
MEASURING POLICY PERFORMANCE:

Can we do better than the World Bank?

April 2013

Julia Cagé¹²

Department of Economics
Littauer Center
Harvard University

Abstract

This article questions the relevance of the different measures of policy performance that are currently used by international organizations. It evaluates more especially the pertinence of the World Bank's CPIA and of the various alternatives that have been proposed. Using a cross-country panel dataset over 146 “developing” countries between 1977 and 2008, I show that the CPIA is a blunt and biased tool that can and should be improved upon. In particular, I show that while the CPIA is correlated with current growth, it is not a good predictor for future growth. I thus argue in favor of other measures of policy performance. First, I underline the need of introducing new criteria when measuring policy performance, in particular proxies for the development of fiscal capacity (e.g. domestic tax revenues) and the quality of industrial policy (e.g. export promotion strategies). This is of particular importance to bring sustained growth to sub-Saharan countries. Second, focusing more specifically on the allocation of development aid, I show that performance-based measures—as opposed to measures implying *ex ante* conditionality like the CPIA—are more accurate instruments for aid allocation. Finally, I make concrete proposals for the development of new performance indicators: the idea is to use “aid effectiveness” to allocate aid selectively.

¹ I am particularly thankful to Ravi Kanbur, Akbar Noman, Thomas Piketty and Joseph Stiglitz for helpful comments and suggestions. Participants at the Initiative for Policy Dialogue Task Forces also provided valuable advice. A first draft of this paper was prepared for the Tokyo International Conference on African Development (TICAD). The usual disclaimer applies.

² cage@fas.harvard.edu

1. INTRODUCTION

This article questions the relevance of the different measures of policy performance that are currently used. It especially evaluates the pertinence of the World Bank's Country Policy and Institutional Assessment (CPIA) and of the various alternatives that have been proposed in the literature. It argues in favor of new measures of policy performance.

Measuring policy performance is of particular importance today. First, the current global economic crisis has been very harmful for developing countries, in particular sub-Saharan countries, with a slowdown of capital flows, trade flows, flows of remittances and development aid flows. Development aid is both a limited and needed resources for developing countries. Finding criteria to allocate it selectively is thus of great concern for donors.

Second, finding an accurate measure of policy performance is of particular importance in the context of sub-Saharan countries which, despite the revival of growth in the past decade, have made little progress on the path to "sustained" growth. No important structural changes have taken place in the majority of sub-Saharan countries. Striking is the fact that the share of manufacturing and formal sector employment are still declining in these countries. A good measure of policy performance has thus to be concerned, with policies that would bring about economic transformations, structural changes and sustained growth. I discuss the relevance of existing indicators and propose a new approach in light of these concerns.

Using a yearly panel dataset over 146 developing and emerging countries between 1977 and 2008, I first show that the CPIA is correlated with current growth rates. This contemporaneous correlation can be explained in part by the fact that the assessments of the World Bank staff are colored by perceptions of countries current performances. Next I show that the CPIA is not a good predictor for future economic growth. I also find a positive and statistically significant correlation between developing countries' votes in the United Nations General Assembly (UNGA) with those of the United States and their CPIA scores. This is obviously subject to a variety of interpretations about causality, but it can be seen as an indication of the influence of a pro-United States disposition in foreign policy on the CPIA. At the very least, this shows that whatever matters, it is more strongly correlated to UN votes than to future growth prospects, which seems odd.

I thus argue in favor of other measures of policy performance. First, I underline the need for introducing new criteria when measuring policy performance. In particular, I show that more weight has to be given to the development of state capacities, which supposes to take into account its fiscal capacity. Fiscal capacity is indeed of crucial importance for raising domestic tax revenues. One has to go further than just underlying the need for "good institutions" when emphasizing the role of the state. Moreover, I underline the importance of the quality of industrial policy and especially of export promotion strategies on the path towards sustained growth.

I then focus more specifically on the allocation of development aid, which is of great importance given both the scarcity and need for aid. I show that performance-based measures, as opposed to measures implying *ex ante* conditionality, are more accurate instruments for aid

allocation. In particular, I discuss the relevance of the Kanbur's proposal (Kanbur, 2005) of introducing some "outcome criteria" in the CPIA. I show that introducing straightforward outcome variables will be a significant improvement on the CPIA but will leave some difficulties unsolved.

Finally, I make concrete proposals for the development of new performance indicators. The basic idea beyond these indicators would be to use "aid effectiveness" to allocate aid selectively. Such indicators were supposed to compute one "aid effectiveness" coefficient per country and year. I show how this can be done using the "local Gaussian-weighted ordinary least squares" econometric method.

1.1. RELATED LITERATURE

This paper is first related to the literature which evaluates the relevance of the existing indicators measuring policy performance. The main focus of this literature is the World Bank.³ Kurtz and Schrank (2007) evaluate the World Bank's coding of "good governance" by exploiting the time-dimension of the data. Using Granger-style causality tests, they found weak support for the notion that "better governance" was connected with successive improvements in growth. Other studies focus more specifically on the CPIA. Since the CPIA data was not disclosed until recently, they mainly emanate from the World Bank. For example, Gelb Ngo and Ye (2004) show that CPIA ratings have been quite strongly associated with medium-run growth performance. On the contrary, in a review of the performance-based allocation system, the World Bank (World Bank, 2001) underlines that, on average, CPIA ratings may be considerably affected by contemporaneous growth, with only modest predictive power with respect to future growth or poverty reduction. This is consistent with the empirical findings I obtain in this paper. Kraay and Nehru (2004) use CPIA ratings and find a significant inverse correlation between the quality of a country's policies and institutions on the one hand and its probability of debt distress on the other hand. Outside of the World Bank, however, other articles are much more critical toward the CPIA. For example, Herman (2004) calls for appreciating the weaknesses of this indicator, especially its low ability to discriminate among countries or over time. This paper contributes to this debate by analyzing both conceptually and empirically the relevance of the CPIA.

This paper is also related to the literature which proposes alternatives to the CPIA or other performance indicators.

Kanbur (2005) argues in favor of introducing some outcome variables in the CPIA. His proposal is in the spirit of Collier *et. al.* (1997)'s outcomes-based allocation. They indeed propose a basis for aid allocation in terms of retrospective assessment of a few major outcomes such as growth. Similarly, Barder and Birdsall (2006) defend the idea of *payments for progress*. I

³ An exception is Stuckler *et. al.* (2009) who show, using the EBRD's own data, that the EBRD's indices of progress in market reforms are biased in the direction of positive growth.

contribute to this literature by opening the way for a new indicator that improves upon the previous proposals that have been done.

Finally, this paper is related to a growing literature on optimal aid allocation, which emphasizes the necessity to take into account the level of policies as a selectivity criterion. This necessity was first highlighted by Burnside and Dollar (2000) and further examined in the World Bank report *Assessing Aid* (World Bank, 1998). Using these findings, Collier and Dollar (1999) derive an effective allocation of aid in terms of poverty reduction and compare it to the current allocation. They find out that the current allocation is radically different from the allocation which would be effective on poverty reduction. They stress the fact that an optimal allocation of aid not only depends on levels of poverty but also on the political environment. Moreover, they further develop this idea by applying their approach to the dynamic question of poverty reduction Collier and Dollar (2001). Cogneau and Naudet (2007) propose an alternative allocation based on the principle of equality of opportunity: they take into account structural growth handicaps rather than the quality of past policies (see also Llavador and Roemer, 2001). In the same spirit, Wood (2007) presents a more general model of optimal allocation of aid, in which donors take into account future poverty as well as current poverty. Finally, Amprou *et al.* (2006) argue in favor of considering vulnerability to exogenous shocks and low level of human capital as selectivity criteria. All this literature takes the level of policy or of performance in the effective use of development assistance as given. On the contrary in this article, I try to determine the optimal performance indicator one can use in the selectivity formula for aid allocation.

The rest of the article is organized as follows. Section 2 presents descriptive evidence about the CPIA and assesses its relevance. Section 3 provides new ways to improving upon the CPIA. Section 4 concludes.

2. THE CPIA, A GOOD MEASURE OF POLICY PERFORMANCE?

One of the most influential tools for measuring policy performance as of today is the Country Policy and Institutional Assessment (CPIA), which is the World Bank's tool. In this section, I analyze the relevance of this tool as a measure of policy performance.

2.1. SOME HISTORY OF THE CPIA

Since 1977, the World Bank has carried out an annual performance assessment of its client countries' capacity to effectively absorb development assistance. This assessment, the CPIA, is one of the main criteria used to allocate International Development Assistance (IDA) resources between low-income developing countries. The CPIA is an assessment tool for the Bank, to gauge the likely return to development assistance in specific countries and to guide IDA allocations to countries below the income threshold. CPIA assessments do not directly reflect specific "outcome" criteria as set out in the Millennium Development Goals (MDGs)—e.g. poverty reduction, school enrollment, maternal health, etc.—neither do they directly rest on

proxy outcome variables such as GDP, export or investment growth rates. The emphasis is on policy actions and institutional effectiveness. They rely on the judgments of technical analysts to assess how well a country's policy and institutional framework fosters poverty reduction, sustainable growth and the effective use of development assistance (Gelb, Ngo and Ye, 2004).⁴ Ratings are against specific criteria but are subjective. Indeed, the CPIAs are produced by the Bank's own staff, i.e. its country teams.

In the past, CPIA results were not made available to the public. Only recently have governments themselves, whose policies are assessed in a particular CPIA, come to be informed of the numerical ratings on a confidential basis. Since 2000, there has been a public quintile-based disclosure.⁵ The exact numerical values of the CPIA have been disclosed starting with the results of the 2005 CPIA exercise. They are fully available to the public today.

The criteria used in the CPIAs have evolved over the years, in response to new analytical insights and lessons the Bank feels it has learned from experience. Originally called Country Performance Ratings (CPR), the assessment exercise acquired the name CPIA with the 1998 redesign to emphasize that it was the policy and institutional environment that was being assessed, not economic outcomes. The definition of the criteria, their relative importance, the rating and disclosure procedures have undergone important changes over the years (van Waeyenberge, 2006).

(A) THE 1980'S

Significantly, during the 1980's, the emphasis moved from an initial concern with both policy inputs and economic performance indicators (growth and savings rates), to a predominant concern with policy inputs. By the early 1990s, an exclusive emphasis on policy inputs prevailed.

