WPS6186

Policy Research Working Paper 6186

MTEFs and Fiscal Performance

Panel Data Evidence

Francesco Grigoli
Zachary Mills
Marijn Verhoeven
Razvan Vlaicu

The World Bank
Poverty Reduction and Economic Management Network
Public Sector & Governance Unit
September 2012

Abstract

In the last two decades more than 120 countries have adopted a version of a Medium-Term Expenditure Framework (MTEF). These are budget institutions whose rationale it is to enable the central government to make credible multi-year fiscal commitments. This paper analyzes a newly-collected dataset of worldwide MTEF adoptions during 1990–2008. It exploits within-country variation in MTEF adoption in a dynamic panel framework to estimate their impacts. The analysis finds that MTEFs strongly improve fiscal discipline, with more advanced MTEF phases having a larger impact. Higher-phase MTEFs also improve allocative efficiency. Only top-phase MTEFs have a significantly positive effect on technical efficiency.

This paper is a product of the Public Sector & Governance Unit, Poverty Reduction and Economic Management Network. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at zmills@worldbank.org and mverhoeven@worldbank.org.

MTEFs and Fiscal Performance: Panel Data Evidence*

Francesco Grigoli† Zachary Mills‡ Marijn Verhoeven§ Razvan Vlaicu¶

JEL Classification: E62, H51, H62.
Keywords: budget institutions, MTEF, fiscal discipline, allocative efficiency, technical efficiency, dynamic panel data analysis.

* We would like to thank Abdul Abiad, Nataliya Biletska, Jim Brumby, Stefania Fabrizio, Richard Hemming, Malcolm Holmes, Sam Ouliaris, Philip Keefer, Steve Knack, Ajay Tandon, and participants at the PREM Knowledge & Learning Forum and the World Bank MTEF Workshop. Haishan Yuan provided excellent research assistance. The views expressed in this paper are those of the authors and should not be attributed to the World Bank or its affiliated institutions, the IMF, its Executive Board, or its management.
† IMF, Fiscal Affairs Department, 1900 Pennsylvania Ave. NW, Washington, DC 20431, USA. E-mail: fgrigoli@imf.org.
‡ The World Bank, Europe and Central Asia Department, 1818 H Street NW, Washington, DC 20043, USA. E-mail: zmills@worldbank.org.
§ The World Bank, Poverty Reduction and Economic Management Vice Presidency, 1818 H Street NW, Washington, DC 20043, USA. E-mail: mverhoeven@worldbank.org.
¶ University of Maryland, Department of Economics, and Department of Government and Politics, 3105 Tydings Hall, College Park, MD 20742, USA. E-mail: vlaicu@econ.umd.edu.

1 Introduction

It is generally accepted that fiscal performance is a key factor in a country's long-run growth prospects.[1] Understanding the determinants of fiscal performance has thus become a central topic of research. A recent strand of literature has emphasized the role of *budget institutions* in affecting fiscal performance. Budget institutions are the formal rules and procedures according to which budgets are drafted, approved, and implemented. They can take the form of either (i) laws establishing ex ante constraints on the government's conduct of fiscal policy, such as balanced budget and debt ceiling provisions, or (ii) bargaining rules between the executive and the legislature, such as line-item executive veto or legislative amendment rules.

In the last two decades more than 120 countries have adopted laws instituting multiyear fiscal targets, known as *Medium-Term Expenditure Frameworks* (*MTEFs*). First introduced in a small set of developed countries to contain expenditure overruns in the 1980s, MTEFs spread rapidly during the 1990s and 2000s, being in operation in 132 countries at the end of 2008 (see Figure 1 and Table 2). MTEFs translate macro-fiscal objectives and constraints into broad budget aggregates as well as detailed expenditure plans by sector. The rationale of this budget institution is to enable the central government to more adequately incorporate future fiscal challenges into the annual budgets, thereby reducing an undue emphasis on short-term goals.

The key public finance problem that MTEFs are intended to overcome is dynamic fiscal inefficiency.[2] Whether it takes the form of strategic obstruction of future political opponents (Alesina and Tabellini 1990), or electoral manipulation through budget cycles (Drazen 2000, Brender and Drazen 2005), dynamic common pool (Velasco 1999), or time-inconsistent voters (Bisin, Lizerri, and Yariv 2011), government spending and borrowing deviates from the social planner level, resulting in suboptimally high deficits and debt.

Dynamic inefficiency seems particularly inherent in a *yearly* cycle of budget planning and implementation. Wildavsky (1986, p. 317) makes this point as follows:

> "Multiyear budgeting has long been proposed as a reform to enhance rational choice by viewing resource allocation in a long-term perspective. One year, it has been argued, leads to short-sightedness, because only the next year's expendi-

[1] See, e.g., Fischer (1993), Easterly and Rebelo (1993), Easterly, Irwin, and Serven (2008).

[2] Other public finance distortions include static common pool problems (Weingast, Shepsle and Johnsen 1982), rent seeking (Persson, Roland and Tabellini 2000, Besley and Smart 2007), and clientelism (Keefer and Vlaicu 2008).

tures are reviewed; overspending, because huge disbursements in future years are hidden; conservatism, because incremental changes do not open up large future vistas; and parochialism, because programs tend to be viewed in isolation rather than in comparison with their future costs in relation to expected revenue."

At the basis of an MTEF is a commitment by the budget actors to a medium-term, typically two to four years, fiscal trajectory. Thus, it can be seen either as an ex ante constraint on the government, similar to a balanced budget requirement (Alesina and Perotti 1996), or as a "contract approach" to centralizing the budget process through a broad-based political agreement, as opposed to a "delegation approach" where the executive receives enhanced powers (von Hagen and Harden 1995). While the theoretical underpinnings of this institution are well understood, the empirical evidence on its impact is scarce. The main obstacle has been the shortage of data on MTEF adoption. An additional impediment has been the lack of sources of exogenous variation in national budget institutions (Acemoglu 2005).[3]

This paper reports on the first large-sample empirical study of the MTEFs' impacts on fiscal performance. As part of a larger World Bank (2012) study, we collect and analyze MTEF adoption data for a panel of 181 countries over the period 1990-2008, the most comprehensive dataset to date on worldwide MTEF adoption. Following World Bank (2012) we classify MTEFs into three phases, based on their level of sectoral disaggregation, and generate testable hypotheses about the effects of each phase on multiple dimensions of fiscal performance.

The rich time variation in the data allows us to model the dynamics of the fiscal adjustment process, as well as address the endogeneity of MTEF to fiscal performance. We use a Difference Generalized Method of Moments (D-GMM) approach to estimate dynamic panel data models of fiscal performance. These models are designed for "small T large N" panels and, when correctly applied, generate valid internal instruments that hold the promise of overcoming identification issues due to the absence of strong external instruments, a typical hurdle with country-level data.[4]

When an MTEF is implemented well we should observe (i) spending that is limited by resource availability (fiscal discipline), (ii) budget allocations that reflect spending pri-

[3] One way to circumvent econometric identification issues has been to study similar institutions operating at *sub-national* levels of government. See Besley and Case (2003) for a review of the literature that has employed U.S. state-level data.

[4] The D-GMM approach was first proposed by Holtz-Eakin, Newey, and Rosen (1988) and later developed by Arellano and Bond (1991). Recent refinements include Windmeijer (2005) and Roodman (2009).

orities (allocative efficiency), and (iii) public goods delivery that is cost effective (technical efficiency). We classify MTEFs into three *phases*: Medium-Term Fiscal Framework (MTFF, which establishes the aggregate resource envelope), Medium-Term Budget Framework (MTBF, which focuses on the allocation of spending across sectors, programs and agencies) and Medium-Term Performance Framework (MTPF, which sets sectoral performance targets). These three MTEF phases are "nested": an MTPF contains an MTBF, which in turn contains an MTFF.[5]

The new data reveal patterns in the timing of MTEF adoption across regions and levels of development. OECD countries were the first to adopt MTEFs, and by the early 1990s most countries in this group had an MTPF in place. The bulk of MTEF reforms in Sub-Saharan African countries took place in the 1990s, Latin American countries adopted MTEFs in the late 1990s, and Eastern Europe and the former soviet republics joined the trend in the 2000s. We exploit differential patterns of MTEF adoption across regions to construct external instruments that complement the standard GMM-style internal instruments.

The empirical results show that MTEFs strongly improve fiscal discipline and that the effect is larger the more advanced the MTEF phase. The effect varies between 1 and 3 percentage points of central government balance as a percentage of GDP. We also find that MTBFs and MTPFs improve allocative efficiency, measured by the volatility in per capita health spending in purchasing power parity dollars (PPP$), and that MTPFs contribute to technical efficiency, measured as technical efficiency scores from a stochastic frontier model of public health delivery. These results are robust to excluding highly autocratic and highly developed countries.

Our results are more favorable about MTEF effectiveness than the conclusions of prior work. Bevan and Palomba (2000), La Houerou and Talierco (2002), Holmes and Evans (2003), and Oyugi (2008), based on case studies of about a dozen African countries, conclude that the budget process has generally not improved after the adoption of an MTEF, while McNab, Martinez-Vasquez, and Boex (2000) and Oxford Policy Management (2000) raise questions of adequate implementation. However, Gleich (2003) and Ylaoutinen (2004) find that MTEFs in Central and Eastern Europe alleviated the deficits and debts that emerged in the second half of the 1990s. Wescott (2008) and Filc and Scartascini (2010), using data from Central and Latin America, found mixed results and emphasized the importance of piloting MTEFs in areas where they are likely to deliver the largest payoff.[6]

[5]This taxonomy follows World Bank (2012) and borrows from Castro and Dorotinsky (2008) with the nesting concept as an added innovation.

[6]Drawing on extensive operational experience with MTEF implementation in developing countries,

Alesina, Hausmann, Hommes, and Stein (1999) and Fabrizio and Mody (2006) include MTEFs in aggregate indexes of budget institutions using data from Latin America and Eastern Europe, respectively.[7] Our results complement this empirical literature by providing evidence of MTEF impacts from a worldwide sample. While previous studies rely on small samples and either cross-sectional or static panel models our empirical methodology takes advantage of the time variation to estimate dynamic panel models. We also propose a new classification of MTEF phases, based on the level of disaggregation of the central government's fiscal objectives.

Our paper is also related to the broader empirical literature on budget institutions.[8] The most studied institutions have been balanced budget amendments, debt ceilings, tax and expenditure limitations, and supermajority requirements for tax increases. Examples include: Bayoumi and Eichengreen (1995), von Hagen and Eichengreen (1996), Stein, Talvi, and Grisanti (1999), Hallerberg and von Hagen (1999), Perotti and Kontopoulos (2002), Fatas and Mihov (2003), von Hagen and Wolff (2006), and Hallerberg and Ylaoutinen (2010), all using country-level data; and Poterba (1994), Bohn and Inman (1996), Kiewert and Szakaly (1996), Poterba and Rueben (1999), Knight (2000), Knight and Levinson (2000), and Fatas and Mihov (2006) using state-level data. Important lessons from this literature are that numerical constraints have limited effectiveness because they can be circumvented, that the effect of reduced fiscal discretion on macroeconomic volatility remains an open question, and that the political environment matters for the effectiveness of budget institutions.

We contribute to this literature by proposing a dynamic panel approach (difference GMM) that models the fiscal adjustment process while at the same time addressing the issue of institutional endogeneity. Apart from employing standard internal instruments, we also propose external instruments that improve estimation efficiency, based on the time-varying degree of MTEF penetration in a country's geographic region. Also, in addition to macroeconomic effects, which have been the focus of this literature, we provide evidence of sectoral effects.

The structure of the paper is as follows. Section 2 provides a background discussion of the MTEFs, their adoption trends, and their expected effects on fiscal performance. Section 3 discusses the data and the empirical strategy. Section 4 presents the empirical results. Section 5 summarizes the paper and suggests directions for future research.

Schiavo-Campo (2009) puts forward conceptual arguments supporting a gradual introduction of these institutions and emphasizes the potential downsides of instant reform.

[7] In these two papers the MTEF component of the index is weighted by 1/10 and 1/12, respectively.

[8] See the NBER volume edited by Poterba and von Hagen (1999), as well as von Hagen (2006), for reviews of the budget institutions literature.

