Multiyear Budgets and Fiscal Performance: Panel Data Evidence*

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

December 2013

Abstract

In the last two decades more than 120 countries have adopted a multiyear budget process (Medium-Term Framework, or MTF) that enables the central government to set multiyear fiscal targets. This paper analyzes a newly-collected dataset of worldwide MTF adoptions during 1990-2008. It exploits within-country variation in adoption in a dynamic panel framework to estimate MTFs' impacts on aggregate as well as sectoral measures of fiscal performance. We find that on average multiyear budgeting improves budget balance by about 2 percentage points with more advanced MTF phases having a larger impact. Higher-phase MTFs also reduce health spending volatility, while only the top-phase MTF has a measurable impact on health sector technical efficiency.

JEL Classification: E62, H51, H62.

Keywords: budget institutions; medium-term framework; fiscal discipline; health sector; technical efficiency; dynamic panel data analysis.

---

*We would like to thank Abdul Abiad, Samuel Bazzi, Nataliya Biletska, Jim Brumby, Rodolphe Desbordes, Stefania Fabrizio, Jessica Goldberg, Kazuhiko Hayakawa, Richard Hemming, Malcolm Holmes, Sam Ouliaris, Ethan Kaplan, Philip Keefer, Steve Knack, Jeff Smith, Ajay Tandon. 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.

†University of Maryland, Department of Economics, and Department of Government and Politics, 3114 Tydings Hall, College Park, MD 20742, USA. E-mail: vlaicu@econ.umd.edu.

‡The World Bank, Poverty Reduction and Economic Management, 1818 H Street, NW, Washington, DC 20043, USA. E-mail: mverhoeven@worldbank.org.

§IMF, Fiscal Affairs Department, 1900 Pennsylvania Ave., NW, Washington, DC 20431, USA. E-mail: fgrigoli@imf.org.

¶The World Bank, Public Sector and Governance Unit, 1818 H Street, NW, Washington, DC 20043, USA. E-mail: zmills@worldbank.org.

1 Introduction

It is generally accepted that fiscal performance is a key factor in a country's long-run growth.[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 the majority of the world's nations have adopted laws instituting multiyear fiscal targets, known as *Medium-Term Frameworks* (*MTFs*). First introduced in a small set of developed countries to contain expenditure overruns in the 1980s, MTFs spread rapidly during the 1990s and 2000s. From 11 countries in 1990 the number of adopters grew to 132 at the end of 2008; see Figure 1. MTFs 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 multiyear budgets are designed to overcome is dynamic fiscal inefficiency. This can manifest itself as strategic obstruction of future political opponents (Alesina and Tabellini 1990), 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). Through dynamic distortions government spending and borrowing deviate from the social planner level, resulting in suboptimally high deficits and debt.[2] 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 expenditures are reviewed; overspending, because huge disbursements in future years are hidden; conservatism,

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

[2] To the extent that MTFs increase budget transparency and accountability they may also alleviate static public finance distortions such as free riding (Weingast, Shepsle, and Johnsen 1982), rent seeking (Persson, Roland, and Tabellini 2000, Besley and Smart 2007), and clientelism (Keefer and Vlaicu 2008).

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 MTF is a commitment by the budget actors to a medium-term, typically two to four years, fiscal trajectory.[3] 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 (von Hagen and Harden 1995).[4] While the theoretical underpinnings of multiyear budgeting are well understood, the empirical evidence on its impact is scarce. The main obstacle has been the shortage of data on MTF adoption. An additional impediment has been the elusiveness of plausible sources of exogenous variation in country-level institutions in general (Acemoglu 2005) and budget institutions in particular.[5]

When an MTF is implemented well we should observe (i) spending that is limited by resource availability (aggregate fiscal discipline), (ii) predictable budget allocations across sectors (sectoral stability), and (iii) cost effective public goods delivery (technical efficiency). We classify MTFs into three *phases*: Medium-Term Fiscal Framework (MTFF, which establishes the aggregate resource envelope), Medium-Term Budgetary Framework (MTBF, which focuses on the allocation of spending across sectors, programs, and agencies), and Medium-Term Performance Framework (MTPF, which sets within-sector performance targets). The three MTF phases are "nested": an MTPF contains an MTBF, which in turn contains an MTFF.[6]

This paper is the first large-sample empirical study of the MTFs' impacts on fiscal performance. We collect and analyze adoption data for a panel of 181 countries over the period 1990-2008, the most comprehensive dataset on worldwide MTF adoption to date. We provide a system for classifying MTFs into three phases, based on their level of sectoral disaggregation, and estimate the effects of each phase on aggregate as well as sectoral measures of fiscal performance.

[3]See Auerbach (2006) for a political economy model of optimal budget window length.

[4]As opposed to a "delegation approach" where one actor, typically the executive, receives enhanced powers.

[5]Fabrizio and Mody (2006), discussing the country-level literature, notice that "Identifying an 'instrument,' or a variable that influences the fiscal institutions but is not itself influenced by budgetary outcomes, is a hurdle that no one has yet crossed." (p. 703). One way to circumvent econometric identification issues has been to study similar institutions operating at sub-national levels of government. The state-level literature does propose instruments that help identify budget institutions' impacts; see, e.g., Knight (2000).

[6]This taxonomy borrows from that of Castro and Dorotinsky (2008), although the nesting concept is an innovation of our dataset.

The rich time variation in the data allows us to model the dynamics of the fiscal adjustment process as well as address potential endogeneity of MTF 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 in the absence of valid and/or strong external instruments, a typical hurdle with country-level data.[7]

Our empirical results show that multiyear budgeting improves fiscal discipline by about 2 percentage points of budget balance (surplus/deficit) to GDP, on average. The effect is larger for the more advanced phases. We also find that an MTBF and an MTPF improve sectoral stability, by reducing the volatility in per capita health spending (in purchasing power parity dollars, PPP$), and that an MTPF also contributes to health sector technical efficiency, measured as technical efficiency scores from a stochastic frontier model of public health delivery. We probe the credibility of these estimates by carrying out a detailed analysis of instrument strength and instrument validity and by exploring sensitivity to alternative specifications. We keep the empirical model simple and transparent by restricting the number of moment conditions through a parsimonious lag structure and collapsed instruments.

