



## Does conditionality work? A test for an innovative US aid scheme

Hannes Öhler<sup>a</sup>, Peter Nunnenkamp<sup>b</sup>, Axel Dreher<sup>c,a,d,e,f,\*</sup>

<sup>a</sup> Georg-August University Goettingen, Platz der Goettinger Sieben 3, D-37073 Goettingen, Germany

<sup>b</sup> Kiel Institute for the World Economy, Hindenburgufer 66, D-24105 Kiel, Germany

<sup>c</sup> Heidelberg University, Alfred-Weber-Institut for Economics, Bergheimer Strasse 58, D-69115 Heidelberg, Germany

<sup>d</sup> CESifo, Germany

<sup>e</sup> IZA, Germany

<sup>f</sup> KOF Swiss Economic Institute, Switzerland

### ARTICLE INFO

#### Article history:

Received 12 October 2010

Accepted 6 May 2011

Available online 26 May 2011

#### JEL classification:

F35

O17

#### Keywords:

Foreign aid

Corruption

Millennium Challenge Corporation

MCC effect

### ABSTRACT

Performance-based aid has been proposed as an alternative to the failed traditional approach whereby donors make aid conditional on the reform promises of recipient countries. However, hardly any empirical evidence exists on whether ex post rewards are effective in inducing reforms. We attempt to fill this gap by investigating whether the Millennium Challenge Corporation (MCC) was successful in promoting better control of corruption. We employ a difference-in-difference-in-differences (DDD) approach, considering different ways of defining the treatment group as well as different time periods during which incentive effects could have materialized. We find evidence of strong anticipation effects immediately after the announcement of the MCC, while increasing uncertainty about the timing and amount of MCC aid appears to have weakened the incentive to fight corruption over time.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

Performance-based aid has received increasing attention in the international development community recently, particularly in the health sector (e.g., Eichler and Glassman, 2008). However, empirical assessments of its effectiveness are still largely lacking. We attempt to fill this gap by investigating the so-called MCC Effect.<sup>1</sup> The Millennium Challenge Corporation, established by the Bush administration in 2004, has been deliberately shaped in such a way as to grant aid as ex post rewards for proven achievements.

The move towards performance-based aid reflects increasing doubts in the donor community that traditional forms of conditional aid have been effective. Ex ante conditionality appears to have failed in international development cooperation, in particular to the extent that donors had used foreign aid as a means to “buy” policy reform in recipient countries (Collier, 1997). A similar verdict appears to apply to less intrusive motives of conditionality such as ensuring loan recovery. International financial institutions, notably the World Bank and the International Monetary Fund (IMF), have achieved little by attaching (a typically large number of) conditions to adjustment loans. Dreher (2009: 256) concludes

\* Corresponding author at: Heidelberg University, Alfred-Weber-Institut for Economics, Bergheimer Strasse 58, D-69115 Heidelberg, Germany. Tel.: +49 551 39 10614.

E-mail addresses: [hoehler@uni-goettingen.de](mailto:hoehler@uni-goettingen.de) (H. Öhler), [peter.nunnenkamp@ifw-kiel.de](mailto:peter.nunnenkamp@ifw-kiel.de) (P. Nunnenkamp), [mail@axel-dreher.de](mailto:mail@axel-dreher.de) (A. Dreher).

<sup>1</sup> See <http://www.mcc.gov/mcc/panda/activities/mcceffect/index.shtml>.

from a comprehensive survey of IMF-related research that “there is no empirical evidence showing that conditions enhance ownership or make program success more likely.”

By contrast, conditionality is widely perceived to have been “very effective” (Sedelmeier, 2008: 806) in the European Union’s enlargement strategy. The desirability of EU membership appears to have prompted candidate countries to adhere to a host of conditions contained in the so-called *Acqui Communautaire*. The EU concept is based on ex post selectivity according to which performance standards have to be met prior to accession and the associated financial rewards. It is open to question whether the EU experience offers relevant lessons for international development cooperation. All the same, the review of the literature in Section 2 leads us to hypothesize that the effectiveness of conditionality depends on the underlying concept of providing reform incentives.

We focus on the effects of MCC aid on corruption in candidate countries. The international aid community has been aware that corruption is a major bottleneck to the effectiveness of aid, and development prospects in general, since the second half of the 1990s (Easterly, 2007). Corruption features most prominently among MCC’s eligibility criteria. We employ a difference-in-difference-in-differences (DDD) approach to assess whether the treatment groups fought corruption more effectively than the control groups after the Bush administration announced the creation of the MCC and its performance-based aid allocation approach (Section 4). We consider different variants of defining the treatment group as well as different time periods during which incentive effects could have materialized. We find evidence of strong anticipation effects immediately after the announcement of the MCC, while increasing uncertainty about the timing and amount of MCC aid appears to have weakened the incentive to fight corruption over time (Section 5). Section 6 summarizes the MCC experience and offers policy conclusions on how conditionality can be designed in such a way that it becomes more effective.

## 2. The debate on conditionality

There are different reasons why international financial institutions and bilateral donors attach conditions to adjustment loans and development aid. Most obviously, temporary balance-of-payments support in the form of loans from the International Monetary Fund (IMF) has to be repaid by the borrowing country, and conditionality appears to be a means for the IMF to monitor the borrowing country’s adjustment efforts and, thereby, ensure timely repayment. Drazen (2002) argues that conditionality implies the presence of conflicts of interest. Conflicts of interest may prevail between the IMF and the borrowing country, or within the country. Drazen stresses that reform-minded local authorities may be interested to implement an adjustment program, but IMF conditionality may still be needed to overcome internal opposition. Moreover, the incentives of local authorities to actually implement policy changes tend to weaken once the adjustment loans have been received, even if they agreed to policy adjustments in earlier negotiations with the IMF.

The IMF typically provides its financial support in successive tranches contingent on progress in achieving policy adjustments in order to mitigate time inconsistency problems (IMF, 2001: 11). According to Vreeland (2006), however, the frequency of waivers from program obligations indicates that noncompliance is not rigorously punished by the IMF. It appears that the IMF is not credibly committed to disbursing successive tranches only if policy conditions are actually implemented by the borrowing countries.<sup>2</sup>

Conditionality – and its failure – could also be due to conflicts of interest between international financial institutions and major member countries. On the one hand, the funding of these institutions may be easier to sell to national constituencies in major member countries if borrowers are supported only when principally agreeing to strict conditions. On the other hand, the lax enforcement of traditional conditionality can also be attributed to political pressure from major member countries. For instance, Stone (2004) shows that the punishment for noncompliance with IMF conditions is significantly weaker for countries that are considered to be important to the United States.

Compared to adjustment loans by the IMF, donors of official development aid are less likely to use conditionality as a means to ensure the repayment of transitory financial support. Aid reported by the OECD’s Development Assistance Committee consisted mainly of outright grants, while repayable loans contributed just 20 percent to total aid in 2005–2009.<sup>3</sup> Furthermore, aid transfers are motivated by different donor objectives, some of which are unrelated to policy reform in the recipient countries. For instance, conditionality is unlikely to be an issue if donors pursue selfish aims such as export promotion through aid. The same applies to emergency relief. All the same, Collier (1997) posits that policy reform is an important criterion on which aid might reasonably be judged: Aid might be “remarkably effective if it induces governments to adopt growth-inducing and poverty-reducing policies. This is indeed the core of what conditionality is supposedly about—aid buys reform. Unfortunately, it does no such thing” (Collier, 1997: 56).

The example of Kenya has repeatedly been used to demonstrate the failure of traditional conditionality in international development cooperation (Collier, 1997; Svensson, 2003). Kenya has outmaneuvered even major donors such as the World Bank. The country agreed to agricultural reform on four occasions in 15 years but backtracked each time after having received the aid money. Ex ante threats by donors that they will not disburse committed aid if the recipient does not fulfill

<sup>2</sup> The IMF’s Independent Evaluation Office (IEO, 2007) finds little to suggest that compliance has improved since the IMF attempted to streamline its conditionality.

<sup>3</sup> See <http://www.oecd.org/dataoecd/50/17/5037721.htm> (last accessed January 13, 2011).

reform promises are hardly credible. Time inconsistency problems loom large considering the incentives for aid agencies and specific country desks to spend overall budgets and fully exhaust country quotas. Indeed, Svensson (2003: 383) finds “no link between a country’s reform effort, or fulfillment of ‘conditionality’, and the disbursement rate [of aid funds].” According to Heckelman and Knack (2008), higher aid even slowed policy reform over the 1980–2000 period.

Focusing on corruption, as we do in the following, Alesina and Weder (2002: 1136) do not find “any even weak evidence” suggesting that efficient and honest governments have been rewarded by more bilateral or multilateral aid during the 1975–1994 period. In 1997, the German government responded to a parliamentary inquiry that “no development cooperation contracts were annulled due to proof of corruption” (as quoted by Cremer, 2008: 122). Isopi and Mattesini (2010) find that Germany gives more aid to more corrupt countries. The same holds for Finland, France, Japan, and the Netherlands, while Canada, Norway and the United Kingdom give less aid to more corrupt recipients.

A vicious circle of aid inducing more corruption and further eroding governance may follow from the failure of traditional conditionality. According to Alesina and Weder (2002), increases in aid tend to be associated with an increase in corruption.<sup>4</sup> Knack (2001) provides cross-country evidence that higher aid levels erode the quality of governance more generally. These rather bleak findings have fueled the debate on how to redesign aid conditionality. Collier (1997) as well as Collier et al. (1997) observed a move towards “short-leash” conditionality in the 1990s; i.e., donors disbursed program aid in several tranches depending on progress in reform implementation. Collier and co-authors argued, however, that donors taking these first steps towards performance-based aid allocation still attempted to buy reforms from recipients that were frequently unwilling to sustain reforms. These authors called for a fundamental change in donor behavior to reward reform-minded recipient countries and enhance the effectiveness of aid: Aid allocation would be based on retrospective performance appraisals, rather than being conditional on reform promises.

