



*Office of Inspector General*

September 26, 2014

Mr. Thomas Kelly  
Acting Vice President, Department of Policy and Evaluation  
Millennium Challenge Corporation  
875 15th Street, NW  
Washington, DC 20005

Subject: Can Foreign Aid Create an Incentive for Good Governance? Evidence from the Millennium Challenge Corporation (MCC)

Dear Mr. Kelly:

This letter transmits a report on an empirical study of the MCC effect: that is, the incentive effect that availability of MCC funding has on policy makers in developing countries (M-000-14-005-S).

The MCC awards large grants, known as compacts, to low- and lower middle-income countries that score well on 20 indicators that measure the quality of candidate countries' political, economic, and social policies. While there is little doubt that availability of MCC funding has influenced some decision makers in developing countries to undertake policy reforms, it is uncertain how large or widespread the MCC incentive effect is.

To contribute to a better understanding of the significance and extent of the MCC incentive effect, OIG engaged the services of an economist to investigate the MCC effect using statistical estimation techniques. This study did not produce convincing statistical evidence of the existence of the MCC effect.

The paper itself accompanies this transmittal letter as Appendix I. Our evaluation of management comments is included as Appendix II, and the full text of management comments appears in Appendix III. The data base and procedures used to conduct the analysis are available at <https://senseplatform.com/tristan/mcc-effect>, best viewed using Google Chrome.

Millennium Challenge Corporation  
1401 H Street, NW  
Suite 770  
Washington, DC 20005  
<https://oig.usaid.gov>

## Prior Work

MCC has reported instances in which specific countries have taken actions to improve their standing on the scorecards. For example, MCC reported that the Sierra Leonean Government worked with MCC and others to improve its scorecard performance after passing only six indicators in fiscal year 2005. According to MCC, the government expanded the mandate and resources of its anticorruption agency, increased public health expenditures and immunization coverage, and reduced average tariff rates while expanding access to credit. Sierra Leone passed the MCC scorecard in fiscal year 2013, but in fiscal year 2014 it did not pass MCC's control of corruption indicator. MCC noted that this indicator must be passed before a compact is approved and encouraged Sierra Leone to continue its efforts to improve.

Similarly, MCC reported that the Government of Niger formed a committee that worked with MCC and others to address policy performance and data quality issues. According to MCC, the committee helped promote the government's efforts, such as increased environmental protection and progress on gender equality. Niger passed the scorecard in fiscal year 2012 and was selected to develop a proposal for a compact.

In 2013 researchers at the College of William and Mary reported the results of a survey designed to assess the influence of MCC's eligibility criteria on governments' reform efforts.<sup>1</sup> The researchers surveyed more than 600 policy makers and practitioners, including heads of governments, ministers, deputy ministers, and senior officials in 100 low- and lower middle-income countries. The survey respondents cited MCC's eligibility criteria as one of the three most influential external assessments of government performance. Other influences included the United Nations' Millennium Development Goals and the World Bank's Doing Business indicators.

Based on the survey results, the researchers concluded that the MCC effect existed but was not evenly distributed among countries, time, or policy areas. The effect was especially prevalent in the areas of corruption and fiscal policy, and less apparent in areas such as political rights and civil liberties. The survey also found that officials in countries already affiliated with MCC said they were influenced by the eligibility criteria. The survey documented that MCC influenced policy makers' decisions and actions to some extent.

A limited number of quantitative analyses have attempted to understand the MCC effect. The earliest of these was conducted by Harvard researchers in 2006.<sup>2</sup> The researchers attempted to ascertain whether countries increased their reform efforts following the announcement and subsequent creation of MCC. The analysis found that candidate countries reformed approximately 25 percent more indicators after the announcement of MCC than before it. However, because the analysis was conducted soon after MCC's creation and with a limited amount of data, the researchers determined that the results were not conclusive.

---

<sup>1</sup> Bradley Parks and Zachary Rice, *Measuring the Policy Influence of the Millennium Challenge Corporation: A Survey-Based Approach*, Technical Report, Williamsburg, VA: College of William and Mary, 2013.

<sup>2</sup> Doug Johnson and Tristan Zajonc, "Can Foreign Aid Create an Incentive for Good Governance? Evidence from the Millennium Challenge Corporation," Working Paper No. 11, Center for International Development at Harvard University, Cambridge, MA, April 2006.

An analysis in 2011 focused exclusively on the control of corruption indicator and whether MCC had successfully promoted better control of corruption.<sup>3</sup> The analysis compared countries that had not passed the control of corruption indicator but were within reach of this goal to countries that had either passed or were very unlikely to pass. The assumption was that countries near the goal would have the greatest incentive to reform.

The researchers concluded that MCC had had an effect on countries' efforts to battle corruption, and that countries with a good chance of passing this indicator addressed corruption more effectively than others, especially in MCC's early years. However, the researchers suggested that MCC did not offer an incentive to countries that had a remote chance of obtaining MCC assistance, and that once countries passed the corruption indicator, they did not attempt to reform further.

In 2013 a research team at the University of Wisconsin-Madison conducted an analysis of the MCC effect but found no compelling evidence of its existence.<sup>4</sup> The researchers studied countries' actions following the creation of MCC in 2004 and analyzed results for five different groups: countries near MCC eligibility; countries that passed the control of corruption indicator; countries participating in MCC threshold programs (designed to help countries become more competitive for compacts); countries in the middle range of government expenditures; and countries that had been identified as having experienced the MCC effect. The researchers conducted hundreds of comparisons, but did not find strong evidence for an MCC effect, as they found a roughly equal number of positive and negative statistically significant results.

## **Current Approaches and Results**

The investigators used two approaches to estimate the MCC effect: a "difference in differences" approach and a "regression discontinuity design" approach. These approaches and the results they produced are discussed below.

### **Difference in Differences**

The difference in differences approach used in the paper estimates the MCC effect by comparing policy reform rates in candidate countries (low- and lower middle-income countries) before and after the establishment of MCC in 2004. To isolate the MCC effect from other factors that influence the rate of policy reform, this approach subtracts out the policy reform rates in noncandidate middle-income countries. The assumption is that candidate countries would have the same change in policy reform rates as middle-income countries, except that they are eligible for MCC funding and therefore

---

<sup>3</sup> Hannes Ohler , Peter Nunnenkamp, and Axel Dreher, "Does Conditionality Work? A Test for an Innovative US Aid Scheme," CESifo Working Paper 3454, CESifo Group Munich, 2011.

<sup>4</sup> Ingrid Aune, Yanyan Chen, Christine Miller, and Joshua Williams, *The MCC Incentive Effect: Quantifying Incentives for Policy Change in an Ex-Post Rewards System*, policy report prepared for MCC, La Follette School of Public Affairs, University of Wisconsin-Madison, October 2013.

experience the MCC incentive effect.<sup>5</sup> If this assumption is not valid, then the estimate will be biased.

The results of this difference in differences analysis do not provide statistically significant evidence of an MCC incentive effect. A combined analysis of MCC indicators that existed in 2008 or earlier (presented graphically in Figure 4 of the attached report) shows a pattern that is suggestive of an MCC effect. In examining point estimates of reform rates in candidate countries, a noisy, quasi-random pattern of reform rates is observed prior to establishment of MCC in 2004; after 2004, point estimates of reform rates rather consistently exceed the 2004 baseline rate. However, the 90 percent and 95 percent confidence intervals include the 2004 baseline rate; so it is not possible to reject the possibility that post-2004 reform rates are equal to the 2004 baseline rate.

Analysis of the individual indicators (presented graphically in Figures 5 through 8 of the attached report) includes some statistically significant results. But the coefficients for several of the indicators show unexpected signs (i.e., reform rates are lower after 2004), contributing to a conclusion that the results, as a whole, do not provide convincing evidence of an MCC incentive effect.

### **Regression Discontinuity Design**

A regression discontinuity design uses a cutoff point to assign members to a treatment group and a control group, and uses the size of the discontinuity in the regression line at the cutoff point to estimate the effectiveness of the treatment. In the current research, the cutoff point is a passing score on an MCC policy indicator in a particular year. Countries that score below the cutoff point are in the treatment group (because they presumably face a greater incentive to improve their score to become eligible for MCC assistance), while countries that score above the cutoff are in the control group.

The local linear regression discontinuity design used in the attached report estimates a separate linear equation for every point on the X-axis, with points on the X-axis representing distance from the cutoff point on a particular policy indicator: that is, the amount by which a country exceeded or fell short of a passing score on that indicator. The regression lines relate distance from the cutoff point as an independent variable to each country's subsequent 3-year reform rate as the dependent variable for each point on the X-axis. When these lines are joined, they form a smooth curve relating distance from the cutoff point to subsequent reform rates.

If the MCC effect exists and countries just below the passing score for a particular indicator have higher subsequent reform rates than countries that already have a

---

<sup>5</sup> More formally,  $\hat{\tau}_{ATT} = \frac{1}{N_1} \sum_{i:W_i=1} (Y_{i1} - Y_{i0}) - \frac{1}{N_2} \sum_{i:W_i=0} (Y_{i1} - Y_{i0})$ , when  $\hat{\tau}_{ATT}$  is the estimated average treatment effect on the treated group. Or for the estimated MCC effect on policy reform rates in candidate countries  $N_1$  is the number of countries in the treatment group,  $N_2$  is the number of countries in the control group,  $\Sigma$  signifies summation,  $W_i = 1$  indicates that a country is in the treatment group,  $W_i = 0$  indicates that a country is in the control group,  $Y_{i1}$  is the reform rate for a country between any 2 years after establishment of MCC in 2004, and  $Y_{i0}$  is the reform rate for a country before establishment of MCC.

passing score on that indicator, we would expect the curve to jump down at the cutoff point. This “jump” is the discontinuity in a regression discontinuity design.

The analysis does not provide statistically convincing evidence of an MCC effect. An analysis of all of the indicators combined (Figure 10 in the attached report) produces an estimated jump of just 0.01 standard deviations, with a 95 percent confidence interval of plus or minus 0.2 standard deviations. That is, with 95 percent confidence, the existence of an MCC effect greater than 0.2 standard deviations at the cutoff point can be rejected.

When the indicators are analyzed individually (Figures 11 and 12 in the attached report), there are two statistically significant results, for the government effectiveness indicator and the inflation indicator. But the sign for government effectiveness is the opposite of what is expected, indicating that countries facing an MCC incentive effect reform more slowly than countries that are not incentivized. Taken as a whole, these results do not provide persuasive statistical evidence of an MCC effect.

## **Conclusion**

This work should be read and interpreted in conjunction with other available evidence on the MCC incentive effect. Prior work cited in this transmittal letter persuades us that an MCC incentive effect is operative in certain specific instances, but this effect is not large enough or widespread enough to be measured with the statistical approaches that were used in the attached report.

Please direct any questions to Melinda Dempsey.

Sincerely,

/s/

Melinda Dempsey  
Deputy Assistant Inspector General  
Office of Audit  
Millennium Challenge Corporation

cc: Karla Chryar, Compliance Officer, MCC  
LeJuan Butler, Senior Executive Assistant, MCC  
Robert Fry, Performance Audits Director, OIG/MCC  
Paula Mathews, Program Analyst, OIG/MCC

# Can Foreign Aid Create an Incentive for Good Governance? Evidence from the Millennium Challenge Corporation\*

Doug Johnson

Gene Goldstein-Plessner

Tristan Zajonc<sup>†</sup>

February 27, 2014

## Abstract

The Millennium Challenge Corporation (MCC) awards aid to countries that perform well on a set of independently compiled governance indicators. Proponents of this new form of aid argue that 1) aid will be more effective when given to well-governed countries and 2) countries will respond to such incentives by improving their policies. While significant qualitative evidence exists supporting an MCC incentive effect, the quantitative evidence is inconclusive. This paper attempts to estimate the MCC incentive effect by exploiting discontinuities in MCC eligibility rules and reform patterns before and after the MCC was created in 2004 for candidate and non-candidate countries. Using data from 2000 to 2012, we find candidate countries reform significantly faster than non-candidate middle-income countries in the years following the MCC's creation, but that these effects become statistically insignificant when controlling for pre-MCC reform trends. With the exception of inflation, we find no evidence that countries that fall just below eligibility thresholds, where the incentive effect would presumably be stronger, reform faster than countries that just pass these thresholds. While we do not find evidence the MCC has spurred dramatic reform in a large number of candidate countries, we also cannot rule out economically significant effects or effects in specific countries. When pooling all indicators and years, our 95% confidence intervals include effects on reform rates as high as 0.2 standard deviations. For individual indicators, our confidence intervals include many effects as large as 0.5 standard deviations. The best evidence for an MCC incentive effect therefore remains qualitative. Our results do not speak to the overall merits of the MCC.

## JEL Classifications: O1

**Keywords:** MCC Effect, Millennium Challenge Corporation, aid policy, foreign assistance, good governance, aid conditionality.

---

\*This paper is a substantially updated version of a working paper with the same title released in 2006 by Tristan Zajonc and Doug Johnson that used preliminary data from the MCC. The current version is supported by a contract from the Inspector General's Office of USAID. The results and conclusions are the authors alone. We are grateful for feedback from Alberto Abadie, Tim Cox, Jishnu Das, Robert Fry, Andria Hayes-Birchler, Asim Khwaja, Paula Mathews, Brad Parks, Michael Walton, and seminar participants at Harvard.

<sup>†</sup>Corresponding author: [tristan@senseplatform.com](mailto:tristan@senseplatform.com).