In the early 1980s, four criteria were cited in the following order of priority as affecting IDA's resource allocation:

1. National poverty as measured through income per capita;
2. Creditworthiness;
3. Economic performance to be assessed in terms of macro indicators including growth and saving rates, but also in terms of the quality of "administration and economic management" together with "the speed and direction of change;"
4. Project readiness.

⁴ Indeed, according to the World Bank, the aim of the CPIA is to assess «how conducive [a country's policy and institutional] framework is to fostering poverty reduction, sustainable growth and the effective use of development assistance» (World Bank, 2007).

⁵ With the quintile-based results for the CPIA, its four clusters, the country portfolio, and the IDA Country Performance (ICP) rating posted on the Bank's external website, as well as the criteria and the methodology of the performance-based allocation system.

Guidelines on the allocation of lending among IDA-eligible countries issued in 1989 were characterized by a shift in emphasis towards greater consideration of policy performance.

According to Kapur *et. al.* (1997), Bank staff was instructed to rate a country's performance in each of three policy categories: (i) short-run economic management (mainly of demand); (ii) long-run economic management (mainly supply side restructuring); and (iii) the country's poverty-alleviation record as characterized by its delivery of social services, together with reforms removing "distortions" from labor markets and from rural-urban terms of trade. As a result, the 1991 CPIA exercise had three component clusters: (i) short-term economic management; (ii) long-term economic management; and (iii) poverty alleviation policies.

(B) THE 1990'S AND 2000'S

In 1997, criteria covering governance-related issues were added, and in 1998 the CPIA process was revised to add a benchmarking phase.⁶

In 2001, several changes were introduced that included establishing a written record, providing detailed guidance for criteria with several subcomponents, broadening the set of benchmark countries, revising the content of the criteria and defining the different rating levels (previously only the 2 and 5 rating levels were fully defined). A review took place in 2004, when the Bank commissioned an external panel to review the CPIA ratings and methodology. The panel made a number of recommendations: (i) simplify CPIA criteria from 20 to 16; (ii) undertake analytic work to better inform the weighting of the various criteria; (iii) reconsider the weight given to the "governance factor"; (iv) provide country authorities with an opportunity for comment on the assessments; (v) establish an independent committee to review the CPIA methodology every three years; and (vi) fully disclose the numerical ratings of the 2005 CPIA exercise for IDA borrowers. The criteria were revised in 2004 to take the recommendations made by the panel into account.

As shown in Table 4 in the Data Appendix, the CPIA currently comprises 16 criteria divided into four clusters. It is split into two groups, the CPIA Cluster A-C (Economic Management; Structural Policies; Policies for Social Inclusion/Equity), and the CPIA Cluster D (Governance Rating: Public Sector Management and Institutions). The CPIA Cluster A-C includes 11 items and the CPIA Cluster D includes 5 items.

The economic management cluster comprises three specific criteria: (i) macroeconomic management; (ii) fiscal policy; and (iii) debt policy.

⁶ "The benchmarking phase helps ensure that, given the criteria, the ratings are set at the right level and are consistent across countries and Regions. The Bank's six Regions, the Networks, and Central Departments assist in selecting a representative sample of countries that covers all six Regions, includes IBRD and IDA-eligible borrowers, good as well as poor performers, and has a ratings distribution similar to the overall distribution of the CPIA country scores. The set of benchmark countries is reviewed every year, taking into account the need to both maintain some continuity in the sample and to refresh it. At the onset of each year's exercise, the set of benchmark countries is communicated to the Regions and Networks, along with the timetable for the exercise." (World Bank, 2007)

The structural policies cluster contains three criteria: (i) policies and institutions for economic cooperation, regional integration and trade; (ii) financial sector; (iii) business regulatory environment.

The policies for social inclusion/equity cluster have five criteria: (i) gender equality; (ii) equity of public resource use; (iii) building human resources; (iv) social protection and labor; and (v) environmental policies and regulations.

Finally, the governance rating cluster comprises five criteria: (i) property rights and rule-based governance; (ii) quality of budgetary and financial management; (iii) efficiency of revenue mobilization; (iv) quality of public administration; and (v) transparency, accountability, and corruption in the public sector (World Bank, 2008).

Moreover, each criterion includes a series of sub-indicators through the guidelines that I will not detail here but that are available online. For example, the macroeconomic management criterion can be divided into three sub-indicators: (i) monetary/exchange rate policy with clearly defined price stability objectives; (ii) aggregate demand policies focus on maintaining short and medium-term external balance; (iii) avoiding crowding out private investment.

Using the guidelines, the Bank's country team gives a score to every country comprised between 1 and 6 for each of the 16 criteria and gives each cluster the same weight in producing the overall country assessment.⁷

(C) THE 2010'S

What is striking is that despite the economic crisis and the evaluation of the CPIA released by the Independent Evaluation Group (IEG) in 2009 (IEG, 2009), the criteria used in the CPIA as of today (2012) are exactly the same than those of 2008. The only difference is that “*macroeconomic management*” is now called “*monetary and exchange rate policies*.”

The main recommendation of the IEG was to “*strengthen the use of financial indicators in the CPIA write-ups*” (IEG, 2009) as if finance was but the only important thing to bring about sustained growth. This is especially striking given the fact that we know from the literature that there is a key trade-off between safe and sound finance on the one hand and the risk-taking in financial sectors’ intermediation between savers and investors on the other hand. Moreover the pattern of the financial sector maturation varies considerably among countries. It has been widely showed that financial instability can lead to poor economic growth. For example, Williamson and Mahar (1998) and Kaminsky and Reinhart (2001) have shown that financial opening preceded most crises. Griffith-Jones (2000) similarly underlines that international markets are inherently unstable due to information asymmetries. Hence it seems that the World Bank, at least from the CPIA point of view, has learned little from the crisis.

As I emphasize below, there is nothing in the CPIA that is related to what could bring sustained growth to developing countries.

⁷ More precisely, for each criterion, countries are rated on a scale of 2 (weak) to 5 (strong), and a country is rated a 1 if it is very weak for two years or more and a 6 if it is very strong for three years or more.

2.2 DESCRIPTIVE EVIDENCE

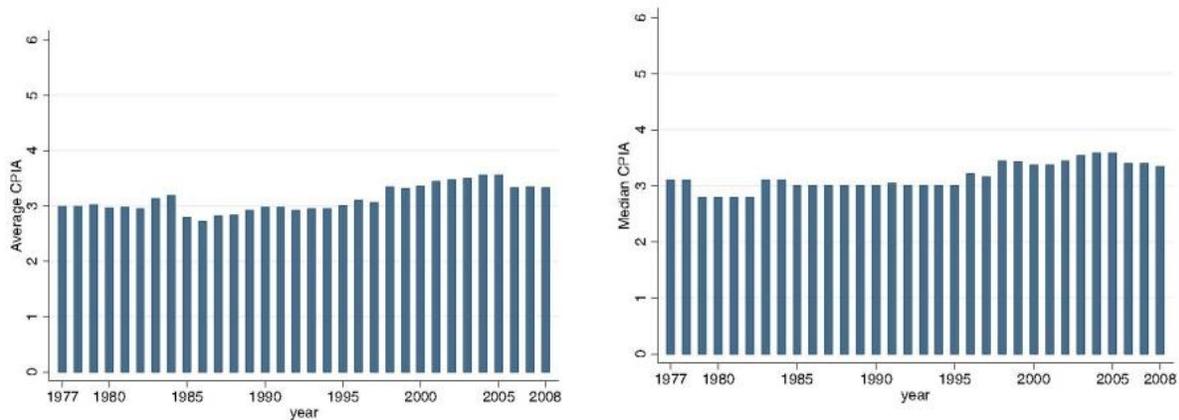
In this paper, I use the annual series of the CPIA for 146 countries between 1977 and 2008.

Interestingly enough, despite the fundamental changes in the CPIA design described above and despite a general understanding that developing country policies have improved on average since 1977, average CPIA results across countries have remained remarkably steady between 1977 and 2008. This appears clearly in Figure 1(a).⁸

Moreover, CPIA results have been concentrating increasingly around the median (Figure 1(b)). One can argue that this steadiness comes from the fact that some developing countries have improved substantially while others have declined. However, in this case, the standard deviation of the CPIA ratings should have risen, whereas it appears that it has decreased since 1985 (Figure 1(c)). According to Gelb, Ngo and Ye (2004), some inertia is to be expected in the ratings because they assess institutions and capacity to implement policies rather than just “stroke-of-the-pen” policy changes. This can cause CPIA scores to lag reform efforts, as better policies can require time to become properly reflected in the ratings. However, the concentration of the results around the same median for thirty years cannot be explained by such a lag, which leads us to question the current relevance of the CPIA. Indeed, if the CPIA criteria have changed while at the same time the scores have remained the same, one can argue that there is some “hidden conditionality” in the CPIA. I come back to this point below. However, one can also propose alternative explanations to this empirical fact, especially the fact that the CPIA is a way to grade countries in a “relative” rather than an “absolute” way, the goal being to sort countries for aid allocation purposes. But this does not explained the decrease in standard deviations, or the higher concentration around the median.

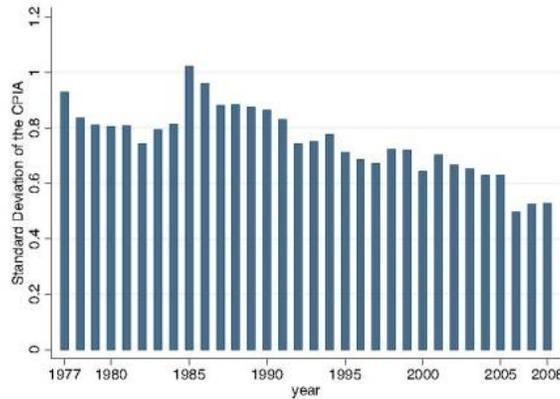
⁸ Similarly, Herman (2004) emphasizes that “although the index was substantially revised in 1998 (and again in 2001) and smaller revisions are made each year, neither the changes in the structure of the CPIA nor in the definitions of individual items seemed to cause significant changes in the rating scores, at least through 2000.” See also World Bank, 2010.

Figure 1: Descriptive Evidence on CPIA, 1977-2008



(a) Average CPIA

(b) Median CPIA



(c) Standard Deviation of CPIA

Note: These figures present some annual descriptive statistics of the CPIA between 1977 and 2008. Figure 1(a) plots the average annual value of the CPIA. Figure 1(b) plots the annual median value of the CPIA. Figure 1(c) plots the annual standard deviation of the CPIA.

Source: CPIA data, World Bank.