As detailed below, the so-called MCC Effect is based on the reasoning that the reform-mindedness of recipient countries will generally increase once aid is disbursed according to the relative performance of competing countries. However, whether rewarding relatively good performance *ex post* will result in stronger ownership of reforms by aid recipients remains the subject of controversy and debate. According to Mosley et al. (2004), selectivity may provide incentives to improve policies prior to receiving aid, but recipients would still have the option to reverse reforms after having been selected by donors. Conflicting hypotheses on whether selectivity promotes sustained reform efforts have hardly been subjected to systematic empirical tests in the aid-related literature.

The plausibility of conflicting hypotheses may be assessed at least tentatively by referring to the extensive literature on the role of preconditions that must be fulfilled by candidate countries for EU membership. Clearly, the situation of EU accession countries differs considerably from the situation of aid recipients in the developing world. The conditions imposed in the context of EU accession, ranging from minority rights and judicial reforms to social security systems and corporate governance, extend far beyond the typical coverage of conditions in aid and lending contracts that developing countries enter with donor countries and international financial institutions. At the same time, financial transfers in the EU context are not only much higher than those to developing countries; they are also rules-based in the sense that net-payers – i.e., the richer EU members – cannot reduce or discontinue the transfers at their own discretion, unlike aid donors who are notorious for granting aid as it fits the current domestic budget situation and political environment. As a consequence, EU transfers are reliable and predictable, whereas aid flows are volatile and hard to predict.

Keeping these important qualifications in mind, the EU approach of granting financial rewards only at the end of a fairly long process of accession invites several hypotheses concerning the effects of performance-based aid. First, the literature on the EU’s enlargement strategy suggests that the incentive effects of conditionality are likely to depend on initial conditions in the treatment group of prospective EU member countries or, respectively, aid recipient countries. Unfavorable initial conditions imply higher costs of compliance. They may weaken the incentive effects of conditionality, especially if selection occurs according to relative performance, as is the case under MCC rules (see below). The reform efforts of candidate countries that lie closer to the current threshold would raise the stakes for all candidates so that those with unfavorable initial conditions might abstain from reform if they consider the reward to be out of reach.

Second, for given costs of compliance the incentive effects can be expected to depend on the expected size of the reward and the likelihood that it will actually materialize before too long. In the EU context, this point has been made with respect to the Eastern enlargement as well as with respect to current candidate countries.<sup>5</sup> Especially Turkey’s incentives to fulfill EU conditions are weakened considerably by the distance and lacking credibility of prospective membership. As we will argue below, uncertainty about prospective rewards may also have weakened the MCC Effect.

Third, one might suspect that the materialization of the reward puts the sustainability of reforms at risk. Studies addressing this hypothesis in the EU context reveal ambiguous findings. In the light of our focus on corruption, it is most interesting to note how Pridham (2008) assesses the fight against corruption in Latvia and Slovakia during the first three years of EU membership. This study concludes that “there is no common pattern whereby conditionality loses momentum and becomes unscrambled.” It remains open to question whether this finding carries over to aid contracts between donor and recipient countries. Phenomena such as status-quo bias and social learning – offered by Pridham (2008) as possible

<sup>4</sup> Svensson (2000) presents similar results for ethnically diverse recipient countries in which social groups compete over common-pool resources, whereas Tavares (2003) finds aid to be associated with less corruption.

<sup>5</sup> According to Grabbe (2001: 1025), the EU’s influence on governance in Central and Eastern Europe has been diluted by “an uncertain linkage between fulfilling particular tasks and receiving particular benefits.”

explanations for sustained reforms in new EU countries – may play a minor role in relatively “casual” donor–recipient relations.

### 3. The MCC's approach and conditions

The failure of traditional forms of conditionality provides one major pillar of the MCC concept. The second pillar is given by Burnside and Dollar's (2000) highly influential analysis which suggests that aid promotes growth only in an environment characterized by “good” policies. Following Burnside and Dollar, it has been subject to controversial debate whether certain conditions have to be met by recipients for aid to be effective, and exactly which of these conditions might be most relevant.<sup>6</sup> However, even strong critics of Burnside and Dollar agree on the relevance of selectivity. Easterly (2007: 645) acknowledges that “the idea that aid money directed to governments would be more productive if those governments had pro-development policies and institutions is very intuitive.” Mosley et al. (2004) share the view that aid could be more effective if it were reallocated to less corrupt recipient countries.

The failure of traditional conditionality and the discussion on the need for an appropriate local environment for aid to be effective may have shaped the aid allocation procedures of various multilateral development agencies and bilateral donors. Yet the MCC clearly stands out so that the effects of the new donor approach should be most visible in the allocation of MCC aid. Instead of merely adjusting or extending the mandate of an existing aid agency such as USAID, the Bush administration explicitly established the MCC as a new aid agency in order to prevent institutional legacies from undermining innovative performance-based allocation rules (Radelet, 2003; Rieffel and Fox, 2008). Given the unresolved debate on the incentive effects of ex post conditionality and sparse empirical evidence, it would be most useful to know whether and to what extent MCC-type conditionality affects the policies of the potential beneficiaries. The MCC (2008: 1) itself is fairly confident of having improved the incentives for reform in potential recipient countries:

The MCC Effect is the positive impact that MCC is having on developing countries beyond its direct investments. To date, the most significant impact has been the incentive created for countries to adopt legal, policy, regulatory, and institutional reforms related to the MCC eligibility criteria. Eligibility for MCC funding can lead to international recognition and increased private sector investment, which has encouraged many countries to implement significant political, social, and economic reforms with tangible results on the ground. In areas as diverse as women's rights, anti-corruption and governance, and business registration, countries are taking it upon themselves to re-evaluate their laws, policies, regulations, and ways of ‘doing business’.

The MCC signed so-called compacts – the multi-year aid programs concluded between the MCC and eligible countries – with 20 strictly selected recipient countries until early 2010.<sup>7</sup> In various cases, the MCC offered remarkable financial rewards. It is not only in small recipient countries such as Cape Verde, Vanuatu and Lesotho that MCC aid played an important role. MCC's aid obligations agreed to in compacts exceeded 20 percent of total aid commitments of all donors during the five previous years in eight out of 20 countries with compacts, and 10 percent in another six countries.<sup>8</sup> Furthermore, Dreher et al. (2010) find that other donors followed the MCC by granting more aid to countries with compacts, pointing to additional indirect MCC effects. MCC aid is also relevant relative to the recipient countries' GDP in the year of signing compacts. On average, this ratio amounted to 6 percent in the group of 20 countries, ranging from less than 1 percent in Morocco to 23 percent in Lesotho.<sup>9</sup> The hypothesis that the MCC creates incentives to improve policies and institutions is thus plausible for most recipient countries.

The MCC's eligibility criteria leave little doubt about the strictness of selectivity in granting aid to needy and well performing recipients only. Eligibility is restricted to relatively poor countries. The “hurdles approach” (Radelet, 2003: 24) requires potential recipient countries to score higher than the median on at least half of the eligibility criteria (in each of three broad categories) across peers in the same income category, with control of corruption being the only mandatory prerequisite (see also below).<sup>10</sup> The MCC groups the 16 indicators, all taken from independent sources, into three categories: Ruling Justly, Investing in People, and Encouraging Economic Freedom. Ruling Justly comprises the Civil Liberties and Political Rights indicators from Freedom House, and four indicators from the World Bank's Governance Indicators (Kaufmann et al., 2009), including the control of corruption. Economic Freedom consists of indicators on regulatory quality, a country's credit rating, inflation, as well as fiscal, regulatory, and trade policies. Investing in People refers to public expenditures on health and primary education, immunization rates, and primary education completion rates.

It is open to question, however, whether the MCC's selective aid allocation has the desired impact on the incentives of potential recipient countries to improve their policies and institutions. To our knowledge, the only independent and systematic study on this issue provides preliminary evidence supporting positive MCC effects (Johnson and Zajonc, 2006).

<sup>6</sup> For instance, Doucouliagos and Paldam (2010) show that the interactions of aid with various conditioning variables in growth models are generally fragile, sensitive to small changes in the data set or in the model specification, and dependent on author affiliation and ideology.

<sup>7</sup> See Appendix B in the working paper version of this article, Öhler et al. (2011).

<sup>8</sup> See Fig. 1 in Öhler et al. (2011).

<sup>9</sup> See Fig. 2 in Öhler et al. (2011).

<sup>10</sup> Specific eligibility criteria have been slightly modified over time. We use the indicators as shown in the MCC's scorebook of 2004 ([http://www.mcc.gov/mcc/bm.doc/score\\_fy04\\_all.pdf](http://www.mcc.gov/mcc/bm.doc/score_fy04_all.pdf)).

Specifically, Johnson and Zajonc find that candidate countries improved their indicators by 25 percent more than non-candidate countries after the MCC had been announced. It is interesting to note, however, that this study does not find support in justification of MCC claims about having induced better control of corruption, even though this criterion figures most prominently among the eligibility criteria. As indicated above, a country must score above the median with regard to the control of corruption, regardless of how well it performs on all other eligibility criteria.