2 Background

This section takes a first look at the worldwide MTEF adoption data collected for this paper by presenting stylized facts of MTEF global growth and regional penetration during 1990-2008. It also discusses the rationale behind MTEFs as budget institutions designed to improve fiscal performance. This discussion helps generate theoretical expectations. We subject these conjectures to empirical scrutiny in the next section.

2.1 MTEF Phases

We classify MTEFs into three phases, based on the following criteria.[9]

- Medium-Term Fiscal Framework (MTFF): the government has rolling aggregate, expenditure, revenue, and other fiscal forecasts. Features include the availability of a macro-fiscal strategy, macroeconomic and fiscal forecasts, and debt sustainability analysis.
- Medium-Term Budgetary Framework (MTBF): budget, spending agency or other reports explain aggregate and sectoral expenditure objectives and strategies, budget circulars detail medium-term expenditure ceilings and revenue forecasts, and budget documents contain some detail about medium-term estimates.[10]
- Medium-Term Performance Framework (MTPF): budget, spending agency or other reports explain program objectives and strategies, and list specific agency and/or program output or outcome targets, as well as results.

These three phases are "nested" in the sense that a higher-phase MTEF contains the lower-phase MTEF just below it.

2.2 Stylized Facts

Although some forms of medium-term expenditure projections existed in OECD countries as early as the 1960s, the first application of a coherent system of multiyear budgeting occurred in Australia, where an MTEF was introduced in the 1980s (see Folscher 2007). MTEFs

[9] We use these definitions to code each country-year in our sample as falling into each of these mutually exclusive categories; see Section 3 and the Data Appendix for a description of variable construction.

[10] We coded countries that introduced a "pilot" MTBF in a few sectors as MTFF since the health sector, our focus in the analysis below, might not be one of the piloted sectors.

have since been adopted by a large number of low and middle-income countries as a central element of public financial management reform.

While MTEFs began to spread across industrial countries and Africa in the early 1990s, it was not until the late 1990s and 2000s that they took off in emerging market economies; see Figure 1. An average of 10 countries per year introduced an MTEF between 1996 and 2008. By the end of 2008, 132 countries, or about two-thirds of the globe, had an MTEF. Figure 2 shows temporal patterns of MTEF growth by continent. The regional trends are evident, with Europe leading the pack, followed by Africa and the Americas.

Initially, most MTEFs were of the first phase, or MTFF, and until recently about two-thirds of the increase in MTEFs has been in the form of new MTFFs. However, there has been a recent uptick in the number of MTBFs and MTPFs. In 2008 there were 71 MTFFs, 42 MTBFs, and 19 MTPFs. Table 2 shows that the shift to MTBFs and MTPFs has been mainly through transitions from a lower MTEF phase to a higher one.[11]

MTEF coverage varies significantly. Advanced economies had achieved almost complete coverage (96%) by the end of our sample period. MTEF adoption in advanced countries occurred in two waves. In the late 1980s and early 1990s only a few advanced economies followed Australia's lead in MTEF adoption. Then, in the late 1990s MTEFs were introduced in the European Union to support budgetary targets set as pre-conditions for monetary union. By the end of 2008, 46% of the MTEFs in advanced economies were MTPFs. The relatively low fraction of the second-phase MTEF (the MTBFs) in these countries suggests that when advanced economies decide to move beyond an MTFF introducing a performance focus is a natural development, reflecting their more sophisticated budgeting systems.

MTEFs have also achieved broad coverage in Europe and Central Asia. The spread was more rapid and consistent in Central and Eastern Europe than in the Former Soviet Union. This may reflect efforts made in Central and Eastern Europe for quick integration with Western Europe.[12] Building on an early start in Botswana and Uganda, MTEFs spread rapidly across Sub-Saharan Africa in the 2000s. MTEFs are more numerous in Francophone Africa than Anglophone Africa. MTEFs have also been adopted by most countries in South Asia, with Nepal and Sri Lanka having implemented MTBFs.

MTEFs are less widespread in other regions, despite a recent spurt of adoptions in East Asia and the Pacific, including MTBFs in Cambodia and Thailand. The picture is similar

[11] Three countries (Bulgaria, Canada and Norway) performed a full transistion - from an MTFF to an MTBF to an MTPF - during this period.

[12] The countries remaining without an MTEF in this region are: Azerbaijan, Belarus, Montenegro, and Turkmenistan.

in Latin America and the Caribbean, where a number of countries have introduced MTFFs following years of managing fiscal policy under IMF programs. Only four countries have moved beyond an MTFF and introduced an MTBF: Argentina, Colombia, Nicaragua and St. Lucia, although Brazil's budgeting system has recognizable MTBF characteristics. In the Middle East and North Africa MTEFs are a very recent innovation. Only Algeria and Jordan have an MTBF, while major oil exporting countries such as Saudi Arabia and United Arab Emirates, as well as Egypt, had not yet adopted MTEFs.

Despite pronounced differences between regions patterns of MTEF adoption have been relatively uniform across income and development levels. Apart from the widespread adoption of MTEFs in high-income countries, there is little difference in penetration across upper middle, lower middle, and lower-income countries. MTEF adoption does not appear to follow a monotonic relationship with respect to income per capita or the human development index; see Figures 3 and 4.

2.3 MTEF Objectives

MTEFs represent a multiyear approach to budgeting that addresses the shortcomings of annual budgeting noted above in the Introduction. Most public programs require funding and yield benefits over a number of years, but annual budgeting largely ignores future costs and benefits. Annual budgets take as their starting point the previous year's budget and modify it in an incremental manner, making it difficult to re-prioritize policies and spending.[13] MTEFs take a strategic forward-looking approach to establishing spending priorities and resource allocation. They also look across sectors, programs and projects to see how spending can be restructured to best serve national objectives, which contrasts with the narrow self-interest of spending agencies and beneficiaries that dominates resource allocation under annual budgeting (World Bank 1998).

Insofar as an MTEF constrains spending to resource availability, makes budget allocations reflect spending priorities, and generates cost effectiveness in the delivery of public goods and services, it should contribute directly to fiscal discipline, allocative efficiency, and technical efficiency.[14] Moreover, there are synergies among these three dimensions of fiscal

[13]While incremental budgeting can work well in times of revenue growth, it comes under particular pressure when revenue falls, becomes more volatile, or reaches its natural limit. In these instances expenditure prioritization takes on increased importance.

[14]There is also a link to broader economic development. With improved fiscal outcomes, growth should be higher, inflation lower, and macroeconomic volatility reduced. Moreover, as the quality of spending improves, higher incomes should be accompanied by lower poverty rates, while better infrastructure should contribute to even higher growth and further poverty reduction.

performance. With fiscal discipline secured, governments should be free to focus on the microeconomic challenges of improving spending efficiency and not preoccupied with having to address the adverse macroeconomic consequences of persistent fiscal imbalances.[15] It should also be easier to maintain fiscal discipline when improvements to both allocative and technical efficiency reduce abuse and waste. Moreover, against a background of fiscal discipline, new expenditure needs are more likely to prompt spending reallocations as opposed to requests for additional funding. Finally, both fiscal discipline and expenditure efficiency create fiscal space for productive spending on economic and social infrastructure, and for responding to fiscal risks.

MTFFs can promote *fiscal discipline* by addressing the root causes of deficit bias. By specifying an overall "top-down" resource constraint, an MTFF reins in the political tendency to over-commit resources (the common pool problem). By imparting a medium-term perspective to budgeting and taking into account the future fiscal costs of government policies and programs, an MTFF can fill information gaps that allow politicians to renege on commitments to implement affordable policies (the time consistency problem). A medium-term perspective also encourages governments to conduct discretionary stabilization in a symmetric, counter-cyclical manner, rather than asymmetrically which leads to rising deficits and debt (Kumar and Ter-Minassian 2007).[16]

Since MTBFs and MTPFs incorporate an MTFF, they should have a stronger effect on fiscal discipline compared to an MTFF alone. This is in part because countries that have the administrative capacity to implement these higher phases will likely also have greater fiscal discipline. But it is also a consequence of better prioritization and more emphasis on performance, which can bring the payoff to fiscal discipline into sharper focus.

Prioritization guided by longer-term sector strategies should improve *allocative efficiency*. Insofar as spending agencies prepare sector strategies, identify their resource needs, and allocate their budgets according to strategic priorities, this "bottom-up" prioritization should produce a shift to spending with higher economic and social returns. However, the full payoff to prioritization requires that choices are also made as to how resources should be

[15] It can be argued that, in fact, large fiscal imbalances prompt better expenditure prioritization; however, the lessons from fiscal adjustments around the world is that spending cuts are borne disproportionately by high-priority spending, and especially public investment in infrastructure, with adverse consequences for future growth (Easterly, Irwin, and Serven 2008). Lewis and Verhoeven (2010) report that the growth of social spending has dipped as the global financial crisis has put fiscal positions under pressure, which risks setting back achievement of human development goals, because these depend on the rapid spending increases achieved in the 1990s and the earlier part of the 2000s.

[16] On the downside, if spending agencies view MTEFs as minimum entitlements, rather than constraints, ceilings, or forward estimates, MTFFs could actually be a source of fiscal indiscipline and deficit bias (Schick 2010).

allocated across sectors, which is done as part of the reconciliation between the "top-down" and "bottom-up" approaches involving a lead agency, normally the Ministry of Finance, and spending agencies, and in connection with which less strategic guidance may be available, especially in the absence of national medium-term planning.[17]

The outcome of effective prioritization should be a change in the allocation of spending. In the short term, spending volatility by sector may increase following MTEF implementation as spending is reallocated to more productive sectors and programs. Thereafter, insofar as spending decisions are guided by strategic priorities with a longer-term focus, sectoral spending should become less volatile, especially in the high-priority areas of health and education. The payoff coming from an MTBF should be even higher with an MTPF since this last phase goes further by setting within-sector and within-program performance targets.[18]

A third dimension of fiscal performance is *technical efficiency*. The better the economic and social outcomes achieved by spending programs from a given amount of budget resources, or the fewer resources used to achieve given outcomes, the more technically efficient is government spending. Improved technical efficiency may follow from an MTFF, but is more likely a consequence of an MTBF and MTPF, with the latter possibly having the largest effect as budgets are linked to results in the form of outcomes or outputs.

Based on these considerations we state the expected MTEF effects on fiscal performance in the following hypotheses (see also World Bank 2012):

(H1) MTFF, MTBF, and MTPF improve fiscal discipline, with higher-phase MTEFs having larger effects.

(H2) MTBF and MTPF improve allocative efficiency, with MTPF having a larger effect.

(H3) MTPF improves technical efficiency.

The rest of the paper examines the evidence for these conjectures.

3 Data and Empirical Strategy

This section discusses the choice of variables for the empirical analysis and takes a first look at the statistical properties of our data. It then outlines our empirical strategy for identifying and estimating the MTEFs' effects on fiscal performance.

[17] Moreover, difficult decisions have to be made to cut low-priority, but often politically sensitive, spending.

[18] A shift away from unproductive spending should also be observed. Poor-quality investment, distortionary and untargeted subsidies, bloated civil services, and the like should not survive scrutiny under the MTEF, while productive spending on economic and social infrastructure, health and education services, and other growth- and development-promoting activities should be favored.

3.1 Data

The dataset contains both cross-sectional and time-series variation in MTEF presence. The sample consists of 181 countries over the period 1990-2008. The country sample reflects data availablility on MTEF status. The period sample reflects data availability on public finances. Here we briefly discuss the key variables. Section A.3 of the Appendix contains the complete list of variables together with their data sources.