Our results are more supportive of MTF effectiveness than the conclusions of prior work. Bevan and Palomba (2001), La Houerou and Taliercio (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 MTF, while Boex, Martinez-Vasquez, and McNab (2000) and Oxford Policy Management (2000) raise questions of adequate implementation. Wescott (2008) and Filc and Scartascini (2010), using data from Central and Latin America, found mixed results and emphasized the importance of piloting MTFs in areas where they are likely to deliver the largest payoffs.[8]

However, Gleich (2003) and Ylaoutinen (2004) find that MTFs in Central and Eastern Europe alleviated the deficits and debts that emerged in the second half of the 1990s. Alesina, Hausmann, Hommes, and Stein (1999) and Fabrizio and Mody (2006) include MTFs in broader indexes of budget institutions using data from Latin America and Eastern Europe, respectively, and find positive effects of the resulting "fiscal institutions index" on the primary-balance-to-GDP ratio.[9]

[7] 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).

[8] Drawing on extensive operational experience, Schiavo-Campo (2009) puts forward conceptual arguments for a gradual introduction of these institutions and emphasizes the potential downsides of instant reform.

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

Our results complement this empirical literature by providing evidence of MTF impacts from a global sample. While previous studies rely on small regional samples and either cross-sectional or static panel models our empirical methodology takes advantage of the sample coverage and the time variation in MTF presence to estimate dynamic panel models. We also propose a new classification of MTF 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, e.g., balanced budget amendments, debt ceilings, tax and expenditure limitations. 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.[10]

We contribute to this literature by, first, providing new data and measurement on an important budget institution; second, by proposing a dynamic panel approach (difference GMM) that models the fiscal adjustment process while making progress in addressing difficult identification issues such as institutional endogeneity; and third, by studying in addition to aggregate effects, which have been the focus of this literature, sectoral effects as well, namely fiscal performance in the key health sector.

The structure of the paper is as follows. Section 2 provides a background discussion of the MTFs, 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.

2 Background

This section takes a first look at the worldwide MTF adoption data collected for this paper by presenting stylized facts on MTF global and regional growth during 1990-2008. It also discusses the rationale behind MTFs as budget institutions designed to improve fiscal performance. This discussion helps generate theoretical expectations. We subject these hypotheses

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

to empirical scrutiny in the next section.

MTF Phases. We classify MTFs into three phases, based on the following criteria.[11]
- 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.[12]
- 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.

The three phases are "nested" in the sense that a more advanced phase contains the lower phase just below it.

Stylized Facts. Although some forms of medium-term fiscal 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 MTF was introduced in the 1980s (see Folscher 2007). MTFs have since been adopted by a large number of low and middle-income countries as a central element of public financial management reform.

An average of 10 countries per year introduced an MTF between 1996 and 2008. By the end of 2008, 132 countries, or about two-thirds of the globe, had an MTF. Initially, most multiyear budgets were of the first phase, or MTFF, and until recently about two-thirds of the increase in multiyear budgeting has been in the form of new MTFFs. However, there has been a recent uptick in the number of MTBFs and MTPFs; see Figure 1. In 2008 there were 71 MTFFs, 42 MTBFs, and 19 MTPFs. Table 2 shows that the shift to MTBFs and MTPFs has been mainly, and in the case of MTPF, exclusively, through transitions from a lower phase to an advanced one. In this sense the nesting concept underlying our dataset is reflected in observed adoption patterns.[13]

---

[11] 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 Online Appendix for a description of variable construction, and World Bank (2012) for further details.

[12] We coded countries that introduced a "pilot" MTBF in a few sectors as MTFF since we cannot systematically determine if the health sector, our focus in the analysis below, is one of the piloted sectors.

[13] Three countries (Bulgaria, Canada, and Norway) performed a full transition - from an MTFF to an MTBF to an MTPF - during our sample period.

Advanced economies had achieved almost complete coverage (96%) by the end of our sample period. MTF 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 MTF adoption. Then, in the late 1990s MTFs 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 MTFs in advanced economies were MTPFs. The relatively low fraction of the second-phase 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.

While MTFs had begun 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. The end result has been that adoption patterns have been relatively uniform across income and development levels. Apart from the widespread adoption of MTFs in high-income countries, there is little difference in penetration across upper middle, lower middle, and lower-income countries. MTF adoption does not appear to follow a monotonic relationship with respect to income per capita or the human development index.[14]

MTF Objectives. MTFs 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.[15] MTFs 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 multiyear budgeting 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, sectoral stability, and technical efficiency.[16]

[14] Figures A1 and A2 in the Online Appendix illustrate this point.

[15] 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.

[16] There is also a link to broader economic development. With improved fiscal outcomes, growth should be

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).[17]

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 *sectoral stability*. Insofar as spending agencies prepare sector strategies, identify their resource needs, and allocate their budgets according to strategic priorities, this "bottom-up" prioritization should increase the predictability of spending with higher economic and social returns. IMF (2007) argues that expenditure decisions not anchored in a medium-term framework could generate "mismatches between available resources and planned spending and could translate into expenditure volatility with adverse consequences for economic and social outcomes." (p. 22). In the health sector in particular spending volatility can lead to stock-outs, higher per item costs, as suppliers start charging premiums to protect against payment and production uncertainties, and additional emergency costs.[18]

The outcome of effective prioritization should be a change in the allocation of spending. In the short term, spending volatility by sector may increase 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

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.

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

[18] See Lane and Glassman (2008) on how unpredictable health spending negatively affects health outcomes.

MTBF should be even higher with an MTPF since this last phase goes further by setting within-sector and within-program performance targets.[19]

A third dimension of fiscal performance is *technical efficiency* at the sector level. 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 now include within-sector performance targets. Of all government functions, the outcomes of government-provided health care are probably easier to compare across nations, which has spawned a significant literature in estimating technical efficiency scores in the health sector cross-nationally (see, e.g., Greene 2004).

Based on these considerations we state the expected MTF effects on fiscal performance in the following hypotheses:

(H1) Multiyear budgeting improves fiscal discipline, with higher-phase MTFs having larger effects.

(H2) MTBF and MTPF improve sectoral stability, with an MTPF having a larger effect.

(H3) MTPF improves sectoral technical efficiency.