If we now look at the average CPIA score by regions, some differences across regions appear (Figure 2). At the end of the 1970s, South Asia had the lowest score with an average score below 3 (Figure 2(e)), followed by Sub-Saharan Africa (Figure 2(a)). However, Asia improved its rating since the end of the 1980s much more than Sub-Saharan Africa. Latin America and the Caribbean also improve their rating during the 1990s (Figure 2(b)). The region which has the best average CPIA score today is Eastern Europe and Central Asia which, despite a slight decrease during the 1990s, is back to its end of the 1970s average score (Figure 2(d)).

However, despite these small changes, it is important to underscore that the average CPIA score for each region has been incredibly stable during the 1977-2008 period. This again goes in the direction of those arguing that there may be today some hidden conditionality that comes with the CPIA.

2.3. THE RELEVANCE OF THE CPIA: REVIEW OF THE EXISTING EVALUATIONS

(A) EX ANTE CONDITIONALITY

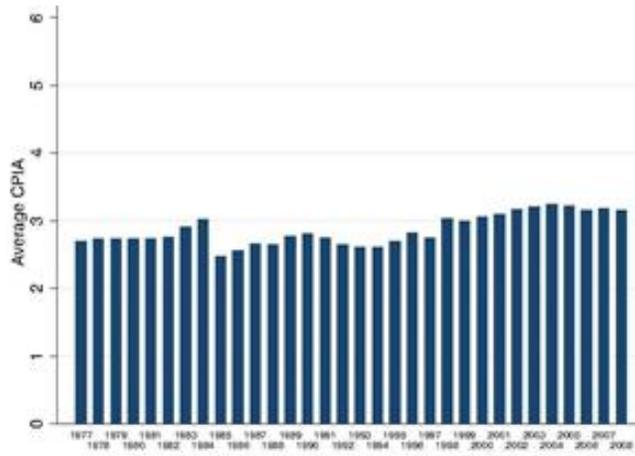
The first caveat of the CPIA is that it relies on policies rather than on outcomes. From this point of view, it corresponds to a model of “*ex ante* conditionality.” The main problem of *ex ante* conditionality is that until now it has never worked. As underlined by Stiglitz (1999), “*good policies cannot be bought.*”

What is often heard in the public opinion, the criteria on which the CPIA relies are for a very large majority considered as important determinants of growth, poverty reduction and effective use of aid in the literature and are not object of current controversies. A possible interpretation of this finding is that the CPIA may have experienced remarkable changes during the last years with the disappearance of the explicit mention of a set of policy imperatives. For example, it appears that certain lessons of the post-Washington Consensus, such as the hazards of capital account liberalization and the fragility of the financial sector in developing Stiglitz (1998(b)) have filtered through into Bank practice. From this point of view, the CPIA seems to have acquired more relevance.

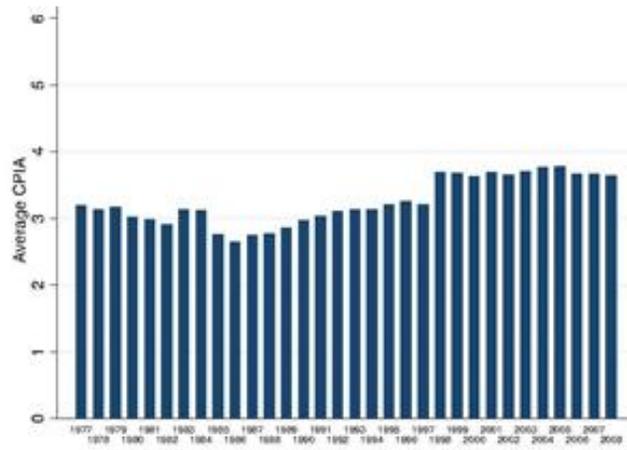
However, as van Waeyenberge (2006) suggests, the meaning of these changes is open to a second possible interpretation. The question could be raised as to whether those imperatives that have disappeared from the narrative guidelines of the CPIAs may have somehow become “embedded” and now steer the CPIA exercise in less visible ways. Van Waeyenberge (2006) illustrates this point with the assessment of trade policy in the 2004 CPIA questionnaire (World Bank, 2004). This questionnaire focuses exclusively on the policy framework regarding trade in goods, without reference to the rules and regulations affecting capital flows. The narrative guidelines on the assessment of the financial sector, in turn, do not make explicit reference to issues regarding foreign investors, state ownership or directed credit. However, van Waeyenberge (2006) argues that closer scrutiny of the guideposts that accompany the narrative guidelines for these respective policy/institutional categories reveals how these specific policy imperatives have in fact been subsumed in the “diagnostic reports” that now serve as guideposts to staff’s assessment: “*these typically embody a bias in favor of foreign investment and trade, and are anchored in a framework of traditional welfare economics where government intervention is tolerated only in the context of static market failure.*”

This point follows the lines of those arguing that there is hidden conditionality in the CPIA. This could help explain at least in part the surprising steadiness of the CPIA scores despite the important changes in the criteria.

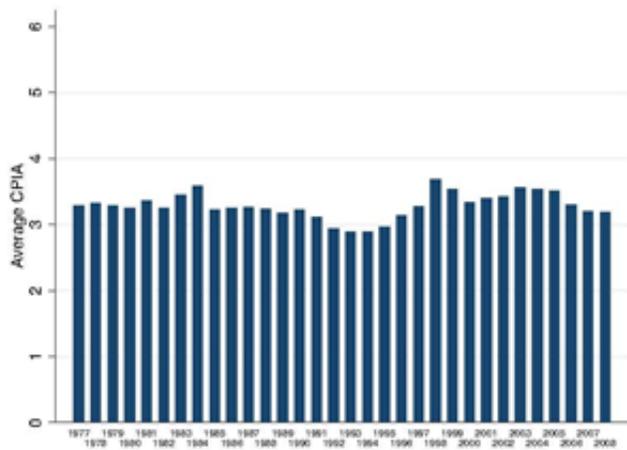
Figure 2: Descriptive Evidence (Average) on CPIA by Region, 1977-2008



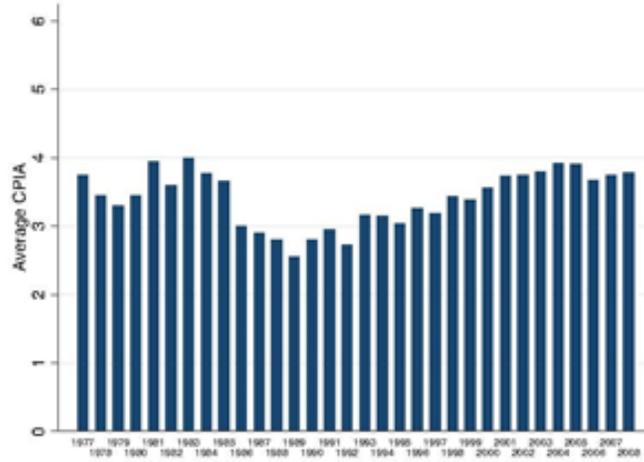
(a) Sub-Saharan Africa



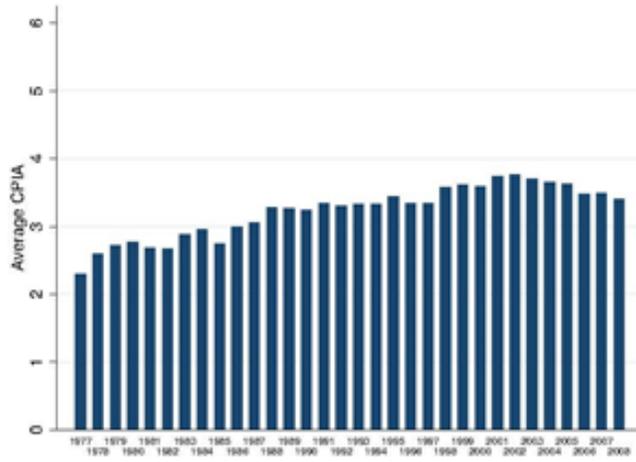
(b) Latin America and the Caribbean



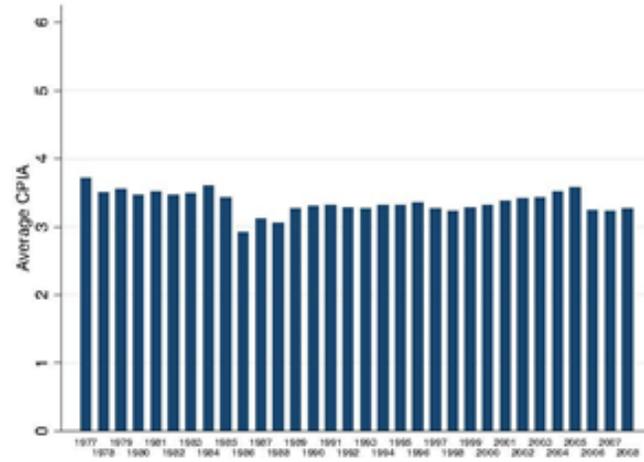
(c) Middle East and North Africa



(d) Eastern Europe and Central Asia



(e) South Asia



(f) East Asia and the Pacific

Note: These Figures present the average annual value of the CPIA by region. Figure 2(a) plots the average annual value of the CPIA for Sub-Saharan Africa; Figure 2(b) for Latin America and the Caribbean; Figure 2(c) for the Middle East and North Africa; Figure 2(d) for Eastern Europe and Central Asia; Figure 2(e) for South Asia and Figure 2(f) for East Asia and the Pacific.

Source: CPIA data, World Bank.

(B) A ONE-SIZE-FITS-ALL APPROACH TO DEVELOPMENT

Another caveat of the CPIA is that it relies too much on a “one-size-fits-all” approach to development while it has been shown in the literature that “*there is no universal recipe*” (Bardar and Birdsall, 2006).⁹ This point is made in Kanbur (2005) as well as in Cage (2009 (b)). For example as to growth, Cage (2009 (b)) underlines that growth-promoting policies tend to be context-specific: one has to take into account individual country experiences when analyzing the determinants of sustained growth. Even the World Bank seems to have accepted this new emphasis on country specificities. For example, drawing the lessons of the 1990s, the World Bank underlines that “*there is no one right way to achieve development*” and that “*which options should be chosen depends on initial conditions, the quality of existing institutions, the history of policies, political economy factors, the external environment, and last but not least, the art of economic policymaking*” (World Bank, 2005).