The construction of the MTEF indicators relied upon an extensive data collection effort as no single type of document sufficiently describes the existing institutional arrangements for all countries or even individual countries. Thus, the data were compiled from a large number of sources, including IMF Article IV country reports, IMF Reports on the Observance of Standards and Codes (ROSC), fiscal transparency modules, World Bank Public Expenditure Reviews (PERs), World Bank Country Financial Accountability Assessments (CFAAs), OECD documents, donor case studies, and country websites. Additionally, World Bank and IMF public sector specialists supplemented the above information with technical details.[19]

We measure fiscal discipline, an indicator of macro fiscal performance, by the central government's overall balance. Although the literature suggests alternative indicators, e.g., primary balance and debt, data availability limited the choice to the overall balance. Moreover, by including government borrowing, the overall balance is a good indicator of the state of public finances.[20]

Allocative efficiency does not have a universally accepted definition. Potential proxies for allocative efficiency are budget composition volatility and volatility of core spending (health and education). Since volatility in these sectors jeopardizes long-term objectives, health care and public education spending should be largely unaffected by short-term fluctuations in GDP. In other words, allocative efficiency implies that spending in core sectors where needs are fairly constant does not behave in a volatile manner. Given data constraints and the requirement that the public good category should be reasonably comparable across countries, we choose to work with the volatility of per-capita health spending, in PPP$. We define volatility of a time series $y_{i,t}$ for country i as the absolute yearly growth rate of the

[19]For the purposes of this paper, we refrain from making judgments to distinguish between an MTEF present in the law (de jure) and a well-functioning MTEF (de facto). Such a distinction would introduce a significant amount of subjectivity into the analysis.

[20]It could be argued that the overall balance does not account for the effect of inflation on interest payments and that interest payments are a function of the accumulated debt and not the present fiscal stance.

detrended series:

$$Volatility_{i,t} = \left| \log \frac{\tilde{y}_{i,t}}{\tilde{y}_{i,t-1}} \right| \times 100 \qquad (1)$$

where $\tilde{y}_{i,t} = y_{i,t} - \left(t - \frac{1}{T}\sum_{k=1}^{T} k\right)$ is the detrended series for $y_{i,t}$.

Technical efficiency is typically measured using technical efficiency scores from a Stochastic Frontier Analysis (SFA). This is the approach we adopt. The SFA approach relies on a reduced-form relationship between inputs and outputs. The country with the highest health output after controlling for inputs is the most efficient, and the efficiency level of the other countries is measured with respect to the most efficient country.[21]

In particular, we compute technical efficiency scores in the health sector using a parsimonious version of the model estimated in Greene (2005). The outcome of interest is life expectancy, and the input is health spending per capita. The model is:

$$\begin{aligned} \log(Life\_Exp_{i,t}) &= \beta_0 + \beta_1 \log(Health\_Spend_{i,t}) + \\ &\quad + \beta_2 Density_{i,t} + \beta_3 OECD_{i,t} + \tau_t + v_{i,t} - u_{i,t} \end{aligned} \qquad (2)$$

where τ_t's are year fixed effects, $v_{i,t} \sim N(0, \sigma_v)$, and $u_{i,t} = |U_{i,t}| \sim Exp(0, \sigma_u)$. The controls are population density and OECD membership.[22]

The parameters are estimated by maximum likelihood. The estimates of $v_{i,t} - u_{i,t}$ are translated into an estimate of $u_{i,t}$ using the standard Jondrow, Materov, Lovell, and Schmidt (1982) formula. Technical efficiency then is simply:

$$Tech\_Efficiency_{i,t} = e^{-\hat{u}_{i,t}} \qquad (3)$$

where $\hat{u}_{i,t}$ is the maximum likelihood estimate of $u_{i,t}$. Table 4 presents the estimation results. The coefficients follow the same pattern noticed in prior work using different data. The asymmetry parameter $\lambda = \sigma_u/\sigma_v$ is also within the range of variation reported previously. Estimated mean efficiency is 86.48.

Following Baltagi, Demetriades, and Law (2009), who study the effect of financial openness on banking sector development, we introduce MTEF regional penetration as external

[21] The SFA was inspired by Farrell (1957), who defined technical efficiency as the ability to produce the maximum possible output from a given set of inputs, and measured it as the difference between maximum attainable output and observed output. Inefficiencies might arise from waste or because the most cost-effective set of programs is not implemented.

[22] Greene (2005) also includes education spending per capita, as an input, and controls for government voice and accountability, government effectiveness, share of government financing, the Gini coefficient, and GDP per capita.

instruments to supplement the internal GMM-style instruments. In a region consisting of n countries, penetration for country i is defined as the fraction of countries in the surrounding region that already have an MTEF:

$$MTEF\_Regional\_Penetration_{i,t} = \frac{\sum_{j=1,j\neq i}^{n} MTEF_{j,t}}{n-1} \tag{4}$$

for each of the MTEF phases. We use the classification of the world into the twenty-two geographic regions defined by the United Nations Statistics Division.

Table 1 reports summary statistics. All variables display considerable variation both between and within countries, justifying the use of panel estimation techniques. An exception is MTPF, which has small within variation due to the few adoptions of this top phase, and OECD membership. Table 3 reports pairwise correlations between the main variables. The correlation coefficients are within plausible ranges and support our choices of regressors in the next subsection.

3.2 Empirical Strategy

Estimating the impact of a budget institution on fiscal performance requires that we address several identification challenges: reverse causality, omitted variable bias, and errors-in-variables.

First, reverse causality arises because fiscal stress, e.g., a financial crisis, may have prompted a country to restrain spending, adopt an MTEF, or strengthen an existing one. Von Hagen (2006, p. 474) notes that "Historical experience suggests that governments make efforts to centralize the budget process to overcome sharp fiscal crises." If MTEFs have positive effects on fiscal performance, and poor fiscal performance increases the probability of adopting an MTEF, then the reverse causality bias is probably negative. In this case, the estimates are still useful as a lower bound for the actual effect.[23]

Second, omitted variable bias arises due to the failure to account for a factor that affects both the adoption of an MTEF and fiscal performance. For instance, strong economic growth may reduce the pressure on a government to reform budget institutions, and, at the same time, improve the government's fiscal outcomes, thus leading to negative omitted variable bias. As suggested by Fabrizio and Mody (2006), a partial solution to this problem is to

[23] The endogeneity of budget institutions with respect to fiscal performance is extensively discussed in the cross-country literature (see Alesina and Perotti 1999, Stein, Talvi, and Grisanti 1999, Perotti and Kontopoulos 2002, and Fabrizio and Mody 2006) yet none of these papers proposes an instrument that influences the probability of fiscal reform while being exogenous to fiscal performance.

use within-country variation in fiscal institutions. This approach, in effect, eliminates the unobservable country specific fixed effects that may influence budget deficits. The problem of omitted variables is thus alleviated; however it is not eliminated.[24,25]

Finally, if some of the variables in the analysis are not measured accurately, there is the potential for errors-in-variables bias, which usually dampens the effect of interest. Although in our empirical model the primary explanatory variable, MTEF status, can be observed with a reasonably high degree of precision, there is still scope for measurement error in the other explanatory variables.

Our empirical strategy exploits within-country variation in fiscal institutions. Annual data allow us to account for the possibility that observed fiscal performance in a given year may not represent long-run equilibrium values, because of incomplete adjustment in other variables. For example, as revenues cannot be perfectly anticipated budget balance in a given year fluctuates around the equilibrium balance level. To allow for the possibility of partial adjustment, we use a dynamic specification with a lagged dependent variable:

$$y_{i,t} = \alpha + \sum_{l=1}^{L} \beta_l y_{i,t-l} + \delta_1 MTFF_{i,t} + \delta_2 MTBF_{i,t} + \delta_3 MTPF_{i,t} + \boldsymbol{\gamma} \mathbf{x}_{i,t} + \varepsilon_{i,t} \tag{5}$$

where $y_{i,t}$ is a measure of fiscal performance, L is the number of lags of the dependent variable, $MTFF_{i,t}, MTBF_{i,t}, MTPF_{i,t}$ are indicators of the three MTEF phases, $\mathbf{x}_{i,t}$ is a vector of covariates, and $\varepsilon_{i,t}$ is an error term that contains country and year fixed effects:

$$\varepsilon_{i,t} = \varphi_i + \tau_t + \epsilon_{i,t} \tag{6}$$

with the idiosyncratic error $\epsilon_{i,t}$ assumed to be mean zero.

While this dynamic panel formulation allows for a richer model of fiscal adjustment, the presence of a lagged dependent variable introduces new sources of endogeneity. First, without controlling for the fixed effects, the model in equation (5) has a built-in positive bias in the first lag of the dependent variable.[26] Second, the differenced version of the equation eliminates this bias, but has a built-in negative bias due to the fact that $y^*_{i,t} = \Delta y_{i,t}$ is negatively correlated with $y^*_{i,t-1} = \Delta y_{i,t-1}$. Thus an unbiased estimate of the lagged

[24] Most studies have not been able to use this method because either budget institutions do not change much over time or because changes are difficult to measure. When it has been implemented with U.S. state-level data (e.g., Knight and Levinson 2000) the results are typically different, indicating that the problem of omitted variables is relevant.

[25] Additional omitted variables could include political institutions. Evidence from Europe shows that institutional design responds to political factors and events (Hallerberg, Rolf, and von Hagen 2009).

[26] Nickell (1981) has shown that this "dynamic panel bias" disappears only when T approaches infinity.

dependent variable coefficient should lie in the range between the FE estimate and the OLS estimate. This bracketing range thus provides a natural specification check (Bond 2002).

The D-GMM solution to this endogeneity problem is to instrument for $\Delta y_{i,t-1}$ using $y_{i,t-2}$ and possibly earlier lags. Notice that the lagged values of the dependent variable are useful instruments unless $y_{i,t}$ is close to a random walk, in which case past levels convey little information about future changes. Table 5 presents unit root test results for the three measures of fiscal performance. The IPS test statistic safely rejects the null of unit root in each case.[27]

In the same fashion, one can instrument $\Delta MTFF_{i,t}$, which is also endogenous to $\Delta y_{i,t}$, with $MTFF_{i,t-2}$ and possibly earlier lags (and similarly for the higher MTEF phases). In this way D-GMM generates internal instruments for the budget institutions. To improve estimation efficiency we supplement the internal instruments with external instruments based on MTEF penetration in a country's geographic region, computed according to equation (4). The external instruments are inspired by the regional penetration patterns noted in the previous section; see Subsection 2.2.[28]

Because our panel is unbalanced, we use the orthogonal deviations transform (Arellano and Bover 1995) in the baseline specifications: $y^*_{i,t} = \sqrt{\frac{T_{i,t}}{T_{i,t+1}}}\left(y_{i,t} - \frac{1}{T_{i,t}}\sum_{s>t} y_{i,s}\right)$, where $T_{i,t}$ is the number of available future observations; we also report estimates based on the difference transform $y^*_{i,t} = y_{i,t} - y_{i,t-1}$ in the alternative specifications. Orthogonal deviations, instead of subtracting the previous observation from the current observation, subtracts the average of all future available observations from the current observation. This maximizes sample size in panels with gaps.

The moment conditions are based on orthogonality between the transformed errors and the lagged values of the dependent variable. To test for this condition one can run two diagnostics, namely first-order and second-order serial correlation in the idiosyncratic error $\epsilon_{i,t}$. The test should reject the null of no first-order serial correlation, and not reject the absence of second-order serial correlation. We include year fixed effects in all specifications to strengthen the case for the assumption of no correlation in the idiosyncratic errors across countries.

[27] An alternative approach would be to use the System GMM estimator proposed by Arellano and Bover (1995) and Blundell and Bond (1998); this estimator requires an additional identifying assumption, namely that first differences of the instrumenting variables are exogenous to the fixed effects. This is the strategy we employ in World Bank (2012). There we also include a larger set of covariates which should mitigate omitted variable bias, but may also increase the possibility of additional endogeneity bias. The consistency in findings between the two approaches adds to the robustenss to our results.

[28] By themselves the external instruments are not strong enough to justify a diff-in-diffs IV strategy. The lack of strength is driven in particular by insufficient variation in MTPF adoption.

The number of moment conditions increases with T. To test for over-identifying restrictions we use a Hansen J test. Too many moment conditions introduce bias while increasing efficiency. Thus, it is important to keep the number of internal instruments in check. In the baseline specifications we use one lag for each lagged dependent variable, and two lags for each MTEF indicator. In the alternative specifications we also report results with collapsed instrument matrix, as recommended by Roodman (2009).[29]

To restrain the number of instruments we only include covariates that satisfy two conditions: (i) have the potential to act as conduits in transmitting regional trends, and (ii) are not endogenous to the dependent variable, at least in the short run. The first condition is needed because our external instruments are regional variables. The second condition simply insures that we do not have to introduce internal instruments for the covariates as well. Openness and conflict are one economic and one political variable that satisfy these criteria and are available for our full sample, thus $\mathbf{x}_{i,t} = (Openness_{i,t}, Conflict_{i,t})$ in equation (5).