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

3 Data and Empirical Strategy

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

Data. Our sample consists of 181 countries over the period 1990-2008. The country sample reflects availability of MTF adoption data. The period sample reflects availability of public finance data. Here we briefly discuss the key variables. The Online Appendix contains the complete list of variables together with their data sources.

The construction of the MTF indicators relied upon an extensive data collection effort as no single type of document fully describes the existing institutional arrangements for all countries or individual countries. The data were compiled from a large number of sources

[19] 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, while productive spending on economic and social infrastructure, health and education services, and other growth- and development-promoting activities should be favored.

that include: IMF Article IV country reports, fiscal transparency modules from IMF Reports on the Observance of Standards and Codes (ROSC), World Bank Public Expenditure Reviews (PERs), World Bank Country Financial Accountability Assessments (CFAAs), OECD documents, donor case studies, and national government websites. Additionally, IMF and World Bank public sector and country specialists supplemented the above information with technical details.[20]

It is important to note that MTF reform is very rarely an abrupt switch and more commonly a prolonged and gradual process: "big bang adoption of MTF [...] in not the norm." (World Bank 2012, p. 88). Countries that used the gradual approach to reform have often prefaced full implementation by piloting the MTF in a limited number of sectors, e.g., education, health, infrastructure. We have chosen to code the year of adoption as the year when the MTF is fully functional. That means a country may have MTF components in place even before it appears as an adopter in our data. Figure 2 provides a visual illustration. The figure plots the average budget balance before/after the year of MTF adoption. The series suggest that MTF adoption follows after several years of persistent deficits in the neighborhood of 3% of GDP. In the two years before full implementation the average deficit starts to shrink; after adoption the average balance improves further and the deficit hovers around 1.5% of GDP.[21]

We measure fiscal discipline, an indicator of aggregate fiscal performance, by the ratio of total budget balance (surplus/deficit) to GDP. The literature suggests alternative indicators, e.g., primary balance and the debt-to-GDP ratio. We preferred the total budget balance because we observe it for 23 more countries than the primary balance (and 33 more countries than debt-to-GDP) affording us 18-22% more panel observations. We also believe that by including government borrowing costs the total budget balance is a more complete indicator of the general state of public finances. For ease of comparison with prior work we also present results with the primary balance; see Table 10.[22]

[20] For the purposes of this paper, we refrain from making judgments to distinguish between an MTF present in the law (de jure) and a well-functioning MTF (de facto). Such a distinction would introduce subjectivity into the analysis. A significant fraction of poorly-implemented MTFs would attenuate our empirical estimates.

[21] Other interpretations are that smaller-scale reforms, e.g., limits on deficit or debt, are attempted before moving on to the more comprehensive multiyear approach, or, that prior to adoption the executive, which typically opposes the constraints implicit in an MTF, may try to diffuse legislative momentum for reform with temporary cost-cutting deficit-reducing measures.

[22] In our sample the correlation between total budget balance and primary balance is 0.90. It could be argued that the total balance does not account for the effect of inflation on interest payments and that interest payments are a function of accumulated debt and not a good reflection of the government's current fiscal stance. Thus the MTF effects on the total balance should be comparatively smaller. Our estimates in Tables 9 and 10 are consistent with this logic.

Sectoral stability does not have a universally accepted definition. However, a plausible indicator seems to be the volatility of sectoral spending, e.g., health or education. Since volatility in these sectors jeopardizes long-term economic and social objectives, health care and public education spending should be largely unaffected by short-term fluctuations in GDP. Given data constraints and the requirement that the public good category should be plausibly comparable across countries, we choose to work with the volatility of per-capita health spending, measured in PPP$. We define volatility of a time series $y_{i,t}$ for country i as the absolute yearly growth rate of the detrended series: $Volatility_{i,t} = |\log(\tilde{y}_{i,t}/\tilde{y}_{i,t-1})| \times 100$, where $\tilde{y}_{i,t}$ is the linearly detrended series for $y_{i,t}$.

Technical efficiency in the health sector 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 countries with the highest health output after controlling for health inputs are the most efficient and delineate the frontier. The efficiency level of the other countries is measured relative to this benchmark.[23]

We compute technical efficiency scores in the health sector using a parsimonious version of the model estimated in Greene (2004). The outcome of interest is life expectancy at birth, and the input is health spending per capita. The covariates are population density and OECD membership. The model is:

$$\begin{aligned} \log(Life\_Exp_{i,t}) &= \beta_0 + \beta_1 \log(Health\_Spend_{i,t}) + \\ &\quad + \beta_2 Density_{i,t} + \beta_3 OECD_{i,t} + \tau_t + v_{i,t} - u_{i,t} \end{aligned} \tag{1}$$

where τ_t's are year fixed effects, $v_{i,t} \sim N(0, \sigma_v)$, and $u_{i,t} = |U_{i,t}| \sim Exp(0, \sigma_u)$.[24]

The parameters are then 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 can be shown to be:

$$Tech\_Efficiency_{i,t} = e^{-\hat{u}_{i,t}} \tag{2}$$

[23] 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 realized output. Inefficiencies might arise from waste or because the most cost-effective set of programs is not implemented.

[24] Greene (2004) also includes education spending per capita as an input, and controls in the inefficiency distribution for time invariant measures of government voice and accountability, government effectiveness, share of government financing, the Gini coefficient, and GDP per capita.

where $\hat{u}_{i,t}$ is the maximum likelihood estimate of $u_{i,t}$.

Table 3 presents estimates of the production function in (1). 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. Our estimated mean efficiency is 86.48.

Figures 3 and 4 give a graphical intuition for the efficiency scores by plotting country averages against control of corruption and a measure of development, together with a quadratic fit curve. As expected, technical efficiency increases in both of these measures of effective governance. In Figure 5 we plot the time series of technical efficiency for four countries that had an MTPF reform during our sample period. The newly democratic countries on the right (both of whom adopted democratic constitutions in the early 1990s) experience more volatility in their efficiency scores, likely due to their ongoing institutional reforms.[25]

Table 1 reports summary statistics for our main variables. 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. Table A1 in the Online Appendix reports pairwise correlations between the main variables. The correlation coefficients are within plausible ranges and their signs are consistent with the data patterns noted above.