The CPIA does not correspond to the empirical reality of development. Being the same for every country, it relies too heavily on a uniform model of what works in development policy (Kanbur, 2005). Even if this model were valid “on average,” the variations around the average make it an unreliable sole guide to the country-specific productivity of aid in achieving the final objectives of development.

Moreover, the CPIA does not only rely on a uniform model of what works in development policy, but it does so by underlying very specific policies. This clearly appears if one considers the “Policies and Institutions for Economic Cooperation, Regional Integration & Trade” criterion and the “Quality of Budgetary and Financial Management” criterion. They are much more detailed than whatever can be found in the literature, relying on too specific quantitative policies. Moreover, whereas the literature insists on the necessity to set priorities, the CPIA seems to put everything on the same plan.

Relying too heavily on a uniform model, the CPIA thus does not appear to be a good tool for allocating aid, at least conceptually. We show that it is even more the case empirically, since it is not a good predictor for economic growth.

⁹ Similarly, Hoff and Stiglitz (2011) emphasize that “*there are clearly no surefire formulas for success; if there were, there would be more successes.*”

2.4. IS THE CPIA A GOOD PREDICTOR FOR ECONOMIC GROWTH? AN EMPIRICAL ANALYSIS

In this section, I test empirically whether the CPIA is a good predictor for economic growth, using an annual panel dataset over 146 countries between 1977 and 2008.

(A) DATA DESCRIPTION AND DESCRIPTIVE STATISTICS

In order to determine whether the CPIA is a good predictor for economic growth, I run growth regressions with the CPIA score and the annual change in the CPIA as control variables in a panel of 146 developing countries over the 1977-2008 period.¹⁰ As a dependent variable, I use the growth rate of per capita GDP. Usual controls in cross-country growth equations, used for example in Levine and Renelt (1992), Ramey and Ramey (1995) and Aizenman and Marion (1999), are the initial log level of real GDP per capita; the initial fraction of the relevant population in secondary schools; the initial growth rate of the population; and the average share of trade in GDP over the period. However, since all these controls are fixed at the country level, I choose to introduce directly country fixed effects in all the specifications for robustness reasons. I also control for M2 as a share of GDP lagged one period and aid flows. Panel data on aid flows are taken from the OECD Development Assistance Committee (DAC) annual series. Following Roodman (2006), I use the Net Aid Transfers (NAT) variable for measuring aid flows.¹¹ Table 1 provides summary statistics for a few key statistics.

Table 1: Summary Statistics¹²

	Mean /
sd	
CPIA Score	3.16 (0.77)
CPIA Change	2.10 (8.13)
Per Capita Growth Rate	1.74 (5.15)

¹⁰ For the description and the sources of the data in more details see the Data Appendix.

¹¹ NAT is a net transfer concept, net of both principle payments received on ODA loans and of interests received on such loans. Moreover, NAT excludes cancellation of old non-ODA loans since such cancellation generates little or no additional net transfers.

¹² Numbers in parentheses are standard deviations and the others are averages. Variables are described in the Data Appendix.

Aid/GDP	7.95 (10.93)
Observations	1095

(B) EMPIRICAL SPECIFICATION

Equations are estimated using a panel of eight four-year periods from 1977-1980 through 2005-2008. Thus, an observation is a country's performance average over a four-year period. The averaging over four-year, which is usual in the literature, allows me to avoid the non-stationarity problem for the growth rate.

The baseline empirical specification is:

$$(1) \quad g_{it} = \alpha + \beta \text{CPIA Score}_{it} + \theta \text{CPIA Change}_{it} + \lambda X_{it} + \vartheta_i + \gamma_t + \epsilon_{it}$$

Where i index the countries and t stands for the eight four-year periods (from 1977-1980 to 2005-2008). g is the growth rate of per capita GDP; "CPIA Score" is the average CPIA score over the period; and "CPIA Change" is the average of the annual change in CPIA rating over the period. X_{it} is a vector of control variables that vary with the specification considered. ϑ_i are country fixed effects; γ_t period fixed effects; and ϵ_{it} is a country-period shock.

I estimate equation (1) using both OLS with robust standard errors and two-step Arellano-Bond GMM (Tables 2 and 3). The use of two-step Arellano-Bond GMM is driven by possible endogeneity concerns. The advantage of the system GMM method is that it helps to overcome endogeneity concerns in the absence of any strictly exogenous explanatory variables or instrument.¹³ The results are robust to both methodologies.

(C) RESULTS

Tables 2 and 3 presents the results of the impact of performance as measured by the CPIA score and the annual changes in this score on the growth rate of per capita GDP (estimation of equation (1)). I find a positive and statistically significant coefficient for the CPIA score using both OLS and two-step Arellano-Bond GMM (columns 1 and 2). I also find a positive impact of the annual change in the CPIA rating but it is not statistically robust to the use of two-step Arellano-Bond GMM. However, these estimates do not prove causality. As

¹³ When I estimate my regressions using two-step system GMM, I thus use the forward orthogonal deviations transform instead of first differencing because it maximizes the sample size in panel with gaps (Roodman (2006)). The forward orthogonal deviations transform is an alternative to differencing proposed by Arellano and Bover (1995) that preserves sample size in panel with gaps. Indeed, instead of subtracting the previous observation from the contemporaneous—what does the first-difference transform which thus magnifies gaps in unbalanced panels—it subtracts the average of all future available observations of a variable. No matter how many gaps, it is computable for all observations except the last for each individual, so it minimizes data loss. And since lagged observations do not enter the formula, they are valid instruments.

acknowledged by Gelb Ngo and Ye (2004) despite the use of clear benchmarks to derive CPIA ratings, it is possible that assessments are colored by perceptions of “how well the country is doing” which are influenced by recent growth trends. In this case, the positive coefficient I obtain for the CPIA score would simply reflect the fact that CPIA scores themselves respond to observed growth rates and so is not indicative of the fact that this score can be interpreted as a good predictor for growth rates.¹⁴

In the “Staff Guidance Notes” used by the World Bank staff in order to measure policy performance, it is strongly emphasized that the only thing that has to be taken into account is the short term: *“The write-up should focus on the developments of the past one calendar year (...). Unless absolutely necessary, staff are not expected to report developments of more than two years ago in the write-up.”*; *“Policy performance should be rated against the CPIA criteria, rather than the degree of improvements from the previous year, and in relation to the ratings of the benchmark countries.”* This emphasis on the short term can be explained while the assessments are colored by the current growth rates.

In order to test whether or not the CPIA score can be interpreted as a good growth predictor, I introduce as a control variable the CPIA score lagged one period. Obviously, this cannot be determined by the current growth rate. But if I were to find a strong positive correlation between the CPIA score lagged one period and the following period growth rate, then the CPIA score could be interpreted as a good predictor of future growth.

When I estimate equation (1) with this new control, I find a negative and statistically significant coefficient for the CPIA score lagged one period, using both OLS and two-step Arellano-Bond GMM (column 3). This means that the contemporaneous correlation between the CPIA score and the growth rate comes from the fact that assessments are colored by current perceptions of country performances. More importantly, this means that the CPIA score is a very bad predictor for future growth rates since the countries with the lowest CPIA scores one period ago are those that do better in terms of growth during the following period. Moreover, the coefficient for the CPIA score lagged one period is negative (and statistically significant for Arellano-Bond GMM) whether or not I include in the regression the current CPIA score (column 4).

One can argue that these estimates do not distinguish between the effect of performance as measured by the CPIA and other influences on growth that may themselves reflect the CPIA rating. For example, growth in high-performing countries may be partly driven by increased AID flows in response to higher CPIA scores. In order to control for these other influences, I include aid flows normalized by GDP as a control variable (column 5) as well as the square of these flows to control for decreasing returns of aid (column 6). The introduction of these controls, whether I use OLS and Arellano-Bond GMM estimations, or 2SLS and 2-step feasible

¹⁴ Similarly, Glaeser *et al* (2004) underline that indicators measuring the quality of institutions, supposed to explain economic growth, are in fact the result and not the cause of economic growth.

IV/GMM in order to control for aid endogeneity, does not change the results (columns 5, 6 and 7).¹⁵

Finally one can claim that the negative and statistically significant coefficient I obtain for the CPIA score lagged one period comes from the presence of some outliers in the sample. In order to check whether this is the case, I identify influential observations using the method of Hadi (1992), which classifies nine observations as outliers at the 5 percent level.¹⁶ Removing these observations does not change the results (column 8).

The negative coefficient I obtain for the CPIA score lagged one period thus seems to be robust. It appears that the CPIA is not a good predictor of economic growth.

Table 2 : OLS and 2SLS Estimation¹⁷

	1	2	3	4	5
	OLS	OLS	OLS	OLS	OLS
	b/se	b/se	b/se	b/se	b/se
CPIA Score	1.819***	1.561***	1.954***		1.957***
	-0.352	-0.408	-0.406		-0.41
CPIA Change		0.060*	0.015		0.015
		-0.032	-0.036		-0.035
(One-Period) Lag of CPIA Score			-0.923**	-0.348	-0.981**
			-0.423	-0.419	-0.435
Period and Country FE	Yes	Yes	Yes	Yes	Yes
Aid	No	No	No	No	Yes
Aid Square	No	No	No	No	No
Outliers Included	Yes	Yes	Yes	Yes	Yes
R-sq	0.48	0.53	0.54	0.51	0.54
Observations	766	748	734	734	722

¹⁵ In columns 6 and 7 aid flows are instrumented by the correlation of the country votes with those of the US in the UNGA lagged one period.

¹⁶ Angola in the 2005-2008 period; Azerbaijan in the 1993-1996 and 2005-2008 periods; Equatorial Guinea in the 1997-2000 and 2001-2004 periods; Liberia in the 1989-1992 and 1997-2000 periods; Tajikistan in the 1993-1996 period; and Ukraine in the 1993-1996 period.

¹⁷ * p<0.10, ** p<0.05, *** p<0.01. Table 2 reports OLS and 2SLS estimates. Table 3 reports Arellano-Bond GMM estimates. The unit of observation is a country/period. Standard errors in parentheses are robust. The dependent variable is the growth rate of per capita GDP. The endogenous variable for the 2SLS estimations are the aid flows. The excluded exogenous variable for the 2SLS estimations is the correlation of the country votes with those of the US in the UNGA lagged one period. All the regressions include M2 as a share of GDP as a control. The controls are described in more details in the text.