Our baseline specification computes two-step D-GMM estimates with standard errors corrected with the Windmeijer (2005) procedure.[30] We also report alternative specifications with one-step D-GMM estimates, in which case we report cluster-robust standard errors, i.e., robust to heteroskedasticity and arbitrary patterns of correlation within countries.

4 Estimation Results

The main results of the paper are contained in Tables 6 through 11. The tables contain estimates of fiscal discipline (Tables 6 and 7), allocative efficiency (Tables 8 and 9), and technical efficiency (Tables 10 and 11) regressions using the dynamic panel model discussed in the previous section. In each case the first table presents the baseline results, namely two-step D-GMM with Windmeijer corrected standard errors. In these specifications we treat the MTEF indicators as endogenous. The second table presents alternative specifications, namely one-step robust standard errors, collapsed instruments, difference transform, and treatment of the MTEF variables as lagged and predetermined instead of contemporaneously endogenous. We preserve the structure of both the baseline and the alternative specifications across all three measures of fiscal performance.

[29] In large samples collapsing the instruments reduces statistical efficiency, however in small samples it may alleviate the bias created when the number of instruments times the number of periods approaches the number of panel units.

[30] The two-step standard error correction is needed because the original formula for the variance produces standard errors that are severely downward biased when the number of instruments is large.

4.1 Fiscal Discipline

Table 6 presents the baseline regressions for fiscal discipline, measured as the central government's budget balance as a percent of GDP. Columns (1) and (2) start with OLS and fixed effects estimates, which determine the bracketing range for the lagged dependent variable coefficient, namely 0.379−0.481. As expected, government balance follows a distinct adjustment process as evidenced by the highly significant coefficients on the two lags of the dependent variable.

Columns (3) and (4) present the D-GMM estimates of the MTEF effects. The internal instruments are the second and third lags of the budget balance and of the MTEF indicators. Column (4) augments column (3) with the external regional penetration instruments. In both specifications the effect of each MTEF phase is positive and statistically significant at conventional levels. Moreover, the effects are economically meaningful, ranging between 1.305 and 4.577 percentage points. In our sample the average fiscal balance among countries without an MTEF is −2.87%. Taking the estimates in column (4) at face value implies that only introducing the top-phase MTPF will put this average country in the black. The increase in the coefficients with more advanced MTEF phases lends support to hypothesis (H1).

These baseline specifications have lagged dependent variable coefficients in the FE-OLS bracketing range, not raising any specification issues. The model with external instruments is somewhat more precise than the one without. The D-GMM estimates of the MTEF effects are larger than the FE and OLS estimates, suggesting that the latter are depressed by a potential negative reverse causality bias, as when countries tend to adopt an MTEF in response to a fiscal crisis.[31]

For both models the diagnostic tests are satisfactory. The absence of first-order serial correlation in errors is rejected, while the absence of second-order serial correlation is not. The Hansen test does not reject the over-identification restrictions. We conclude that D-GMM is an internally consistent estimator in this model, and can be relied upon to carry out statistical inference for the hypotheses of interest.

Table 7 reports alternative specifications. Overall the results uphold the conclusions drawn from the baseline specifications of Table 6 columns (3) and (4). One exception is the model in column (3) that uses the difference transform instead of the orthogonal deviations

[31] This is consistent with the fact that international assistance organizations such as the World Bank, the UK's Department for International Development (DFID), the Asian Development Bank (ADB), and to a lesser extent the IMF, have recommended these reforms as part of a sound public financial management strategy.

transform. The lagged dependent variable coefficient is below the bracketing range, suggesting that just differencing the data produces a model with specification problems. The estimates are also very stable across specifications, with the exception of column (4) where the MTEFs enter lagged, increasing the size of their coefficients by a factor of between 1.5 and 2. As expected, collapsing the lagged dependent variable instruments in column (2) reduces statistical efficiency, yet the coefficients remain close to the baseline. We also note that in some specifications the conflict variable has a significant and large negative coefficient, consistent with the notion that armed conflict is costly on the budget. The openness variable has a very small effect and does not attain statistical significance.

4.2 Allocative Efficiency

Table 8 presents the baseline regressions for allocative efficiency, measured as the volatility in per-capita health spending in PPP$; see equation (1) above. The table structure mimics the previous fiscal discipline regressions. The sample period shortens due to the lack of health spending data in the first half of the 1990s, as well as the loss of one year to compute the volatility measure. The FE-OLS bracketing range for spending volatility, based on columns (1) and (2), is $0.009-0.308$. None of the MTEF coefficient estimates in these first two columns are significant, moreover they seem small relative to the sample average spending volatility of 6.69%. Based on hypothesis (H2) above, we expect the advanced micro MTEF phases (MTBF and MTPF) to reduce spending volatility, with the top-phase reducing it more.[32]

Columns (3) and (4) present the D-GMM estimates of these effects. The internal instruments are the second lag of budget balance and the second and third lags of the MTEF indicators. Compared to the previous two columns, the estimated MTEF effects become more negative and increase in magnitude. The MTBF effect is around -6 percentage points and statistically significant at conventional levels. The MTPF effect is about twice larger, although only weakly significant. Overall, these results provide support for hypothesis (H2).[33]

The statistical properties of these baseline models do not raise model specification issues. The lagged dependent variable coefficient estimate is in the bracketing range, although not reaching the significance threshold. The first-order and second-order serial correlation tests support the model's assumptions about the idiosyncratic error term. The Hansen test does

[32] Some countries chose to pilot an MTBF in the health sector before extending it to other sectors. We have been unable to systematically identify the countries that follow this particular sequencing of reform.

[33] This result can be seen as a sectoral counterpart of the finding in Fatas and Mihov (2003) that "discretionary" fiscal policy, i.e., variation in spending unrelated to economic fundamentals, increases aggregate output volatility.

not reject the over-identification restrictions.

Table 9 reports alternative specifications of the allocative efficiency regressions. Overall the results uphold the conclusions drawn from the baseline specifications of Table 8 columns (3) and (4). The MTBF and MTPF coefficient estimates maintain their prior patterns, although again the difference transform model in column (3) performs somewhat poorly. The covariates openness and conflict do not display any significant effects on health spending volatility.

4.3 Technical Efficiency

Table 10 presents the baseline regressions for technical efficiency, measured as efficiency scores from a stochastic frontier model of health delivery; see equations (2), (3), and Table 4. As expected, this indicator of fiscal performance is much more persistent than the previous ones. The FE-OLS bracketing range, based on columns (1) and (2), climbs to 0.858−0.999, indicating strong persistence. National health delivery is a complex system that may take decades to fully internalize the benefits of a given reform. The MTEF coefficient estimates are all negative, small and far from statistically significant. Hypothesis (3) above predicts that the top-phase MTEF (the MTPF) increases technical efficiency.

Columns (3) and (4) present the D-GMM estimates of the MTEF effects. The internal instruments are the second lag of technical efficiency and the second and third lags of the MTEF indicators. Compared to the previous two columns, the estimated MTEF effects turn positive and increase in magnitude. The specification with external instruments is more precise. The MTPF effect is 1.015 and statistically significant at the 10 percent level. The lack of precision should be expected since the within-country variation in technical efficiency is much smaller relative to the first two measures of fiscal performance (1.67 vs. 9.59, 6.35). As before, the diagnostic tests perform well, increasing confidence in the model specification. Overall, these baseline results provide moderate support for hypothesis (H3).

Table 11 reports alternative specifications of the technical efficiency regressions. The Hansen test suggests that the difference transform and predetermined MTEFs in columns (3) and (5) respectively are not adequate specifications. The remaining three models have satisfactory diagnostics and display a pattern of coefficients similar to the baseline. Among the alternative specifications only the specification with collapsed instruments attains statistical significance in the MTPF coefficient (p-value 0.059). As in the baseline results, the covariates openness and conflict do not display measurable effects on technical efficiency.

4.4 Democracy and Development

Finally, to rule out the possibility that the results are driven by subgroups of countries with extreme characteristics, we restrict the sample in two ways. First, we exclude highly autocratic countries, defined as those whose Polity IV score in 1990 takes the extreme value –10 ("strongly autocratic" in the language of the score producers). Second, we exclude highly developed countries, defined as those that are classified by UNDP in 1990 as having "very high human development" based on the country's Human Development Index (HDI). The two subgroups of countries are listed in Table 14. Note that the most developed countries are also the most democratic.

Tables 12 and 13 present the results with the restricted samples. The tables report the two early D-GMM baseline specifications for each of the three measures of fiscal performance. Overall, the prior patterns are preserved, strengthening the support for hypotheses (H1)-(H3). The government balance regression in Table 12 column (2) shows stronger MTEF effects both in magnitude and in statistical significance. This result seems consistent with the notion that MTEFs as commitment devices are particularly well suited for democratic settings which Mueller and Stratmann (2003) argue are more prone to fiscal indiscipline.[34] The results in Table 13 excluding the most developed countries are comparable to the baseline results in Tables 6, 8, and 10, both in magnitude and statistical significance, providing evidence that the MTEFs' positive effects on fiscal performance are not a phenomenon specific only to the most developed countries.[35]

5 Conclusion

In the last two decades more than 120 countries have moved toward a multiyear budget process. Although there has been much debate in the literature as to whether MTEFs are a worthwhile budget institution, a systematic empirical analysis of their impacts has been lacking due to insufficient data on MTEF adoption around the world. This paper is the first to empirically investigate the MTEFs' impacts on fiscal performance in a large sample of

[34] Ideally, to study how democracy and development condition MTEF impacts we would have interacted each MTEF indicator with measures of democracy and development. However, that would introduce three new endogenous variables, increasing the number of internal instruments above the number of available countries, and biasing the estimates toward their OLS counterparts.

[35] We also explored how the results change when we exclude more countries. Generally, the results are more sensitive when excluding more autocratic countries than when excluding more developed countries. The estimation strategy limits the degree to which we can restrict the sample through the requirement that the number of instruments times the number of periods remain below the number of panel units. See Roodman (2009).

countries. In order to disentangle the effects of the different MTEF phases (MTFF, MTBF, and MTPF) from other factors and to correct for reverse causality we apply a dynamic panel data approach to a newly-collected panel dataset of 181 countries over the period 1990-2008, the most comprehensive dataset to date on worldwide MTEF adoption.

The econometric findings suggest that, unlike in previous small-sample and case-study analyses, MTEF adoption is associated with strong improvement in fiscal discipline, the effects increasing with each successive MTEF phase. The adoption of an MTBF and an MTPF decrease the volatility of health spending per capita, which we interpret as an improvement in allocative efficiency. Finally, the MTPF seems to be the only MTEF phase that exerts a significant effect on technical efficiency in the health sector, although due to insufficient within-country variation in technical efficiency this effect is less precisely estimated. Overall these results are more favorable than the conclusions of prior work, and suggest that budget institutions that restrain short-term incentives to manipulate the budget can have measurable benefits for fiscal performance.

Our analysis may be limited by the fact that an MTEF could be in place only in law (de jure) and not in practice (de facto). However, if this phenomenon were widespread it would induce an attenuation bias and our estimates could still be regarded as a lower bound on the actual effect. Being in effect commitment mechanisms, transparency and enforcement are critical components of MTEFs. Studying which features of the broader civic, juridical, and political environment enhance MTEF effectiveness may lead to a better understanding of these institutions. Also, our analysis of MTEF impacts on allocative efficiency and technical efficiency is limited to the health sector. Whether the results carry over to other categories of productive spending is an interesting question for future research. To create more precise measures of these dimensions of fiscal performance one needs to identify disaggregated categories of spending that are comparable across a broad range of countries.

References

Acemoglu, D. (2005) "Constitutions, Politics, and Economics," *Journal of Economic Literature* 43(4): 1025-1048.

Alesina, A., R. Hausmann, R. Hommes, and E. Stein (1999) "Budget Institutions and Fiscal Performance in Latin America," *Journal of Development Economics* 59: 253-273.

Alesina, A. and R. Perotti (1996) "Fiscal Discipline and the Budget Process," *American Economic Review Papers and Proceedings*: 86(2), 401-407.

Alesina, A. and R. Perotti (1999) "Budget Deficits and Budget Institutions," in J. Poterba and J. von Hagen (eds.) *Fiscal Institutions and Fiscal Performance*, Chicago: U. of Chicago Press.