Empirical Model. Governments use the national budget as a tool for achieving economic objectives. In a steady-state equilibrium the trajectory of the government's budget balance thus reflects the benefits and costs of financing the goverment's economic policy, given the budget institutions in place and the characteristics of the economy. Let $Y_{i,t}$ denote the government's target level of a fiscal variable like budget balance. It can be modeled as:

$$Y_{i,t} = \mathbf{x}'_{i,t}\mathbf{b} + \mathbf{w}'_{i,t}\mathbf{d} + \varepsilon_{i,t} \tag{3}$$

where $\mathbf{x}_{i,t}$ are fiscal institutions, $\mathbf{w}_{i,t}$ are covariates for country i at time t that may include an intercept, and $\varepsilon_{i,t}$ is a mean zero error term that captures unobserved heterogeneity.

In practice, observed fiscal performance in a given year $y_{i,t}$ may deviate from its target

[25] We experimented with alternative specifications of the stochastic frontier model, such as half-normal and truncated normal inefficiency distributions, as well as controlling for time invariant income per capita and government effectiveness, or country fixed effects, either in the production function or in the inefficiency distribution. We chose the model that produced summary statistics closest to those of previous studies. Available time series for education spending and schooling are much shorter than for health spending, preventing us from including them as inputs in the production function (2). The results of this analysis are available from the authors upon request.

level $Y_{i,t}$ due to adjustment costs, e.g., legislative delays in passing executive initiatives, unpredictability of budgeted revenues, unforseen expenditures due to cost overruns or inflation. To capture these factors we can specify a target adjustment model:

$$y_{i,t} - y_{i,t-1} = (1-\gamma)\left(Y_{i,t} - y_{i,t-1}\right) \tag{4}$$

where $\gamma \in [0,1]$ is the adjustment cost. This means for example that if $\gamma = 0$ then $y_{i,t} = Y_{i,t}$ namely the adjustment toward the target takes place immediately. Combining equations (3) and (4) we can write a model for the observables:

$$y_{i,t} = \gamma y_{i,t-1} + \mathbf{x}'_{i,t}\boldsymbol{\beta} + \mathbf{w}'_{i,t}\boldsymbol{\delta} + u_{i,t} \tag{5}$$

where $(\boldsymbol{\beta}, \boldsymbol{\delta}, u_{i,t}) = (1-\gamma) \times (\mathbf{b}, \mathbf{d}, \varepsilon_{i,t})$. Incomplete adjustment ($\gamma \neq 0$) thus leads to a form of state dependence where last period's $y_{i,t-1}$ determines this period's $y_{i,t}$.[26]

A more general version of model (5) would include further lags of $y_{i,t}$ as well as year effects τ_t and time-invariant unobserved country-specific heterogeneity c_i, which likely plays an important role, leading to the following dynamic specification:

$$y_{i,t} = \mathbf{y}'_{i,t-s}\boldsymbol{\gamma} + \mathbf{x}'_{i,t}\boldsymbol{\beta} + \mathbf{w}'_{i,t}\boldsymbol{\delta} + \tau_t + c_i + u_{i,t} \tag{6}$$

with $s = 1, 2, ..., L$ and $\mathbf{x}'_{i,t} = MTF_{i,t}$ or $\mathbf{x}'_{i,t} = (MTFF_{i,t}, MTBF_{i,t}, MTPF_{i,t})$ and the idiosyncratic error $u_{i,t}$ is assumed to be mean zero. Besides providing a more realistic model of fiscal adjustment, this dynamic panel model also partly controls for possible reverse causality. For example, if past budget performance $y_{i,t-1}$ affects current fiscal institutions $\mathbf{x}_{i,t}$ then this feedback effect is accounted for in model (6).

Identification. Consistently estimating the impact of a budget institution on fiscal performance requires that we address two key identification challenges: reverse causality and omitted variable bias. First, reverse causality arises because fiscal stress may have prompted a country to restrain spending, adopt an MTF, 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 MTFs have positive effects on fiscal performance, and poor fiscal performance increases the probability of adopting an MTF, then the reverse causality bias is likely negative.[27]

[26] Alternatively, correlation between $y_{i,t}$ and $y_{i,t-1}$ could be induced by unobserved heterogeneity.

[27] The endogeneity of budget institutions with respect to fiscal performance is extensively discussed in

Second, omitted variable bias arises due to failure to account for a factor that affects both the adoption of an MTF and fiscal performance. For instance, strong macro performance 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 use within-country variation in fiscal institutions. This approach in effect controls for time invariant country-level heterogeneity that may influence budget deficits. The problem of omitted variables is thus alleviated; however it is not eliminated.[28]

The dynamic panel formulation in (6) while partly addressing reverse causality, may however introduce an additional source of endogeneity. To see this notice that without controlling for the fixed effects (i.e., assuming $c_i = 0$ for all i) the dynamic model has a built-in positive bias in the coefficient on the first lag of the dependent variable γ_1, since $\mathbf{y}_{i,t-s}$ are correlated with c_i. Explicitly controlling for the fixed effects using an within FE estimator eliminates this dynamic bias but γ_1 now contains a negative bias due to the fact that the within-group transform $y^*_{i,t-1} = y_{i,t-1} - \frac{1}{T-L}\sum_{k=L+1}^{T} y_{i,k}$ is negatively correlated with $u^*_{i,t} = u_{i,t} - \frac{1}{T-L}\sum_{k=L+1}^{T} u_{i,k}$, where L is the number of observations lost by including L lags of the dependent variable. Thus an unbiased estimate of the lagged dependent variable coefficient $\hat{\gamma}_1$ should lie in the range between the FE estimate and the OLS estimate. This "bracketing range" then provides a natural specification check (Bond 2002) and we use it in our own results below to assess the validity of our chosen specifications.[29]

An alternative approach to controlling for the fixed effects is to first-difference the model in (6) to eliminate the fixed effects c_i :

$$y^*_{i,t} = \sum_{s=1}^{L} \gamma_l y^*_{i,t-s} + \beta MTF^*_{i,t} + \mathbf{w}^{*\prime}_{i,t}\boldsymbol{\delta} + \tau^*_t + u^*_{i,t} \tag{7}$$

where $y^*_{i,t} = y_{i,t} - y_{i,t-1}$. Because our panel is unbalanced, in the baseline specifications we

the country-level literature (see Alesina and Perotti 1999, Stein, Talvi, and Grisanti 1999, Perotti and Kontopoulos 2002, and Fabrizio and Mody 2006). None of these papers, however, proposes an instrument that influences the probability of fiscal reform while being exogenous to fiscal performance.