#### 4. Method and data

The significance of the control of corruption among the MCC's eligibility criteria and the widespread support that it commands within the aid community provide the reasons for us to focus on the effects of MCC conditionality on corruption in potential recipient countries. In accordance with the MCC convention, we use the index of Kaufmann et al. (2009) on the control of corruption. The index is constructed by Kaufmann et al. using an unobserved components model, based on a large number of different surveys of perceived corruption from various independent organizations. Perceived corruption is defined as the abuse of public power for private gains. Index values range from  $-2.5$  to  $+2.5$  with higher values indicating better control of corruption.<sup>11</sup>

Most of our sample countries clearly rank below the mean for all 208 countries (in 2008) included in Kaufmann et al. (2009).<sup>12</sup> This applies to both sub-samples, i.e., the treatment and control groups used for the baseline estimations reported in Table 1 in Section 5.1 below.<sup>13</sup> However, countries in the control group score over a much wider range of the index than countries in the treatment group. While the best-placed country in the treatment group consistently ranks below zero throughout the period of observation, the distribution is less lopsided for the countries in the control group.

Relying on the index of Kaufmann et al., we ask whether the substantial increases in US aid provide sufficient incentives for candidate countries to fight corruption more effectively. President Bush announced in 2002 that the MCC would command over US\$ 5 billion annually and that MCC aid would be “above and beyond existing aid.” Hence, recalling the discussion on sufficiently large rewards in Section 2, MCC effects on corruption could differ considerably from the earlier findings of Alesina and Weder (2002).

Our focus on corruption implies that our approach is rather modest in comparison with that of Johnson and Zajonc (2006) who consider all eligibility criteria, even though we do take into account that the incentives for candidate countries to fight corruption may depend on how they score on other eligibility criteria (see below). On the other hand, we extend the analysis of Johnson and Zajonc in several ways. Most crucially, we overcome the drawback that the MCC was still in its infancy when the earlier study was presented; it was probably too early to draw conclusions with data only extending into the first year (2004) of MCC operations. The considerably longer period of observation available in the present analysis also allows us to test several of the specific hypotheses raised in Section 2 above. First, we can address the disputed issue of whether reforms (i.e., more effective control of corruption) are likely to be sustained once the reward (i.e., MCC aid) has been granted. Second, we can assess whether the incentive to reform weakens once the reward becomes less compelling, either because the MCC's overall budget fell short of initial announcements or because the actual disbursement of aid was delayed in the context of compacts.<sup>14</sup> Finally, by controlling for initial conditions in the candidate countries, we evaluate whether higher costs of compliance with MCC conditions undermine the MCC Effect.

In order to address these questions empirically, we rely on the difference-in-differences (DDD) approach. Our identification strategy rests on the observation that there can be no incentive effect prior to 2002 when the MCC was announced. Moreover, we limit our sample of MCC candidates to countries with per-capita GDP equal to or less than US\$ 1415, as countries above this threshold were ineligible for MCC aid when operations started in 2004.<sup>15</sup>

A simple approach to test for the potential effect of the MCC would be to use the before–after approach which involves comparing the level of perceived corruption in eligible countries before and after 2002. Clearly, the strong assumption that no other omitted variable might have changed corruption after 2002 is unlikely to hold. The alternative with–without approach would simply entail comparing the changes in corruption in countries with incentives to qualify for MCC aid and countries without, subsequent to the announcement of the MCC. Again, this would require a strong assumption that is unlikely to hold, namely that no other factors affect eligible and non-eligible countries systematically in the period of observation.

<sup>11</sup> Composite corruption indexes such as the index of Kaufmann et al. aggregate and synthesize information from various third-party data sources (for details, see, e.g., UNDP, 2008). In the case of Kaufmann et al., the underlying data are partly objective and partly subjective. The index is also hybrid in that input-based (*de jure* on institutions and rules) as well as output-based (*de facto* on impact) indicators are included. It focuses on corruption in the public sector but also uses information on corruption in the private sector. Kaufmann et al. construct the index in such a way that it follows a normal distribution with a mean of zero across all countries and a standard deviation of one. See Öhler et al. (2011) for a summary of the limitations of composite corruption indexes and detailed references.

<sup>12</sup> See Table 1 in Öhler et al. (2011).

<sup>13</sup> The treatment group consists of countries with control of corruption in the second quartile, while countries with control of corruption in the 1st, 3rd or 4th quartile are in the control group. See below for details.

<sup>14</sup> See, for example, Rieffel and Fox (2008) on budget cuts and delayed disbursements of MCC aid.

<sup>15</sup> Eligibility was extended in 2006 to lower–middle income countries with a per-capita income of up to US\$ 3255. Countries subject to legal provisions that prohibit them from receiving United States economic assistance are excluded. See Appendix A in Öhler et al. (2011) for the list of eligible countries.

Combining the before–after approach and the with–without approach has considerable merit for alleviating the problem of drawing correct inferences regarding the MCC Effect (Johnson and Zajonc, 2006). By applying the difference-in-difference-in-differences estimator to the levels of perceived corruption, identification is based on the change in the perceived corruption differentials between the treatment group and the control group that occurred between the periods before and after the announcement of the MCC. Formally, the DDD estimator for our base specification amounts to

$$DDD = ((Corr_{2004}^T - Corr_{2002}^T) - (Corr_{2002}^T - Corr_{2000}^T)) - ((Corr_{2004}^C - Corr_{2002}^C) - (Corr_{2002}^C - Corr_{2000}^C)) \quad (1)$$

with *Corr* being the level of perceived corruption in treatment group *T* and control group *C*, respectively, in the years indicated. For a start, we choose the periods 2004–2002 and 2002–2000 as the MCC was announced in 2002 and became operational in 2004. However, we consider alternative periods below in order to test the hypotheses introduced before.

The standard errors and *t*-statistics come from a regression where the change in corruption in the different periods is specified as the dependent variable. Dummies for the treatment group and the second period (2004–2002) and an interaction term between the two dummies are included as independent variables. The coefficient on the interaction term corresponds to the DDD estimate. Formally, the regression is as follows:

$$\text{Change in corruption} = \alpha + \beta \text{Treat} + \gamma \text{2nd Period} + \delta (\text{Treat} * \text{2nd Period}) + \varepsilon \quad (2)$$

Most importantly, we have to decide how to allocate MCC candidate countries to the treatment and control groups. Our baseline estimation assumes, in line with the reasoning in Section 2, that only countries that do not lie too far below the median with respect to the control of corruption indicator have an incentive to improve on that indicator in order to become eligible for MCC aid. Countries below, but relatively close to the median (second quartile) thus qualify for our treatment group. The underlying argument is that the costs of compliance are relatively low for this quartile, while the prospect of being rewarded by MCC aid is relatively favorable. All other countries qualify for the control group in our baseline estimation, either because their position far below the median (first quartile) implies high costs of compliance or because their position above the median (third and fourth quartiles) does not preclude them from being eligible for MCC aid.<sup>16</sup> In additional estimations, we redefine the treatment and control groups to account for more complex incentive effects. We introduce these alternatives in the next section, following the presentation of baseline results.

## 5. Results

### 5.1. Baseline results

The upper panel of Table 1 presents the difference (*D*) in the levels of the control of corruption index over the two-year periods in our sample, averaged for the treatment and control groups, respectively. We include countries in the treatment group if their level of corruption is in the second quartile and in the control group otherwise. This split places 14 countries in the treatment group and 48 countries in the control group.<sup>17</sup> Countries' positions on other performance indicators are disregarded in Table 1, but are introduced further below.

The difference (*D*) results reported in panel (1) show that, prior to the announcement of the MCC (2002–2000), control of corruption decreased in the treatment group (i.e., perceived corruption became more pervasive), while it increased in the control group. This pattern is reversed in the 2004–2002 period, with an increase in control of corruption in the treatment group and a decrease in the control group.

The second panel in Table 1 reports the DDD results for Eq. (1) above. The DDD is positive and significant at the one percent level, implying that the countries in the treatment group did indeed react to the incentives offered by getting access to the MCC. In other words, it appears that the MCC had positive incentive effects even before becoming operational, probably because prospective candidate countries anticipated that the proposal by President Bush would reward reform efforts with considerable amounts of additional aid.<sup>18</sup>

The third panel of Table 1 slightly modifies Eq. (1), changing the periods employed for comparison. Instead of capturing anticipation effects (2004–2002), we now consider the first two years of MCC operations (2006–2004). The DDD is significantly positive at the five percent level, but substantially smaller in size. This weakening of the MCC Effect may be somewhat surprising. It becomes fairly plausible, however, once the “rough start” of the MCC is taken into account (Rieffel and Fox, 2008: 6 and 7). It took longer than expected for President Bush to sign the MCC-related legislation, and funding proposals met with “resistance from the Congress immediately.” The first compact (with Madagascar) only came into effect in July 2005. Consequently, the prospect that reform efforts would be rewarded with additional US aid might have been diluted.

Comparing the 2008–2006 period with the 2002–2000 period (panel 4), the DDD is significant at the ten percent level only and its size is reduced further. Arguably, ongoing MCC operations led to increasing uncertainty about the timeliness and the amount of expected aid rewards. The number of signed compacts remained fairly small. It became increasingly

<sup>16</sup> The allocation of the MCC candidate countries to the treatment and control groups is based on the control of corruption indicator of the year 2002 (reported in the scorebook of 2004).

<sup>17</sup> See Appendix C in Öhler et al. (2011) for the countries included in the treatment and control groups according to the different variants of our DDD analysis.

<sup>18</sup> Note that the Bush administration announced the 16 performance indicators already at the end of 2002.