Alesina, A., Tabellini G. (1990), "Voting on the Budget Deficit," *American Economic Review*: 80(1): 37-49.

Arellano, M. and S. Bond (1991) "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies* 58: 277-297.

Arellano, M. and O. Bover (1995) "Another Look at the Instrumental Variables Estimation of Error-Components Models," *Journal of Econometrics* 68: 29–51.

Baltagi, B.H., P.O. Demetriades, and S.H. Law (2009) "Financial Development and Openness: Evidence from Panel Data," *Journal of Development Economics* 89: 285-296.

Bayoumi, T. and B. Eichengreen (1995) "Restraining Yourself: The Implications of Fiscal Rules for Economic Stabilizations," *IMF Staff Papers* 42: 32-67.

Besley, T. and A. Case (2003) "Political Institutions and Policy Choices: Evidence from the United States," *Journal of Economic Literature* 41(1): 7-73.

Besley, T. and M. Smart (2007) "Fiscal Restraints and Voter Welfare," *Journal of Public Economics* 91(3-4): 755-773.

Bevan, D. and G. Palomba (2000) "Uganda: The Budget and Medium-Term Expenditure Framework Set in a Wider Context," background paper for Poverty Reduction Support Credit with DFID finance.

Bisin, A., A. Lizzeri, and L. Yariv (2011) "Government Policy with Time Inconsistent Voters," working paper, NYU.

Blundell, R. and S. Bond (1998) "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics* 87: 115–143.

Bohn, H. and R.P. Inman (1996) "Balanced-Budget Rules and Public Deficits: Evidence from the U.S. states," *Carnegie-Rochester Conference Series on Public Policy* 45(1): 13-76.

Bond, S. (2002) "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice," working paper, Institute for Fiscal Studies.

Brender, A. and A. Drazen (2005) "Political Budget Cycles In New Versus Established Democracies," *Journal of Monetary Economics* 52: 1271-1295.

Castro, I. and W. Dorotinsky (2008) "Medium-Term Expenditure Frameworks: Demystifying and Unbundling The Concept," *Journal of Public Budgeting*, forthcoming.

Drazen, A. (2000) "The Political Business Cycle after 25 Years," *NBER Macroeconomic Annual* 15: 75-117.

Easterly, W. and S. Rebelo (1993) "Fiscal Policy and Economic Growth: An Empirical Investigation," *Journal of Monetary Economics* 32(3): 417-458.

Easterly, W., T. Irwin and L. Servén (2008) "Walking up the Down Escalator: Public Investment and Fiscal Stability," *World Bank Research Observer* 23(1): 37-56.

Fabrizio, S. and A. Mody (2006) "Can Budget Institutions Counteract Political Indiscipline?" *Economic Policy* 21: 689-739

Farrell, M.J. (1957) "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society* 120(3): 253-282.

Fatas, A. and I. Mihov (2003) "The Case for Restricting Fiscal Policy Discretion" *Quarterly Journal of Economics* 118(4): 1419–1447.

Fatas, A. and I. Mihov (2006) "The Macroeconomic Effects of Fiscal Rules in the U.S. States," *Journal of Public Economics* 90(1-2): 101-117.

Filc, G. and C. Scartascini (2010) "Is Latin America on the Right Track? An Analysis of Medium-Term Frameworks and the Budget Process," working paper, Inter-American Development Bank, Research Department.

Fischer, S. (1993) "The Role of Macroeconomic Factors in Growth," *Journal of Monetary Economics* 32(3): 485-512.

Folscher, A. (2007) "Budget Methods and Practices," in A. Shaw (ed.) *Budgeting and Budgetary Institutions*, Washington, DC: World Bank.

Gleich, H. (2003) "Budget Institutions and Fiscal Performance in Central and Eastern European Countries," European Central Bank research paper 215.

Greene, W. (2005) "Distinguishing between Heterogeneity and Inefficiency: Stochastic Frontier Analysis of the WHO's Panel Data on National Health Care Systems," *Health Economics* 13: 959-980.

Hallerberg, M., S. Rolf, and J. von Hagen (2009) *Fiscal Governance: Evidence from Europe*, Cambridge: cambridge University Press.

Hallerberg, M. and J. von Hagen (1999) "Electoral Institutions, Cabinet Negotiations, and Budget Deficits in the European Union," in J. Poterba and J. von Hagen (eds.) *Fiscal Institutions and Fiscal Performance*, Chicago: U. of Chicago Press.

Hallerberg, M. and S. Ylaoutinen (2000) "Political Power, Fiscal Institutions and Budgetary Outcomes in Central and Eastern Europe," *Journal of Public Policy* 30: 45-62.

Holmes, M. and A. Evans (2003) "A Review of Experience in Implementing Medium-Term Expenditure Frameworks in a PRSP Context: A Synthesis of Eight Country Studies," working paper, ODI.

Holtz-Eakin, D., W. Newey, and H.S. Rosen (1988) "Estimating Vector Autoregressions with Panel Data," *Econometrica* 56: 1371-1395.

Jondrow, J., I. Materov, K. Lovell, and P. Schmidt (1982) "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model," *Journal of Econometrics* 19(2/3): 233-238.

Keefer, P. and R. Vlaicu (2008) "Democracy, Credibility and Clientelism," *Journal of Law, Economics and Organization* 24(2): 41-67.

Knight, B. (2000) "Supermajority Voting Requirements for Tax Increases: Evidence from the States," *Journal of Public Economics* 76: 191-222.

Knight, B. and A. Levinson (2000) "Fiscal Institutions in U.S. States," in R. Strauch and J. von Hagen (eds.) *Institutions, Politics, and Fiscal Policy*, Boston: Kluwer.

Kumar, M. S. and T. Ter-Minassian (2007) "Fiscal discipline: key issues and overview," in M. S. Kumar and T. Ter-Minassian (eds.) *Promoting Fiscal Discipline*, Washington, DC: IMF.

Le Houerou, P. and R. Taliercio (2002) "Medium-Term Expenditure Frameworks: From Concept to Practice (Preliminary Lessons from Africa)," working paper, World Bank.

Lewis, M. and M. Verhoeven (2010) "Financial Crisis and Social Spending: The Impact of the 2008-2009 Crisis," *World Economics* 11(4): 79-110.

Marshall, M.G., K. Jaggers, and T.R. Gurr (2010) *Polity IV Project: Political Regime Characteristics and Transitions 1800-201*0, Center for Systemic Peace.

McNab, R.B., J. Martinez-Vazquez, and L.F.J. Boex (2000) "Multi-Year Budgeting: A Review of International Practices and Lessons for Developing and Transitional Economies," *Public Budgeting and Finance* 20(2): 91-112.

Mueller, D. and T. Stratmann (2003) "The Economic Effects of Democratic Participation," *Journal of Public Economics* 87(9-10): 2129-2155.

Nickell, S. (1981) "Biases in Dynamic Models with Fixed Effects," *Econometrica* 49(6):

1417-26.

Oyugi, L. N. (2008) "Experiences with Medium-Term Expenditure Framework in Selected Southern and Eastern African Countries," SEAPREN Working Paper No. 7.

Oxford Policy Management (2000) "Medium-Term Expenditure Frameworks—Panacea or Dangerous Distraction?" *Oxford Policy Management Review* 2: 1-4.

Perotti, R. and Y. Kontopoulos (2002) "Fragmented Fiscal Policy," *Journal of Public Economics* 86(2): 191-222.

Persson, T., G. Roland, and G. Tabellini (2000) "Comparative Politics and Public Finance," *Journal of Political Economy* 108(6): 1121-1161.

Poterba, J. (2004) "State Responses to Fiscal Crises: The Effects of Budgetary Institutions and Politics," *Journal of Political Economy* 102(4): 799-821.

Poterba, J. and J. von Hagen (1999) (eds.) *Fiscal Institutions and Fiscal Performance*, Chicago: U. of Chicago Press.

Roodman, D. (2009) "A Note on the Theme of Too Many Instruments," *Oxford Bulletin of Economics and Statistics* 71: 135-158.

Schiavo-Campo, S. (2009) "Potemkin Villages: 'The' Medium-Term Expenditure Framework in Developing Countries," *Public Budgeting and Finance* 29(2): 1-26.

Stein, E., E. Talvi, and A. Grisanti (1999) "Institutional Arrangements and Fiscal Performance: The Latin American Experience," in J. Poterba and J. von Hagen (eds.) *Fiscal Institutions and Fiscal Performance*, Chicago: U. of Chicago Press.

Schick, A. (2010) "Fiscal and Budget Institutions: Consolidators or Accommodators?" Washington, DC: IMF Fiscal Forum.

Velasco, A. (1999) "A Model of Endogenous Fiscal Deficits and Delayed Fiscal Reforms," in J. Poterba and J. von Hagen (eds.) *Fiscal Institutions and Fiscal Performance*, Chicago: U. of Chicago Press.

von Hagen, J. (2006) "Political Economy of Fiscal Institutions," in B.R. Weingast and D.A. Wittman (eds.) *The Oxford Handbook of Political Economy*, Oxford, UK: Oxford University Press.

von Hagen, J. and I. Harden (1995) "Budget Processes and Commitment to Fiscal Discipline," *European Economic Review* 39: 771-779.

von Hagen, J. and G. Wolff (2006) What Do Deficits Tell Us about Debt? Empirical Evidence on Creative Accounting with Fiscal Rules in the EU," *Journal of Banking and Finance* 30(12): 3259-3279.

Wescott, C. G. (2008) "World Bank Support for Public Financial Management: Conceptual Roots and Evidence of Impact," IEG working waper, World Bank.

Weingast, B.R., K. Shepsle, and C. Johnsen (1982) "The Political Economy of Benefits and Costs: A Neoclassical Approach to Distributive Politics," *Journal of Political Economy* 89(4): 642-664.

Wildavsky, A. (1986) *Budgeting: A Comparative Theory of Budgetary Processes*, New Brunswick, NJ: Transaction Books.

Windmeijer, F. (2005) "A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators," *Journal of Econometrics* 126: 25-51.

World Bank (1998) *Public Expenditure Management Handbook*, Washington, DC: World Bank Publications.

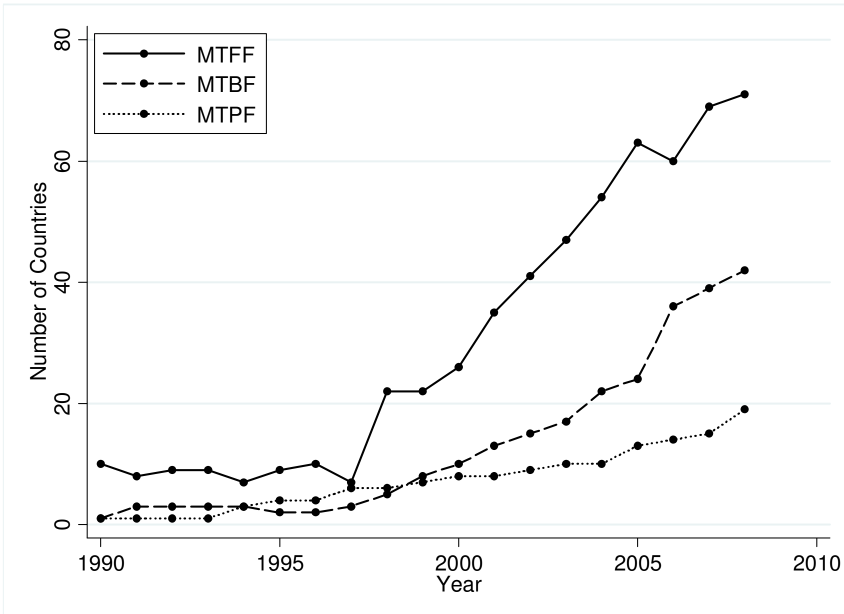
World Bank (2012) *Beyond the Annual Budget: Review of Global Experience with Medium Term Expenditure Frameworks*, "Directions in Development" Series, Washington, DC: World Bank Publications, forthcoming.

Ylaoutinen, S. (2004) "Fiscal Frameworks in Central and Eastern European Countries," Finnish Ministry of Finance discussion paper 72.