[28] 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. 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).

[29] Nickell (1981) shows that the "dynamic panel bias" disappears only when T approaches infinity. Since in our data $T \leq 19$, it is necessary to apply methods that correct for this bias.

use instead a related transformation, namely forward orthogonal deviations defined as:

$$y^*_{i,t} = \sqrt{\frac{T_{i,t-1}}{T_{i,t-1}+1}} \left(y_{i,t-1} - \frac{1}{T_{i,t-1}} \sum_{s \geq t} y_{i,s} \right) \tag{8}$$

where $T_{i,t}$ is the number of available future observations.[30] Orthogonal deviations, instead of subtracting the previous observation from the current observation, subtract the average of all future available observations. This transformation maximizes sample size in panels with gaps (Arellano and Bover 1995). Note that using either transformation makes $y^*_{i,t-1}$ correlated with $u^*_{i,t}$ in model (7). Thus, the OLS estimator would still be inconsistent for this model. However, under appropriate conditions, a panel GMM estimator for (7) would be consistent.[31]

Our baseline identification assumptions are that, given $t = 2, ..., T$:

$$\begin{aligned} E(u_{i,t-1}u_{i,t}) &= 0 && (9) \\ E(MTF_{i,t-s}u_{i,t}) &= 0, \text{ for } s = 1, 2, ..., t-1 && (10) \end{aligned}$$

and we treat $\mathbf{w}_{i,t}$ as strictly exogenous: $E(\mathbf{w}'_{i,t\pm s}u_{i,t}) = \mathbf{0}$ for all s.[32] The above equations say that the idiosyncratic error is serially uncorrelated and that past values of the MTF are not correlated with the current error. Notice that assumption (10) is weak in the sense that it allows for both contemporaneous correlation between MTF and the error, for instance due to omitted time-varying factors or simultaneity, as well as for feedback effects from $u_{i,t}$ to future MTF. These assumptions generate the following moment conditions (see, e.g., Cameron and Trivedi 2005):

$$\begin{aligned} E(y_{i,t-s}u^*_{i,t}) &= 0, \text{ for } s = 2, ..., t-1 && (11) \\ E(MTF_{i,t-s}u^*_{i,t}) &= 0, \text{ for } s = 2, ..., t-1 && (12) \end{aligned}$$

which mean that: (a) one can use second and later lags of $y_{i,t}$ as instruments for $y^*_{i,t-1}$ and potentially other lags, and (b) one can use second and later lags of $MTF_{i,t}$ as instruments

[30] The subscript $t-1$ on the right side of (8) reflects the standard practice of dating ortogonal deviations one period late, for consistency with the first-difference transform.

[31] Hayakawa (2009a) presents Monte Carlo simulations that suggest that the GMM estimator for a dynamic panel data model transformed by orthogonal deviations tends to perform better than for one transformed by first differences in terms of biases, standard deviations, and RMSE.

[32] We include all covariates $\mathbf{w}_{i,t}$ lagged one period to minimize the potential for endogeneity with the dependent variable (Baltagi, Demetriades, and Law 2009).

for $MTF^*_{i,t}$.

Below we provide auxiliary evidence that these conditions hold in our setting and study the sensitivity of our results to their potential violation. Here we mention two ways to empirically examine (9)-(10). First, a basic test for the no serial correlation condition in (9) can be implemented as a second-order serial correlation test for the differenced residuals (Arellano and Bond 1991) and we report this test with all our specifications. To strengthen the case for lack of serial correlation we include year fixed effects τ_t in all specifications.[33] Second, to check for potential violations of (10) we have examined the lagged correlation between MTF and several factors that prior literature has linked to fiscal discipline: macro, international, political, and institutional; see Table 4. Based on the two first panels of this table we select the statistically significant correlates as covariates for our baseline model. These are: openness, conflict, balanced budget rule, and debt rule.[34]

Estimation and Inference. We implement the instrumentation strategy described above using the Difference Generalized Method of Moments (D-GMM) approach by constructing GMM-style instruments, i.e., replace missing observations with zeros. We denote them as $y^{GMM}_{i,t-2}$, $MTF^{GMM}_{i,t-2}$ and similarly for other lags. The D-GMM estimator is consistent under the identification assumptions (9)-(10); see, e.g., Wooldridge (2002). The number of moment conditions increases with T and negatively affects the performance of the estimator in finite samples. Too many moment conditions can also overfit endogenous variables, failing to adequately address the aforementioned endogeneity problems. Thus, it is important to keep the number of instruments in check (Roodman 2009). Two ways of achieving this goal are: limiting the number of lags and collapsing the instrument matrix. In our baseline specifications we use both: we limit the instrument lag structure to one lag for each endogenous variable and we collapse the instrument matrix, unless otherwise noted.[35]

Our baseline specification computes two-step D-GMM estimates with standard errors corrected with the Windmeijer (2005) procedure. We use the two-step standard error correction because the original variance formula has been shown to produce two-step standard

[33] Notice that $y_{i,t-2}$ and later lags of the dependent variable are relevant instruments for $y^*_{i,t}$ unless $y_{i,t}$ is close to a random walk, in which case past levels convey little information about future changes. Table A2 in the Online Appendix presents unit root test results for our four fiscal performance measures. The IPS test statistic safely rejects the null hypothesis of unit root in each case.

[34] The case for strict exogeneity of the covariates balanced budget and debt rule is perhaps less strong. However, when we omit them, or treat them as predetermined, instrumenting their transformation with their second lag, the results reported below change little.

