0.125	0.160	0.140	0.168
	(0.02)	(0.00)	(-0.74)	(-0.21)	(0.11)	(0.12)	(0.31)	(0.41)	(0.34)	(0.42)
Muslim pop.	0.594***	0.649***	0.625***	0.656***	0.691**	0.689**	0.690**	0.768**	0.814***	0.420
	(2.97)	(2.99)	(2.90)	(3.09)	(2.48)	(2.45)	(2.36)	(2.75)	(2.85)	(1.29)
Buddhist pop.	0.461	0.492	0.518*	0.515**	0.608*	0.609*	0.711*	1.194***	1.157***	0.919**
	(1.49)	(1.51)	(1.76)	(2.07)	(1.82)	(1.78)	(1.80)	(3.25)	(3.00)	(2.27)
Gov. expenditure	0.000	0.003	0.009	-0.003	-0.007	-0.010	-0.014	-0.025	-0.022	-0.035
	(0.02)	(0.10)	(0.38)	(-0.16)	(-0.30)	(-0.32)	(-0.45)	(-0.73)	(-0.62)	(-1.32)
RER distortions	-0.006***	-0.006***	-0.005**	-0.006***	-0.005*	-0.004	-0.005	-0.003	-0.003	-0.001
	(-2.96)	(-2.83)	(-2.58)	(-2.81)	(-1.86)	(-1.64)	(-1.70)	(-1.10)	(-0.88)	(-0.49)
Years open	0.527**	0.568**	0.609**	0.658***	0.627*	0.607*	0.657*	0.528	0.481	0.256
	(2.11)	(2.17)	(2.54)	(2.82)	(1.74)	(1.76)	(1.90)	(1.56)	(1.34)	(0.72)
Pop. density	-0.002	-0.002	-0.002	-0.003**	-0.004**	-0.004**	-0.005**	-0.005**	-0.005**	-0.004*
	(-1.28)	(-1.17)	(-1.27)	(-2.03)	(-2.40)	(-2.22)	(-2.14)	(-2.24)	(-2.17)	(-1.85)
Tropic land	0.247	0.223	0.223	0.080	0.132	0.125	0.243	0.340	0.287	0.275
	(1.04)	(0.95)	(1.07)	(0.38)	(0.50)	(0.47)	(0.79)	(1.11)	(0.89)	(0.93)
LATAM		0.099	0.124	0.788	0.825	0.819	0.685	0.594	0.673	0.458
		(0.50)	(0.65)	(1.66)	(1.48)	(1.44)	(0.87)	(0.74)	(0.81)	(0.63)
Confucian pop.			2.393***	2.304***	1.179	0.635	2.025	2.389	1.607	0.000
			(4.98)	(5.53)	(0.09)	(0.04)	(0.13)	(0.16)	(0.11)	(.)
Frm Spanish colony				-0.822*	-0.896*	-0.908*	-1.009*	-1.011*	-1.015*	-0.726
				(-1.78)	(-1.88)	(-1.73)	(-1.92)	(-1.86)	(-1.84)	(-1.30)
Linguistic diffs.					0.265	0.251	0.274	0.316	0.354	0.247
					(0.58)	(0.54)	(0.55)	(0.68)	(0.75)	(0.52)
Fiscal policy volatility						-0.042	-0.004	0.013	0.014	-0.042
						(-0.16)	(-0.01)	(0.04)	(0.05)	(-0.15)
Genetic homogeneity							-35.669	-41.567	-32.948	-34.793
							(-0.94)	(-1.07)	(-0.74)	(-0.69)
Gen. homogeneity (sq)							56.436	66.966	53.103	57.254
							(0.93)	(1.08)	(0.74)	(0.72)
East-Asia								-0.935**	-0.909**	-0.725
								(-2.32)	(-2.24)	(-1.69)
Mining (%GDP)									-1.259	-0.430
									(-0.63)	(-0.20)
Historical plough use										0.558
										(1.50)
Obs.	57	57	57	57	48	48	48	48	48	47
Adj.-R^2	0.520	0.512	0.594	0.653	0.418	0.398	0.370	0.395	0.379	0.447
F-test	10.538	9.828	157.362	234.927	14.515	14.733	12.280	21.754	19.346	13.268
Prob> F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Pooled cross-section OLS estimations using White's consistent estimator. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the average growth rate for the period 1965-2007. Income data from the Maddison Project. Detail for growth determinants can be found in the Appendix. $sk_{c,i}$ is the standardized i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.8: Panel regressions for long-term growth with country fixed effect

Dep. var.:	10-avg growth rate.					
	(1)	(2)	(3)	(4)	(5)	(6)
GDPpc (logs)	-0.330***	-0.329***	-0.056	-0.424***	-0.424***	-0.378***
	(-4.77)	(-4.72)	(-0.98)	(-6.19)	(-6.12)	(-4.52)
Investment (%GDP)		-0.000	0.005**		0.000	-0.001
		(-0.22)	(2.02)		(0.10)	(-0.41)
Primary enroll.			0.001			0.000
			(1.13)			(0.03)
skc18	0.012	0.014	-0.054	0.029	0.028	-0.044
	(0.24)	(0.27)	(-1.53)	(0.63)	(0.61)	(-0.63)
skc19				0.086***	0.086***	0.066***
				(5.28)	(5.21)	(3.79)
Obs.	403	403	1651	403	403	133
Countries	124	124	115	124	124	72
Adj.-R^2	0.196	0.195	0.036	0.294	0.293	0.334
F-test	16.151	10.742	2.060	17.604	13.405	7.325
Prob> F	0.000	0.000	0.091	0.000	0.000	0.000

Notes: Panel estimations with country fixed-effects. Heteroskedasticity robust t-statistics in parentheses. Dependent variable is the 10-year average growth rate for each country starting in 1965, 1975, 1985 and 1995. Controls are initial level per capita gdp (in logs), investment share of GDP and enrollment in primary education. $sk_{c,i}$ is the standardized i-th component of the k_c vector. An omitted constant was also included in all specifications. * significant at 10%; ** significant at 5%; *** significant at 1%.