# 1 Introduction

Created by the U.S. Congress in January 2004, after being proposed two years earlier by President Bush, the Millennium Challenge Corporation (MCC) is an independent U.S. foreign aid agency with a mission to reduce poverty and promote economic growth through an innovative model of development assistance. Rather than delivering aid to a broad collection of impoverished nations, the MCC grants assistance only to countries deemed most committed to ruling justly, investing in their citizens, and promoting economic freedom as measured by a set of objective and transparent governance indicators provided by non-governmental entities. Proponents of this new form of aid argue that 1) aid will be more effective when given to well-governed countries and 2) countries will respond to such incentives by pursuing sound policies. This latter incentive effect – often called the “MCC Effect” – is the focus of this paper. We seek to answer the question: have the MCC eligibility rules cause candidate countries to improve their governance as measured by the MCC indicators?

Existing quantitative evidence for an MCC incentive effect is limited and inconclusive. The first empirical study of the MCC effect is a 2006 version of this paper. Using data from 2000 to 2004, Johnson and Zajonc (2006) find that MCC candidate countries reformed faster after the MCC was announced in 2002 than they had in preceding years and than higher-income non-candidate countries. However, due to the infancy of the MCC at the time of publication, the paper only covers a short window between the MCC’s announcement in 2002 and the beginning of its operations in 2004. The results are therefore only suggestive.

Several more recent papers update this analysis along various lines. Ohler, Nunnenkamp and Dreher (2011) report evidence for an incentive effect on the MCC’s control of corruption indicator, particularly in early years. Aune et al. (2013) provide a comprehensive exploration of the MCC effect using independently compiled data and variations on the empirical strategy outlined in Johnson and Zajonc (2006). They find no compelling evidence for an MCC incentive effect. The statistically significant results they do report are picked out of 645 comparisons, with roughly an equal number of positive and negative statistically significant results.

The most compelling evidence for an MCC incentive effect comes from anecdotal reports and qualitative surveys. MCC (2013*a*) and World Bank (2007) document policymakers citing the MCC as motivation for their reform efforts. In a comprehensive survey of MCC stakeholders, Parks and Rice (2013) find that the MCC eligibility criteria are one of the three most influential external assessments of government performance, along with the Millennium Development Goals and World Bank’s Doing Business indicators. Parks (forthcoming) expands this analysis and documents MCC-inspired policy responses in a constructed database of policy reforms coded from archival data. These qualitative results are persuasive in one respect: MCC eligibility rules clearly factor in the calculations made by policymakers, even if the magnitude and exact nature of these calculations are unclear. Section 3 provides a fuller review of existing quantitative and qualitative evidence for an MCC incentive effect.

This paper attempts to estimate the MCC effect quantitatively by updating and significantly expanding on the earlier work by Johnson and Zajonc (2006). Using data from 2000 to 2012, we explore two empirical strategies. First, we compare reform before and after the MCC was created for low and low-middle income candidate countries compared to non-candidate middle-income countries that do not face an MCC incentive. Second, we compare reform for

candidate countries just below the MCC eligibility thresholds, where the incentive for reform is presumably stronger, with candidates that just pass the thresholds and presumably have a weaker, although potentially still positive, incentive to reform. The first approach is similar to the difference-in-differences strategy followed by Johnson and Zajonc (2006), Ohler, Nunnenkamp and Dreher (2011), and Aune et al. (2013). The second approach is an example of regression discontinuity (RD) design (Thistlethwaite and Campbell, 1960; Hahn, Todd and der Klaauw, 2001).

Looking at reform rates between 2004 and 2012, low income and low-middle income candidate countries reformed significantly faster than middle income non-candidate countries — countries with GNI per capita below \$10,000 — across a number of indicators. Averaging across the original indicators used in 2004, candidate countries increased their performance by roughly 0.25 standard deviations more than middle income non-candidate countries in the years following the MCC’s creation. However, middle income countries are substantially different in their baseline income, indicator levels, and reform patterns. To correct for this, Johnson and Zajonc (2006) propose a difference-in-differences strategy that uses pre-MCC trends to control for differences between candidate and non-candidate countries. That is, Johnson and Zajonc (2006) assume that, absent the MCC, differences in reform rates between candidate and non-candidate countries after 2004 would have been the same as the differences prior to 2004. Ohler, Nunnenkamp and Dreher (2011) and Aune et al. (2013) follow similar strategies although with different definitions of the control group and time periods. Given the sensitivity of the results in Johnson and Zajonc (2006), Ohler, Nunnenkamp and Dreher (2011) and Aune et al. (2013) to the exact specifications used, in this paper we report the majority of results graphically, and adjust for baseline differences in reform by controlling for pre-MCC trends rather than using a simple differencing approach.

Controlling for baseline differences in reform largely eliminates the differences in post-MCC reform rates, although the aggregate point estimates remain positive in our default specification (Figures 4, 5, 6, 7, and 8). For specific years and indicators there are statistically significant effects, but the adjusted estimates include zero effect in the vast majority of cases and are not consistently signed in the direction of a positive improvement. Averaging across the original 2004 indicators (Figure 4), we cannot reject no effect for both low and low-income countries in any year between 2004 and 2012. Nevertheless there is considerable uncertainty. Even when pooling indicators, our 95% confidence intervals include effects as large as 0.25 standard deviations.

A similar analysis for GNI per capita growth — the ultimate goal of MCC-driven reform — also finds no statistically significant effect (Figure 9). However the confidence intervals for these estimates cover all plausible changes in growth. We include these results only for completeness.

While the comparisons across time and between candidate and non-candidate countries is a useful exercise, it is possible that events other than the MCC’s creation differentially affected candidate countries after 2004. Our RD results address this potential concern. Using the published eligibility cutoffs for each year and original data from MCC report cards, we estimate the effect of just failing the eligibility criteria on countries’ subsequent three years of reform. If the MCC incentive effect was large, countries just failing the eligibility cutoff would presumably reform faster than countries that already pass the eligibility threshold. Given these cutoffs are



arbitrary outside the context of the MCC, any change in reform around these cutoffs can be credibly attributed to the MCC incentive effect.

With the exception of inflation, we find no statistically significant positive effects (Figures 10, 11, and 12). Averaging across all indicators and years, the RD estimate for the incentive effect on reform is 0.01 standard deviations (0.1 standard error). In the aggregate RD result there is no visible discontinuity (Figure 10) . While the positive result for inflation is interesting, we cannot reject that it is a statistical artifact given the number of comparisons we report. Any formal multiple comparisons procedure (e.g. a Bonferroni correction) would eliminate the statistical significance of both findings. The inflation indicator is also one of the least susceptible indicators to policy influence according to surveys reported in Parks and Rice (2013). We also find a negative effect for government effectiveness, although this result is not robust to varying RD bandwidths.

While we do not find any evidence of large, broad-based reform spurred by the MCC's eligibility rules, we cannot rule out economically significant effects or effects for specific countries. For both our difference-in-differences and regression discontinuity based estimates, our confidence intervals are large, including many effects between 0.25 and 0.5 standard deviations. Effects of this magnitude would represent a major success. We also only study aggregate effects — the effect for all candidate countries in the case of our difference-in-differences estimates and the effect for countries near a single eligibility threshold for our RD estimates. We cannot rule out that the MCC eligibility rules have a large impact on a handful of truly committed reform-minded governments.

More broadly, the results of Parks and Rice (2013) and Parks (forthcoming) suggest that the MCC incentive effect varies by countries, year, and indicator, and that it operates over multiple stages of the policymaking process. In this paper, we focus exclusively on reform as measured by changes in the MCC indicators. Yet many of these indicators are difficult to change, lag significantly behind policy reforms, and include substantial noise, particularly when examining changes. The lack of a measurable effect on MCC indicators does not preclude the possibility of other significant effects on the inputs and processes that drive reform.

Our results do not speak to the overall merits of the MCC. While the MCC incentive effect has received considerable attention, the MCC is based on the premise that aid is more effective when well-governed countries receive significant grants to pursue country-led solutions that promote economic growth and reduce poverty, followed by rigorous evaluations of what works and what doesn't. If the MCC compacts generate a positive economic rate of return or significant learning, the MCC incentive effect, if it exists, would be an extra benefit, rather than the solitary feature that determines the MCC's success.

This paper is organized as follows: Section 2 summarizes the history of the MCC, the manner in which the MCC distributes aid, and the current status of the MCC. Section 3 discusses previous qualitative and quantitative research on the MCC incentive effect. Sections 4 and 5 outline our empirical strategy and data. Section 6 reports our main results and Section 7 concludes.

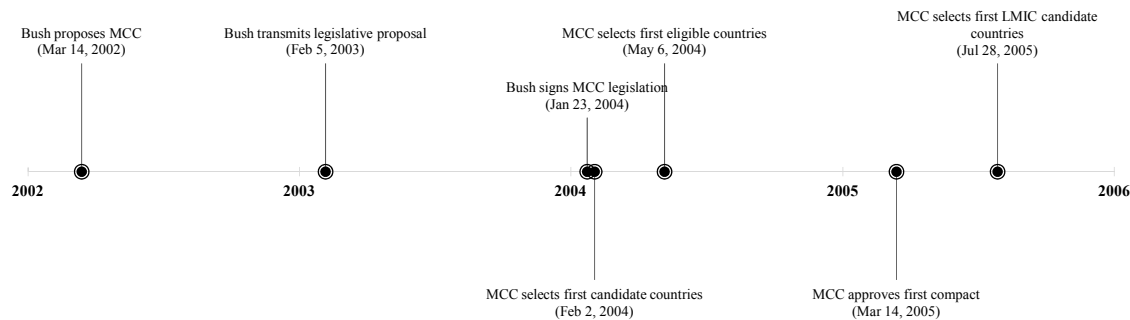


Figure 1: Timeline of MCC legislation and implementation.

## 2 Background

### 2.1 History and Motivation for MCC

In March 2002, the Bush administration committed to increase the US budget for overseas development assistance by \$5 billion annually — the largest single increase in overseas assistance in over forty years. This increase included the creation of a new government agency, the Millennium Challenge Corporation (MCC), to oversee the distribution of these funds. Two years after being proposed, the MCC was created by the U.S. Congress in January 2004 and began operations shortly thereafter. Figure 1 gives an overview of the MCC’s timeline, which is relevant for our difference-in-differences estimates and interpreting the results in Johnson and Zajonc (2006) and Ohler, Nunnenkamp and Dreher (2011).

The principles underlying the MCC and those of traditional aid agencies are different. Rather than distributing aid based on a country’s need or U.S. geopolitical interests, the MCC selectively helps those countries judged committed to governing justly, investing in their citizens, and promoting economic freedom, as measured by a set of objective, transparent, and independently compiled indicators (Table 1).

There were two primary motivations for MCC’s design. First, because MCC funds would only go to the best governed poor countries, proponents argued that the funds would be better spent. This idea received support not only from the long-held intuition of many development professionals, but also from timely empirical research. In an influential paper, Burnside and Dollar (2000) used historical data on aid, growth, and various measures of governance to argue that aid, when disbursed to countries that are governed effectively, causes more growth. They conclude, “aid has a positive impact on growth in developing countries with good fiscal, monetary, and trade policies but has little effect in the presence of poor policies.” While this conclusion has since been questioned by other research, such as Easterly, Levine and Roodman (2004), it was a powerful statement in support of the MCC’s eligibility rules.

Second, by basing aid on past performance, proponents of the MCC argued that the MCC would create an incentive for good governance. In a speech outlining his vision for the MCC, newly installed MCC CEO John Danilovich expressed the belief that “the only way that [the MCC] can be transformational is to incentivize these countries to make the reforms that are necessary” (Danilovich, 2006). This belief that the MCC would alleviate poverty by creating an

	<b>Indicator</b>	<b>Sign</b>	<b>Source</b>	<b>Year Added</b>	<b>Year Dropped</b>
<b>Ruling Justly</b>	Political Rights	(+)	Freedom House	2004	
	Civil Liberties	(+)	Freedom House	2004	
	Voice and Accountability	(+)	World Bank/Brookings Institution	2004	2012
	Freedom of Information	(-)	Freedom House	2012	
	Government Effectiveness	(+)	World Bank/Brookings Institution	2004	
	Rule of Law	(+)	World Bank/Brookings Institution	2004	
	Control of Corruption	(+)	World Bank/Brookings Institution	2004	
<b>Investing in People</b>	Child Health	(+)	Columbia/Yale	2012	
	Girls' Primary Education Completion Rate	(+)	UNESCO	2004	
	Girls' Secondary Education Enrollment Ratio	(+)	UNESCO	2012	
	Immunization Rate	(+)	World Health Organization	2004	
	Natural Resource Management	(+)	Columbia/Yale	2008	2012
	Natural Resource Protection	(+)	Columbia/Yale	2012	
	Public Expenditure on Health (% GDP)	(+)	World Health Organization	2004	
	Public Primary Education Spending (% GDP)	(+)	UNESCO	2004	
<b>Economic Freedom</b>	1-Year Consumer Price Inflation	(-)	International Monetary Fund	2004	
	Access to Credit	(+)	International Finance Corporation	2012	
	Business Start-Up	(+)	International Finance Corporation	2008	
	Cost of Starting a Business	(-)	World Bank Group	2006	2008
	Days to Start a Business	(-)	World Bank Group	2004	2008
	Fiscal Policy	(+)	International Monetary Fund	2004	
	Gender in the Economy	(-)	World Bank	2012	
	Land Rights and Access	(+)	International Fund for Agricultural Development	2010	
	Regulatory Quality	(+)	World Bank/Brookings Institution	2004	
	Trade Policy	(+)	Heritage Foundation	2004	

Table 1: Summary of MCC indicators. Sign denotes direction of change in indicator that is associated with an improvement. Political rights, civil liberties, and trade policy indicators were rescaled in 2007 to be positive. All three are positive in 2012. Girls' secondary education indicator replaced primary education indicator for LMIC countries only.

incentive for countries to pursue better policies was based on theory rather than evidence. Just as individuals respond to incentives, proponents argued, so too do countries. Yet the MCC was the first aid agency to put this assumption to the test using a form of ex-post conditionality.