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

errors that are implausibly small.[36]

Our empirical strategy generates credible estimates as long as the instruments are sufficiently strong predictors of the endogenous variables. A standard test for instrument strength in dynamic panel GMM regressions does not currently exist, so gauging instrument strength empirically is not trivial.[37] We follow an approach for assessing instrument strength based on Hayakawa (2009b) and Bun and Windmeijer (2010), and recently applied to the case of multiple endogenous regressors by Bazzi and Clemens (2013). Specifically, we construct the instrument matrix and estimate the difference equation (7) in a 2SLS framework. This allows us to compute an LM statistic for underidentification, which tests the null that the excluded instruments are uncorrelated with the multiple endogenous regressors, as well as an F statistic for the predictive strength of GMM-style instruments in our D-GMM framework.[38]

An alternative estimation approach is the System Generalized Method of Moments (S-GMM) proposed by Blundell and Bond (1998); this estimator may increase efficiency by further instrumenting for the endogenous variables in the levels equation with their lagged differences, but requires an additional identifying assumption, namely that the instruments are exogenous to the fixed effects. Here we choose D-GMM over S-GMM for two reasons: (i) our analysis of internal instrument strength (see Tables 5b and 7) suggests that weak instrument bias, the typical justification for S-GMM, may not be so serious in our context, since MTF once adopted persists, so past levels can predict future changes reasonably well, and (ii) the additional identification assumption required by S-GMM, namely that MTF reform is uncorrelated with time-invariant country characteristics, is untestable and may be difficult to defend, raising instrument validity concerns. However, for comparison purposes, we also present a S-GMM version of our baseline specification (compare columns 1 and 6 in Table 7).[39]

[36] See Windmeijer (2005) for D-GMM regressions on simulated panels that support this assertion. We also computed one-step D-GMM estimates, in which case we use cluster-robust standard errors, i.e., robust to heteroskedasticity and arbitrary patterns of correlation within countries. Results are very similar and are available form the authors upon request.

[37] See Stock and Wright (2000) for why weak-instrument diagnostics for linear IV regressions do not apply to the more general setting of GMM.

[38] Bun and Windmeijer (2010) argue that the weak-instrument test statistics derived in cross-section are informative about instrument strength in the panel 2SLS equations of the type we estimate.

[39] In addition, recent research has challenged the perceived superiority of S-GMM in contexts with weak internal instruments. Hayakawa (2009b) and Bun and Windmeijer (2010) find that S-GMM may not be as robust to weak instrument bias as previously thought.

4 Estimation Results

We present our estimation results in the following order. We start with a benchmark model, namely equation (6) with $\mathbf{x}'_{i,t} = MTF_{i,t}$, i.e., not distinguishing between the three phases of multiyear budgeting. Our main fiscal performance measure is fiscal discipline, measured by total budget balance as a percent of GDP. We first estimate a simple specification in Table 5a and then replicate it in a 2SLS framework to assess the strength of its instruments in Table 5b. In Table 6 we explicitly model the simultaneity between budget balance and MTF. In Tables 7 and 8 we study the robustness of the results to potential violations of our identification assumptions. We then move to examining the extent to which the different multiyear budgeting phases $\mathbf{x}'_{i,t} = (MTFF_{i,t}, MTBF_{i,t}, MTPF_{i,t})$ have differential effects on fiscal discipline in Table 9. Finally, in Table 10 we study alternative measures of fiscal performance, namely primary budget balance, health spending volatility, and health technical efficiency. These estimates provide additional insight into which multiyear budgeting phases affect which dimensions of fiscal performance.

Benchmark Model. In Table 5a we start with a static model of budget balance and do a side-by-side comparison with its dynamic counterpart. Columns (1) and (2) report OLS and FE estimates of the static model. The OLS coefficient for MTF is 0.945 percentage points and not statistically significant while the FE coefficient is negative, closer to zero, and also not statistically significant. In column (3) we instrument for MTF adoption using $MTF^{GMM}_{i,t-2}$ and estimate the model by D-GMM. The MTF coefficient is now 2.569 percentage points and statistically different from zero. In columns (4)-(6) we estimate a dynamic version of the model, consistent with the one derived above in equation (6). As discussed above, the dynamic model partly addresses reverse causality from budget balance to MTF: within-country, an experience of past deficits may increase the likelihood of adopting an MTF. Some evidence of this effect can be seen by comparing column (2) to column (5) where the FE coefficient of the dynamic model turns positive, suggesting a negative bias in the static model.[40]

If MTF positively affects balance and the dynamic model is correct, so past balance up to two lags is positively correlated with current balance, then the static model in column (3) violates the exclusion restriction for the instrument $MTF^{GMM}_{i,t-2}$ and would likely overestimate the MTF coefficient. Indeed, in column (3) the MTF coefficient is half of a percentage point higher compared to its dynamic counterpart in column (6), namely 2.569 vs. 2.047.

[40] Across countries, a reverse causality effect of a similar type might be counteracted by institutional costs for adopting this budgeting system.

Table 5b probes the strength of the GMM-style instrument by replicating the D-GMM regressions in a standard 2SLS framework, following Hayakawa (2009b), Bun and Windmeijer (2010), and Bazzi and Clemens (2013). We produce orthogonal deviations transformations of the variables, marked with an asterisk (*), based on equation (8), and then compute the first stage and second stage least squares estimates of the transformed model; see columns (1),(4) and (2),(5), respectively. This approach allows us to use standard underidentification and weak identification statistics as indicators for the relevance and strength of our instruments. The first-stage estimates suggest significant correlations between the instrument $MTF^{GMM}_{i,t-2}$ and the endogenous variable $MTF^{*}_{i,t}$.[41] In both the first and second stages the underidentification test rejects the null of no correlation between the instruments and the endogenous regressors. Moreover, the weak identification F statistic has reasonably large values, alleviating concerns about instrument strength. Columns (3) and (6) then present the reduced-forms for the dependent variable.[42]

In Table 6 we explicitly model the potential simultaneity between MTF and budget balance hinted at by the results of Table 5a. Columns (1),(2) report the OLS estimates of a simple simultaneous equations model between $Balance_{i,t}$ and $MTF_{i,t}$ that includes country and year fixed effects. Columns (3),(4) report 2SLS estimates. Columns (5),(6) present the reduced-forms. As instruments we use the prior covariates as well as checks and balances and control for GDP growth in both equations. Diermeier and Vlaicu (2011) show that political environments with strong checks and balances have low levels of legislative success; this should reduce the chances that a fiscal reform such as an MTF can pass. While the OLS estimates are close to zero, the 2SLS estimates show the pattern of causality conjectured above. MTF improves budget balance, while a higher budget balance reduces the probability of MTF adoption. 3SLS estimates, not reported, are comparable. The reduced-forms in columns (5),(6) verify that at least one of the variables excluded from one equation has a nonzero population coefficient in the other equation.