Appendix

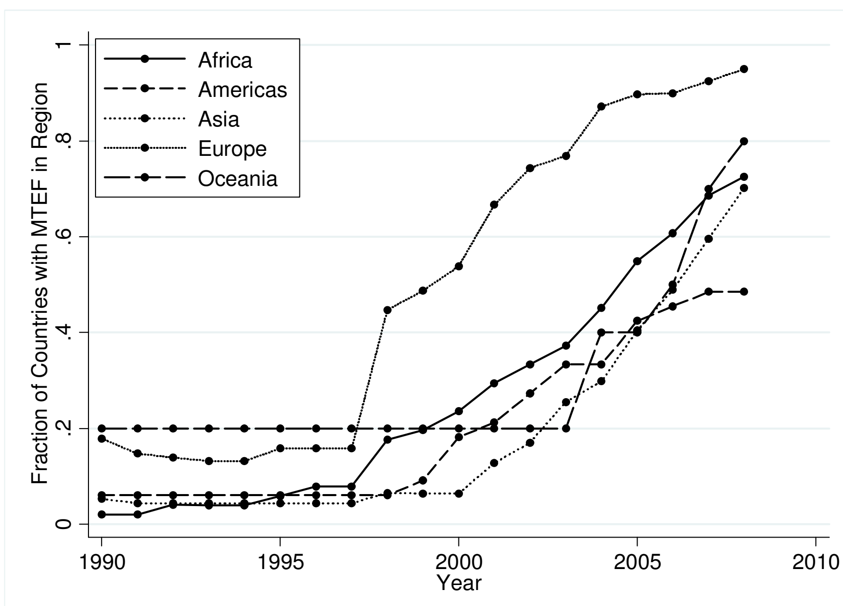
A1. Figures

Figure 1: MTEF Growth Worldwide, 1990-2008



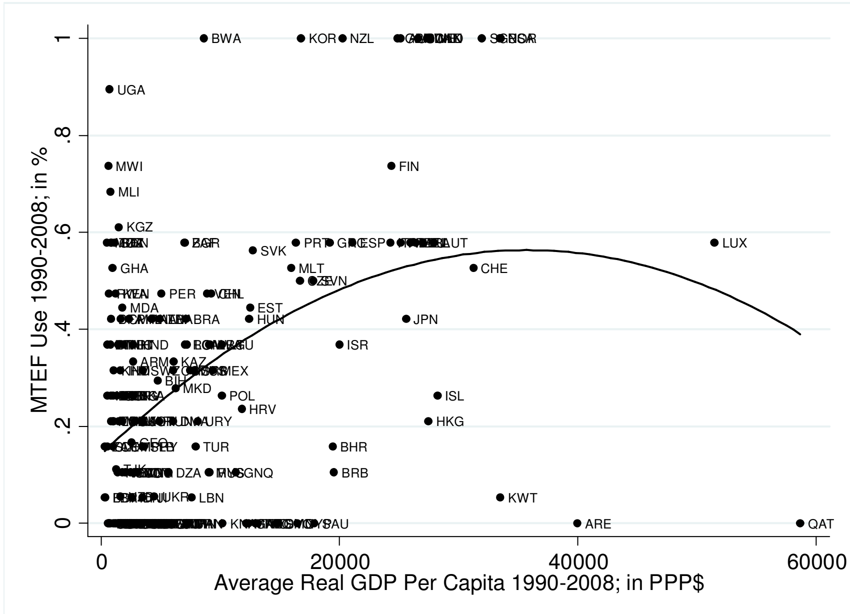
Notes: Authors' calculations based on the sample of 181 countries during 1990-2008 described in the Data Appendix.

Figure 2: MTEF Growth by Region, 1990-2008



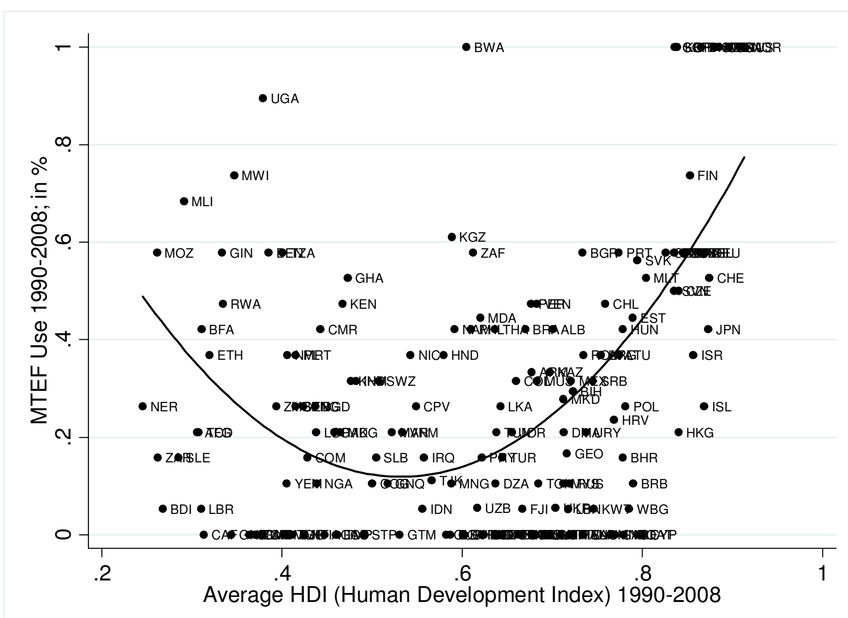
Notes: Authors' calculations based on the sample of 181 countries during 1990-2008 described in the Data Appendix. MTEF status indicates the presence of one of the three MTEF phases. Geographical regions as defined by the UN Statistics Division.

Figure 3: MTEF and Income



Notes: Authors' calculations based on the sample of 181 countries during 1990-2008 described in the Data Appendix. Country codes adjacent to each scatter point. Scatter plot quadratic fit shown.

Figure 4: MTEF and Development



Notes: Authors' calculations based on the sample of 181 countries during 1990-2008 described in the Data Appendix. Country codes adjacent to each scatter point. Scatter plot quadratic fit shown.

A2. Tables

Table 1: Summary Statistics

Variable	Obs.	Mean	Std. Deviation		Min	Max
			Across	Within		
Government Balance	2,991	–2.24	11.87	9.59	–151.33	384.15
Spending Volatility	2,282	6.69	7.62	6.35	0.002	83.19
Technical Efficiency	2,359	86.48	12.11	1.67	39.21	99.00
MTFF	3,378	0.17	0.38	0.32	0	1
MTBF	3,378	0.07	0.26	0.22	0	1
MTPF	3,378	0.04	0.20	0.13	0	1
MTFF Regional Penetration	3,359	0.17	0.21	0.18	0	1
MTBF Regional Penetration	3,359	0.07	0.14	0.12	0	1
MTPF Regional Penetration	3,359	0.04	0.14	0.08	0	1
Openness	3,069	85.28	48.72	16.89	0.31	456.64
Conflict	3,439	0.05	0.21	0.17	0	1
Health Spending per Capita	2,460	669.04	983.71	284.64	7.09	7,536.2
Life Expectancy	3,331	66.09	10.39	2.13	26.41	82.58
Population Density	3,304	188.89	650.61	60.86	1.43	6,943.2
OECD Membership	3,439	0.16	0.36	0.07	0	1

Notes: The summary statistics are based on the sample of 181 countries during 1990-2008 described in the Data Appendix. The appendix contains details on the data sources, units of measurement and construction of variables. The differences in number of observations across variables reflect data availability in the different data sources.

Table 2: MTEF Growth, 1990-2008

	1990	2008	Adoptions	Transitions	Reversals
MTFF	9	71	104	–41	–1
MTBF	1	42	21	23	–3
MTPF	1	19	0	18	0
Total MTEF	11	132	125	0	–4

Notes: The summary statistics are based on the sample of 181 countries during 1990-2008 described in the Data Appendix. Of the eighteen transitions to MTPF nine are from MTFF and nine from MTBF. The MTFF reversal is Argentina. The MTBF reversals are Argentina, Estonia, and the United States.

Table 3: Correlations Matrix

	Gov Bal	*Spend Volat*	*Tech Eff*	*MTFF*	*MTBF*	*MTPF*	*Openn.*	*Confl.*
Gov_ Bal	1							
Spend_ Volat	.01	1						
Tech_ Eff	.02	–.05**	1					
MTFF	.06***	.04*	.10***	1				
MTBF	.01	.05**	–.16***	–.13***	1			
MTPF	.06***	–.05**	.14***	–.09***	–.06***	1		
Openn.	.09***	–.02	.17***	.04**	–.05***	.05***	1	
Confl.	–.09***	.05**	–.17***	–.05***	–.00	–.05***	–.13***	1

Notes: The correlations are based on the sample of 181 countries during 1990-2008 described in the Data Appendix. The number of observations varies between 2,038 and 3,378 depending on the pair of variables. ***, **, * indicate statistical significance at the 1%, 5%, 10% levels, respectively.

Table 4: Stochastic Frontier Model for Public Health Delivery

Dependent Variable:	$\log(Life\_Exp_{i,t})$ (Log of Life Expectancy at Birth)		
	Coefficients	Model Statistics	
$\log(Health\_Spend_{i,t})$	0.00348*** (0.00099)		
$Pop\_Density_{i,t}$	0.00002*** (0.00000)	Mean Efficiency:	μ = 86.4783***
$OECD_{i,t}$	0.07829*** (0.00402)	Std. Deviations:	σ_u = 0.1567*** σ_v = 0.0247***
Constant	4.28722*** (0.00747)	Ratio:	λ = 6.3398***
Year Effects	Yes	Log Likelihood:	$\log(L)$=1686.66
Sample Period	1995-2008		
Countries	177		
Observations	2,359		

Notes: The unit of observation is a country-year from the sample described in the Data Appendix. Standard errors in parentheses. The table reports maximum likelihood (ML) estimates of a stochastic frontier model for life expectancy with time-varying inefficiency term u_{it}. The model assumes an exponential distribution for the inefficiency term. ***, **, * indicate statistical significance at the 1%, 5%, 10% levels, respectively.

Table 5: Unit Root Tests

	Government Balance	Spending Volatility	Technical Efficiency
IPS Test Statistic $Z_{\tilde{t}bar}$	–9.348***	–17.066***	–2.047**
p-value	0.000	0.000	0.020
Average Panel Length	18.55	12.94	13.83
Countries	159	175	167

Notes: The tests are performed on the sample of 181 countries during 1990-2008 described in the Data Appendix. The table reports Im-Pesaran-Shin (IPS) unit root test results for the dependent variables. The null hypothesis is H_0: All panels contain unit roots. The varying number of countries reflects the unbalanced nature of our panel and the requirement that the minimum length of each individual panel has to be at least ten. Panel means included. Time trends or lags not included.

Table 6: MTEFs and Fiscal Discipline: Baseline

Dependent Variable:	$Gov\_Balance_{i,t}$ (Central Government Balance, % of GDP)			
Model:	OLS	FE	D-GMM	D-GMM-IV
	(1)	(2)	(3)	(4)
$Gov\_Balance_{i,t-1}$	0.481***	0.379***	0.421***	0.423***
	(0.045)	(0.041)	(0.040)	(0.040)
$Gov\_Balance_{i,t-2}$	0.174***	0.116***	0.101***	0.101***
	(0.051)	(0.038)	(0.037)	(0.038)
$MTFF_{i,t}$	0.070	0.018	1.936**	1.305**
	(0.268)	(0.318)	(0.929)	(0.605)
$MTBF_{i,t}$	–0.187	0.274	2.068*	2.427**
	(0.352)	(0.448)	(1.133)	(1.147)
$MTPF_{i,t}$	0.897***	1.066*	4.577**	3.375**
	(0.305)	(0.609)	(2.103)	(1.359)
$Openness_{i,t}$	0.002	0.005	0.007	0.009
	(0.001)	(0.007)	(0.008)	(0.008)
$Conflict_{i,t}$	–1.296*	–1.736**	–1.193*	–1.055
	(0.685)	(0.838)	(0.693)	(0.670)
Year Effects	Yes	Yes	Yes	Yes
Internal Instruments	No	No	Yes	Yes
External Instruments	No	No	No	Yes
AR(1) Test p-val.			0.000	0.000
AR(2) Test p-val.			0.902	0.911
Hansen J Test p-val.			0.681	0.792
Sample Period	1990–2008	1990–2008	1990–2008	1990–2008
Countries	162	162	161	161
Observations	2,478	2,478	2,316	2,316

Notes: The unit of observation is a country-year from the sample described in the Data Appendix. Columns (1) and (2) report standard errors clustered at the country level. Columns (3) and (4) report two-step estimates and standard errors with the Windmeijer correction. GMM models use the orthogonal deviations transform. The internal instruments are: the second and third lags of $Gov\_Balance_{i,t}$, and of $MTFF_{i,t}$, $MTBF_{i,t}$, $MTPF_{i,t}$. The external instruments are: $MTFF\_Regional\_Penetration_{i,t}$, $MTBF\_Regional\_Penetration_{i,t}$, and $MTPF\_Regional\_Penetration_{i,t}$. The internal instruments enter uncollapsed. ***, **, * indicate statistical significance at the 1%, 5%, 10% levels, respectively.