While the World Bank and the other international financial institutions have long attempted to link aid to policy, these institutions do so mainly by placing conditions on the use of money after its disbursement rather than selecting potential recipients on the basis of their policies before disbursement. In the case of the World Bank, future tranches of loans may, in theory, be cancelled if governments fail to meet the conditions specified in the original disbursement. However, the World Bank has opted to enforce this rule only once — the case of Senegal in the 1980s (Mosley, Hudson and Verschoor, 2004).

Moreover, even though many organizations claim to support countries with good policies, this does not appear to be true. Alesina and Weder (2002) find that there is no relationship between bilateral and multilateral aid and the corruption level of recipient countries. For the United States they actually find a positive relationship between corruption and aid. This stands in stark contrast to the MCC allocation rule; countries cannot receive MCC aid unless they are above the median on the control of corruption indicator.

## **2.2 Funding and Distribution of MCC Aid**

Although the Bush Administration originally requested \$5 billion to be allocated to the MCC for disbursement each year, Congress approved only \$1 billion, \$1.5 billion, and \$1.75 billion for fiscal years 2004, 2005, and 2006 respectively. In its first year, despite certifying sixteen countries as eligible to receive MCC funds, the MCC signed no compacts with any of the eligible countries, leading to criticism that it was too slow ramping up its operations. In its second year, MCC signed five separate compacts, incurring a grant obligation of over \$900 million over five years. Subsequent years have seen a grand total of 29 full compacts signed, along with 24 Threshold Programs, small aid packages designed to help a country improve its governance practices in order to meet MCC qualifications.

Although the Obama Administration has expressed support for the MCC's methods, foreign aid budgets have been trimmed amid recent budget negotiations in Congress, with only \$898 million allocated to the MCC in FY2014 ([foreignassistance.gov](http://foreignassistance.gov), 2013). Thus, the MCC accounts for only a small fraction of total U.S. foreign assistance, 5% in FY2011 (Figure 2).

To receive funding from the MCC as of FY2014 a country must meet four independent criteria(MCC, 2013*b*):

1. It must qualify as a Low Income Country (LIC) or Low-Middle Income Country (LMIC) by having a GNI per capita under \$1,965 or \$4,085 respectively.
2. It must not be restricted from receiving U.S. aid by Congressional statute. Countries may be restricted for poor human rights records or for being designated a state sponsor of terror.
3. It must perform sufficiently well on a set of twenty (originally sixteen) independently compiled governance indicators.

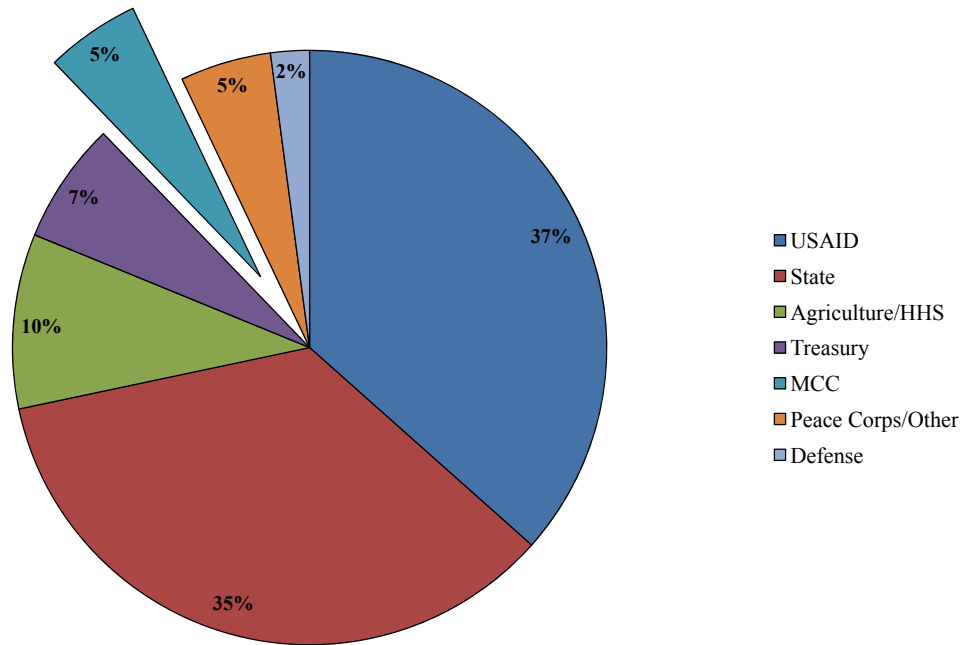


Figure 2: Sources of US foreign assistance. Data reflects total disbursements for FY2011 out of \$30.964 billion spent on foreign assistance (USAID Economic Analysis and Data Services, 2012).

4. It must submit a compact proposal deemed worthy by the MCC’s board of directors.

A country is determined to be a “candidate country” if it passes the first two criteria by qualifying as a LIC/LMIC and not being subject to congressional restriction. The performance of these candidate countries on twenty governance indicators (Table 1) is gauged against the performance of income group peers, with LICs being compared against other LICs and LMICs against other LMICs.

Although transparent, the process for determining eligibility is complicated and has changed over the years. In FY2014, the twenty indicators are split into three distinct groups: Economic Freedom, Ruling Justly, and Investing in People. To be eligible, a country must pass certain threshold values on at least ten of the twenty indicators, and must pass at least one threshold value in each category. These threshold values may be absolute minimums or median values. Additionally, countries must score over the median on the control of corruption indicator, and above the absolute threshold on either the civil liberties or the political rights indicator. Missing indicator values are judged to be failing, although the MCC historically has made efforts to fill in missing values from a variety of sources.

In addition to the objective eligibility rules, the MCC reserves the right to exercise its own discretion in determining the final list of countries which pass this stage. In the past, the MCC has used this discretionary power sparingly to exclude countries that, for political reasons, are deemed inappropriate recipients of US aid such as China, India, and Bhutan and to include countries which are very close to qualifying.

Once deemed eligible for MCC funds, countries are invited to submit funding proposals. Submitted proposals are reviewed by the MCC and those found suitable are then approved and signed by the MCC's board of directors along with the leaders of the country and turned into compacts — legally binding documents specifying the amount of aid to be given, the uses to which it will go, and the performance criteria that the country must meet in its use of the money. According to the MCC's stated guidelines, countries may submit proposals for any projects which are both growth enhancing and poverty reducing.

In addition to compacts, the MCC also distributes a limited portion of the funds appropriated to it to countries at or near the threshold for receiving funds with the purpose of improving these countries' scores on their MCC indicators. Access to funds through this program, known as the "Threshold Program," is solely at the discretion of the MCC. Table 2 summarizes the history of MCC eligibility and compacts.

Since the MCC began operations, several changes have been made to the selection methodology. By far the most significant changes involve the addition and exclusion of indicators. A business start-up hybrid indicator was introduced in 2008, combining the cost of starting a business and days to start a business indicators, along with a natural resource management indicator. Several indicators were adjusted in 2012, with voice and accountability dropped in favor of freedom of information, natural resource protection replacing natural resource management, and indicators tracking child health and gender in the economy added. Other major revisions in FY2012 included the addition of a democracy hard hurdle, requiring that countries pass either the political rights or civil liberties indicator, and the loosening of the requirement that countries pass half the indicators in each thematic category. As of FY2012, countries need only pass ten of the twenty indicators overall in order to obtain eligibility (Dunning, 2013; MCC, 2013*b*). While the exact rules have changed several times since 2004, the MCC has always used sharp eligibility thresholds per indicator.

### **3 Existing Evidence of MCC Effect**

#### **3.1 Plausibility of the MCC Effect**

For the MCC effect to pass the basic litmus test of plausibility, the size of the reward must be large enough that countries modify their behavior to gain access to the funds. Furthermore, the probability of receiving the reward must be sufficiently high that countries deem the possibility of receiving the award worth the effort of reforming. While a subjective judgement, the MCC easily passes these two hurdles. Looking at the sample of compact countries that report total aid and government expenditure in the World Development Indicators (World Bank, 2013), Figure 3 shows MCC funds have caused a substantive impact in compact countries. MCC annual funding averages 4% of total government expenditure and 14% of net aid and development assistance received. Furthermore, with a relatively small pool of compact eligible countries, the odds of gaining access to MCC funds once eligible appear high enough to merit the attention of candidate countries. As of FY2014, 25 of 28 eligible countries have signed compacts with the MCC (Table 2).

Country	First Fiscal Year Eligible		Year Signed	
	For Compact	For Threshold	Compact	Threshold Program
Albania		2004		2006
Armenia	2004		2006	
Benin	2004		2006	
Bolivia	2004			
Burkina Faso	2006	2005	2008	2005
Cape Verde	2004		2005	
Colombia	2009			
East Timor	2006	2004		2010
El Salvador	2006		2006	
Georgia	2004		2005	
Ghana	2004		2006	
Guyana		2005		2007
Honduras	2004		2005	2013
Indonesia	2009	2006	2011	2006
Jordan	2007	2006	2010	2006
Kenya		2004		2007
Kyrgyzstan		2006		2008
Lesotho	2004		2007	
Liberia	2013	2009		2010
Madagascar	2004		2005	
Malawi	2007	2005	2011	2005
Mali	2004		2006	
Mauritania		2008		
Moldova	2007	2006	2010	2006
Mongolia	2004		2007	
Morocco	2005		2007	
Mozambique	2004		2007	
Namibia	2006		2008	
Nepal		2012		
Nicaragua	2004		2005	
Niger	2013	2007		2008
Paraguay		2005		2006
Peru		2007		2008
Philippines		2005	2010	2006
Rwanda		2007		2008
Sao Tome and Principe		2004		2007
Senegal	2004		2009	
Sierra Leone	2013			
Sri Lanka	2004			
Tanzania	2006	2004	2008	2006
Gambia	2006			
Uganda		2004		2007
Ukraine	2007	2006		2006
Vanuatu	2004		2006	
Yemen, Rep.		2004		
Zambia	2009	2005	2012	2006

Table 2: Countries eligible for MCC compacts and threshold program. Bolivia and Ukraine were dropped from the list of eligible countries in FY2009 after being selected in previous years. East Timor was downgraded from compact eligible to threshold eligible in FY2012. Niger had its threshold status revoked in FY2010, but was selected as compact eligible in FY2013 after improvement on many indicators. Yemen had its threshold eligible status revoked in FY2009. The compacts of Madagascar and Mali were terminated prematurely due to non-democratic developments. Honduras had its 2005 compact partially terminated in 2009, but began a threshold program in 2013.

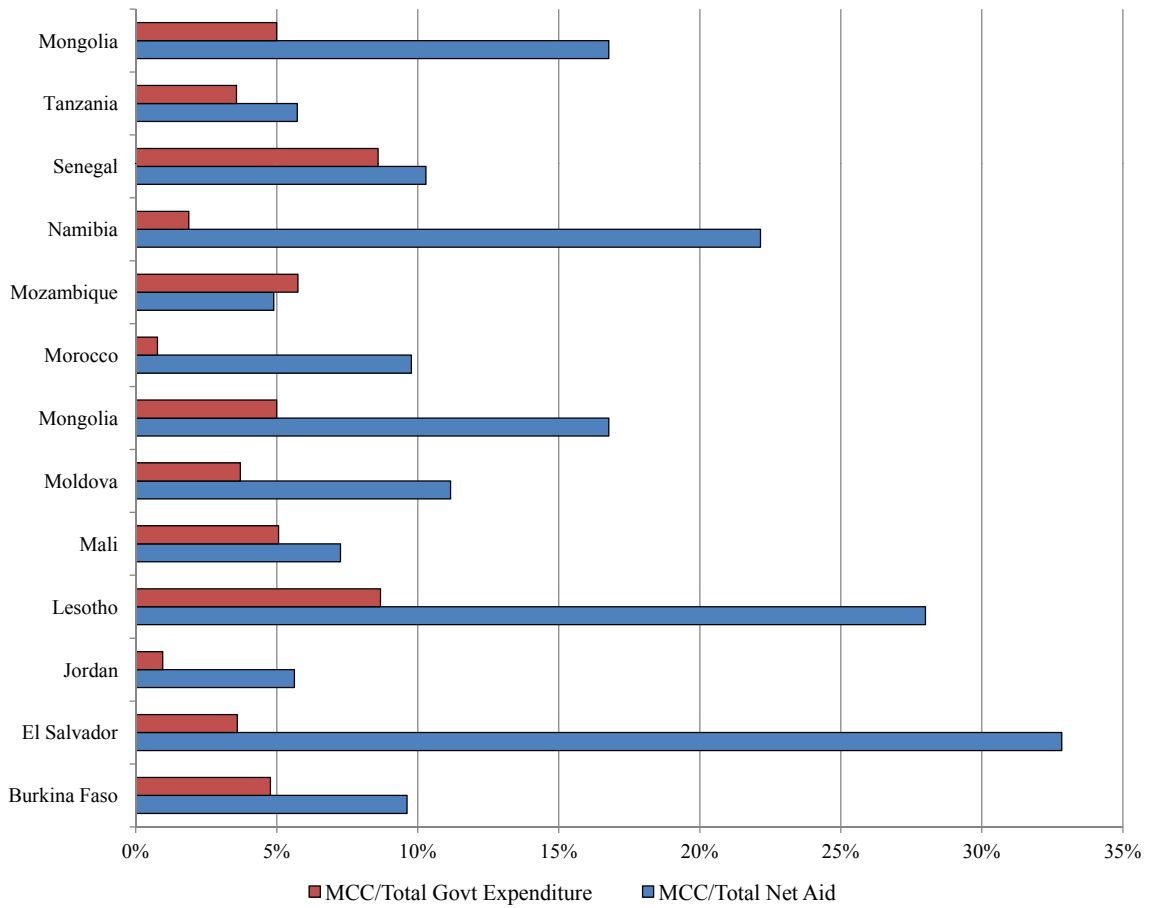


Figure 3: MCC compacts as a percentage of recipient total aid and government expenditures. MCC compacts calculated as total compact budget divided by five-year horizon to create an annual number (MCC, World Bank 2013). Only compact countries with aid data reported in the World Development Indicators are included.