Robustness. In Table 7 we report estimates where the two lags of the dependent variable are treated as endogenous. We instrument for them using $Balance^{GMM}_{i,t-2}$ and $Balance^{GMM}_{i,t-3}$. In column (1) the MTF coefficient is 2.007, very close to the 2.047 estimate in the simpler

[41] Lagged MTF status is positively correlated with future changes because $MTF^{GMM}_{i,t-2} = 0$ is likely followed by a change (a negative $MTF^{*}_{i,t}$) while $MTF^{GMM}_{i,t-2} = 1$ is likely followed by no change (zero $MTF^{*}_{i,t}$).

[42] While the second-stage underidentification LM test is analytically exact, the F statistic should be considered heuristic in a panel setting (see Bun and Winmeijer 2010, Bun and de Haan 2010, Bazzi and Clemens 2013). For this reason we do not compute p-values for this test. Moreover, Stock and Yogo (2005) critical values are limited to three endogenous regressors with at least five excluded instruments (see their Table 1).

specification of Table 5a. The implied long-run effect is approximately $2.007/(1-0.417-0.111) = 4.252$ percentage points. The lagged dependent variable coefficient of 0.417 lies in the FE-OLS bracketing range of $[0.392, 0.501]$, see Table 5a, not raising any specification issues. 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. We conclude that D-GMM is an internally consistent estimator for our dynamic panel model and, provided the instruments are valid and strong, can be relied upon to carry out statistical inference for hypothesis (H1).[43]

The remaining columns in Table 7 report alternative specifications that relax or modify our identification assumptions in (10) and (12). One way in which the assumption that past MTF is not correlated with current errors can fail is if governments rather than react to past budget problems are instead forward looking, i.e., put an MTF in place to avoid a future anticipated fiscal crisis. For example, if the government adopts MTF in anticipation of a large deficit one year ahead, then $E(MTF_{i,t-1}u_{i,t}) \neq 0$ and so $MTF^{GMM}_{i,t-2}$ is no longer a valid instrument for $MTF^{*}_{i,t}$. However, if forward looking reform occurs no earlier than a year ahead, $MTF^{GMM}_{i,t-3}$ and longer lags still remain valid instruments. In column (2) we report estimates using $MTF^{GMM}_{i,t-6}$ which is valid if forward looking behavior is no longer than four years, a typical election cycle length. The MTF coefficient is now 1.983, with p-value 0.108. Estimates based on each intermediate lag between two and six, not reported, are all statistically significant and of magnitude similar to the result in column (1). Another way in which assumption (10) may fail is if MTF affects budget balance with a lag. This possibility should be less likely at lower data frequencies, e.g., three-year averaged data. In column (3) we reestimate the model based on data averaged within three-year non-overlapping windows. The MTF coefficient is similar to the baseline in column (1) in terms of sign, size, and significance. The lagged dependent variable coefficient is smaller, as would be expected, since the lagged dependent variable now covers a longer time horizon and cyclical fluctuations are dampened through averaging.

One can also probe the validity of the MTF instrument by exploring possible violations of the moment condition (12). In Table 4 we study the within-country correlation between the instrument $MTF^{GMM}_{i,t-2}$ and potential determinants of budget balance excluded from our benchmark model. The coefficients indicate that macro factors and international organiza-

[43] Estimation results using one lag of the dependent variable or using the first-difference transform instead of the orthogonal deviations transform produce coefficients on the first lag of the dependent variable that lie outside the bracketing range, above in the first case, below in the second case.

tion interventions are associated with the instrumnent, conditional on covariates.[44] In Table 7, columns (4) and (5) we explicitly include these factors as additional covariates and find that the baseline result in column (1) is upheld.[45] In column (6) we also present S-GMM estimates, i.e., MTF in levels is additionally instrumented by its lagged differences. The MTF coefficient magnitude remains close to the D-GMM estimate in column (1). This was expected, since the instrument in the transformed model is sufficienty strong (see Table 5b) precluding the need for additional moment conditions.[46]

A more formal approach to checking instrument validity is the Hansen test for overidentifying restrictions as well as the difference-in-Hansen test for instrument subsets. While these tests are limited in that they hinge on the untestable assumption that at least one instrument is valid, they can be useful in spotting gross violations of validity. We also adopt a strategy from Bazzi and Clemens (2013) of tracking changes in the Hansen test after removing key instruments. If the original instruments were invalid, removing them would raise the p-value of the Hansen test, indicating that the validity of the remaining instruments has improved. The results of this analysis are in Table 8. In columns (1) and (3) the Hansen test does not reject instrument validity for lagged balances and MTF, respectively. Removing $Balance^{GMM}_{i,t-2}$ and $Balance^{GMM}_{i,t-3}$ in column (2) and $MTF^{GMM}_{i,t-2}$ in column (4) decrease, rather than raise, the statistical support for the removed instruments' validity. Finally in column (5) we run a specification with the minimal number of instruments 4/4 that allow difference-in-Hansen tests for both categories of instruments. In both cases, the difference-in-Hansen tests cannot reject the null that the excluded set of four instruments is valid. We use this final specification as the basis for placebo tests, i.e., we bring the adoption date forward or backward a number of years and reestimate the model of column (5) with the altered $MTF_{i,t}$ variable. We graph the resulting placebo MTF coefficients in Figure 6. The placebo

[44] International organizations such as the World Bank, the UK's Department for International Development (DFID), the Inter-American Development Bank (IDB), the Asian Development Bank (ADB), and to a lesser extent the IMF, may recommend, but cannot impose, these reforms as part of a sound public financial management strategy. Yet in countries that rely heavily on foreign aid (poorly indebted countries) these recommendations may carry some weight. See Holmes and Evans (2003), World Bank (2012).

[45] The reported results treat the additional covariates as exogenous, i.e., they are instrumented with themselves. Treating them as predetermined, i.e., instrumenting for their transformation using their second lag, does not materially affect the MTF coefficient estimate. We omit $Unemployment_{i,t-1}$ since it would cut sample size almost in half.