**Table 1**  
MCC eligibility and control of corruption (2nd quartile).

| (1)         |              |       |            |       |     |         |
|-------------|--------------|-------|------------|-------|-----|---------|
| D Treatment |              |       |            |       |     |         |
| Period      | 02–00        | 04–02 | 06–04      | 08–06 |     |         |
| Mean        | –0.12        | 0.09  | 0.06       | 0.02  |     |         |
| SE          | 0.05         | 0.08  | 0.04       | 0.04  |     |         |
| SD          | 0.20         | 0.28  | 0.16       | 0.14  |     |         |
| Countries   | 14           | 14    | 14         | 14    |     |         |
| D Control   |              |       |            |       |     |         |
| Period      | 02–00        | 04–02 | 06–04      | 08–06 |     |         |
| Mean        | 0.04         | –0.03 | 0.02       | 0.03  |     |         |
| SE          | 0.03         | 0.03  | 0.04       | 0.02  |     |         |
| SD          | 0.20         | 0.23  | 0.25       | 0.15  |     |         |
| Countries   | 48           | 48    | 48         | 48    |     |         |
| Period      | DD Treatment |       | DD Control |       | DDD | t-value |
| (2)         |              |       |            |       |     |         |
| 2004–2002   | Mean         | 0.21  | –0.06      | 0.27  |     | 2.84*** |
| 2002–2000   | SE           | 0.08  | 0.05       | 0.09  |     |         |
| (3)         |              |       |            |       |     |         |
| 2006–2004   | Mean         | 0.18  | –0.02      | 0.20  |     | 2.12**  |
| 2002–2000   | SE           | 0.08  | 0.04       | 0.09  |     |         |
| (4)         |              |       |            |       |     |         |
| 2008–2006   | Mean         | 0.13  | –0.01      | 0.14  |     | 1.87*   |
| 2002–2000   | SE           | 0.07  | 0.04       | 0.08  |     |         |
| (5)         |              |       |            |       |     |         |
| 2008–2004   | Mean         | 0.11  | 0.04       | 0.07  |     | 0.60    |
| 2004–2000   | SE           | 0.10  | 0.06       | 0.12  |     |         |
| (6)         |              |       |            |       |     |         |
| 2006–2002   | Mean         | 0.31  | –0.05      | 0.36  |     | 2.54**  |
| 2002–1998   | SE           | 0.12  | 0.07       | 0.14  |     |         |

Notes: Treatment group: corruption 2nd quartile; control group: corruption 1st, 3rd or 4th quartile. Panel 1: difference in the level of corruption for the treatment and control groups (D Treatment and D Control, respectively) in the periods indicated in the head row. Panels 2–6: DD represents the difference between the D's for the two periods indicated in the front column; DDD represents the difference between the corresponding DD's for the treatment and control groups; for details, see Eqs. (1) and (2) in the text. Panels 2–4 compare two-year periods, while panels 5 and 6 compare four-year periods; \* (\*\*,\*\*\*): significant at the ten (five, one) percent level.

obvious that the actual MCC budget would persistently fall short of the originally proposed additional US aid of US\$ 5 billion per annum.<sup>19</sup> At the same time, the MCC “has been extraordinarily slow in disbursing the sizeable amount of funding appropriated to it, raising questions about the efficacy of this new model of performance and ownership-based aid giving” (Lancaster, 2008: 8). Sustained reform efforts were further eroded by rumors that the MCC might not survive as a distinct aid agency under the Obama administration.<sup>20</sup> In total, our results imply that the effect was strongest directly after the announcement of the MCC and decreased in the following years of operations.<sup>21</sup>

In the remaining panels of Table 1, we compare four-year periods in order to assess whether the findings reported so far are sensitive to the choice of two-year periods for the DDD calculations. Panel (5) separates the years 2008–2004 and 2004–2000. By taking 2004 as cut-off, this specification disregards any anticipation effects and focuses on possible incentive effects during the full period of MCC operations for which data are available. The DDD is not significant at conventional levels. This is no longer surprising when recalling that we found strong anticipation effects before, while enthusiasm among candidate countries about prospective rewards appears to have cooled progressively in the course of MCC operations.

The strongly positive incentive effects found in panel (6) fit perfectly into this pattern. By comparing the years 2006–2002 and 2002–1998 we capture anticipation effects as well as the effects during the first phase of MCC operations. Recent years are excluded during which the sustainability of reform efforts has been eroded by the uncertainty of rewards. The quantitative difference given in panel (6) is relevant. It is, e.g., similar to the difference in 2007 between Uganda

<sup>19</sup> The budget volume for 2010 amounts to only US\$ 1.1 billion (<http://www.mcc.gov/mcc/press/releases/release-uscongressapproves-121309.shtml>).

<sup>20</sup> For instance, Rieffel and Fox (2008: 11) speculate that the MCC “is still ‘small potatoes’” that may be moved into “a beefed up USAID.”

<sup>21</sup> One might worry that mean reversion may drive the results. That is because treatment and control group are not balanced and symmetric: The countries in the treatment group all start below the median while two thirds of the countries in the control group start above the median. We gauge the relevance of this issue by estimating an alternative specification: The control of corruption indicator of 2000 was used to define the treatment and control groups with 2000–1998 being the ex ante period and 2002–2000 the ex post period. Since we do not find any effect with this specification, mean reversion is unlikely to be a problem.

(which is in our treatment group) and Nigeria (which is not). Of all the countries for which the control of corruption indicator is calculated, only 20 improved their indicator by at least this difference over the 2002–2006 period.

As the treatment group comprises only 14 countries, a visible presentation may be useful to identify potentially influential observations. Fig. 1 shows the change in the control of corruption of each candidate country prior to the announcement of the MCC (2002–1998) and after the announcement of the MCC (2006–2002). Countries between the two vertical lines correspond to the treatment group (second quartile). The other countries constitute the control group. Looking at the treatment group, the Solomon Islands (SLB) seems to be influential, with a decrease in control of corruption of 0.74 between 1998 and 2002 and an increase of 1.07 between 2002 and 2006. The DDD decreases to 0.24 if this country is excluded from the estimation. Importantly, the effect remains significant, though at the ten percent level only.

The estimations shown in Table 1 are based on the composite index values of control of corruption as reported in the 2009 release of Kaufmann et al.'s data. We used both an earlier and a more recent release of this database (2008 and 2010, respectively) in order to assess whether the findings reported so far are sensitive to changes in the composition of Kaufmann et al.'s index.<sup>22</sup> This does not appear to be the case. Rather, the results based on the earlier and most recent composition of the index reveal essentially the same pattern as before.<sup>23</sup> In particular, the evidence on the strong anticipation effects is corroborated for all releases of the index. In another (unreported) robustness test of our baseline results, we excluded all countries for which the overall index on control of corruption relied on less than three sources. This applied to seven (mostly small) countries in 1998, the beginning of our period of observation.<sup>24</sup> The precision of the index is probably weakest for these countries. Moreover, compositional changes are most likely to matter when the number of underlying sources increases considerably over time.<sup>25</sup> Yet excluding these countries affects our results only modestly. The size of the DDDs and significance levels are somewhat lower, but the general pattern of effects remains the same.

## 5.2. Refined treatment groups

It has so far been assumed that the incentive effects of the MCC are restricted to candidate countries that do not lie too far below the median score (second quartile) for the control of corruption indicator. The allocation to the treatment and control group was based on the assumption that candidate countries far below the median (first quartile) had no reasonable chance to get their reform efforts rewarded by MCC aid and, thus, would not increase attempts to fight corruption effectively. We now test whether this assumption is valid.

In Table 2, we perform estimations with the treatment group comprising candidate countries falling into the first quartile rather than the second quartile, and the control group comprising candidate countries above the median. In this way, we can test directly whether candidate countries with unfavorable initial conditions nevertheless embark on reforms as a response to the announcement of the MCC. Our results clearly contradict such a proposition. The fairly remote chances that candidate countries falling far below the median score will receive MCC aid appear to have eroded the incentives to fight corruption. In most panels of Table 2, the DDDs are not significant at conventional levels. The only exception is the comparison between the 2002–1998 and 2006–2002 periods, where the difference is marginally significant. This marginal significance is driven by one influential observation, Georgia. Georgia is at the border to the second quartile, and substantially reduced corruption over the relevant periods. When we exclude it from the treatment group, the DDD in Panel (6) of Table 2 is no longer significant. Note that the (previously strong) anticipation effects are no longer observed when initial conditions rendered it unlikely to have reform efforts rewarded through MCC aid.

Table 3 modifies the analysis of Table 1 to make the results strictly comparable with those reported for the first quartile in Table 2. The treatment group includes those countries with control of corruption in the second quartile (as in Table 1), but the control group is restricted to countries with control of corruption above the median (as for the first quartile in Table 2). As can be seen, the results are very similar to those in Table 1. In other words, the incentive effects of the MCC differ considerably depending on initial conditions; the effect is strong under favorable conditions and at best weak otherwise.

In the next step, we take into account that candidate countries scoring just slightly above the median (third quartile) may also have an incentive to control corruption more effectively. Even though eligibility for MCC aid is based on 2004 scores, candidates in the third quarter may have to “defend” their position by further reform efforts relative to reformers in the second quartile in order to retain eligibility in the future. Therefore, we enlarge the treatment group in Table 4, including all candidates in the second and third quartiles (with the first and fourth quartiles representing the control group). This split places a roughly equal number of countries in the treatment group and the control group.<sup>26</sup>

<sup>22</sup> The sources used in constructing the index do not only vary from country to country; they may also change over time for one particular country. Kaufmann et al. (2009: 21) address this point by assessing the extent to which changes in the composite index were driven by new data sources. They find that “compositional effects are not large” during the period of observation underlying our analysis (1998–2008).