	6	7	8
	OLS	2SLS	2SLS
	b/se	b/se	b/se
CPIA Score	1.889***	1.958***	2.297***
	-0.425	-0.568	-0.426
CPIA Change	0.014	0.005	-0.032
	-0.035	-0.038	-0.032
(One-Period) Lag of CPIA Score	-0.993**	-0.954**	-0.855**
	-0.436	-0.468	-0.361
Period and Country FE	Yes	Yes	Yes
Aid	Yes	Yes	Yes
Aid Square	Yes	No	No
Outliers Included	Yes	Yes	No
R-sq	0.54	0.49	0.56
Observations	722	658	653

Table 3: GMM Estimation¹⁸

	1	2	3	4	5
	GMM	GMM	GMM	GMM	GMM
	b/se	b/se	b/se	b/se	b/se
CPIA Score	1.766***	1.516***	1.446***		1.491***
	-0.365	-0.343	-0.338		-0.339
CPIA Change		0.042	0.006		0.008
		-0.031	-0.036		-0.037
(One-Period) Lag of CPIA Score			-	-	-
			1.205***	1.433***	1.271***
			-0.42	-0.362	-0.427
Period and Country FE	Yes	Yes	Yes	Yes	Yes
Aid	No	No	No	No	Yes
Aid Square	No	No	No	No	No
Outliers Included	Yes	Yes	Yes	Yes	Yes
Observations	591	571	559	563	547

¹⁸ ¹⁸ * p<0.10, ** p<0.05, *** p<0.01. Table 2 reports OLS and 2SLS estimates. Table 3 reports Arellano-Bond GMM estimates. The unit of observation is a country/period. Standard errors in parentheses are robust. The dependent variable is the growth rate of per capita GDP. The endogenous variable for the 2SLS estimations are the aid flows. The excluded exogenous variable for the 2SLS estimations is the correlation of the country votes with those of the US in the UNGA lagged one period. All the regressions include M2 as a share of GDP as a control. The controls are described in more details in the text.

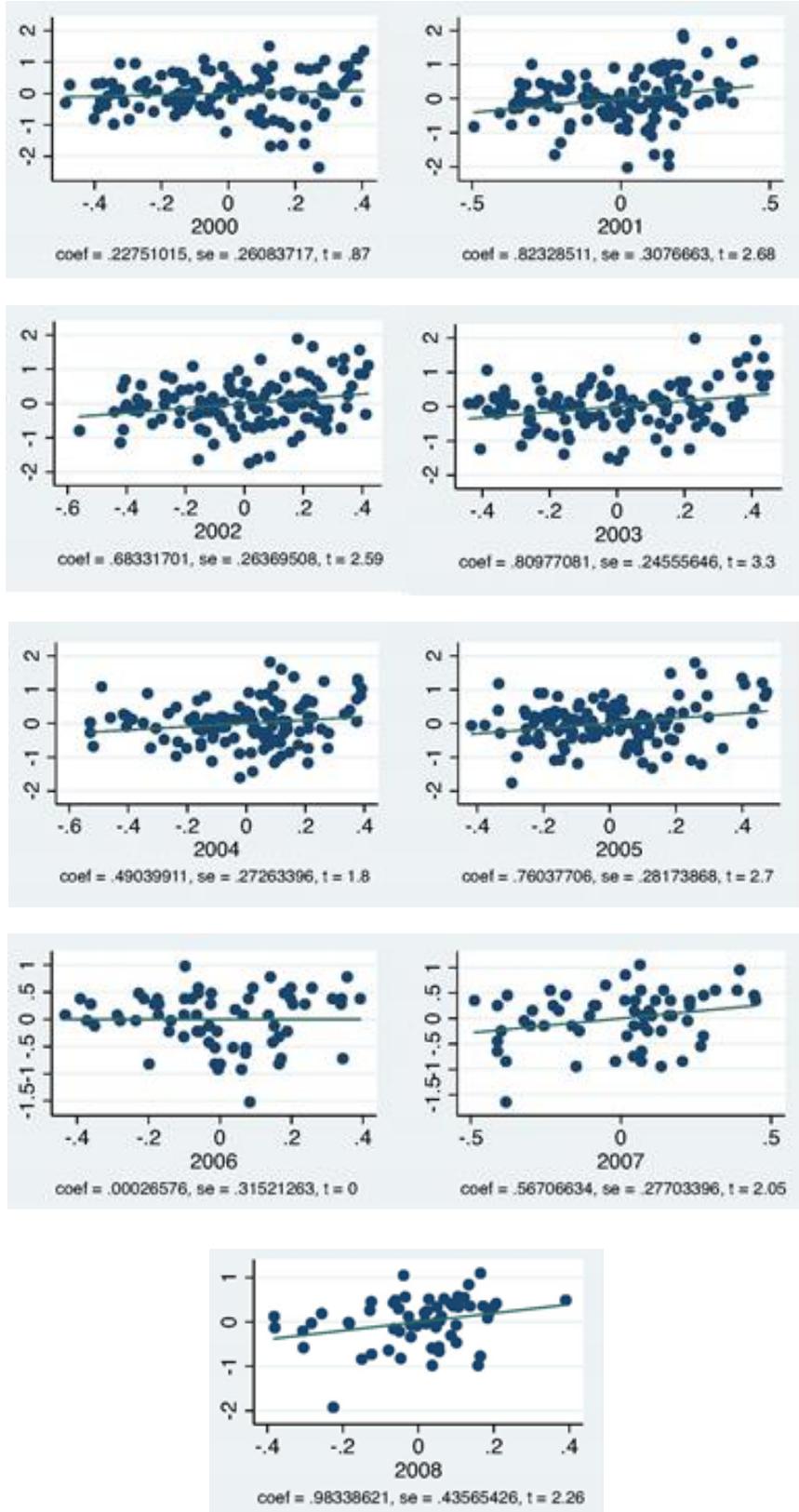
	6	7	8
		IV	IV
	GMM	GMM	GMM
	b/se	b/se	b/se
CPIA Score	1.487***	1.958***	2.297***
	-0.345	-0.561	-0.42
CPIA Change	0.008	0.005	-0.032
	-0.037	-0.038	-0.032
(One-Period) Lag of CPIA Score	-1.280***	-0.954**	-0.855**
	-0.427	-0.463	-0.357
Period and Country FE	Yes	Yes	Yes
Aid	Yes	Yes	Yes
Aid Square	Yes	No	No
Outliers Included	Yes	Yes	No
Observations	547	656	651

(a) VOTES IN THE UNGA

I finally find that there is a strong and statistically significant positive correlation between the CPIA scores of developing countries and the correlation of their votes with those of the US in the UNGA. This appears clearly in Figure 3. For each year between 2000 and 2008, I plot the relationship between the correlation of the votes with those of the US in the UNGA and the CPIA score of the countries. It appears clearly that this relationship is positive and statistically significant: the higher the correlation of the votes, the higher the CPIA score.

Obviously, correlation is not causality, but it seems hard to find an intuitive causal link going from the CPIA score to the correlation of the vote in the UNGA. On the contrary, one can argue that CPIA scores are biased in favor of countries having political links with the US. Moreover, if one remembers that the literature on aid allocation has shown that aid may be used to buy political support from the recipients of aid (Alesina and Dollar, 2000; Alesina and Weder, 2002; Schraeder *et. al.*, 1998; Kuziemko and Werker, 2006), another interpretation could be that countries to which the World Bank is willing to give more aid received a higher CPIA score in order for these countries to effectively receive more aid through the Bank allocation formula. Whatever the precise sense in which that might or might not be the case, this questions the relevance of the CPIA.

Figure 3: CPIA Score and the Correlation of the Votes with those of the US in the UNGA, 2000-2008



These Figures present for each year between 2000 and 2008 correlation between the CPIA scores of developing countries and the correlation of their votes with those of the US in the UNGA.

Source: CPIA data are from the World Bank; Correlation of the Votes with those of the US in the UNGA data has been constructed by the author using UNGA Votes data from Erik Voeten.

The CPIA, despite all the recent modifications and improvements that have been made and need to be acknowledged, thus still appears as a blunt and biased tool that can and should be improved upon.

3. HOW CAN WE IMPROVE UPON THE CPIA?

In this section I first underline the fact that other criteria, and in particular the role of the state and the quality of industrial policy need to be taken into account into the CPIA. Focusing more specifically on the allocation of development aid, I then consider the alternative proposals to the CPIA that have been formulated in the literature and in particular the one of Kanbur (2005). I finally make concrete proposals for the development of new possible allocations based on the idea of using “aid effectiveness” as an allocation tool.

3.1. INTRODUCING NEW CRITERIA INTO THE CPIA

There is still nothing in these criteria related to what could bring sustained growth to developing countries, for example fiscal capacity or industrial policy. The words “fiscal capacity” or “industrial policy” do not even appear in the 103 pages of the “Staff Guidance Notes” for the 2012 CPIA. As I will underline it below, there is similarly nothing on export promotion strategies despite the fact that it is widely acknowledged that all successful liberalizations either explicitly or implicitly promoted export growth. The CPIA only emphasizes the necessity to remove trade restrictions.