Table 7: MTEFs and Fiscal Discipline: Alternative Specifications

Dependent Variable:	$Gov\_Balance_{i,t}$ (Central Government Balance, as % of GDP)				
Model:	D-GMM-IV (1)	D-GMM-IV (2)	D-GMM-IV (3)	D-GMM (4)	D-GMM-IV (5)
$Gov\_Balance_{i,t-1}$	0.400***	0.405***	0.321***	0.420***	0.414***
	(0.037)	(0.035)	(0.041)	(0.043)	(0.040)
$Gov\_Balance_{i,t-2}$	0.110***	0.055	0.033	0.099***	0.104***
	(0.038)	(0.038)	(0.047)	(0.038)	(0.039)
$MTFF_{i,t}$	1.204*	1.149*	1.539*	2.484***	1.045*
[$MTFF_{i,t-1}$ in col. (4)]	(0.643)	(0.681)	(0.918)	(0.881)	(0.571)
$MTBF_{i,t}$	1.965**	2.103*	3.679**	3.192**	2.162**
[$MTBF_{i,t-1}$ in col. (4)]	(0.997)	(1.235)	(1.676)	(1.495)	(1.028)
$MTPF_{i,t}$	3.600**	2.783*	4.757**	6.106***	3.227**
[$MTPF_{i,t-1}$ in col. (4)]	(1.757)	(1.598)	(2.156)	(2.210)	(1.555)
$Openness_{i,t}$	0.003	0.003	0.005	0.006	0.007
	(0.007)	(0.008)	(0.014)	(0.008)	(0.008)
$Conflict_{i,t}$	–1.804**	–1.436	–1.469	–1.369**	–0.807
	(0.776)	(0.889)	(0.972)	(0.653)	(0.623)
Change from Baseline	One-Step Robust	Collapsed Instrum.	Difference Transform	MTEFs Lagged	MTEFs Predeterm.
Year Effects	Yes	Yes	Yes	Yes	Yes
Internal Instruments	Yes	Yes	Yes	Yes	Yes
External Instruments	Yes	Yes	Yes	No	Yes
AR(1) Test p-val.	0.000	0.000	0.000	0.000	0.000
AR(2) Test p-val.	0.923	0.517	0.597	0.829	0.970
Hansen J Test p-val.	0.792	0.991	0.484	0.867	0.367
Sample Period	1990–2008	1990–2008	1990–2008	1990–2008	1990–2008
Countries	161	161	161	161	161
Observations	2,316	2,316	2,313	2,316	2,316

Notes: The unit of observation is a country-year from the sample described in the Data Appendix. Each table column reports the estimates of a variation on the baseline specification in Table 6 column (4). Column (1) reports one-step estimates with robust standard errors instead of two-step estimates with Windmeijer corrected standard errors. Column (2) collapses the instruments for the lagged dependent variable. Column (3) uses the difference transform instead of the orthogonal deviations transform. Column (4) enters the three MTEF variables lagged one period and drops the external instruments. The bracketing range for this specification is 0.379–0.482. Column (5) treats the three MTEF variables as predetermined. The internal instruments are now the first and second lags of $MTFF_{i,t}$, $MTBF_{i,t}$, $MTPF_{i,t}$. ***, **, * indicate statistical significance at the 1%, 5%, 10% levels, respectively.

Table 8: MTEFs and Allocative Efficiency: Baseline

Dependent Variable:	$Spend\_Volatility_{i,t}$ (Per Cap. Health Spending Volatility)			
Model:	OLS	FE	D-GMM	D-GMM-IV
	(1)	(2)	(3)	(4)
$Spend\_Volatility_{i,t-1}$	0.308***	0.009	0.087	0.105
	(0.055)	(0.050)	(0.067)	(0.065)
$MTFF_{i,t}$	0.342	0.277	–1.935	–0.092
	(0.620)	(0.647)	(1.946)	(1.889)
$MTBF_{i,t}$	0.338	0.196	–6.439**	–5.823**
	(0.577)	(1.120)	(2.829)	(2.415)
$MTPF_{i,t}$	–0.950	–3.386	–14.879*	–10.530
	(1.122)	(2.078)	(8.053)	(6.699)
$Openness_{i,t}$	–0.002	0.006	0.003	0.005
	(0.003)	(0.012)	(0.013)	(0.013)
$Conflict_{i,t}$	–0.556	–0.898	–0.668	–0.918
	(0.769)	(1.103)	(1.445)	(1.492)
Year Effects	Yes	Yes	Yes	Yes
Internal Instruments	No	No	Yes	Yes
External Instruments	No	No	No	Yes
AR(1) Test *p*-val.			0.000	0.000
AR(2) Test *p*-val.			0.888	0.769
Hansen *J* Test *p*-val.			0.688	0.582
Sample Period	1996–2008	1996–2008	1996–2008	1996–2008
Countries	172	172	170	170
Observations	1,870	1,870	1,698	1,698

Notes: The unit of observation is a country-year from the sample described in the Data Appendix. Columns (1) and (2) report standard errors clustered at the country level. Columns (3) and (4) report two-step estimates and standard errors with the Windmeijer correction. GMM models use the orthogonal deviations transform. The internal instruments are: the second lag of $Spend\_Volatility_{i,t}$, and the second and third lags of $MTFF_{i,t}$, $MTBF_{i,t}$, $MTPF_{i,t}$. The external instruments are: $MTFF\_Regional\_Penetration_{i,t}$, $MTBF\_Regional\_Penetration_{i,t}$, and $MTPF\_Regional\_Penetration_{i,t}$ The internal instruments enter uncollapsed. ***, **, * indicate statistical significance at the 1%, 5%, 10% levels, respectively.

Table 9: MTEFs and Allocative Efficiency: Alternative Specifications

Dependent Variable:	$Spend\_Volatility_{i,t}$ (Per Cap. Health Spending Volatility)				
Model:	D-GMM-IV (1)	D-GMM-IV (2)	D-GMM-IV (3)	D-GMM (4)	D-GMM-IV (5)
$Spend\_Volatility_{i,t-1}$	0.079	0.115	0.044	0.072	0.046
	(0.068)	(0.072)	(0.065)	(0.043)	(0.063)
$MTFF_{i,t}$	0.897	0.425	0.047	–1.764	–0.112
[$MTFF_{i,t-1}$ in col. (4)]	(1.156)	(2.390)	(1.816)	(1.638)	(1.189)
$MTBF_{i,t}$	–5.328*	–5.913**	–2.904	–6.977*	–2.671
[$MTBF_{i,t-1}$ in col. (4)]	(2.846)	(2.914)	(3.546)	(3.914)	(2.344)
$MTPF_{i,t}$	–11.086*	–7.583	–8.651	–14.697**	–12.578**
[$MTPF_{i,t-1}$ in col. (4)]	(6.615)	(7.957)	(5.870)	(7.041)	(5.699)
$Openness_{i,t}$	0.004	0.005	–0.000	0.004	0.004
	(0.012)	(0.013)	(0.027)	(0.013)	(0.015)
$Conflict_{i,t}$	–0.353	–1.306	0.187	–0.203	0.294
	(1.136)	(1.462)	(1.578)	(1.431)	(1.141)
Change from Baseline	One-Step Robust	Collapsed Instrum.	Difference Transform	MTEFs Lagged	MTEFs Predeterm.
Year Effects	Yes	Yes	Yes	Yes	Yes
Internal Instruments	Yes	Yes	Yes	Yes	Yes
External Instruments	Yes	Yes	Yes	No	Yes
AR(1) Test p-val.	0.000	0.000	0.000	0.000	0.000
AR(2) Test p-val.	0.808	0.721	0.939	0.944	0.956
Hansen J Test p-val.	0.582	0.743	0.762	0.732	0.386
Sample Period	1996–2008	1996–2008	1996–2008	1996–2008	1996–2008
Countries	170	170	170	170	170
Observations	1,698	1,698	1,694	1,697	1,698

Notes: The unit of observation is a country-year from the sample described in the Data Appendix. Each table column reports the estimates of a variation on the baseline specification in Table 8 column (4). Column (1) reports one-step estimates with robust standard errors instead of two-step estimates with Windmeijer corrected standard errors. Column (2) collapses the instruments for the lagged dependent variable. Column (3) uses the difference transform instead of the orthogonal deviations transform. Column (4) enters the three MTEF variables lagged one period and drops the external instruments. The bracketing range for this specification is 0.011–0.309. Column (5) treats the three MTEF variables as predetermined. The internal instruments are now the first and second lags of $MTFF_{i,t}$, $MTBF_{i,t}$, $MTPF_{i,t}$. ***, **, * indicate statistical significance at the 1%, 5%, 10% levels, respectively.

Table 10: MTEFs and Technical Efficiency: Baseline

Dependent Variable:	$Tech\_Efficiency_{i,t}$ (Estimated, see Stoch. Frontier Table 4)			
Model:	OLS	FE	D-GMM	D-GMM-IV
	(1)	(2)	(3)	(4)
$Tech\_Efficiency_{i,t-1}$	0.999***	0.858***	0.920***	0.934***
	(0.004)	(0.015)	(0.066)	(0.047)
$MTFF_{i,t}$	–0.072	–0.091	–0.049	0.009
	(0.052)	(0.060)	(0.137)	(0.119)
$MTBF_{i,t}$	–0.048	–0.071	0.303	0.427
	(0.079)	(0.080)	(0.423)	(0.281)
$MTPF_{i,t}$	–0.065	–0.060	0.616	1.015*
	(0.079)	(0.216)	(0.693)	(0.588)
$Openness_{i,t}$	–0.002***	–0.002	–0.001	–0.001
	(0.001)	(0.001)	(0.002)	(0.001)
$Conflict_{i,t}$	0.176**	–0.028	–0.011	–0.000
	(0.080)	(0.044)	(0.070)	(0.048)
Year Effects	Yes	Yes	Yes	Yes
Internal Instruments	No	No	Yes	Yes
External Instruments	No	No	No	Yes
AR(1) Test p-val.			0.040	0.039
AR(2) Test p-val.			0.375	0.384
Hansen J Test p-val.			0.182	0.647
Sample Period	1995–2008	1995–2008	1995–2008	1995–2008
Countries	169	169	169	169
Observations	1,970	1,970	1,801	1,801

Notes: The unit of observation is a country-year from the sample described in the Data Appendix. Columns (1) and (2) report standard errors clustered at the country level. Columns (3) and (4) report two-step estimates and standard errors with the Windmeijer correction. GMM models use the orthogonal deviations transform. The internal instruments are: the second lag of $Tech\_Efficiency_{i,t}$, and the second and third lags of $MTFF_{i,t}$, $MTBF_{i,t}$, $MTPF_{i,t}$. The external instruments are: $MTFF\_Regional\_Penetration_{i,t}$, $MTBF\_Regional\_Penetration_{i,t}$, and $MTPF\_Regional\_Penetration_{i,t}$ The internal instruments enter uncollapsed. ***, **, * indicate statistical significance at the 1%, 5%, 10% levels, respectively.