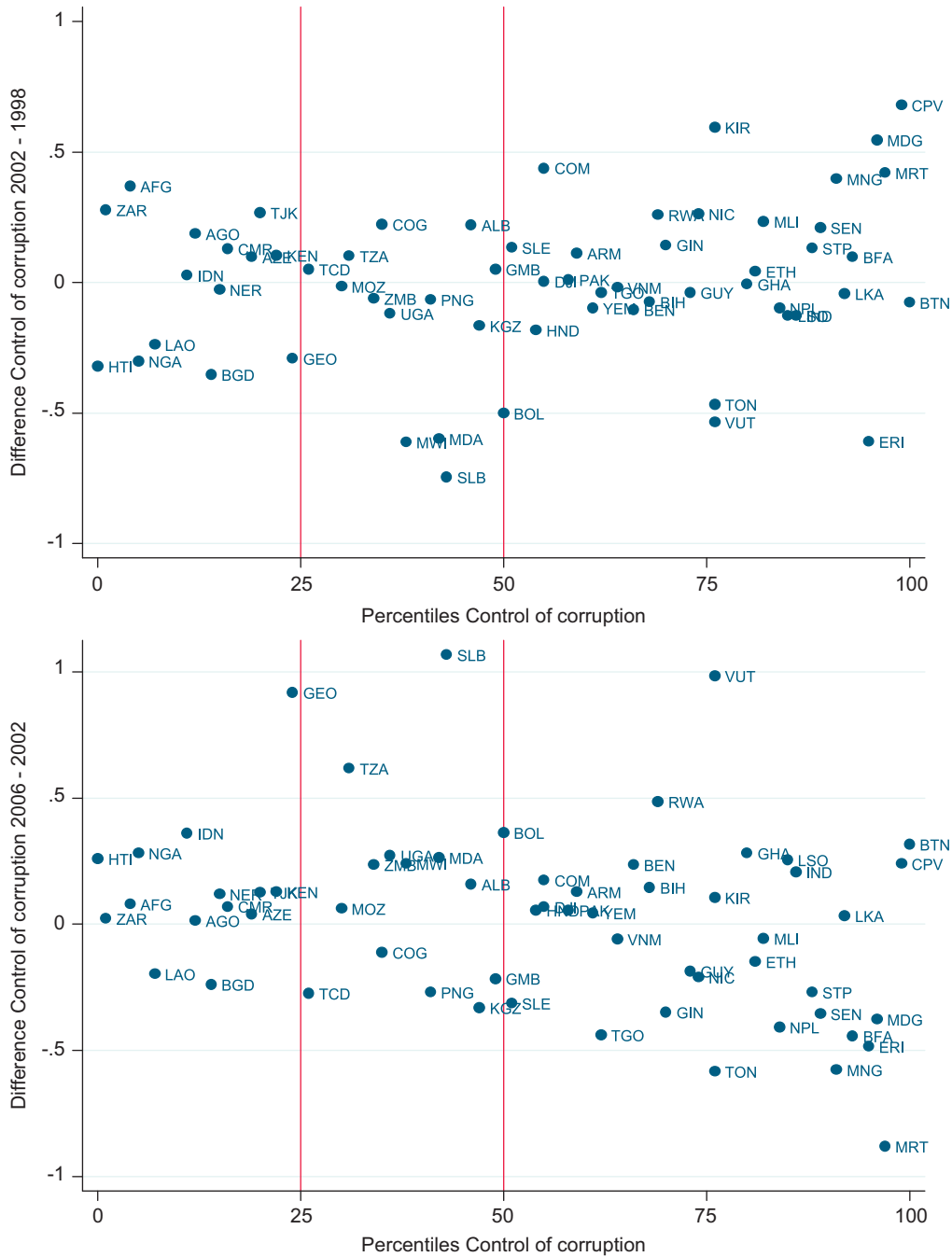
<sup>23</sup> It should be noted that Panels (4) and (5) could not be replicated for the earlier release of the database as the data for 2008 were not yet available at that time. For the sake of brevity, the estimations based on the index values released in 2008 and 2010 are not shown in detail. However, the additional tables corresponding to Table 1 are available on request.

<sup>24</sup> Afghanistan, Bhutan, Cape Verde, Kiribati, Solomon Islands, Tonga, and Vanuatu.

<sup>25</sup> In the case of Afghanistan, the number of sources increased from just one in 1998 to nine in 2008.

<sup>26</sup> Note that the MCC, when calculating the country scores on the different indicators, includes countries that would be considered candidate countries but are subject to legal provisions that prohibit them from receiving US economic assistance. However, these countries are excluded in the present analysis. This explains why the treatment and control groups do not precisely add up to the same number of countries in Table 4.





**Fig. 1.** Change in Control of corruption before and after the announcement of MCC. Notes: Shows the change in control of corruption of each candidate country prior to the announcement of the MCC (2002–1998) and after the announcement of the MCC (2006–2002). Countries between the two vertical lines correspond to the treatment group (second quartile). The other countries constitute the control group.

With the inclusion of Georgia in the control group, none of the DDDs are significant at conventional levels (not shown in table). Given Georgia’s position at the border to the second quartile, we also tested for differences including it in the treatment group, as shown in Table 4. Overall, the effects are considerably weaker than those in Table 1. Nevertheless, there are several similarities. The comparison between the 2006–2002 and 2002–1998 periods in panel (6) again shows the strongest effect, significant at the ten percent level. As before, we find a positive effect – also significant at the ten percent level – when comparing the years 2004–2002 and 2002–2000 in panel (1). On the other three panels, the DDDs are not significant at conventional levels. This underscores the overriding role of anticipation effects in MCC candidate countries; MCC effects were increasingly eroded over the time of actual operations.

**Table 2**  
MCC eligibility and control of corruption (1st quartile).

| (1)         |       |              |            |       |         |
|-------------|-------|--------------|------------|-------|---------|
| D Treatment |       |              |            |       |         |
| Period      | 02–00 | 04–02        | 06–04      | 08–06 |         |
| Mean        | 0.00  | 0.05         | 0.09       | 0.05  |         |
| SE          | 0.05  | 0.06         | 0.03       | 0.04  |         |
| SD          | 0.20  | 0.22         | 0.13       | 0.15  |         |
| Countries   | 14    | 14           | 14         | 14    |         |
| D Control   |       |              |            |       |         |
| Period      | 02–00 | 04–02        | 06–04      | 08–06 |         |
| Mean        | 0.05  | –0.06        | –0.01      | 0.02  |         |
| SE          | 0.03  | 0.04         | 0.05       | 0.03  |         |
| SD          | 0.20  | 0.23         | 0.28       | 0.16  |         |
| Countries   | 34    | 34           | 34         | 34    |         |
| Period      |       | DD Treatment | DD Control | DDD   | t-value |
| (2)         |       |              |            |       |         |
| 2004–2002   | Mean  | 0.05         | –0.11      | 0.16  | 1.65    |
| 2002–2000   | SE    | 0.08         | 0.05       | 0.10  |         |
| (3)         |       |              |            |       |         |
| 2006–2004   | Mean  | 0.08         | –0.06      | 0.14  | 1.41    |
| 2002–2000   | SE    | 0.09         | 0.05       | 0.10  |         |
| (4)         |       |              |            |       |         |
| 2008–2006   | Mean  | 0.05         | –0.03      | 0.08  | 1.01    |
| 2002–2000   | SE    | 0.07         | 0.04       | 0.08  |         |
| (5)         |       |              |            |       |         |
| 2008–2004   | Mean  | 0.08         | 0.02       | 0.07  | 0.53    |
| 2004–2000   | SE    | 0.11         | 0.07       | 0.13  |         |
| (6)         |       |              |            |       |         |
| 2006–2002   | Mean  | 0.15         | –0.13      | 0.28  | 1.96*   |
| 2002–1998   | SE    | 0.12         | 0.08       | 0.14  |         |

Notes: Treatment group: corruption 1st quartile; control group: corruption above the median. Panel 1: difference in the level of corruption for the treatment and control groups (D Treatment and D Control, respectively) in the periods indicated in the head row. Panels 2–6: DD represents the difference between the D's for the two periods indicated in the front column; DDD represents the difference between the corresponding DD's for the treatment and control groups; for details, see Eqs. (1) and (2) in the text. Panels 2–4 compare two-year periods, while panels 5 and 6 compare four-year periods; \* (\*\*, \*\*\*): significant at the ten (five, one) percent level.

When comparing the significant DDDs in Table 4 with those in Table 1, we find that they have almost halved. This seems to imply that it is mainly countries below the median that have fought corruption in response to the announcement of the MCC, while countries above the median have not been incentivized to control corruption more effectively in order to defend their favorable positions. It cannot be ruled out that countries in the third quartile were simply myopic and unaware of the risk of losing eligibility once reformers that previously scored below them climbed up in the ranking. More plausibly, however, it may be argued that their reluctance to step up efforts toward controlling corruption was rational. Performance-based selection of eligible countries notwithstanding, it was open to question whether the MCC would be equally strict in suspending eligibility once it had been granted.<sup>27</sup> According to the MCC's Policy on Suspension and Termination, its CEO “may make a recommendation to the Board” to suspend eligibility if the country has “engaged in a pattern of actions inconsistent with selection criteria.”<sup>28</sup> In a public outreach meeting in early 2007, the (former) CEO of the MCC, John Danilovich, stated<sup>29</sup>:

I sent remediation letters to our partner countries that exhibited certain slippages in our criteria in the areas of ruling justly, investing in people and economic freedom. We are constructively and continuously engaging our Compact-eligible countries to help them create and implement a corrective plan of action to address these areas of slippage.

This suggests that, in contrast to performance-based selection, the credibility of the MCC in suspending eligibility may be no greater than the credibility of traditional aid agencies in imposing sanctions against countries violating previously agreed-upon conditions for aid. This would explain why the MCC Effect weakened considerably once countries had passed the eligibility criterion of scoring above the median with regard to the control of corruption.

<sup>27</sup> Yemen appears to have been the first country (among very few cases up to now) that was suspended from MCC funding in late 2005, “due to policy slippage on a number of MCC's selection indicators” (<http://www.mcc.gov/mcc/bm.doc/mcc-workingpaper-corruption.pdf>).