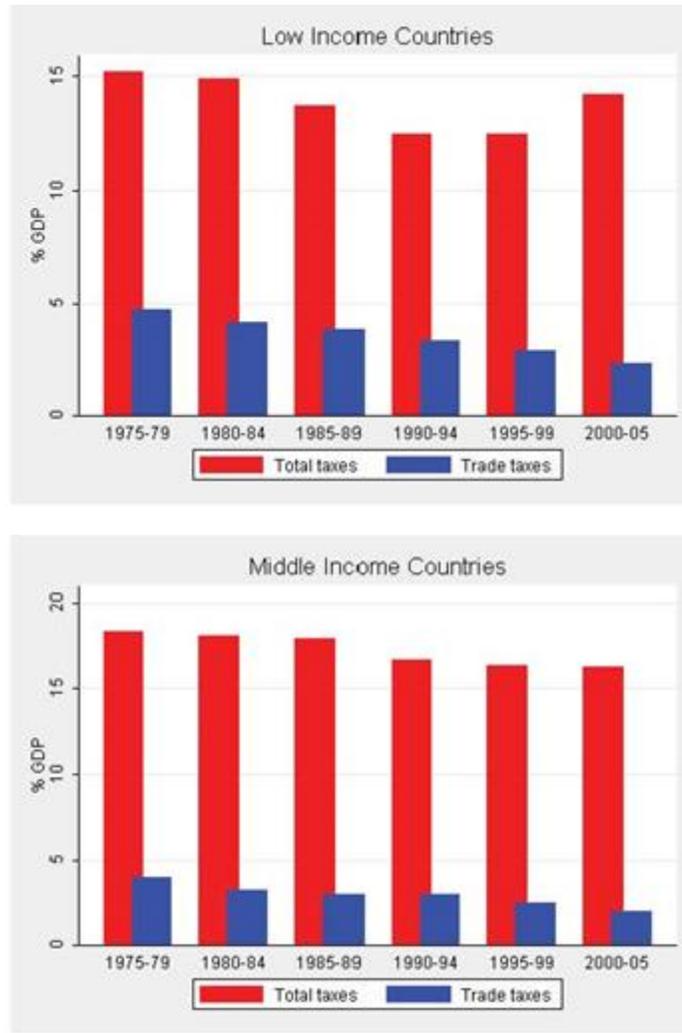
(A) THE ROLE OF STATE BUILDING AND FISCAL CAPACITY

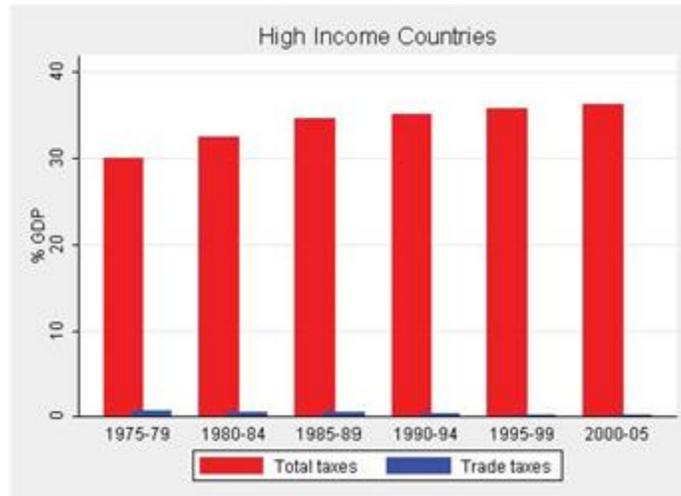
The CPIA clearly underestimates the role that a well-functioning government can and must play in the development process. This is striking when one considers the “trade criteria” in which it is underlined that *“MFN tariffs have been streamlined into a limited number of tariff bands in many countries, so CPIA ratings should reflect how distortionary is the overall structure of trade taxes, including not just tariffs but also other border taxes.”*

From this point of view, the CPIA totally ignores the fiscal consequences of trade liberalization, while trade liberalization can have a very negative impact for developing countries in terms of tax revenues as shown by Cage and Gadenne (2012). Trade taxes are indeed an important source of revenue for developing countries. These revenues have fallen over the past decades as these countries liberalized trade. Many developing countries

simultaneously experienced a decrease in their total tax revenues. These appears clearly in Figure 4 from Cage and Gadenne (2012). Using a novel panel dataset of tax revenues and government expenditures in developing countries between 1945 and 2006, Cage and Gadenne (2012) identify 110 episodes of decreases in tariff revenues. They show that less than half of the countries recover the lost tax revenues 5 years after the start of the episode. Moreover they find a similar picture when they consider government expenditures.

Figure 4: Evolution of tax revenues as a share of GDP, 1975-2005 (Cage and Gadenne, 2012)





Note: All values are median values for the country group and time period considered. The sample includes in each time period 26 low income countries, 40 middle income countries and 32 high income countries.

Source: Cage and Gadenne (2012).

The questions are thus the following: how to deliver a proper education, health and infrastructure system with tax revenues representing less than 15 percent of GDP, creating a clear competitive disadvantage for developing countries? And how to bring sustained growth without a proper education, health and infrastructure system? In order to levy domestic taxes—and so to be able to open itself to international trade—a country needs pre-existing tax capacity. These tax capacities are not taking into account in the CPIA. I argue in favor of their inclusion as one of the main criterion rather than only emphasizing the necessity to remove trade restrictions.

(B) EXPORT PROMOTION POLICIES

Similarly, while the CPIA only emphasizes the necessity to remove trade restrictions, efficient export promotion policies may have an important role to play and should be taken into account. Cage and Rouzet (2012) show for example that export subsidies may have a positive welfare effect on exporting developing countries by improving both the average quality of their exports and their terms of trade. They study the effect of firm and country reputations (the famous “made in” label) on exports when buyers cannot observe quality prior to purchase. Measuring national reputations by analyzing the content of US newspaper articles about foreign countries over the period 1988-2006, they find that more positive news coverage of foreign countries and companies is associated with higher unit values of the exports to the United States, particularly in sectors with larger scope for vertical differentiation. They rationalize this finding in a model in which firm-level demand is determined by expected quality which depends on both past experience with good and country of origin’s reputation. Asymmetric

information acts as barrier to entry for high-quality firms but facilitates sales by “fly-by-night” low-quality firms. Countries with a bad quality reputation can thus be locked into exporting low-quality, low-cost goods. In this case, export subsidies have a positive welfare effect on exporting countries.

The findings of this paper are consistent with the success of some export-led growth strategies for developing countries. East Asian economies in particular pursued a few decades ago strategies consisting on exporting low-quality, low-cost goods and gradually moving up to higher quality, higher unit value goods.¹⁹ China is currently attempting to follow the same path. This is for example the strategy of Lenovo, the only Chinese company to get a worldwide sponsorship for the Beijing Olympics. With a Western sounding name, the legacy of IBM brand name and technology and the chief executive from Dell and NCR, Lenovo Group is not a company that most Americans would assume is Chinese. This is exactly what the company aims for, although Lenovo’s largest shareholder is the Chinese government, because it is aware of fact that the American consumer associates Chinese products with cheap and unreliable.

Without policy intervention, moving up to higher quality exportations may not be feasible if the economy is trapped in a self-fulfilling low equilibrium, in which country’s reputation for low quality prevents high-quality firms from entering the export markets. In this case, a successful export promotion policy would consist in subsidizing exporters’ initial losses or investing public resources into raising country’s perception abroad. This is why the quality of the industrial policy—and for example the fact of having an efficient export promotion strategy—has to be taken into account in the CPIA.

3.2. CRITERIA FOR AID ALLOCATION

One of the main uses of the CPIA is as a criterion in the development aid allocation formula. Indeed, when a country is eligible to the International Development Association (IDA)—the development aid agency of the World Bank—the IDA formula to allocate aid is made of four different terms: (i) the CPIA; (ii) the portfolio performance which is used to determine a rating for each country’s implementation performance; (iii) population; and (iv) per capita income.²⁰ The combination of the CPIA and of the portfolio performance forms the Country Performance Rating (CPR):

¹⁹ At the end of the Second World War, “Made in Japan” goods had the reputation of being cheap low-quality goods. Japanese companies were suffering from an inferior “national brand.” Currently, Japanese cars and electronics ranked among the most reliable in all consumer surveys. Japan’s pattern of specialization in manufactures has evolved dramatically. Japanese companies achieved such a dramatic change by privately imposing strict quality norms. They formed export cartels which provided product quality guarantees. In particular, they set product design and quality standards; established industry brand names; guaranteed delivery schedules; and mediated the disputes between exporters and foreign buyers.

²⁰ In order to be eligible to the IDA resources, a country has to meet two criteria. First, its relative poverty defined as GNI per capita as to be below an established threshold which is updated annually (in fiscal year 2010: \$1,135).

$$\text{CPR} = 0.24 \times \text{CPIA}_{A-C} + 0.68 \times \text{CPIA}_D + 0.08 \times \text{Portfolio}$$

CPIA_{A-C} stands for the clusters A through C of the CPIA; CPIA_D for the cluster D; and Portfolio for the portfolio performance rating.

The IDA allocation formula is then computed as follows:

$$\text{IDA Country Allocation} = F \left[\text{CPR}^{0.5}, \text{Population}^{1.0}, \left(\frac{\text{Gini}}{\text{Population}} \right)^{-0.125} \right]$$

In the last part of this article, I study alternative tools to the CPIA to allocate aid in the most effective way.

(A) EXISTING ALTERNATIVE PROPOSALS TO THE CPIA

(a) KANBUR'S PROPOSAL

Underlying that the CPIA implicitly relies too heavily on a uniform model of what works in development policy, Kanbur (2005) proposes to introduce outcome variables in the development aid allocation formula.

Indeed, as it clearly appears from the formula, the IDA essentially captures needs through the income criterion, and does not go directly to indicators such as infant mortality, maternal mortality, girls' education and other components of the Millennium Development Goals. Moreover, the CPIA itself does not contain any final outcomes variables like poverty, extreme poverty, girls' enrollment, etc. What it has instead is a series of intermediate variables like trade policy, regulatory policy, property rights, corruption, etc.

On the contrary, Kanbur (2005) argues in favor of an outcomes-based aid allocation, or at least in favor of introducing some outcome variables in the CPIA itself.²¹ His main idea is to measure the needs side by side with the levels of the outcomes one is interested in, while measuring the performance side by side with the rate of improvement of these outcome variables over a given period of time up to the point of assessment, suitably normalized by the total aid flow over this period. He gives the following example: a country that has very low levels of girls' enrollment in primary schools should get more aid on grounds of need. But a country that is showing rapid improvements of girls' enrollment from this low level, relative to the aid it is receiving, should get even more. A country that is showing relatively slow rates of improvement should get relatively less on account of this measure of performance. The main advantage of this focus on outcomes is that it prevents the easy temptation of a "one-size-fits-all" approach to development.

Second, it has to lack creditworthiness to borrow on market terms and therefore to need concessional resources to finance its development program.

²¹ "While leaving the current IDA allocation methodology essentially intact, IDA should introduce one new category of scoring in the CPIA. This category should evaluate the evolution of an actual development outcome variable up to the present. The choice of variable is open."

As he acknowledges himself, this proposal is in the spirit of Collier's outcomes-based allocation (Collier *et. al.*, 1997). They propose a basis for aid allocation in terms of retrospective assessment of a few major outcomes such as growth. They show how outcome measures can control for influences on growth over which the government has no control and argue that donors should switch from attempting to "purchase" a pre-specified menu of policy changes to the allocation of aid on the basis of periodic overall assessments of government achievements.

Similarly, Barder and Birdsall (2006) defend the idea of "*payments for progress*," the main objective being to link additional aid to clear evidence on progress already achieved on the ground. In order to do so, payments would be determined as a function of the outcomes and not linked to the implementation of any particular policies, any other intermediate outputs, or tied to purchases from particular suppliers or companies.