## 3.2 Qualitative Evidence

There is significant qualitative evidence supporting the existence of an MCC incentive effect. Senior government officials in MCC candidate countries have stated in interviews that their countries have taken note of the MCC standards and are working towards reform (MCC, 2013*a*; World Bank, 2007; Parks and Rice, 2013). There is also some evidence of an MCC “signaling effect,” where other sources of aid and capital increase their efforts in countries that become MCC eligible (Dreher, Nunnenkamp and Ohler, 2010).

Parks and Rice (2013) and Parks (forthcoming) provide the most comprehensive qualitative evidence on the MCC’s impact on countries’ behavior. Using a cross-country survey of 640 individuals with connections to the MCC’s aid efforts, both in the United States and across 100 low and low-middle income countries, their results indicate a strong MCC effect, particularly in the areas of corruption, education, and fiscal policy. Respondents cite MCC eligibility criteria as one of the three most influential external judgments of government performance, behind the United Nation’s Millennium Development Goals and ahead of the World Bank’s Doing Business indicators. Parks (forthcoming) constructs a database of more than 14,000 country-policy-domain-year observations that measures whether and how governments change their policy behavior in order to achieve or maintain MCC eligibility and finds significant evidence for MCC-inspired policy responses – defined as any discernible impact of the MCC eligibility criteria on the policymaking process

Of note, both Parks and Rice (2013) and Parks (forthcoming) find a positive correlation between MCC eligibility status and attention paid to achieving or maintaining eligibility. Parks (forthcoming) finds that 65% of MCC-inspired policy responses happen in compact and threshold countries and the remaining 35% happen in countries that have not qualified for compact or threshold assistance. Threshold and compact countries therefore focus more closely on maintaining eligibility than candidate countries focus on achieving it. This fact, and the reported importance of the Millennium Development Goals and Doing Business indicators, suggest that the MCC may influence policy priorities in ways not directly related to the sharp eligibility rules originally emphasized by proponents of the MCC’s design. Moreover, if true, our regression discontinuity estimates may understate the true impact of the MCC on reform given the lack of a sharp discontinuity in incentives.

The MCC itself has also gathered a body of anecdotal evidence indicating that it has exerted an influence over the domestic policy agendas of several developing countries (MCC, 2013*a*). For instance, MCC (2013*a*) point to the success of Georgia in overhauling a variety of policies related to the ease of starting and running a business with the goal of improving their business start-up score. The World Bank, which publishes the Doing Business indicators used in calculating business start-up, credits MCC for an increase in awareness of the metric in the developing world (World Bank, 2007). In one notable example of potentially MCC-driven reform, Sierra Leone passed only six MCC indicators in 2006, and reached out to the MCC in order to formulate a strategy to improve its performance. Through an expansion of its anti-corruption agency, an increase in their public health expenditure, and a successful immunization campaign, Sierra Leone passed the MCC scorecard in 2012 and was declared eligible to pursue a compact in fiscal year 2013 (MCC, 2013*a*). While it is impossible to prove these reforms would not have occurred absent the MCC, a wide range of qualitative and anecdotal evidence suggests that spe-

cific countries have responded to MCC incentives. Yet the qualitative evidence says little about the size and breadth of MCC-inspired reform.

### 3.3 Quantitative Evidence

In a 2006 version of this paper, Johnson and Zajonc (2006) perform the first quantitative evaluation of the purported MCC effect. With data available only through 2004, Johnson and Zajonc (2006) select 2002 — the year the MCC was announced — as the earliest year any MCC effect could be present in the data. Using this timeframe, Johnson and Zajonc (2006) evaluate whether low income candidate countries were more likely to reform between 2002 and 2004 than between 2000 and 2002, controlling for changes in time trends using the reform patterns of low-middle income countries that did not face an explicit MCC incentive until 2006 (Figure 1). These difference-in-differences results provide limited evidence of an anticipatory MCC effect. Potential recipients of MCC funds improved 25 percent more indicators after the announcement of the MCC than before it, compared LMICs.

While suggestive, the results in Johnson and Zajonc (2006) are not conclusive. They report a number of positive and negative estimates and only report statistically significant results for some specifications and definitions of reform. The results are also extremely early, especially given that many of the MCC indicators are lagging, with reform measures often taking years to appear in the data. The first scorecards — official evaluations of each country’s performance on the indicators – were not released until 2004, so countries would not have known in 2002 which indicators required improvement. Given these facts, any improvement in the performance of candidate countries in 2003 and 2004 can be described as anticipatory at best, and difficult to link directly to a fledgling MCC. Nevertheless, the framework established by Johnson and Zajonc (2006) provides a useful methodological baseline.

Ohler, Nunnenkamp and Dreher (2011) apply the same difference-in-differences approach as Johnson and Zajonc (2006) to a wider swath of available data, but limit the scope of their analysis to one indicator, control of corruption. They also use a different method of separating countries into treatment and control groups, selecting as a treatment group those countries whose control of corruption scores place them in the second quartile. They argue that second quartile scores define a set of countries that perform below the MCC threshold, but not so poorly that achieving MCC standards would be unreasonably difficult. The control group is thus populated by countries that have already exceeded the MCC standard as well as “hopeless cause” countries, both of which have little incentive to attempt improvement in their indicators. Ohler, Nunnenkamp and Dreher (2011) conclude that the MCC did have a positive impact on the treatment group. However, running the same models over various 2-4 year intervals, they determine that the incentive to fight corruption has been weakened over time. They posit that this weakening is a result of uncertainty about the timing and quantity of MCC aid, especially in the wake of a difficult start-up period for the agency in 2004 and 2005, although this interpretation is speculative. Our results do not confirm an effect on control of corruption (Figures 5 and 11), although we do not replicate the Ohler, Nunnenkamp and Dreher (2011) analysis exactly.

Aune et al. (2013) expand the methodology used by Johnson and Zajonc (2006) to cover subsequent years of data, as well as several alternative definitions of a treatment group. Rather than centering their data around the announcement of the MCC in 2002, they also use the es-

establishment of the MCC in 2004 as their focal point, as we do in this paper, and independently recreate all indicators from original third party sources. They report results from five treatment group definitions: countries bordering on eligibility, countries that pass the control of corruption indicator, countries participating in an MCC threshold program, countries in the middle range of government expenditures, and a hand-selected group of countries determined qualitatively to have experienced the MCC effect. After running hundreds of difference-in-differences comparisons, they are unable to find compelling evidence for an MCC incentive effect. Of 645 possible instances of statistical significance, 38 are positive and statistically significant, while a comparable number, 32, are negative and significant.

The results of Johnson and Zajonc (2006), Ohler, Nunnenkamp and Dreher (2011), and Aune et al. (2013) contain some suggestive evidence in favor of an MCC effect, but are largely inconclusive. All three report statistical significant results only after testing a wide range of specifications and for varying definitions of treatment groups, indicators, and timeframes. This paper provides a comprehensive reanalysis of the MCC effect.

## 4 Empirical Strategy

### 4.1 Setup

Our goal is to estimate the causal effect of MCC incentives on candidate countries reform patterns. That is, had candidate countries not faced an MCC incentive, would they have reformed differently?

We can make this more precise by defining causal effects in terms of potential outcomes, following Neyman (1923) and Rubin (1978). Let  $Y_i(1)$  give the outcome for country  $i$  if they face an MCC incentive and  $Y_i(0)$  give the outcome if they don't. Potential outcomes  $Y_i(1)$  and  $Y_i(0)$  are linked to the observed variables  $Y_i$  by  $Y_i = Y_i(W_i)$ , where  $W_i$  indicates whether a country faces an incentive or not. Assuming all candidate countries face an incentive, we only ever observe  $Y_i(1)$  since  $W_i = 1$ . Because we cannot observe  $Y_i(0)$  and  $Y_i(1)$  simultaneously there is no direct way to observe the MCC incentive effect  $Y_i(1) - Y_i(0)$  for an individual country.

There are various interesting groups of countries for which we might want to estimate the causal effect of the MCC incentive, but a logical starting point is to estimate the average effect on all countries that experience an MCC incentive,

$$\tau_{\text{ATT}} = \mathbb{E}[Y_i(1) - Y_i(0) \mid W_i = 1]. \quad (1)$$

In the causal inference literature this is the *average treatment effect on the treated* — the causal effect on the population that actually received the treatment. In this paper we define that group as all candidate countries, although Ohler, Nunnenkamp and Dreher (2011), Aune et al. (2013), and our regression discontinuity estimates assume different groups of countries experience an incentive effect.

### 4.2 Difference-in-Differences

Johnson and Zajonc (2006) propose estimating (3) using a difference-in-differences strategy. Let the outcome  $Y_{i1}$  be reform on an indicator between any two years post-MCC and  $Y_{i0}$  be the

rate of reform pre-MCC. The difference-in-differences estimate of the MCC effect is then

$$\hat{\tau}_{\text{ATT}} = \frac{1}{N_1} \sum_{i:W_i=1} Y_{i1} - Y_{i0} - \frac{1}{N_2} \sum_{i:W_i=0} Y_{i1} - Y_{i0}.$$

That is, we estimate the MCC effect as the difference in reform for candidate (treated) countries before and after the MCC was created minus the difference in reform rate for non-candidate (control) countries before and after the MCC was created. Depending on the measure of outcome — a rate of reform or indicator level — this can be viewed as a difference-in-differences or difference-in-difference-in-differences estimate. The estimate subtracts any fixed country effects (the change rather than level), any fixed time trends (the first pre-post difference), and any change in time trends (the treatment/control difference).

The key assumption is that absent the MCC, candidate countries would have increased their reform equivalently to non-candidate countries. Formally,

$$\mathbb{E}[Y_{i1}(0) - Y_{i0}(0) \mid W_i = 1] = \mathbb{E}[Y_{i1}(0) - Y_{i0}(0) \mid W_i = 0]. \quad (2)$$

This assumption allows us to solve for the reform rate absent an incentive for candidate countries,  $\mathbb{E}[Y_{i1}(0) \mid W = 1]$ , which is fundamentally unobservable. The assumption would be violated if there were other policies, besides that MCC, that affected candidate and non-candidate countries differentially.

In this paper, we follow a slightly different strategy than a traditional difference-in-differences design in order to make the basis for our calculations clearer. We plot the reform rate for LICs and LMICs between the first year an indicator was used (mostly 2004) and each subsequent year, subtracting out the reform rate for middle income (control) countries — countries with GNI per capital below \$10,000 that do not face an MCC incentive. This allows us to compare pre-post trends visually. We also report adjusted differences that control linearly for two pre-MCC periods of reform. Formally, these plots embed the assumption that reform after 2004 absent the MCC,  $Y_{i1}(0)$ , is independent of candidacy status conditional on observed baseline reform patterns,  $Y_{i1}(0) \perp W_i \mid \mathbf{Y}_{i0}$ .

### 4.3 Regression Discontinuity Design

Our second empirical strategy relies on the MCC’s unique eligibility rules. We compare reform rates for ineligible countries that fall just below and just above the MCC eligibility thresholds and therefore experience different incentives to reform but are otherwise similar. While the exact incentive effects created by the MCC eligibility rules are complex, we examine each indicator individually.

Using discontinuities treatment rules to identify causal effects dates back to Thistlethwaite and Campbell (1960). However the idea has received an explosion of attention in economics after Hahn, Todd and der Klaauw (2001) formalized regression discontinuity designs in the language common to program evaluation. Imbens and Lemieux (2008) and Lee and Lemieux (2009) provide accessible reviews of this literature. Imbens and Zajonc (2011) give methods, to estimate regression discontinuity designs with multiple forcing variables, such as the

multi-dimensional MCC eligibility rules, but these methods come at the expense of significant complexity.