<sup>28</sup> See: <http://www.mcc.gov/mcc/bm.doc/07-suspensionandterminationpolicy.pdf>; emphasis added.

<sup>29</sup> See: <http://www.mcc.gov/mcc/bm.doc/transcript-021507-publicoutreach.pdf>.

**Table 3**  
MCC eligibility and control of corruption (2nd quartile).

| (1)         |              |       |            |       |     |         |
|-------------|--------------|-------|------------|-------|-----|---------|
| D Treatment |              |       |            |       |     |         |
| Period      | 02–00        | 04–02 | 06–04      | 08–06 |     |         |
| Mean        | –0.12        | 0.09  | 0.06       | 0.02  |     |         |
| SE          | 0.05         | 0.08  | 0.04       | 0.04  |     |         |
| SD          | 0.20         | 0.28  | 0.16       | 0.14  |     |         |
| Countries   | 14           | 14    | 14         | 14    |     |         |
| D Control   |              |       |            |       |     |         |
| Period      | 02–00        | 04–02 | 06–04      | 08–06 |     |         |
| Mean        | 0.05         | –0.06 | –0.01      | 0.02  |     |         |
| SE          | 0.03         | 0.04  | 0.05       | 0.03  |     |         |
| SD          | 0.20         | 0.23  | 0.28       | 0.16  |     |         |
| Countries   | 34           | 34    | 34         | 34    |     |         |
| Period      | DD Treatment |       | DD Control |       | DDD | t-value |
| (2)         |              |       |            |       |     |         |
| 2004–2002   | Mean         | 0.21  | –0.11      | 0.32  |     | 3.16*** |
| 2002–2000   | SE           | 0.08  | 0.05       | 0.10  |     |         |
| (3)         |              |       |            |       |     |         |
| 2006–2004   | Mean         | 0.18  | –0.06      | 0.24  |     | 2.33**  |
| 2002–2000   | SE           | 0.09  | 0.06       | 0.10  |     |         |
| (4)         |              |       |            |       |     |         |
| 2008–2006   | Mean         | 0.13  | –0.03      | 0.17  |     | 2.07**  |
| 2002–2000   | SE           | 0.07  | 0.04       | 0.08  |     |         |
| (5)         |              |       |            |       |     |         |
| 2008–2004   | Mean         | 0.11  | 0.02       | 0.09  |     | 0.69    |
| 2004–2000   | SE           | 0.11  | 0.07       | 0.13  |     |         |
| (6)         |              |       |            |       |     |         |
| 2006–2002   | Mean         | 0.31  | –0.13      | 0.44  |     | 2.88*** |
| 2002–1998   | SE           | 0.13  | 0.08       | 0.15  |     |         |

Notes: Treatment group: corruption 2nd quartile; control group: corruption above the median. Panel 1: difference in the level of corruption for the treatment and control groups (D Treatment and D Control, respectively) in the periods indicated in the head row. Panels 2–6: DD represents the difference between the D's for the two periods indicated in the front column; DDD represents the difference between the corresponding DD's for the treatment and control groups; for details, see Eqs. (1) and (2) in the text. Panels 2–4 compare two-year periods, while panels 5 and 6 compare four-year periods; \* (\*\*,\*\*\*): significant at the ten (five, one) percent level.

### 5.3. Complex incentive structure

The separation of treatment and control groups has so far been based exclusively on the control of corruption. However, the incentives for candidate countries to fight corruption may also depend on where a country stands with respect to the other eligibility criteria used by the MCC. Deviation from other requirements can be expected to weaken the incentive to fight corruption, even when a country is close to the median score on the control of corruption indicator. The costs of compliance would clearly be higher for countries that have to reform on various fronts in order to become eligible. At the same time, the prospect of being rewarded with MCC aid would be fairly remote. In the following, we take this into account by refining our classification of the treatment and control groups.

In Table 5 the treatment group consists of 10 countries that meet two requirements simultaneously. As before in Table 1, these countries score in the second quartile with respect to the control of corruption. In addition, they score in the second, third or fourth quartile with respect to each of the three broad MCC categories of Ruling Justly, Investing in People and Economic Freedom. The average scores for these broad categories are calculated according to the aforementioned MCC rule that, in order to become eligible, a country has to score higher than the median on at least half the eligibility criteria in each category. A country attempting to obey by this rule at minimal costs of compliance would probably focus on those eligibility criteria in each category where it is already performing close to the median level. Consequently, we average the scores for the three indicators on which the country performs best in the categories Ruling Justly and Economic Freedom. The two best indicators are considered in the category Investing in People.<sup>30</sup> Countries falling into the first quartile in any of these three broad categories are included in the control group. The related costs of compliance tend to be highest for these countries and the prospects of receiving aid rewards are most unfavorable. Consequently, their incentives to fight corruption might be particularly weak.

<sup>30</sup> Recall that the overall number of indicators in the category Investing in People is only four. In averaging the scores, all scores above the median of 50 are set equal to 50. This is because it has no implications for the costs of compliance whether a country performs considerably, or only slightly better than the median.

**Table 4**  
MCC eligibility and control of corruption (2nd or 3rd quartile).

| (1)         |       |              |            |       |         |
|-------------|-------|--------------|------------|-------|---------|
| D Treatment |       |              |            |       |         |
| Period      | 02–00 | 04–02        | 06–04      | 08–06 |         |
| Mean        | –0.03 | 0.05         | 0.04       | 0.02  |         |
| SE          | 0.04  | 0.05         | 0.04       | 0.03  |         |
| SD          | 0.20  | 0.25         | 0.19       | 0.14  |         |
| Countries   | 30    | 30           | 30         | 30    |         |
| D Control   |       |              |            |       |         |
| Period      | 02–00 | 04–02        | 06–04      | 08–06 |         |
| Mean        | 0.03  | –0.05        | 0.01       | 0.03  |         |
| SE          | 0.04  | 0.04         | 0.05       | 0.03  |         |
| SD          | 0.21  | 0.24         | 0.27       | 0.16  |         |
| Countries   | 32    | 32           | 32         | 32    |         |
| Period      |       | DD Treatment | DD Control | DDD   | t-value |
| (2)         |       |              |            |       |         |
| 2004–2002   | Mean  | 0.08         | –0.08      | 0.16  | 1.93*   |
| 2002–2000   | SE    | 0.06         | 0.06       | 0.08  |         |
| (3)         |       |              |            |       |         |
| 2006–2004   | Mean  | 0.07         | –0.02      | 0.09  | 1.12    |
| 2002–2000   | SE    | 0.06         | 0.06       | 0.08  |         |
| (4)         |       |              |            |       |         |
| 2008–2006   | Mean  | 0.05         | 0.00       | 0.04  | 0.67    |
| 2002–2000   | SE    | 0.05         | 0.05       | 0.07  |         |
| (5)         |       |              |            |       |         |
| 2008–2004   | Mean  | 0.04         | 0.06       | –0.02 | –0.24   |
| 2004–2000   | SE    | 0.07         | 0.07       | 0.10  |         |
| (6)         |       |              |            |       |         |
| 2006–2002   | Mean  | 0.15         | –0.08      | 0.23  | 1.95*   |
| 2002–1998   | SE    | 0.09         | 0.08       | 0.12  |         |

Notes: Treatment group: corruption 2nd or 3rd quartile; control group: corruption 1st or 4th quartile. Panel 1: difference in the level of corruption for the treatment and control groups (D Treatment and D Control, respectively) in the periods indicated in the head row. Panels 2–6: DD represents the difference between the D's for the two periods indicated in the front column; DDD represents the difference between the corresponding DD's for the treatment and control groups; for details, see Eqs. (1) and (2) in the text. Panels 2–4 compare two-year periods, while panels 5 and 6 compare four-year periods; \* (\*\*, \*\*\*): significant at the ten (five, one) percent level.

Accounting for the more complex incentive structure in Table 5 leads to results that are surprisingly similar to those reported in Table 1 above. Again, the differences (D) in panel (1) show that, prior to the announcement of the MCC, control of corruption decreased in the treatment group, while it increased in the control group. This pattern is again reversed in the 2002–2004 period.

The DDD is positive in all sample periods, though (as before) not significant at conventional levels when comparing the years 2008–2004 and 2004–2000 (panel 5). Table 5 also confirms that the effect is largest when comparing the years 2006–2002 and 2002–1998 (panel 6), followed by the effect shown in panel (1) for the years 2004–2002 vs. 2002–2000. In other words, the dominance of positive anticipation effects is corroborated when refining the classification of treatment and control groups according to the more complex incentive structure. Likewise, taking account of all eligibility criteria confirms the previous finding that the MCC Effect weakened over time, even though this development is less pronounced than in Table 1. Rising uncertainty among candidate countries about the timeliness and amount of aid rewards again appears to have undermined incentive effects.

#### 5.4. Robustness tests

Finally, we perform a range of robustness tests. In particular, we assess whether previous results are sensitive to the classification of MCC candidate countries into treatment and control groups. We systematically combine alternative variants of the two dimensions of this classification used before, namely the cut-off with regard to the control of corruption and the cut-off with regard to the three broad categories Ruling Justly, Investing in People, and Economic Freedom. In terms of the three categories, the treatment group either comprises all countries except those in the first quartile (variant 1), or all those above the median (variant 2). In terms of control of corruption, the treatment group comprises countries scoring (a) in the second quartile (as in Tables 1 and 5), (b) in the second quartile, but with the reduced control group (as in Table 3), or (c) in the second and third quartile (as in Table 4).