(b) ADVANTAGES OF KANBUR'S PROPOSAL

One of the main advantages of the Kanbur's proposal is that it relies on performance-based measures—on actual performance—and so does not imply *ex ante* conditionality. Indeed, rating countries according to their rates of improvement of certain outcomes rather than according to their policies corresponds to an "*ex post*" approach of conditionality. From the point of view of this approach, one has to reward countries that used past aid well (*ex post* conditionality) without conditions (*ex ante* conditionality). This is in the spirit of the Paris Declaration (2005) and of the following Accra Agenda for Action (AAA) (2008):

Developing countries and donors will work together at the international level to review, document and disseminate good practices on conditionality with a view to reinforcing country ownership and other Paris Declaration Principles by increasing emphasis on harmonized, results-based conditionality"

This is an important improvement on the CPIA since, as I underline above, using *ex ante* conditionality is a very inefficient way to allocate aid.

Moreover, the Kanbur's proposal is a useful improvement since it underlines that the CPIA relies excessively on a "one-size-fits-all" approach to development and proposes a way to overcome this difficulty with the use of outcome variables. Finally, his proposal is very well argued and he anticipates various criticisms.

(c) LIMITS OF KANBUR'S PROPOSAL

Given all the advantages of the Kanbur's proposal, the only criticism made by Buiter (2007) to this proposal that seems acceptable is the one according to which realized past outcome changes as a measure of future aid productivity: "*the aid could have been looted, diverted or wasted, that is, not even spent on any activity likely to boost the indicator, and the improvement in the indicator could have been produced by domestic or foreign factors that have nothing to do with the aid*

dispensed during the benchmark period, but never mind...."²² In other words, the Kanbur's proposal relies on the implicit assumption that past output indicators are a good guide to future aid productivity. And this is still to be shown.

Similarly, McGillivray (2004) underlines that Kanbur (2005)'s proposal does not really provide an understanding of what makes aid works. This is why he argues in favor of more radical changes to the IDA formula than outlined in the Kanbur's proposal. According to him, what is required is a better knowledge of what makes aid work and the revisions to aid allocation formula should be considered in this light.

(B) THE NEED FOR AN AID PRODUCTIVITY MEASURE

What emerges clearly from the criticisms of both the CPIA and the Kanbur proposal—and more generally of any outcomes-based allocation—is that what is needed is an aid productivity measure. The CPIA is not an aid productivity measure, being excessively focused on a “one-size-fits-all” approach of development. An outcomes-based allocation would not overcome this difficulty. In order to overcome it, one needs to establish a clear statistical link between past outcomes and future aid productivity. This has never been done and seems hardly feasible. Indeed, this supposes first to evaluate the elasticity of this outcome variable with respect to aid (which can be interpreted as aid productivity). Second, this supposes to evaluate the elasticity of aid productivity with respect to past changes in this outcome variable. In both steps of the estimation, one would be faced with endogeneity and omitted variables concerns. Moreover, in case one would like to introduce not one but various measures of outcomes, the estimation would be even more complicated by the fact that these outcomes may be interdependent. And then it remains to determine the optimal weight to give to each of these outcomes.

Well aware of all these difficulties but at the same time of the real need for an aid productivity measure, I discuss below a new way to approaching this issue. The idea is to use directly aid effectiveness as such a measure.

(a) A NEW APPROACH: USING AID EFFECTIVENESS TO ALLOCATE AID EFFECTIVELY

Aid effectiveness has to be defined with respect to a given outcome, which has to be chosen by donors when they establish their selectivity criteria. This can be the growth rate of the economy; the reduction in the poverty rate; the rate of girls' enrollment or other goals depending both on donors' priorities and recipient countries specificities. Aid is said to be very effective if aid elasticity with respect to the outcome is very high. For example, if the outcome is the growth rate, the higher aid elasticity with respect to growth—i.e. the more the growth rate increases for each increase in the aid flows received—the more aid is effective.

²² Similarly, Collier *et. al.* (1997) acknowledge that “one disadvantage with switching from policies to outcomes is that it can reward good luck.”

Generally, in order to compute aid effectiveness, one would like to estimate the following equation for each country i and year t :

$$(2) \text{ Outcome}_{it} = \alpha + \beta_{it} \frac{\text{Aid}}{\text{GDP}_{it}} + \gamma X_{it} + \epsilon_{it}$$

Where $\epsilon_{it} \sim N(0, \sigma_{\epsilon}^2)$

β_{it} measures aid effectiveness. Outcome_{it} is the outcome of interest (for example the growth rate of per capita GDP) and $\frac{\text{Aid}}{\text{GDP}_{it}}$ are the aid flows normalized by GDP. X_{it} is a vector of control variables. ϵ_{it} is a year-country shock. All the coefficients in equation (2) are time-varying, which is why I write β_{it} to denote the coefficient on aid in country i at year t .

The difficulty comes from the fact that with the econometric methods that are usually used the equation one estimates is not (2) but:

$$(3) \text{ Outcome}_{it} = \alpha + \beta_i \frac{\text{Aid}}{\text{GDP}_{it}} + \gamma X_{it} + \epsilon_{it}$$

Where $\epsilon_{it} \sim N(0, \sigma_{\epsilon}^2)$

That is, one only computes one coefficient for each country i and the entire time period (β_i) and not one coefficient country i and year t (β_{it}). In other words, one cannot estimate aid effectiveness annually.

One way to estimate aid effectiveness annually—to implement the estimation of equation (2)—is to use the “local Gaussian-weighted ordinary least squares” method (used for example by Aghion and Marinescu (2007)). The basic idea of this method—which is also called kernel-based nonparametric regression or local smoothing—is to put more weight on the most recent years. For each year, points that are closer in time are given more weight than points that are further away. More precisely, all the observations are weighted by a Gaussian centered at date t but, since to estimate aid effectiveness in t one only wants to use the information available for the years preceding t , I put a zero weight on the years following t .

Under this method, jumps in the coefficient β are mainly due to changes in the immediate neighborhood of date t , as those observations in the immediate neighborhood of date t are given highest weight. Hence, if there is a change in the aid effectiveness coefficient in t —say an increase—this comes from the fact that the country has improved its effective use of aid in t . And so in terms of aid allocation, it has to be rewarded for this improvement.

Using the local Gaussian-weighted ordinary least squares method, one could thus estimate an aid effectiveness coefficient for each country and each year. However, aid can be endogenous which can lead to biased results.²³ In order to deal with these endogeneity problems, one can use Gaussian-weighted two-stage least squares instrumenting for aid. It is

²³ This endogeneity can come from (i) reverse causation: growth causes aid (e.g. the higher its growth rate, the less aid a country receives because it does not need it); or (ii) simultaneous causation: an omitted variable causes both aid and growth.

however important to underline that the relevance of the method is totally independent of the instrument choice.

Using Gaussian-weighted two-stage least squares, one could thus obtain a time-varying measure of aid effectiveness, with one coefficient per year and per country. These coefficients can be interpreted as an estimate of aid elasticity with respect to the outcome of interest, i.e. a measure of the country performance with respect to aid effectiveness. This tool could thus be used to reward good “aid performer.” Doing so, it helps overcome one of the main weaknesses of the outcomes-based allocations that have been proposed until now in the literature: it does not reward good luck. Indeed, a country can have one given year for example a very high growth rate and at the same time obtain a very low coefficient for aid effectiveness if it did not use aid in an effective way. In this case, its aid allocation decreases despite its good growth performance. The tool I propose only relies on the implicit assumption that past aid effectiveness is a good guide to future aid effectiveness. This is a weaker and more relevant assumption than the one according to which past output indicators are a good guide to future aid effectiveness.

One can also choose to normalize the coefficients obtained using local Gaussian-weighted ordinary least squares by the “global” aid effectiveness coefficient obtained by performing equation (3) for all the countries of the sample taken together (cross-countries regression with country fixed effects). Another possibility could be, rather than to normalize by the cross-countries coefficient, to take into account for each country the performance of its neighbors, for example using the geographical distance. Indeed, there may be some externalities created by an increase or a decrease in aid effectiveness in a country for its neighbors.

A possible caveat of such a measure is that it does not take into account how donors can have an impact on aid effectiveness. Implicit here is indeed the assumption that aid performance is only the consequence of decisions of the recipient country itself and not of somebody else. But aid effectiveness does not depend only on the behavior of the recipient countries but also on the donor's behavior and one does not want recipients to suffer from bad behavior of donors. One possibility would be to control for an index of donor performance.

Similarly, aid effectiveness can be affected by events not depending only on the recipient's economic policy, for example exogenous shocks like natural disasters. One would have to be very careful in controlling for these exogenous shocks.

Despite these caveats, I think that using and aid effectiveness coefficient—together with other indicators—to allocate aid would be a relevant tool. Indeed, with such a tool, one does not give aid to countries that are the best performers for example in the sense of having a higher rate of girls' enrollment and so perhaps are not the one which need it the most, but to the countries where aid will be used in the most efficient way. That is to say, to the countries that have a sufficient absorptive capacity for receiving higher aid flows. It could help allocating a scarce resource in the most efficient way.

4. CONCLUSION

In this article I question the relevance of the different measures of policy performance that are currently used. I evaluate more especially the pertinence of the CPIA and of the various alternatives that have been proposed in the literature.

Using a yearly panel dataset over 146 countries between 1977 and 2008, I show that the CPIA is correlated with current growth rates but that it is not a good predictor for future economic growth. I thus argue in favor of other measures of policy performance. I underline the need of introducing new criteria when measuring policy performance. In particular, I show that more weight has to be given to the role of the government in the development process, which supposes to take into account its fiscal capacity. Fiscal capacity is indeed of crucial importance for raising domestic tax revenues. I also underline the importance of the quality of industrial policy and especially of export promotion strategies on the path towards sustained growth.

I then focus more specifically on the allocation of development aid which is of great importance given both the scarcity and need for aid. I show that performance-based measures, as opposed to measures implying *ex ante* conditionality, are more accurate instruments for aid allocation. However, performance-based measures proposed in the literature do not help overcome the difficulty of estimating the elasticity of aid effectiveness with respect to outcome-based performance indicators. They let unsolved the question of whether when a donor rewards a recipient for its good performance with respect to a given outcome variable it is not rewarding “luck” rather than an effective use of aid.