Adopting the classic scalar regression discontinuity setup, let  $X_{it}$  represent a baseline indicator level in year  $t$  and  $Y_{it}$  represent the subsequent reform. In Section 6 we define the outcome as change in the indicator over three subsequent years, normalized by the baseline level. Let  $c_t$  be the eligibility cutoff in year  $t$  for a particular indicator. Then regression discontinuity design identifies the average causal effect of the treatment for countries at the treatment boundary ( $X_{it} = c_t$ ) by comparing countries  $\varepsilon$  above and below the treatment boundary as  $\varepsilon$  goes to zero. Assuming the conditional regression functions of the potential outcomes  $\mathbb{E}[Y_{it}(0) | X_{it} = x]$  and  $\mathbb{E}[Y_{it}(1) | X_{it} = x]$  are continuous in  $x$ , countries just above and below the discontinuity have the same average potential outcomes  $Y_{it}(0)$  and  $Y_{it}(1)$  but differ by the incentive they actually experience,  $W_{it}$ . Due to the complexity of MCC rules,  $W_{it}$  should be interpreted loosely as “experiences a higher MCC incentive.”

The average causal effect of the treatment for countries at the treatment boundary is identified by taking the limits from above and below,

$$\begin{aligned} \tau_{\text{SRD}} &= \mathbb{E}[Y_{it}(1) - Y_{it}(0) | X_{it} = c_t] \\ &= \mathbb{E}[Y_{it} | W_{it} = 1, X_{it} = c_t] - \mathbb{E}[Y_{it} | W_{it} = 0, X_{it} = c_t] \\ &= \lim_{x \downarrow c_t} \mathbb{E}[Y_{it} | X_{it} = x] - \lim_{x \uparrow c_t} \mathbb{E}[Y_{it} | X_{it} = x] \end{aligned} \tag{3}$$

where the final equality follows from continuity (Hahn, Todd and der Klaauw, 2001). It is in general not possible to identify broader effects, such as the average treatment effect on the treated countries.

The practical problem becomes how to estimate the two limits  $\lim_{x \downarrow c_t} \mathbb{E}[Y_{it} | X_{it} = x]$  and  $\lim_{x \uparrow c_t} \mathbb{E}[Y_{it} | X_{it} = x]$  given the limited data near the boundary. Hahn, Todd and der Klaauw (2001) propose estimating the limits by local linear regression, to adjust for differences when including data away from the cutoff. With a rectangular kernel, local linear regression amounts to predicting both limits from a linear regression on a subset of countries within the bandwidth  $h$  of the cutoff. Imbens and Kalyanaraman (2008) recommend local linear regression with an edge (triangular), rather than rectangular, kernel and derive an optimal, data-dependent, bandwidth selection rule. We follow this recommendation and present our results graphically (Figures 10, 11, and 12).

## 5 Data

As described previously, the MCC currently uses 20 indicators drawn from a number of independent sources. Table 1 describes these indicators, their sources, and the direction of improvement. Our primary dataset consists of the historical performance of every country on every MCC indicator currently used, as released in November 2013 by MCC’s Open Data initiative. Where the MCC’s historical data has gaps, particularly around indicators that are no longer used or that have been rescaled, we have augmented their database with datasets obtained directly from the independent bodies that produce the indicators.

<b>Indicator</b>	<b>Obs</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>	<b>First Year</b>
Days to Start a Business	1168	43.58	64.40	2.00	694.00	2003
Cost to Start a Business	1168	66.06	132.75	0.29	2051.46	2003
Voice and Accountability	1344	-0.23	0.79	-2.21	1.31	2000
Civil Liberties	1224	35.64	13.09	1.00	58.00	2002
Control of Corruption	1335	-0.44	0.60	-1.91	1.56	2000
Education Expenditure	510	1.87	0.87	0.16	6.37	2000
Fiscal Policy (Actual year)	1431	-1.76	6.78	-24.00	125.44	2000
Girls' Primary Education Completion Rates	897	82.35	24.12	12.57	138.77	2000
Government Effectiveness	1333	-0.41	0.60	-2.32	1.26	2000
Health Expenditures	1339	3.40	2.33	0.58	21.57	2000
Immunization Rates	1450	83.67	16.17	18.50	99.00	2000
Inflation	1447	7.85	13.49	-10.04	325.03	2000
Political Rights	1224	23.11	11.18	0.00	39.00	2002
Regulatory Quality	1333	-0.36	0.64	-2.68	1.54	2000
Rule of Law	1342	-0.40	0.67	-1.96	1.38	2000
Trade Policy	1367	68.26	12.48	0.00	90.00	2000
GNI Per Capita Growth	1027	3.69	14.98	-49.96	445.97	2000

Table 3: Summary statistics for low, low-middle, and middle income countries. Data reflects FY14 time-series data. LIC/LMIC status determined by LIC/LMIC by MCC in FY06 thresholds and MIC is defined as countries with GNI per capita below \$10,000. Excludes countries prohibited from receiving MCC funds between 2004 and 2012.

In addition to this FY2013 data, we also examine what we call “vintage data,” the datasets used in past years to make determinations about a country’s eligibility. As the values of many indicators are regularly revised after their initial release, we use the original data in order to make a determination on whether a particular country faced an incentive to improve on any particular indicator, even if that incentive was based on incomplete or faulty data. We also consider a supplementary dataset composed of basic country-level data from the World Bank’s World Development Indicators. Table 3 gives summary statistics for the key variables in our dataset.

While the MCC indicators are selected in part for their breadth of coverage, not all indicators cover all countries. There is also considerable variation in the length of the time-series for each variable. For example, Freedom House’s political rights and civil liberties indicators are available back to 1973 and are updated yearly. In contrast, natural resource management is only available back to 2004.

For our analysis, we drop a number of observations. We consider only the set of indicators which existed in 2008 or earlier. We also exclude all countries with GNI per capita above \$10,000 in 2006. We use 2006 as our determining year for this calculation, as this was the first year the MCC formalized a definition of LMICs. Because most data is not available for 2013, we do not include any 2013 data in our analysis. With these restrictions, we are left with 102 countries. Because of missing data, however, fewer countries are available for many indicators. In addition, we generally exclude countries that are statutorily restricted from receiving assistance. These countries neither have an incentive to improve their indicators or are plausible control countries for those that do.

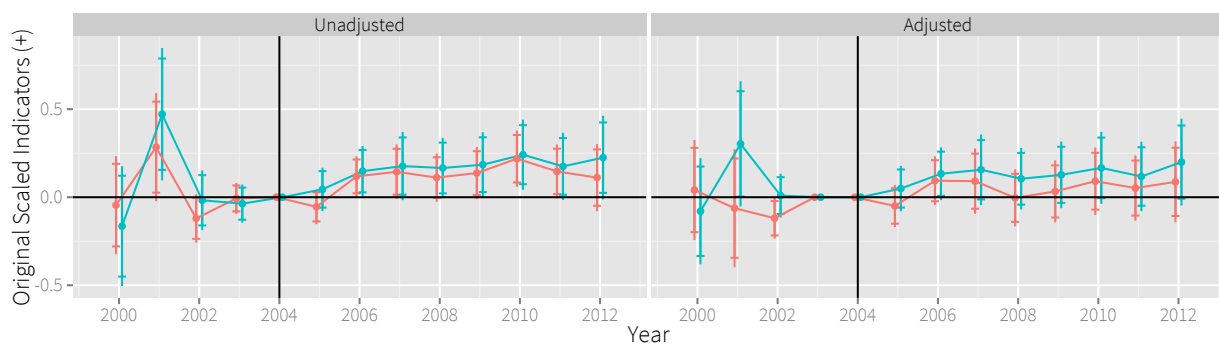


Figure 4: Average reform all FY2004 indicators for low income (red) and low-middle income (blue) candidate countries compared to middle income non-candidate countries ( $y = 0$ ). Reform is the change between the current year and 2004. Before averaging, indicators are normalized to have standard deviation one in the baseline year and signed such that positive values are an improvement. Bars and crosses represent 90% and 95% confidence intervals. Adjusted reform controls for two pre-MCC periods of reform.

## 6 Results

### 6.1 Difference-in-Differences Style Estimates

The large number of indicators and years makes it difficult to concisely summarize our empirical results. Figure 4 gives the top line difference-in-differences style result. We plot the change between 2004 and each year for LICs and LMICs compared to MICs ( $y = 0$ ), averaging across all indicators. To make the indicators comparable, we scale all the indicators to have a standard deviation of one in 2004 and change signs when necessary so that a positive value always represents an improvement. The left hand pane gives the unadjusted estimates and the right hand pane gives adjusted estimates that control for two prior values of reform.

As can be seen, LICs and LMICs reform significantly faster after 2004 than the non-candidate middle income countries (left panel). Two years after the MCC began operations, the unadjusted reform rates are around 0.15 to 0.25 standard deviations higher for both LICs and LMICs than MICs. Both 90% and 95% confidence intervals, represented by ticks and bars, do not cover zero for most years. However, the adjusted results tell a less compelling story. Controlling linearly for two pre-MCC periods of reform (2002, 2003) eliminates statistically significant difference between LIC/LMICs and MICs in the post-MCC years, although the point estimates remain positive. We cannot reject zero effect in any year at the 95% level. However, we also cannot reject economically significant effects. While the point estimates hint at a positive MCC effect, various control strategies, such as controlling linearly for GNI per capita, changed these estimates significantly.

Figures 5, 6, 7, and 8 plot similar analyses for individual indicators. For specific years, the unadjusted results show positive and statistically significant results for some indicators. However, there are few clear patterns in the adjusted estimates. Several point estimates are positive and several are negative. Cost of starting a business, an indicator added as a component of busi-

ness startup in 2006, flips sign when controlling for baseline results, indicating the difficulty of using middle income countries as controls for low income countries.

For completeness, Figure 9 gives a similar plot for GNI per capita growth — an ultimate goal of MCC-driven reform. We find no statistically significant effect, although the confidence intervals cover all remotely plausible effects. The financial crisis of 2008 further complicates this analysis.

## 6.2 Regression Discontinuity Design Estimates

Perhaps the most compelling result in this paper is given by Figure 10. Here we plot regression discontinuity estimates for all years and indicators in a single graph. The x-axis measures the distance to the eligibility threshold for each indicator. The y-axis measures the subsequent three year change in the indicator. Each point represents a currently ineligible country, indicator, and baseline year. Both the vintage indicators (x-axis) and the time-series reform indicators (y-axis) are normalized so that all they have standard deviation one in the baseline year and are signed such that a positive value is an improvement. The smoothed line is a local linear regression for each side of the cutoff with an edge kernel and bandwidth selected following Imbens and Kalyanaraman (2008).

If countries that just fail the eligibility rule have a greater incentive to reform than countries that just pass the eligibility threshold, we'd expect the smooth lines to jump down at 0, moving left to right. We see no such effect. Averaging across indicators and years, our best estimate for the jump is 0.01 standard deviations (0.1 standard error). In other words, we find no evidence that countries that just fail to meet eligibility thresholds reform faster in subsequent years than countries that just pass the eligibility thresholds. We can reject effects greater than 0.2 standard deviations at the 5% level.

Figures 11 and 12 plot similar analyses for individual indicators. The majority of estimates are statistically insignificant from zero. The two statistically significant effects, government effectiveness and inflation, go in opposite directions. Our estimated impact of government effectiveness is negative, while our estimated effect on inflation is positive. These results are likely due to statistical chance and are not robust to larger and smaller bandwidths.

Overall, our regression discontinuity results do not support a large, broad-based MCC incentive effect that is tied closely to the sharp eligibility cutoffs. This could be because the effect is small or present only in specific countries. It could also be because our simplified regression discontinuity design, which focuses only on a single indicator at a time, does not capture the complicated incentive effects generated by MCC eligibility rules. Nevertheless, the lack of any clear effect in the aggregated result given by Figure 10 suggests that the MCC incentive effect is modest.