**Table 5**  
MCC eligibility and control of corruption (complex incentive structure).

| (1)         |       |              |            |       |         |
|-------------|-------|--------------|------------|-------|---------|
| D Treatment |       |              |            |       |         |
| Period      | 02–00 | 04–02        | 06–04      | 08–06 |         |
| Mean        | –0.12 | 0.06         | 0.07       | 0.05  |         |
| SE          | 0.07  | 0.06         | 0.05       | 0.04  |         |
| SD          | 0.21  | 0.19         | 0.17       | 0.13  |         |
| Countries   | 10    | 10           | 10         | 10    |         |
| D Control   |       |              |            |       |         |
| Period      | 02–00 | 04–02        | 06–04      | 08–06 |         |
| Mean        | 0.02  | –0.01        | 0.02       | 0.02  |         |
| SE          | 0.03  | 0.04         | 0.02       | 0.02  |         |
| SD          | 0.20  | 0.25         | 0.24       | 0.15  |         |
| Countries   | 52    | 52           | 52         | 52    |         |
| Period      |       | DD Treatment | DD Control | DDD   | t-value |
| (2)         |       |              |            |       |         |
| 2004–2002   | Mean  | 0.18         | –0.04      | 0.22  | 1.96*   |
| 2002–2000   | SE    | 0.10         | 0.04       | 0.11  |         |
| (3)         |       |              |            |       |         |
| 2006–2004   | Mean  | 0.19         | 0.00       | 0.19  | 1.82*   |
| 2002–2000   | SE    | 0.10         | 0.04       | 0.11  |         |
| (4)         |       |              |            |       |         |
| 2008–2006   | Mean  | 0.17         | 0.00       | 0.17  | 1.98**  |
| 2002–2000   | SE    | 0.08         | 0.04       | 0.09  |         |
| (5)         |       |              |            |       |         |
| 2008–2004   | Mean  | 0.18         | 0.03       | 0.15  | 1.15    |
| 2004–2000   | SE    | 0.12         | 0.05       | 0.13  |         |
| (6)         |       |              |            |       |         |
| 2006–2002   | Mean  | 0.31         | –0.02      | 0.33  | 2.05**  |
| 2002–1998   | SE    | 0.15         | 0.06       | 0.16  |         |

Notes: Treatment group: corruption 2nd quartile, the three categories 2nd, 3rd, or 4th quartile; control group: corruption 1st, 3rd or 4th quartile or at least one category 1st quartile. Panel 1: difference in the level of corruption for the treatment and control groups (D Treatment and D Control, respectively) in the periods indicated in the head row. Panels 2–6: DD represents the difference between the D's for the two periods indicated in the front column; DDD represents the difference between the corresponding DD's for the treatment and control groups; for details, see Eqs. (1) and (2) in the text. Panels 2–4 compare two-year periods, while panels 5 and 6 compare four-year periods; \* (\*\*,\*\*\*): significant at the ten (five, one) percent level.

Table 6 lists the results starting with variant 1 of the cut-off for the three categories combined with alternative cut-offs for corruption in columns 1b and 1c.<sup>31</sup> Columns 2a to 2c present the results for combinations of variant 2 of the cut-off for the three categories with all three variants of corruption. We consider variant 2 in order to test whether our results are sensitive to the implicit assumption under variant 1 that scoring below the median in any category increases the costs of compliance to the same extent. It is hardly possible to relax this assumption by accounting for varying costs of compliance for each category as these costs cannot be observed. Instead, we require in variant 2 that countries in the treatment group perform better than the median in all three categories and, thus, do not incur additional costs of compliance. This “restrictive” refinement results in relatively small treatment groups whose incentives to fight corruption are not affected by varying costs of compliance across the three categories.

Table 6 only reports the DDDs for the five robustness tests, omitting differences and difference-in-differences. As can be seen, the results are similar throughout, corroborating the major findings reported above. First of all, the comparison of the years 2004–2002 and 2002–2000 reveals significantly positive DDDs in all five columns, at the ten percent level in column 1c and the five percent level at least for the remaining specifications. This underscores the prominence of anticipation effects. It also turns out, as before, that the evidence for *sustained* reform efforts is weaker, as indicated by the comparison of the years 2008–2006 and 2002–2000. The DDD for this comparison is not significant at conventional levels in column 1c and significantly positive only at the ten percent level in column 2c. Likewise, the MCC Effect during the first phase of operations appears to be weak in columns 1c and 2c where the treatment groups include candidates scoring in the second and third quartiles with respect to the control of corruption. This resembles the insignificant incentive effects reported in Table 4 that used the same cut-off for corruption. However, the MCC Effect during the first phase of operations turns out to be slightly stronger than the anticipation effect in columns 2a and 2b.

Our previous results based on four-year periods for the DDD calculations, instead of two-year periods, are hardly affected by the redefinition of our treatment and control groups in Table 6. The DDD is largest throughout, and significant

<sup>31</sup> Note that the combination of variant 1 of the three categories and variant *a* of corruption is the same as that presented in Table 5 above.

**Table 6**  
Tests for robustness.

| Period    |      | (1b) |          | (1c) |          | (2a) |          | (2b) |          | (2c) |          |
|-----------|------|------|----------|------|----------|------|----------|------|----------|------|----------|
|           |      | DDD  | <i>t</i> | DDD  | <i>t</i> | DDD  | <i>t</i> | DDD  | <i>t</i> | DDD  | <i>t</i> |
| (1)       |      |      |          |      |          |      |          |      |          |      |          |
| 2004–2002 | Mean | 0.25 | 2.28**   | 0.16 | 1.94*    | 0.31 | 2.45**   | 0.33 | 2.65***  | 0.25 | 2.71***  |
| 2002–2000 | SE   | 0.11 |          | 0.08 |          | 0.13 |          | 0.12 |          | 0.09 |          |
| (2)       |      |      |          |      |          |      |          |      |          |      |          |
| 2006–2004 | Mean | 0.22 | 1.98**   | 0.09 | 1.07     | 0.33 | 2.65***  | 0.35 | 2.81***  | 0.16 | 1.71*    |
| 2002–2000 | SE   | 0.11 |          | 0.08 |          | 0.12 |          | 0.12 |          | 0.09 |          |
| (3)       |      |      |          |      |          |      |          |      |          |      |          |
| 2008–2006 | Mean | 0.19 | 2.16**   | 0.11 | 1.60     | 0.26 | 2.58**   | 0.27 | 2.72***  | 0.14 | 1.80*    |
| 2002–2000 | SE   | 0.09 |          | 0.07 |          | 0.10 |          | 0.10 |          | 0.08 |          |
| (4)       |      |      |          |      |          |      |          |      |          |      |          |
| 2008–2004 | Mean | 0.16 | 1.21     | 0.03 | 0.33     | 0.28 | 1.80*    | 0.29 | 1.92*    | 0.04 | 0.36     |
| 2004–2000 | SE   | 0.14 |          | 0.10 |          | 0.15 |          | 0.15 |          | 0.11 |          |
| (5)       |      |      |          |      |          |      |          |      |          |      |          |
| 2006–2002 | Mean | 0.37 | 2.31**   | 0.27 | 2.15**   | 0.46 | 2.46**   | 0.49 | 2.70***  | 0.35 | 2.53**   |
| 2002–1998 | SE   | 0.16 |          | 0.12 |          | 0.19 |          | 0.18 |          | 0.14 |          |

Notes: (1b) Treatment group: corruption in the 2nd quartile and not in the 1st quartile on any category; control group: corruption above the median or at least one category in 1st quartile. (1c) Treatment group: 2nd or 3rd quartile of corruption and not in the 1st quartile on any category; control group: corruption 1st or 4th quartile or at least one category in the 1st quartile. (2a) Treatment group: corruption in the 2nd quartile and above the median on the three categories; control group: corruption 1st, 3rd or 4th quartile or at least one category below median. (2b) Treatment group: corruption in the 2nd quartile and above the median on the three categories; control group: corruption above the median or at least one category below median. (2c) Treatment group: 2nd or 3rd quartile of corruption and above the median on the three categories; control group: corruption 1st or 4th quartile or at least one category below median. DDD represents the difference between the corresponding DD's for the treatment and control groups; for details, see Eqs. (1) and (2) in the text. Panels 1–3 compare two-year periods, while panels 4 and 5 compare four-year periods;  $t=t$ -value; \* (\*\*,\*\*\*): significant at the ten (five, one) percent level.

at the five percent level at least, when we compare the years 2006–2002 and 2002–1998 which attempts to capture anticipation effects together with the MCC effects during the first phase of operations. In sharp contrast, the DDDs become substantially smaller in size when comparing the four years of MCC operations with the four years before the MCC commenced operations, with three DDDs being completely insignificant.

Interestingly, the DDDs are generally larger in columns 2a and 2b where the definition of the treatment group applies variant 2 of the cut-off for the three broad categories. As noted above, this definition is relatively rigorous and results in comparatively small treatment groups. It is plausible that larger DDDs result from relegating those countries which fall below the median in at least one category to the control group. These countries have weaker incentives to fight corruption effectively as the costs of compliance tend to be higher if reforms are required on several fronts. However, applying variant 2 of the cut-off for the three broad categories yields rather ambiguous results (with most DDDs being smaller compared to column 1b) when treatment in terms of corruption includes candidate countries from the third quartile (column 2c). This corroborates the previous argument that the MCC Effect might weaken considerably once countries have been selected as eligible as the threat to suspend eligibility in the case of policy slippage does not appear particularly credible.

## 6. Summary and conclusion

Performance-based aid has been proposed as an alternative to the failed traditional approach of donors making aid conditional on reform promises of recipient countries. However, empirical evidence on the effectiveness of ex post rewards of reforms hardly exists. We attempt to fill this gap by investigating the so-called MCC Effect. The Millennium Challenge Corporation was explicitly established as a new US aid agency to pursue an innovative approach to the allocation of aid. Strict selectivity in granting aid exclusively to needy and well performing recipients was expected to strengthen the reform-mindedness of possible recipient countries.