Table 11: MTEFs and Technical Efficiency: Alternative Specifications

Dependent Variable:	$Tech\_Efficiency_{i,t}$ (Estimated, see Stoch. Frontier Table 4)				
Model:	D-GMM-IV (1)	D-GMM-IV (2)	D-GMM-IV (3)	D-GMM (4)	D-GMM-IV (5)
$Tech\_Efficiency_{i,t-1}$	0.905***	0.912***	0.916***	0.931***	0.919***
	(0.043)	(0.060)	(0.072)	(0.069)	(0.061)
$MTFF_{i,t}$	–0.021	0.035	–0.105	–0.087	–0.131
[$MTFF_{i,t-1}$ in col. (4)]	(0.147)	(0.133)	(0.129)	(0.105)	(0.121)
$MTBF_{i,t}$	0.366	0.440*	0.095	0.136	0.153
[$MTBF_{i,t-1}$ in col. (4)]	(0.255)	(0.258)	(0.296)	(0.283)	(0.158)
$MTPF_{i,t}$	0.791	1.163*	–0.529	0.615	0.276
[$MTPF_{i,t-1}$ in col. (4)]	(0.565)	(0.617)	(0.553)	(0.700)	(0.363)
$Openness_{i,t}$	–0.002	–0.001	–0.002*	–0.001	–0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
$Conflict_{i,t}$	–0.036	–0.035	–0.031	0.054	–0.028
	(0.047)	(0.053)	(0.044)	(0.043)	(0.046)
Change from Baseline	One-Step Robust	Collapsed Instrum.	Difference Transform	MTEFs Lagged	MTEFs Predeterm.
Year Effects	Yes	Yes	Yes	Yes	Yes
Internal Instruments	Yes	Yes	Yes	Yes	Yes
External Instruments	Yes	Yes	Yes	No	Yes
AR(1) Test p-val.	0.038	0.041	0.040	0.042	0.039
AR(2) Test p-val.	0.376	0.381	0.364	0.408	0.381
Hansen J Test p-val.	0.647	0.735	0.007	0.300	0.143
Sample Period	1995–2008	1995–2008	1995–2008	1995–2008	1995–2008
Countries	169	169	169	169	169
Observations	1,801	1,801	1,798	1,801	1,801

Notes: The unit of observation is a country-year from the sample described in the Data Appendix. Each table column reports the estimates of a variation on the baseline specification in Table 10 column (4). Column (1) reports one-step estimates with robust standard errors instead of two-step estimates with Windmeijer corrected standard errors. Column (2) collapses the instruments for the lagged dependent variable. Column (3) uses the difference transform instead of the orthogonal deviations transform. Column (4) enters the three MTEF variables lagged one period and drops the external instruments. The bracketing range for this specification remains the same 0.858–0.999. Column (5) treats the three MTEF variables as predetermined. The internal instruments are now the first and second lags of $MTFF_{i,t}$, $MTBF_{i,t}$, $MTPF_{i,t}$. ***, **, * indicate statistical significance at the 1%, 5%, 10% levels, respectively.

Table 12: MTEFs and Fiscal Performance: Excluding Highly Autocratic Countries

Dep. Variable ($y_{i,t}$):	$Gov\_Balance_{i,t}$		$Spend\_Volatility_{i,t}$		$Tech\_Efficiency_{i,t}$	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
$y_{i,t-1}$	0.415***	0.418***	0.103	0.117*	0.962***	0.981***
	(0.038)	(0.037)	(0.063)	(0.061)	(0.102)	(0.062)
$y_{i,t-2}$	0.101***	0.100***				
	(0.037)	(0.037)				
$MTFF_{i,t}$	1.905**	1.875***	–1.852	–0.520	–0.089	–0.081
	(0.898)	(0.676)	(1.855)	(1.791)	(0.130)	(0.129)
$MTBF_{i,t}$	1.867*	3.139***	–6.289**	–5.708**	0.201	0.286
	(1.042)	(1.152)	(3.019)	(2.409)	(0.366)	(0.353)
$MTPF_{i,t}$	4.853**	4.894***	–14.77*	–11.14*	0.618	0.888
	(2.156)	(1.709)	(8.225)	(6.630)	(0.623)	(0.665)
$Openness_{i,t}$	0.005	0.007	0.006	0.007	–0.000	–0.001
	(0.007)	(0.008)	(0.013)	(0.013)	(0.001)	(0.001)
$Conflict_{i,t}$	–1.351*	–1.071*	–0.770	–0.907	0.010	0.022
	(0.697)	(0.599)	(1.428)	(1.495)	(0.073)	(0.047)
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
Internal Instrum.	Yes	Yes	Yes	Yes	Yes	Yes
External Instrum.	No	Yes	No	Yes	No	Yes
AR(1) Test p-val.	0.000	0.000	0.000	0.000	0.042	0.039
AR(2) Test p-val.	0.777	0.788	0.904	0.815	0.399	0.391
Hansen Test p-val.	0.644	0.934	0.702	0.663	0.308	0.300
Sample Period	1990–2008		1996–2008		1995–2008	
Countries	155	155	164	164	163	163
Observations	2,228	2,228	1,638	1,638	1,735	1,735

Notes: The unit of observation is a country-year from the sample described in the Data Appendix. The table reports estimation results from running the baseline D-GMM specifications, columns (3)-(4) from Tables 6, 8, and 10, on a restricted sample that excludes the most autocratic countries, see table 14. Two-step estimates with Windmeijer corrected standard errors in parentheses. The models use the orthogonal deviations transform. The internal instruments are: the second and third lags of $y_{i,t}$ in columns (1)-(2), the second lag of $y_{i,t}$ in columns (3)-(6), and the second and third lags of $MTFF_{i,t}$, $MTBF_{i,t}$, $MTPF_{i,t}$. in all columns. The external instruments are: $MTFF\_Regional\_Penetration_{i,t}$, $MTBF\_Regional\_Penetration_{i,t}$, and $MTPF\_Regional\_Penetration_{i,t}$ The internal instruments enter uncollapsed. ***, **, * indicate statistical significance at the 1%, 5%, 10% levels, respectively.

Table 13: MTEFs and Fiscal Performance: Excluding Highly Developed Countries

Dep. Variable ($y_{i,t}$):	$Gov\_Balance_{i,t}$		$Spend\_Volatility_{i,t}$		$Tech\_Efficiency_{i,t}$	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
$y_{i,t-1}$	0.422***	0.425***	0.070	0.087	0.924***	0.937***
	(0.039)	(0.036)	(0.062)	(0.056)	(0.071)	(0.048)
$y_{i,t-2}$	0.103***	0.096***				
	(0.037)	(0.036)				
$MTFF_{i,t}$	1.443	1.273*	–2.632	0.187	–0.035	0.047
	(1.063)	(0.691)	(2.364)	(1.774)	(0.116)	(0.158)
$MTBF_{i,t}$	2.042	2.685**	–5.915	–5.193*	0.286	0.531
	(1.583)	(1.270)	(3.906)	(2.738)	(0.333)	(0.372)
$MTPF_{i,t}$	3.939*	3.685**	–23.63*	–15.70*	0.939	1.321*
	(2.195)	(1.507)	(10.652)	(9.352)	(0.726)	(0.801)
$Openness_{i,t}$	0.012*	0.011	0.002	0.003	–0.001	–0.001
	(0.007)	(0.008)	(0.013)	(0.013)	(0.001)	(0.001)
$Conflict_{i,t}$	–1.013	–0.787	–0.288	–0.246	0.003	0.007
	(0.809)	(0.654)	(1.434)	(1.357)	(0.046)	(0.055)
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
Internal Instrum.	Yes	Yes	Yes	Yes	Yes	Yes
External Instrum.	No	Yes	No	Yes	No	Yes
AR(1) Test p-val.	0.000	0.000	0.000	0.000	0.046	0.043
AR(2) Test p-val.	0.915	0.856	0.959	0.887	0.387	0.379
Hansen Test p-val.	0.785	0.946	0.889	0.821	0.677	0.791
Sample Period	1990–2008		1996–2008		1995–2008	
Countries	151	151	160	160	159	159
Observations	2,173	2,173	1,602	1,602	1,700	1,700

Notes: The unit of observation is a country-year from the sample described in the Data Appendix. The table reports estimation results from running the baseline D-GMM specifications, columns (3)-(4) from Tables 6, 8, and 10, on a restricted sample that excludes the most developed countries, see table 14. Two-step estimates with Windmeijer corrected standard errors in parentheses. The models use the orthogonal deviations transform. The internal instruments are: the second and third lags of $y_{i,t}$ in columns (1)-(2), the second lag of $y_{i,t}$ in columns (3)-(6), and the second and third lags of $MTFF_{i,t}$, $MTBF_{i,t}$, $MTPF_{i,t}$. in all columns. The external instruments are: $MTFF\_Regional\_Penetration_{i,t}$, $MTBF\_Regional\_Penetration_{i,t}$, and $MTPF\_Regional\_Penetration_{i,t}$ The internal instruments enter uncollapsed. ***, **, * indicate statistical significance at the 1%, 5%, 10% levels, respectively.

Table 14: Country Extremes, 1990

Highly Autocratic Countries	Highly Developed Countries
Bahrain	Australia
Bhutan	Belgium
Oman	Canada
Qatar	Japan
Saudi Arabia	Netherlands
Swaziland	New Zealand
	Norway
	Sweden
	Switzerland
	United States

Notes: Authors' calculations using the sample of 181 countries described in the Data Appendix. A country is defined as highly autocratic if its Polity IV score in 1990 takes the extreme value −10. A country is defined as highly developed if classified by UNDP in 1990 as having "very high human development" based on the country's Human Development Index.

A3. Data Appendix

This appendix contains the complete list of variables used in the paper, together with details on measurement and sources.

Fiscal Performance

Government Balance: Ratio of the overall central government fiscal balance to GDP, in percent. *Sources:* IMF World Economic Outlook.

Spending Volatility: Absolute growth rate in health spending per capita, in PPP$. *Sources:* Authors' calculations. See equation (1) in the paper.

Technical Efficiency: Estimations of efficiency scores from a stochastic frontier model that shows life expectancy as output and health spending per capita in PPP$ as input. *Sources:* Authors' calculations. See equations (2) and (3), and Table 4 in the paper.

Budget Institutions

MTFF: Dummy variable that takes the value one if MTFF is the highest MTEF phase adopted, zero otherwise. *Sources:* World Bank and IMF documents and country specialists, case studies.

MTBF: Dummy variable that takes the value one if MTBF is the highest MTEF phase adopted, zero otherwise. *Sources*: World Bank and IMF documents and country specialists, case studies.

MTPF: Dummy variable that takes the value one if MTPF is the highest MTEF phase adopted, zero otherwise. *Sources:* World Bank and IMF documents and country specialists, case studies.

MTFF Regional Penetration: The percentage of MTFF adopters in the country's geographic region. We use the twenty-two geographic regions defined by the United Nations Statistics Division. See equation (4) in the text. *Sources:* Authors' calculations.

MTBF Regional Penetration: The percentage of MTBF adopters in the country's geographic region. We use the twenty-two geographic regions defined by the United Nations Statistics Division. See equation (4) in the text. *Sources:* Authors' calculations.

MTPF Regional Penetration: The percentage of MTPF adopters in the country's geographic region. We use the twenty-two geographic regions defined by the United Nations Statistics Division. See equation (4) in the text. *Sources:* Authors' calculations.

Country Characteristics

Conflict: Dummy variable that takes the value one if there are at least 1,000 battle-related casualties, zero otherwise. *Sources:* UCDP/PRIO Armed Conflict Dataset, Uppsala University.

Health Spending per Capita: Health Expenditure per capita in PPP$ terms. *Sources:* World Health Organization.

HDI: Human Development Index, a composite index measuring average achievement in three basic dimensions of human development: a long and healthy life, knowledge, and a decent standard of living. *Sources:* United Nations Development Programme.

Life Expectancy: Life expectancy at birth, in years. *Sources:* World Bank World Development Indicators.

OECD Membership: Dummy variable that takes the value one if the country belongs to the OECD, zero otherwise. *Sources:* OECD.

Openness: Trade openness measured as the ratio of the sum of imports plus exports to GDP. *Sources:* World Bank World Development Indicators.

Polity Score: Composite score ranging from -10 (strongly autocratic) to 10 (strongly democratic) based on the Polity IV methodology. *Sources:* Marshall, Jaggers, and Gurr (2010).

Population Density: Residents per square kilometer. *Sources:* World Bank World Development Indicators.

Region: One of twenty-two geographical regions. *Sources:* United Nations Statistics Division.