## 7 Conclusion

The MCC is based on the premise that 1) aid will be more effective when given to well-governed countries and 2) countries will respond to eligibility rules by pursuing sound policies. While a large empirical literature has focused on the relationship between policies and aid effectiveness,



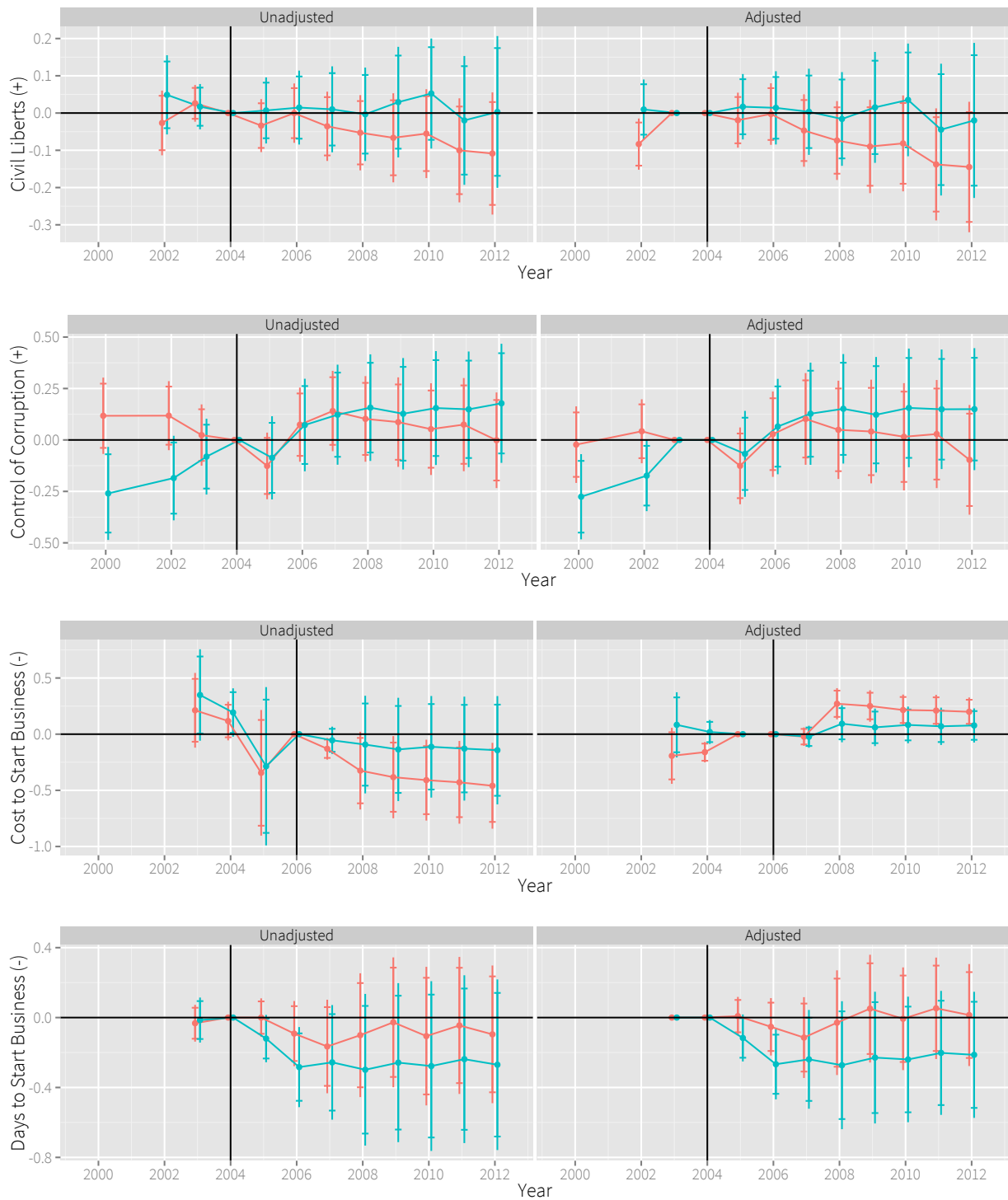


Figure 5: Reform since indicator introduction for low income (red) and low-middle income (blue) candidate countries compared to middle income non-candidate countries ( $y = 0$ ). Reform is the change between the current year and the year the indicator was added. Indicators are normalized to have standard deviation one in the baseline year. The sign of improvement is given in the y-axis label. Bars and crosses represent 90% and 95% confidence intervals. Adjusted reform controls for two pre-MCC periods of reform.

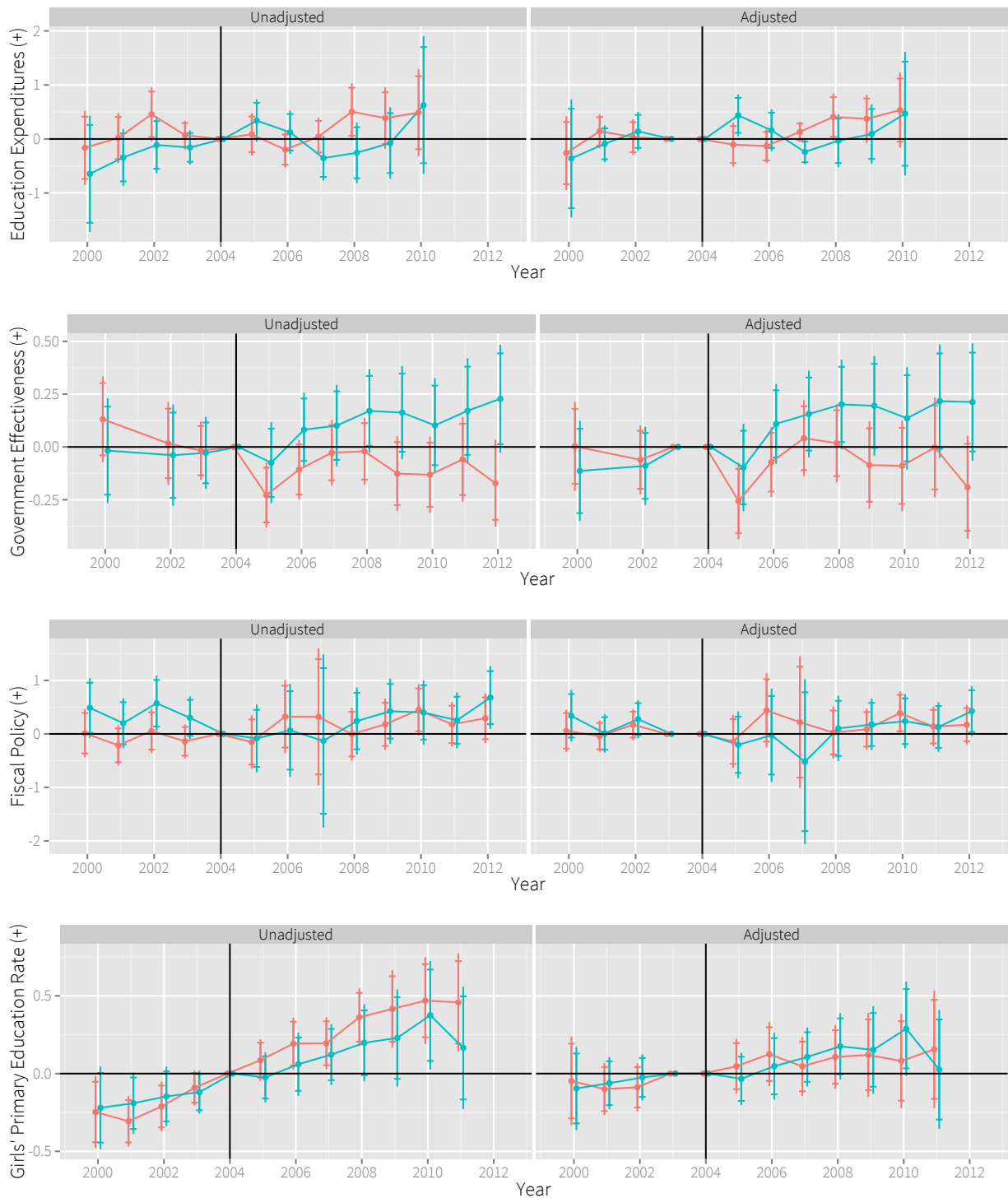


Figure 6: Reform since indicator introduction for low income (red) and low-middle income (blue) candidate countries compared to middle income non-candidate countries ( $y = 0$ ). Reform is the change between the current year and the year the indicator was added. Indicators are normalized to have standard deviation one in the baseline year. The sign of improvement is given in the y-axis label. Bars and crosses represent 90% and 95% confidence intervals. Adjusted reform controls for two pre-MCC periods of reform.

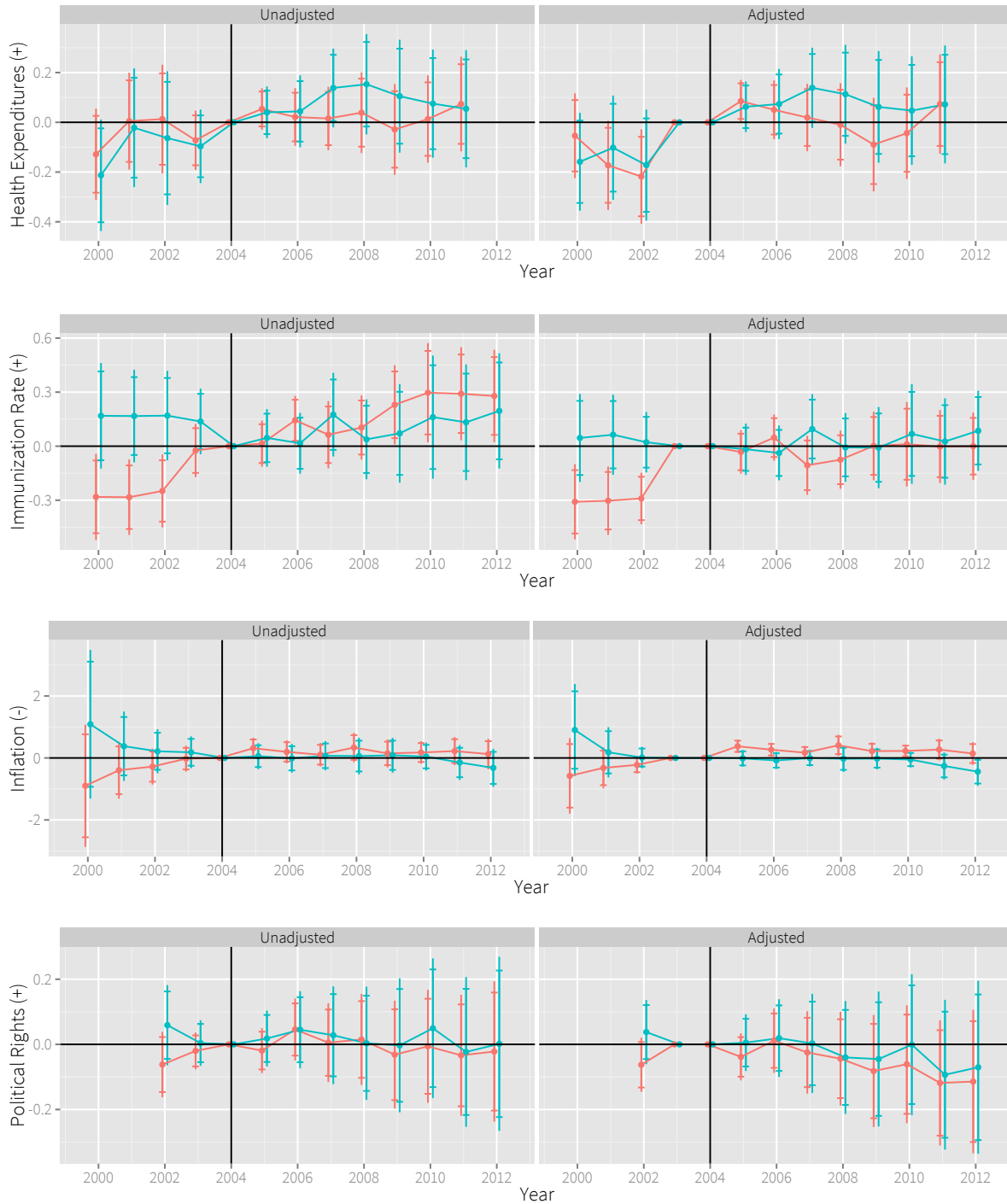


Figure 7: Reform since indicator introduction for low income (red) and low-middle income (blue) candidate countries compared to middle income non-candidate countries ( $y = 0$ ). Reform is the change between the current year and the year the indicator was added. Indicators are normalized to have standard deviation one in the baseline year. The sign of improvement is given in the y-axis label. Bars and crosses represent 90% and 95% confidence intervals. Adjusted reform controls for two pre-MCC periods of reform.

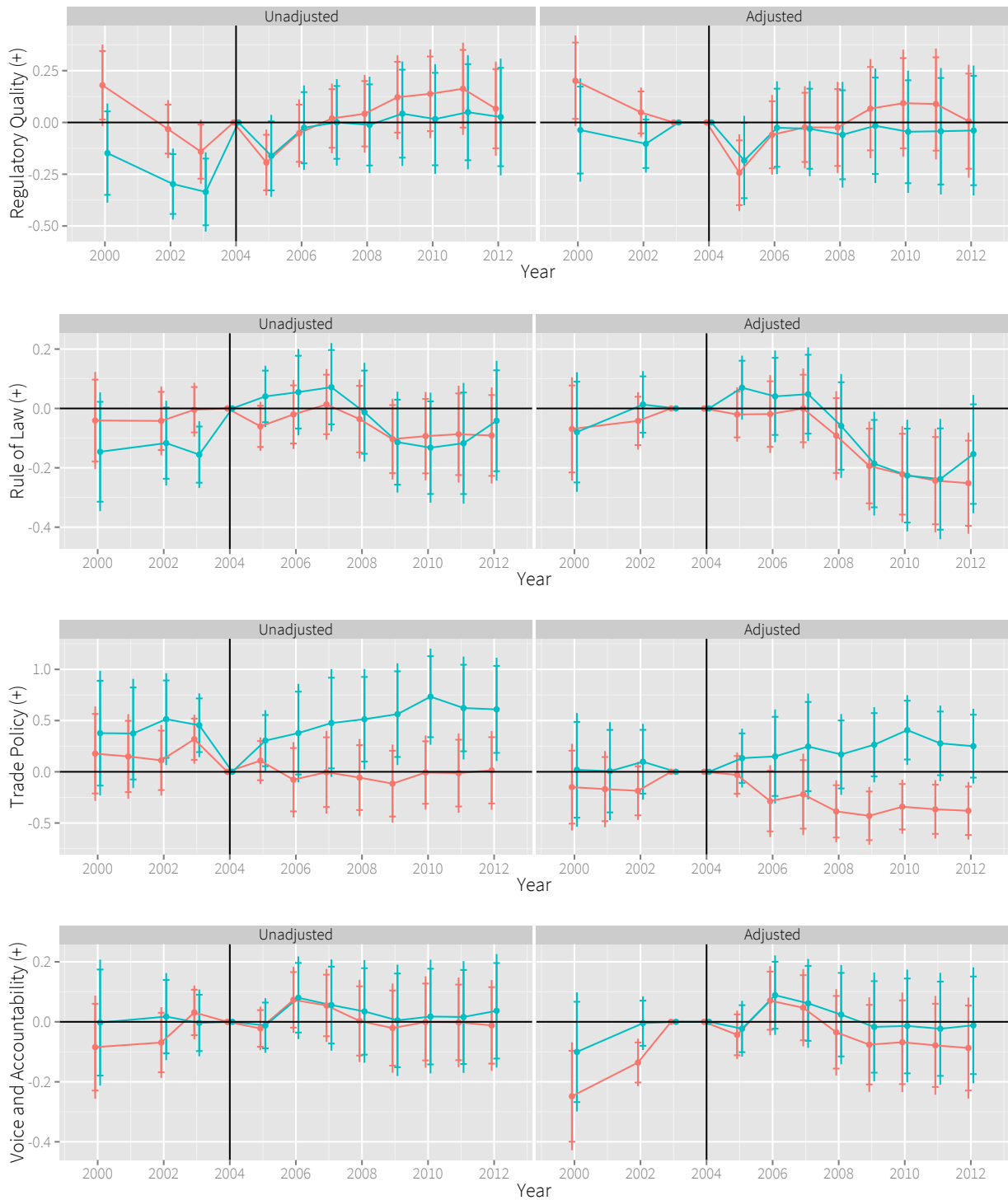


Figure 8: Reform since indicator introduction for low income (red) and low-middle income (blue) candidate countries compared to middle income non-candidate countries ( $y = 0$ ). Reform is the change between the current year and the year the indicator was added. Indicators are normalized to have standard deviation one in the baseline year. The sign of improvement is given in the y-axis label. Bars and crosses represent 90% and 95% confidence intervals. Adjusted reform controls for two pre-MCC periods of reform.