Since in order to allocate aid effectively it appears essential to evaluate the elasticity of aid effectiveness with respect to performance indicators, I discuss a new tool based on this elasticity. Using new econometric methods, I show that one could use a time-varying measure of aid effectiveness as an indicator of the performance of a country with respect to aid efficiency. This tool shares with an outcome-based allocation the advantage of not relying on *ex ante* conditionality. Moreover, it is an improvement upon this outcome-based approach since it is a way to reward good “aid performer” rather than good “luck.”

Needless to say, more research on aid effectiveness indicators is necessary before they can be applied. The tool could indeed be used by different donors with different goals (growth, poverty, education and so on.). This is an important characteristic because different criteria can be used given the complexity of the relationship between aid, growth and poverty reduction, depending on the recipient country specificities and on the donor preferences. The downside is that aid effectiveness coefficients might be too volatile to be used as single indicators. In my view, the most promising avenue is to use them together with other key development indicators such as investment in fiscal capacities.

REFERENCES

- Aghion, P., and I. Marinescu (2007): "Productivity Growth and Countercyclical Budgetary Policy: What Do We Learn from OECD Panel Data?," *NBER Macroeconomic Annual*.
- Aizenman, J., and N. Marion (1999): "Volatility and Investment: Interpreting Evidence from Developing Countries," *Economica*, 66, pp. 157-179.
- Alesina, A., and D. Dollar (2000): "Who Gives Foreign Aid to Whom and Why?," *Journal of Economic Growth*, 5(1), pp. 3-63.
- Alesina, A., and B. Weder (2002): "Do Corrupt Governments Receive Less Foreign Aid?," *The American Economic Review*, 92(4), pp. 1126-1137.
- Amprou, J., P. Guillaumont, and S. Guillaumont-Jeanneney (2006): "Aid Selectivity According to Augmented Criteria" .
- Arellano, M., and O. Bover (1995): "Another Look at the Instrumental Variables Estimation of Error Components Models," *Journal of Econometrics*, 68, pp. 29-51.
- Barder, O., and N. Birdsall (2006): "Payments for Progress: A Hands-Off Approach to Foreign Aid," *Center for Global Development Working Paper* Number 102.
- Buiter, W. H. (2007): "No Bricks without Straw: A Critique of Ravi Kanbur's 'Modest Proposal for Introducing Development Outcomes in IDA Allocation Procedures,'" Note prepared for the meeting at Columbia University on April 5th, 2007, organized by the *Initiative for Policy Dialogue* at Columbia University on the CPIA and Aid Allocation.
- Burnside, C., and D. Dollar (2000): "Aid, Policies and Growth," *American Economic Review*, 90(4), pp. 847-868.
- Cage, J. (2009): "Growth, Poverty Reduction and Governance in Developing Countries: a Survey," CEPREMAP Docweb No. 0904.
- Cage, J., and L. Gadenne (2012): "The Fiscal Cost of Trade Liberalization," *PSE Working Papers* halshs-00705354, HAL.
- Cage, J., and D. Rouzet (2012): "Improving "National Brands": Reputation for Quality and Export Promotion Policies," .
- Cogneau, D., and J.-D. Naudet (2007): "Who Deserves Aid? Equality of Opportunity, International Aid, and Poverty Reduction," *World Development*, 35(1), pp. 104-120.
- Collier, P., and D. Dollar (1999): "Aid Allocation and Poverty Reduction," *World Bank Policy Research Working Paper* No. 2041.
- Collier, P., and D. Dollar (2001): "Can the World Cut Poverty in Half? How Policy Reform and Effective Aid Can Meet International Development Goals," *World Development*, 29(11), pp. 1787-1802.
- Collier, P., P. Guillaumont, S. Guillaumont, and J. W. Gunning (1997): "Redesigning Conditionality," *World Development*, 25(9), pp. 1399-1407.
- Gelb, A., B. Ngo, and X. Ye (2004): "Implementing Performance-Based Aid in Africa: The Country Policy and Institutional Assessment," *World Bank Africa Region Working Paper Series*, No. 77, World Bank, Washington, D.C.

- Glaeser, E., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2004): "Do Institutions Cause Growth?," *Journal of Economic Growth*, 9, pp. 271-303.
- Griffith-Jones, S. (2000): "Proposal for a Better International Financial System," *World Economics*, 2, pp. 111-133.
- Hadi, A. S. (1992): "A New Measure of Overall Potential Influence in Linear Regression," *Computational Statistics and Data Analysis*, 14, pp. 1-27.
- Herman, B. (2004): "How Well Do Measurement of an Enabling Domestic Environment for Development Stand Up?," Draft, XVII Technical Group Meeting of the Group of 24, Geneva, March 8-9, 2004.
- Hoff, K., and J. Stiglitz (2001): "Modern Economic Theory and Development," in *Frontiers of Development Economics: The Future in Perspective*, ed. by G. Meier, and J. E. Stiglitz, pp. pp. 389-485. New York: Oxford University Press.
- Independent Evaluation Group (2009): "The World Bank's Country Policy and Institutional Assessment An Evaluation" .
- Kaminsky, G. L., and C. M. Reinhart (2001): "Financial Markets in Times of Stress," *NBER Working Paper* No. 8569.
- Kanbur, R. (2005): "Reforming the Formula: A Modest Proposal for Introducing Development Outcomes in IDA Allocation Procedures," Paper for presentation at the *AFD-EUDN conference*, Paris, November 25-27, 2004.
- Kapur, D., J. Lewis, and R. Webb (1997): *The World Bank. Its First Half Century. Volume I: History*. Brookings, Washington, DC.
- Kraay, A., and V. Nehru (2004): "When is External Debt Sustainable?," *World Bank Policy Research Working Paper*, No. 3200, World Bank Institute., Washington, D.C.
- Kurtz, M. J., and A. Schrank (2007): "Growth and Governance: Models, Measures, and Mechanisms," *Journal of Politics*, 2, pp. 538-554.
- Kuziemko I. & Werker, E. (2006): "How Much is a Seat on the Security Council Worth? Foreign Aid and Bribery at the United Nations," *Journal of Political Economy*, 114, pp. 905-930.
- Levine, R., and D. Renelt (1992): "A Sensitivity Analysis of Cross-Country Growth Regressions," *American Economic Review*, 82(4), pp. 942-963.
- Llavador, H. G., and J. E. Roemer (2001): "An Equal-Opportunity Approach to the Allocation of International Aid," *Journal of Development Economics*, 64, pp. 147-171.
- McGillivray, M. (2004): "Aid Effectiveness and Recipient Policy Regimes: A Comment on Ravi Kanbur's "Reforming the Formula" Paper," Comment prepared for the 2nd *AFD-EUDN Conference*, Paris, November 25-27, 2004.
- Ramey, G., and V. A. Ramey (1995): "Cross-Country Evidence on the Link Between Volatility and Growth," *The American Economic Review*, 85(5), pp. 1138-1151.
- Roodman, D. (2006a): "An Index of Donor Performance," *Center for Global Development Working Paper* No. 67.
- Roodman, D. (2006b): "How to Do xtabond2: An Introduction to Difference and System GMM in Stata," *Center for Global Development Working Paper* Number 103.

- Schraeder, P., S. Hook, and B. Taylor (1998): "Clarifying the Foreign Aid Puzzle: a Comparison of American, Japanese, French, and Swedish Aid Flows," *World Politics*, 50, pp. 294-323.
- Stiglitz, J. (1998): "More Instruments and Broader Goals: Moving Toward the Post Washington Consensus," *Wider Annual Lecture*, Helsinki, 7 January. Reprinted in Chang, HJ.
- Stiglitz, J. and the World Bank (2001): *The Rebel Within*, Anthem Press, London.
- Stiglitz, J. E. (1999): "The World Bank at the Millennium," *The Economic Journal*, 109(459), pp. F577-F597.
- Stuckler, D., L. King, and G. Patton (2009): "The Social Construction of Successful Market Reforms," *Political Economy Research Institute Working Paper*.
- van Waeyenberge, E. (2006): "The Missing Piece: Country Policy and Institutional Assessments at the Bank," .
- Williamson, J., and M. Mahar (1998): *A Survey of Financial Liberalization (Essays in International Economics)*.
- Wood, A. (2007): "Looking Ahead Optimally in Allocating Aid:," *QEH Working Paper* Number 137.
- World Bank (1998): *Assessing Aid: What Works, What Doesn't and Why*. Washington, D.C., World Bank.
- World Bank (2001): "OED IDA Review: Review of the Performance-Based Allocation System, IDA 10-12,"
- World Bank (2004): "Country Policy and Institutional Assessment. 2004 Assessment Questionnaire. Operations Policy and Country Services.,."
- World Bank (2005): *Economic Growth in the 1990's Learning from a Decade of Reform*. The International Bank for Reconstruction and Development / The World Bank.
- World Bank (2007): "Country Policy and Institutional Assessment 2007: Assessment Questionnaire," .
- World Bank (2008): "The 2008 Country Policy and Institutional Assessments (CPIA). 2008 CPIA Questionnaire," .

APPENDIX: DATA SOURCES AND DESCRIPTION

Aid: Net Aid Transfers (NAT). Source: DAC.

CPIA: Country Policy and Institutional Assessment. Annual performance assessment of its client countries' capacity to effectively absorb development assistance carried out by the World Bank since 1977. Source: World Bank.

Per capita GDP growth rate: Annual percentage growth rate of per capita GDP at market prices based on constant local currency. Source: WDI.

M2 (percent GDP): Money and quasi money comprise the sum of currency outside banks, demand deposits other than those of the central government, and the time, savings, and foreign currency deposits of resident sectors other than the central government. Source: WDI.

UNus: Annual correlation of voting records in UNGA between recipient and the US (-1 to 1). Source: Erik Voeten.

Table 4: CPIA Criteria 2008

<p>A. Economic Management</p> <ol style="list-style-type: none">1. Macroeconomic Management2. Fiscal Policy3. Debt Policy <p>B. Structural Policies</p> <ol style="list-style-type: none">4. Policies and Institutions for Economic Cooperation, Regional Integration & Trade5. Financial Sector6. Business Regulatory Environment <p>C. Policies for Social Inclusion/Equity</p> <ol style="list-style-type: none">7. Gender Equality8. Equity of Public Resource Use9. Building Human Resources10. Social Protection and Labor11. Environmental Policies and Regulations <p>D. Governance Rating: Public Sector Management and Institutions</p> <ol style="list-style-type: none">1. Property Rights and Rule-based Governance2. Quality of Budgetary and Financial Management3. Efficiency of Revenue Mobilization4. Quality of Public Administration5. Transparency, Accountability, and Corruption in the Public Sector
--

Source: World Bank, 2008