The analysis in this paper has focused on the control of corruption – an aspect of reform that commands widespread support in the aid community and features most prominently in the MCC's aid eligibility criteria. Employing a difference-in-difference-in-differences (DDD) approach, we find that the MCC was successful in promoting better control of corruption. Candidate countries that had reasonably good chances of gaining access to the MCC (as they scored relatively close to the selection criteria thresholds) fought corruption more effectively than other candidate countries. The impact of the MCC on corruption is not only statistically significant, but also of quantitative importance. The results reported in panel 6 of Table 1, e.g., imply an effect of the MCC amounting to the difference in corruption (in 2007) between Uganda (which is in our treatment group) and Nigeria (which is not). Of all the countries for which the control of corruption indicator is calculated, only 20 improved their indicator by at least this difference over the 2002–2006 period. This finding

suggests that performance-based aid may succeed in promoting reforms, in contrast to aid conditioned on recipients' promises to reform.

However, the MCC experience also reveals that the incentive effects of performance-based aid are rather weak, or even absent, under several circumstances. While we find surprisingly strong anticipation effects, i.e., candidates fighting corruption even before MCC operations started, the incentive effects weakened over time. Presumably, this was mainly because it became increasingly uncertain whether the timeliness and magnitude of the MCC's aid rewards would sufficiently compensate for the costs of reform efforts. In other words, sustained reform efforts may suffer not only from insufficient willingness on the part of aid recipients, but also from broken promises and delayed aid on the part of donors.

Even if the rewards for reform could be taken for granted, the incentive effects appear to be restricted to a sub-group of candidate countries. According to our results, the MCC failed to give any impetus to the fight against corruption in candidate countries that scored far below the eligibility threshold. Economically, this finding is plainly intuitive: These countries had only remote chances to receive MCC aid as a reward for reforms, while the costs of complying with MCC conditions were comparatively high. The implication is rather troubling, however, as the MCC concept of performance-based aid tends to fail exactly where the need for reforms is most urgent.

Arguably, the MCC concept should be adapted to strengthen the incentive effects for countries where starting conditions are unfavorable. Rather than offering fruits that simply hang too high, the thresholds for eligibility could be designed relative to a country's own performance. A country could be eligible, e.g., if it improved its control of corruption score by a predefined extent over a specific period of time. This modified approach could also take into account that the costs of corruption to a country do not arise from its relative position compared to other countries, but from absolute corruption levels. The modified approach would thus help reduce corruption in those countries suffering from it the most.

Another limitation of the MCC concept relates to those candidate countries that have already passed the eligibility criteria. As long as they are not too far above the threshold, one would expect that, by judging performance relative to the median, the MCC would provide incentives for these candidates to further improve their score on the control of corruption. Our results suggest, however, that candidates having passed the threshold did not intensify reform efforts in an attempt to defend their position against reformers trying to become eligible. More research is necessary in order to identify more clearly the reasons behind this apparently irrational, or at least myopic, behavior.

As it seems, the MCC concept is flawed in its ability to encourage continued reform efforts once a country has passed the gate-keeping stage of becoming eligible. In contrast to strict selectivity in defining eligible countries, the MCC's credibility in suspending eligibility does not appear to be stronger than the credibility of international financial institutions, notably the IMF, and traditional aid agencies in imposing sanctions against countries that violate previously agreed-upon conditions for adjustment loans and development aid. The conclusion is obvious, but may be hard to enforce: Strict selectivity must be complemented with automatic suspension for performance-based aid to have better chances at succeeding across the board.

## Acknowledgments

We thank Michaela Rank for excellent research assistance, Hendrik van Broekhuizen for careful proof-reading, and seminar participants at the University of Goettingen, the 4th Workshop on Political Economy (Dresden 2010), the International Studies Association Annual Convention (Montreal 2011), the Spring Meeting of Young Economists (Groningen 2011), and the European Public Choice Society Meeting (Rennes 2011) for helpful comments. Financial support from the German Research Foundation (DFG GZ: DR 640/2-1) is gratefully acknowledged.

## References

- Alesina, A., Weder, B., 2002. Do corrupt governments receive less foreign aid? *American Economic Review* 92 (4), 1126–1137.
- Burnside, C., Dollar, D., 2000. Aid, policies, and growth. *American Economic Review* 90 (4), 847–868.
- Collier, P., 1997. The failure of conditionality. In: Gwin, C., Nelson, J.M. (Eds.), *Perspectives on Aid and Development*. John Hopkins University Press, Washington, DC, pp. 51–77.
- Collier, P., Guillaumeont, P., Guillaumeont, S., Gunning, J.W., 1997. Redesigning conditionality. *World Development* 25 (9), 1399–1407.
- Cremer, G., 2008. *Corruption and Development Aid: Confronting the Challenges*. Lynne Rienner Publishers, Boulder.
- Doucoulagos, H., Paldam, M., 2010. Conditional aid effectiveness: a meta study. *Journal of International Development* 22 (4), 391–410.
- Drazen, A., 2002. Conditionality and ownership in IMF lending: a political economy approach. *IMF Staff Papers* 49 (special issue), pp. 36–67.
- Dreher, A., 2009. IMF conditionality: theory and evidence. *Public Choice* 141 (1/2), 233–267.
- Dreher, A., Nunnenkamp, P., Öhler, H., 2010. Why it pays for aid recipients to take note of the Millennium Challenge Corporation: other donors do! Kiel Working Paper 1609, Kiel Institute for the World Economy.
- Easterly, W., 2007. Are aid agencies improving? *Economic Policy* 22 (52), 633–678.
- Eichler, R., Glassman, A., 2008. Health systems strengthening via performance-based aid: creating incentives to perform and to measure results. *Brookings Global Economy and Development Working Paper* 3.
- Grabbe, H., 2001. How does Europeanization affect CEE governance? Conditionality, diffusion and diversity. *Journal of European Public Policy* 8 (6), 1013–1031.
- Heckelman, J., Knack, S., 2008. Foreign aid and market-liberalizing reform. *Economica* 75 (299), 524–548.
- IEO (Independent Evaluation Office), 2007. *An IEO Evaluation of Structural Conditionality in IMF-supported Programs*. International Monetary Fund, Washington, DC.
- International Monetary Fund, 2001. *Conditionality in fund-supported programs: policy issues*. IMF Policy Development and Review Department. <<http://www.imf.org/external/np/pdr/cond/2001/eng/policy/021601.pdf>> accessed: January 2011.

- Isopi, A., Mattesini, F., 2010. Aid and corruption: do donors use development assistance to provide the “right” incentives? Paper presented at the AidData Conference, University College, Oxford, March 22–25.
- Johnson, D., Zajonc, T., 2006. Can Foreign Aid Create an Incentive for Good Governance? Evidence from the Millennium Challenge Corporation. Center for International Development, Harvard University, <[http://www.people.fas.harvard.edu/~tzajonc/mcc\\_wp\\_apr06.pdf](http://www.people.fas.harvard.edu/~tzajonc/mcc_wp_apr06.pdf)> accessed: June 2010.
- Kaufmann, D., Kraay, A., Mastruzzi, M., 2009. Governance matters VIII: aggregate and individual governance indicators, 1996–2008. World Bank Policy Research Working Paper No. 4978.
- Knack, S., 2001. Aid dependence and the quality of governance: cross-country empirical tests. *Southern Economic Journal* 68 (2), 310–329.
- Lancaster, C., 2008. George Bush's Foreign Aid: Transformation or Chaos?. Center for Global Development, Washington, DC.
- MCC, 2008. The ‘MCC Effect’: creating incentives for policy reform; promoting an environment for poverty reduction. Millennium Challenge Corporation, Washington, DC.
- Mosley, P., Hudson, J., Verschoor, A., 2004. Aid, poverty reduction and the ‘new conditionality’. *Economic Journal* 114 (June), F217–F243.
- Öhler, H., Nunnenkamp, P., Dreher, A., 2011. Does conditionality work? A test for an innovative US aid scheme, CESifo Working Paper 3454, <<http://www.cesifo-group.de/portal/pls/portal/docs/1/1205692.PDF>> accessed: May 2011.
- Pridham, G., 2008. The EU's political conditionality and post-accession tendencies: comparisons from Slovakia and Latvia. *Journal of Common Market Studies* 46 (2), 365–387.
- Radelet, S., 2003. Challenging Foreign Aid. A Policymaker's Guide to the Millennium Challenge Account. Center for Global Development, Washington, DC.
- Rieffel, L., Fox, J.W., 2008. The Millennium Challenge Corporation: an opportunity for the next president. *Brookings Global Economy and Development Working Paper* 30.
- Sedelmeier, U., 2008. After conditionality: post-accession compliance with EU law in East Central Europe. *Journal of European Public Policy* 15 (6), 806–825.
- Stone, R.W., 2004. The political economy of IMF lending in Africa. *American Political Science Review* 98 (4), 577–592.
- Svensson, J., 2000. Foreign aid and rent-seeking. *Journal of International Economics* 51 (2), 437–461.
- Svensson, J., 2003. Why conditional aid does not work and what can be done about it? *Journal of Development Economics* 70 (2), 381–402.
- Tavares, J., 2003. Does foreign aid corrupt? *Economics Letters* 79 (1), 99–106.
- UNDP, 2008. A users' guide to measuring corruption. <<http://commons.globalintegrity.org/2008/09/users-guide-to-measuring-corruption.html>> accessed: January 2011.
- Vreeland, J., 2006. IMF program compliance: aggregate index versus policy specific research strategies. *Review of International Organizations* 1 (4), 359–378.