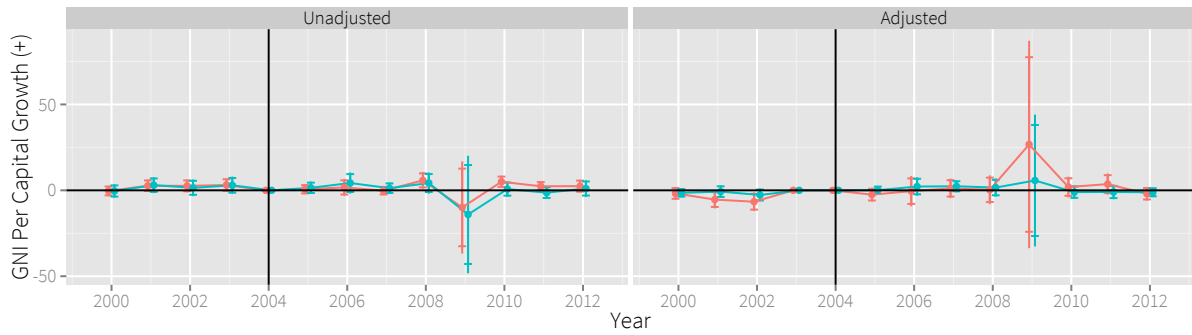


Figure 9: Change in GNI per capital growth for low income (red) and low-middle income (blue) candidate countries compared to middle income non-candidate countries ( $y = 0$ ). Change is the difference in growth between the current year and 2004. Bars and crosses represent 90% and 95% confidence intervals. Adjusted growth controls and two pre-MCC growth rates.

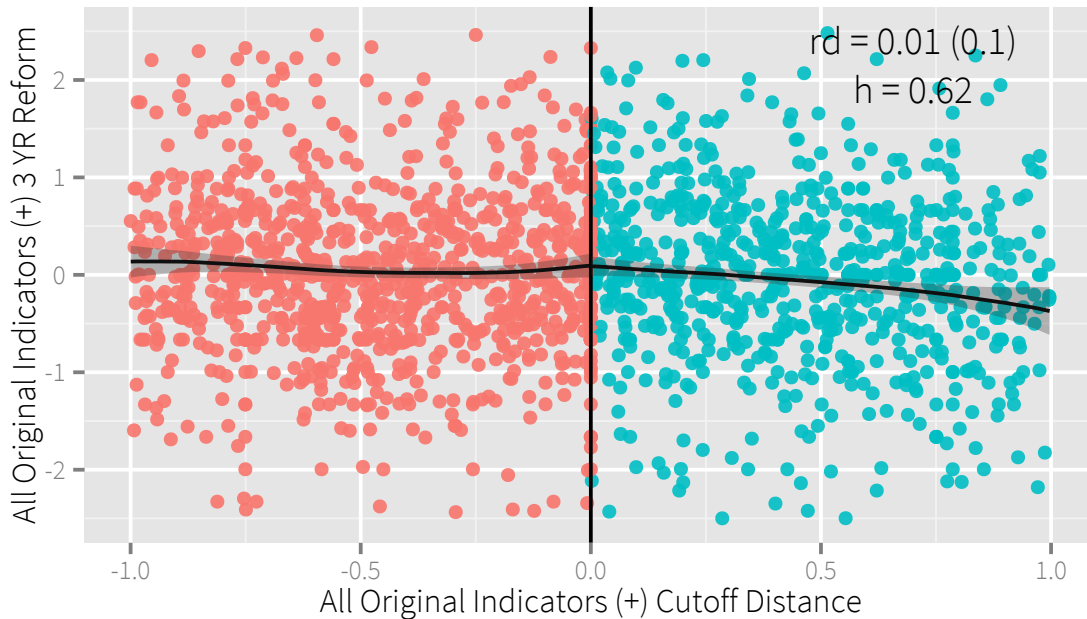


Figure 10: Three year reform pooling all indicators vs. distance to the eligibility cutoff for candidate countries between 2004 and 2010. All indicators are normalized such that 0 is the cutoff and positive values are an improvement. Red dots represent countries failing the eligibility threshold (incentivized) and blue dots represent those already passing. The RD estimate is in the upper right hand corner, with standard error in parentheses. The optimal Imbens and Kalyanaraman (2008) bandwidth of the local linear estimate is given by  $h$ .

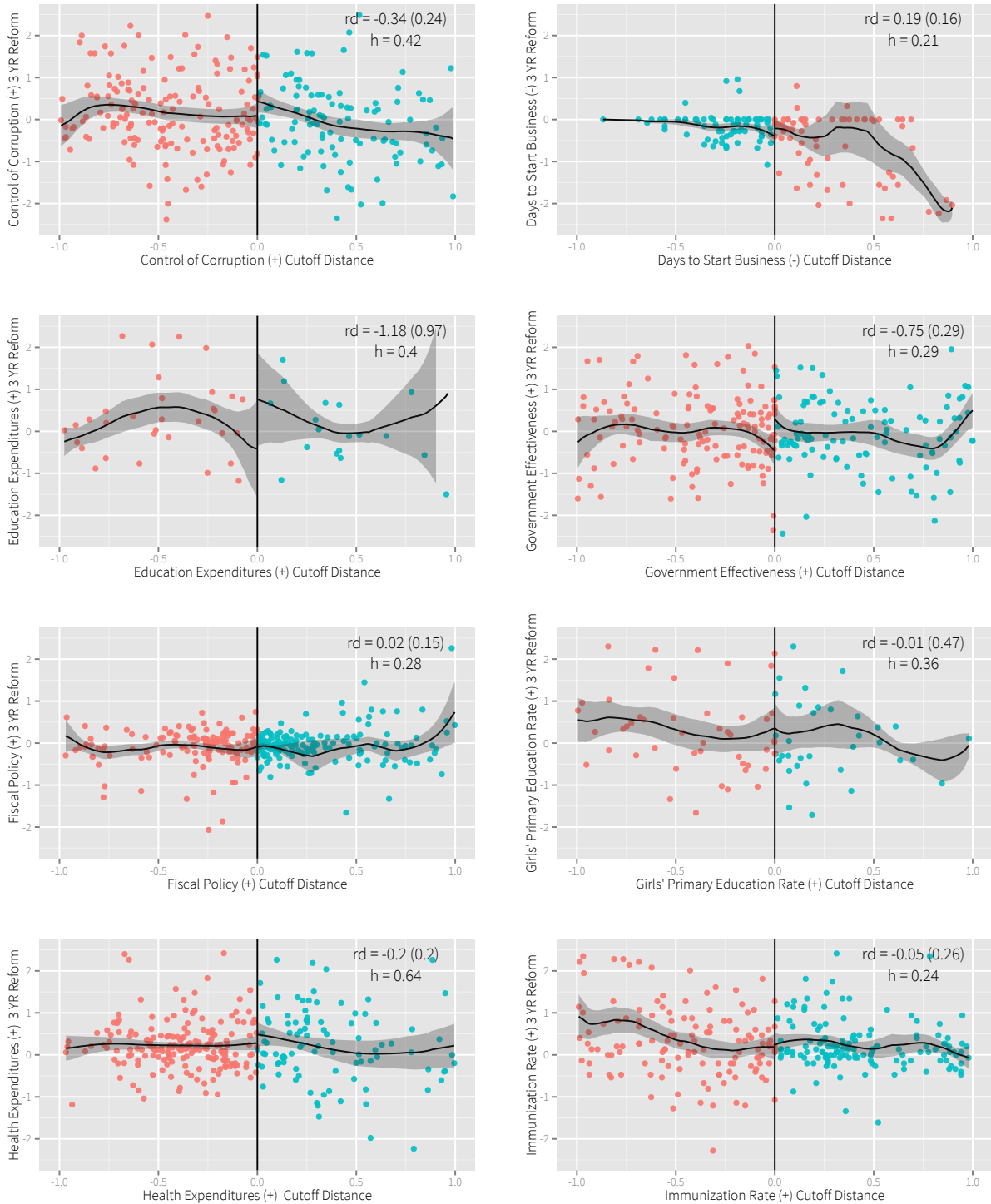


Figure 11: Three year reform by indicator vs. distance to the eligibility cutoff for candidate countries between 2004 and 2010. Indicators are normalized each year such that 0 is the cutoff. Red dots represent countries failing the eligibility threshold (incentivized) and blue dots represent those already passing. The RD estimate is in the upper right hand corner, with standard error in parentheses. The optimal Imbens and Kalyanaraman (2008) bandwidth of the local linear estimate is given by  $h$ .

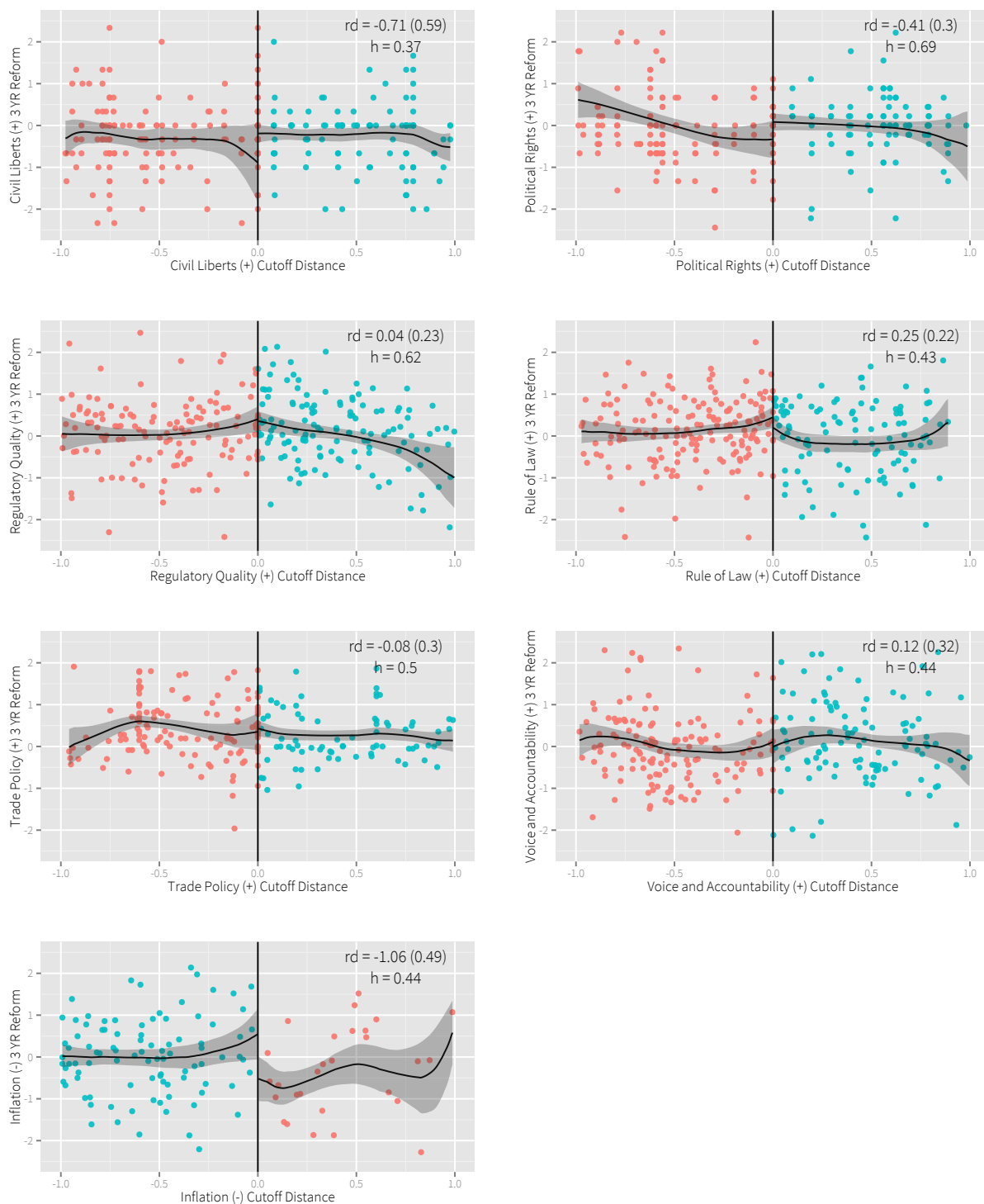


Figure 12: Three year reform by indicator vs. distance to the eligibility cutoff for candidate countries between 2004 and 2010. Indicators are normalized each year such that 0 is the cutoff. Red dots represent countries failing the eligibility threshold (incentivized) and blue dots represent those already passing. The RD estimate is in the upper right hand corner, with standard error in parentheses. The optimal Imbens and Kalyanaraman (2008) bandwidth of the local linear estimate is given by  $h$ .

few quantitative attempts have been made evaluate the second claim. We seek to answer the question: can foreign aid create an incentive for good governance?

By exploiting features of the MCC's candidacy and eligibility rules and variation before and after the MCC was created, we estimate the MCC incentive effect on MCC indicators using a difference-in-differences and regression discontinuity design strategy. While our difference-in-differences estimates are positive, they are not statistically significant once controlling for baseline differences in reform and they rely on strong assumptions. Our regression discontinuity estimates, while potentially not capturing the full MCC effect, show no consistent positive effects. Countries that fall just below eligibility cutoffs reform no faster in subsequent years than countries just passing these cutoffs, on average.

Our results do not support the view that the MCC has triggered dramatic reform across a large number of countries and indicators, but we also cannot reject economically significant effects or effects on specific countries. There is an inherent difficulty in teasing out causal effects given small sample sizes and country-level data. Many of our estimates are associated with large confidence intervals. We cannot rule out average effects across indicators as high as 0.25 standard deviations and individual effects for specific indicators as high as 0.5 standard.

Our results are largely consistent with the comprehensive results reported by Aune et al. (2013), who also report no consistent positive effects. They are weaker than we originally reported in Johnson and Zajonc (2006), but our current data and analysis is far more comprehensive. We fail to document a robust effect on the control of corruption indicator, as reported by Ohler, Nunnenkamp and Dreher (2011), although we do not replicate their empirical strategy exactly or attempt to estimate anticipatory effects. Looking at these results in their totality, the quantitative evidence for a large, broad-based MCC incentive effect appears weak.

The best evidence for an MCC incentive effect remains qualitative and country specific. The results reported by Parks and Rice (2013) and Parks (forthcoming) suggest that MCC eligibility rules are an important metric used by policymakers in developing countries. But the picture may be more complex than the simple view of policymakers pursue MCC eligibility solely to gain financial assistance. Parks and Rice (2013) find the Millennium Development Goals and Doing Business indicators are also highly influential and that already eligible countries cite the MCC as an important influence more often than countries seeking eligibility. Parks (forthcoming) reports that the majority of MCC-inspired policy reforms occur in compact and threshold countries and that the influence of MCC eligibility rules extends from the agenda-setting phase of the policymaking process all the way to the reform implementation phase. The MCC eligibility rules may therefore play an important role in focusing attention on reform and setting a clear mission for MCC compacts, in addition to any direct incentive effect of financial assistance.

Our results do not speak to the overall merit of the MCC. The MCC has attempted to innovate broadly – promoting country-led solutions followed by rigorous evaluations to encourage continued learning. If the economic rate of return on MCC compacts is greater than traditional aid players or the rate of learning is high, any MCC incentive effect, whether on outcomes measured by MCC indicators or the broader policymaking process, would be an extra benefit not the defining metric of success.



## References

- Alesina, Alberto and Beatrice Weder. 2002. "Do Corrupt Governments Receive Less Foreign Aid?" *American Economic Review* 92(4):1126–1137.
- Aune, Ingrid, Yanyan Chen, Christina Miller and Joshua Williams. 2013. "The MCC Incentive Effect: Quantifying Incentives for Policy Change in an Ex-Post Rewards System." *Working Paper: University of Madison* .
- Burnside, Craig and David Dollar. 2000. "Aid, Policies, and Growth." *American Economic Review* 90(4):847–868.
- Danilovich, John. 2006. "Millennium Challenge Corporation Public Outreach Meeting on Benin's Millennium Challenge Compact."  
**URL:** <http://www.mcc.gov/documents/transcripts/transcript-022306-beninoutreach.pdf>
- Dreher, Axel, Peter Nunnenkamp and Hannes Ohler. 2010. "Why it pays for aid recipients to take note of the Millennium Challenge Corporation: Other donors do!" *Center for European, Governance and Economic Development Research Discussion Papers* (99).
- Dunning, Casey. 2013. MCC's New Selection System: Three New Indicators, Two Hard Hurdles, One Muddied Mandate? Technical report Center for Global Development.
- Easterly, William, Ross Levine and David Roodman. 2004. "Aid, Policies, and Growth: Comment." *American Economic Review* 94(3):774–780.
- foreignassistance.gov. 2013.  
**URL:** <http://www.foreignassistance.gov/web/default.aspx>
- Hahn, Jinyong, Petra Todd and Wilbert Van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69(1):201–209.
- Imbens, Guido and Tristan Zajonc. 2011. Regression discontinuity design with multiple forcing variables. Technical report Working Paper.
- Imbens, Guido W. and Karthik Kalyanaraman. 2008. "Optimal bandwidth choice for the regression discontinuity estimator." *Harvard University: Working Paper* .
- Imbens, G.W. and T. Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics* 142(2):615–635.
- Johnson, Doug and Tristan Zajonc. 2006. "Can Foreign Aid Create an Incentive for Good Governance? Evidence from the Millennium Challenge Corporation." *Working Paper* .
- Lee, David S. Lee Thomas Lemieux David S. and Thomas Lemieux. 2009. "Regression Discontinuity Designs in Economics." *NBER Working Paper 14723* .
- MCC. 2013a. "The MCC Effect."  
**URL:** <http://www.mcc.gov/documents/reports/issuebrief-2013002131301-mcc-effect.pdf>

- MCC. 2013*b*. Report on the Criteria and Methodology for Determining the Eligibility of Candidate Countries for Millennium Challenge Account Assistance in Fiscal Year 2014. Technical report Millennium Challenge Corporation.
- Mosley, Paul, John Hudson and Arjan Verschoor. 2004. "Aid, Poverty Reduction and the 'New Conditionality' \*." *The Economic Journal* 114(496):F217–F243.
- Neyman, J. 1923. "On the application of probability theory to agricultural experiments. Essay on principles." *Roczniki Nauk Rolniczych* 10:1–51.
- Ohler, Hannes, Peter Nunnenkamp and Axel Dreher. 2011. "Does Conditionality Work? A Test for an Innovative US Aid Scheme." *CESifo Working Paper No. 3454, Category 2: Public Choice* .
- Parks, Bradley C. forthcoming. Brokering Development Policy Change: The Parallel Pursuit of Millennium Challenge Account Resources and Reform PhD thesis The London School of Economics and Political Science.
- Parks, Bradley C. and Zachary J. Rice. 2013. Measuring the Policy Influence of the Millennium Challenge Corporation: A Survey-Based Approach. Technical report Williamsburg, VA: The College of William and Mary.
- Rubin, D.B. 1978. "Bayesian Inference for Causal Effects: The Role of Randomization." *The Annals of Statistics* 6(1):34–58.
- Thistlethwaite, D. and D. Campbell. 1960. "Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment." *Journal of Educational Psychology* 51(6):309.
- World Bank. 2007. Celebrating Reforms 2007: Doing Business Case Studies. Technical report Washington, DC: The World Bank.
- World Bank. 2013. "World Development Indicators."  
**URL:** <http://data.worldbank.org/data-catalog/world-development-indicators>

# EVALUATION OF MANAGEMENT COMMENTS

MCC provided written comments on the report that are in Appendix III. MCC agreed that an MCC incentive effect is operative in certain specific instances. MCC also agreed that the incentive effect does not affect every developing country, but rather a handful of truly committed, reform-minded governments. MCC noted some of the difficulties and complexities involved with researching the MCC incentive effect.

# MANAGEMENT COMMENTS



August 20, 2014

Office of the Inspector General  
1401 H Street, NW  
Washington, DC 20005

Dear Inspector General:

MCC appreciates the IG's efforts to review previous case studies, empirical research, and surveys on the MCC incentive Effect, as well as conduct a new empirical analysis on whether the MCC Incentive Effect occurs frequently enough to be detected by statistical analysis.

Based on MCC's decade of experience - and review of empirical literature on the topic - MCC agrees with the IG's finding that "an MCC incentive effect is operative in certain specific instances." MCC also agrees the MCC incentive effect does not affect every developing country, but rather a handful of truly committed reform-minded governments. MCC is glad that the IG shares its view that there are some instances of specific countries taking on specific reform efforts, as highlighted in third-party literature, surveys, and MCC's own website (<http://www.mcc.gov/documents/reports/issuebrief-2013002131301-mcc-effect.pdf>.)

Consistent with these points of agreement on the country specific-nature of the MCC-effect, MCC recognizes that it is an extraordinary challenge to try and design research methodologies that empirically isolate an "MCC incentive effect." This is because of the small sample size, the tremendous number of exogenous variables that create noise in the data, and the complex process of how policy reforms occur and how they are measured.

Of the two methodologies employed in the new empirical analysis commissioned by the IG, MCC finds the regression discontinuity - which examines countries just below meeting MCC eligibility requirements to those just above meeting the requirements - more conceptually valid. However, it still has serious limitations.

MCC shares the hypothesis that countries just below meeting the MCC eligibility requirements are most likely to be subject to any MCC Incentive Effect. If countries are very close to being eligible for MCC funding, they have the most incentive to make reforms to become eligible. This mirrors MCC's experiences to date. However, the *second* most likely group of countries to be impacted by the MCC eligibility requirements are countries just above the threshold. They are likely to be the most concerned about sustaining or improving scorecard performance over time. The countries where MCC would least

expect to see MCC Incentive Effect are countries which have been well-governed for years and pass the scorecard solidly or countries that perform so poorly on the scorecard that they are not within the realm of MCC consideration. MCC is therefore concerned that the regression discontinuity approach minimizes the MCC Incentive Effect by comparing the *most likely* candidates for reform to the *second most likely* candidates for reform.

As for the difference-in-differences methodology, which compares MCC candidate countries to upper-middle income countries, MCC has greater concerns about the conceptual validity of this approach. The primary concern is that the sample set of countries and the comparator set are so different that it is next to impossible to “hold constant” all exogenous variables in order to tease out an MCC Incentive Effect (or lack thereof.) When comparing a set of countries whose average GNI per capita is less than \$1,500 to a set of countries whose average GNI per capita is over \$6,500, there are likely too many differences between the basic characteristics of these countries to make them credible comparators. Notwithstanding the invention of the MCC scorecard, one can hypothesize many reasons why countries like South Africa, Costa Rica, and Malaysia have outpaced most of Sub-Saharan Africa on policy reform outcomes.

Setting aside specific methodological concerns, however, there is also a larger question about “what is the MCC incentive effect?” As one can see on MCC’s [Issue Brief](#), when MCC examined its experiences with countries’ governments on policy reform, it came to several important conclusions and areas for future research.

- First, as mentioned above, it appears that some countries are incentivized to reform their policies or update data to improve performance on MCC’s scorecard. However, not all countries are incentivized. So how many countries do actually engage in processes or policy reform to improve performance? Is the impact more than the handful of examples showcased in MCC’s literature, but fewer than the number necessary to be detected in a regression analysis? How might MCC or future researchers determine the breadth and depth of the MCC Incentive Effect?
- Second of all, MCC’s Issue Brief and Dr. Park’s paper based on survey data both assume that the MCC Incentive Effect is a process, which may or may not always result in significant improvements on outcome data. Most empirical studies on the MCC Incentive Effect to date focus only on policy outcomes. That is understandable, given that it is the outcome data that appears on our scorecards, but it doesn’t capture reform efforts that have not yet resulted in significant shifts in the outcome data. The lack of measureable outcomes does not preclude the possibility of significant inputs, processes, or outputs linked to the MCC Incentive Effect. Once again, this is a ripe area for further examination: how does one assess these inputs, process, and outputs in a manner more robust than our current Issue Brief?

Again, MCC appreciates the IG’s research, which expands the methodologies used to test whether the MCC Incentive Effect occurs frequently enough to be detected by statistical analysis. This review further strengthens the literature on this subject. The methodological concerns MCC highlights are not meant to highlight flaws by the authors, but rather the difficulties in asking and answering the most robust and relevant questions possible about the MCC Incentive Effect.

Sincerely,

/s/

Sheila W. Herrling

Vice President

Department of Policy and Evaluation

**U.S. Agency for International Development**  
**Office of Inspector General**  
1300 Pennsylvania Avenue, NW  
Washington, DC 20523  
Tel: 202-712-1150  
Fax: 202-216-3047  
<http://oig.usaid.gov>