[46] To rule out the possibility that the results are driven by subgroups of countries with extreme characteristics we restricted 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 Online Appendix Table A3. The estimates are broadly comparable to the baseline and are available from the authors upon request.

coefficients are always smaller than the real coefficient and never statistically significant. They remain high up to three periods around the coded adoption period, and then fluctuate close to zero. This pattern may reflect the fact that in the majority of countries MTF reform implementation tends to be a gradual process, often with an experimentation period, rather than a sudden switch.[47]

MTF Phases. In Table 9 we further explore the relationship between MTF and fiscal discipline by looking at whether more advanced MTF phases have stronger effects on fiscal discipline, as the arguments presented in Section 2 would have us expect. To check this, we disaggregate the $MTF_{i,t}$ dummy into three dummies $MTFF_{i,t}$, $MTBF_{i,t}$, and $MTPF_{i,t}$, with the base category remaining yearly budgeting. The first three columns of Table 9 report OLS, FE, and D-GMM estimates for the static version of the model in equation (6), while the next three columns present the counterpart estimates for the dynamic version of the model. To improve the efficiency of the GMM estimates we increase the number of instruments for MTFs by using $MTFF^{GMM}_{i,t-2}, MTBF^{GMM}_{i,t-2}, MTPF^{GMM}_{i,t-2}$ uncollapsed. The additional instruments can be used to perform overidentification tests and we report these at the bottom of the table.

We notice two features of the estimates. First, the patterns observed in Table 5a reappear here; compare columns (2) vs. (5) and (3) vs. (6). Second, the three MTF coefficient magnitudes are monotonic, supporting hypothesis (H1). The MTF coefficients for the dynamic model of column (6) increase from 1.717 percentage points for the basic MTFF to 4.401 percentage points for the advanced MTPF. The under/weak identification statistics based on the 2SLS analogs, reported at the bottom of the table, suggest that the instruments are still significantly correlated with the endogenous regressors, however, their strength is now less. This is to be expected, since each individual MTF phase is further from an absorbing state than the aggregate $MTF_{i,t}$ dummy.

The last three columns of Table 9 present specifications where we instrument for the transformed lags of the dependent variable. We maintain the same instruments as before, namely the collapsed second and third lags of budget balance. Column (7) shows that the instrumentation has minimal effects on the estimates. The lagged dependent variable coefficient of 0.405 is inside the FE-OLS range of $[0.391, 0.498]$, supporting the chosen model specification. Column (8) estimates the same model using data averaged within three-year non-overlapping windows. As expected the MTF coefficients increase while maintaining their

[47] As the average MTF adoption year is 2002.81 and the sample ends in 2008, placebo coefficents at +6 years and beyond are identified off of a small number of reforms, making their standard errors very large.

monotonic pattern. In column (9), for comparison, we also present S-GMM estimates.

Alternative Fiscal Performance Measures. In Table 10 we present additional evidence of the impacts of MTF phases by looking at alternative measures of fiscal performance: primary budget balance, health spending volatility, and health technical efficiency. The primary balance is that part of the total balance that excludes interest payments on government debt, and thus reflects the ongoing aggregate state of public finances. The other two measures are sectoral measures of fiscal performance, discussed above in Section 3.

The primary balance regressions are in the first four columns of Table 10. As we noted earlier, the primary balance data is less comprehensive than the total balance data, reducing the size of the panel from 161 to 140. The MTF coefficients estimated by GMM are larger than OLS and FE and display the same monotonic pattern as in the total balance regressions: more advanced phases have larger coefficients. The magnitudes are larger than for total balance: $(1.66, 2.05, 4.30)$ now become $(2.91, 3.08, 5.99)$. This seems plausible since the primary balance contains relatively more discretionary spending, and so should be more responsive to budget institutions. The underidentification and weak identification tests based on analogous 2SLS versions show a reduction in the overall strength of the instruments. Other diagnostic tests appear satisfactory.

The next four columns focus on health spending volatility. The panel length 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. Based on hypothesis (H2) above, we expect the two advanced MTF phases (MTBF and MTPF) to reduce spending volatility, with the top-phase MTPF reducing it more. The pattern of the MTF coefficients in columns (7) and (8) seems consistent with this hypothesis. However, the underidentification test in the analogous 2SLS model fails to reject that the instruments are uncorrelated with the endogenous regressors. This is partly a consequence of a significantly shorter panel but nevertheless raises small sample bias issues.[48]

The last four columns present estimates of MTF effects on technical efficiency, measured as efficiency scores from a stochastic frontier model of health delivery; see equations (1), (2), and Table 3. According to hypothesis (H3) only the most advanced MTF phase, namely MTPF, should materially affect technical efficiency because it is the only one that sets within-sector performance targets. As expected this indicator of fiscal performance is much more persistent than the previous ones. The FE-OLS bracketing range, based on

[48]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.

columns (9) and (10), shifts up to $[0.84, 0.99]$, indicating strong persistence. National health delivery is a complex system that may take decades to fully internalize the benefits of a given reform.[49] While the instrument diagnostics in this smaller sample suggest that they are weak predictors of the endogenous variables, the D-GMM point estimates do show the hypothesized pattern. The MTPF coefficient is sizable at about one percentage point, while the lower MTF phases have coefficients very close to zero. If, as is usually the case in the presence of weak instruments, D-GMM is downward biased (Bun and Windmeijer 2010) then these coefficients may still serve as a lower bound on the actual effect.

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 MTFs are a worthwhile budget institution, a systematic empirical analysis of their impacts has been lacking due to insufficient data on MTF adoption around the world. This paper is the first to empirically investigate the MTFs' impacts on fiscal performance in a large global sample of countries. We collected a panel dataset of 181 countries over the period 1990-2008, the most comprehensive dataset on global MTF adoption to date. In order to disentangle the effects of the MTF and its different phases (MTFF, MTBF, and MTPF) from other factors and to correct for potential reverse causality we use a dynamic panel data approach.

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

Our analysis may be limited by the fact that an MTF could be in place only in law (de jure) and not in practice (de facto). If this phenomenon were widespread it would induce an attenuation bias and our estimates could still be regarded as a lower bound on

[49] Statistically, this implies that lagged levels of the dependent variable are weak instruments for future changes.

the actual effect. Being in effect commitment mechanisms, transparency and enforcement are critical components of MTFs. Studying which features of the broader civic, juridical, and political environment enhance MTF effectiveness may lead to a better understanding of these budget institutions. Also, our analysis of MTF impacts on sectoral performance is limited by the shorter panel available and the corresponding weaker within-country variation in the sectoral performance measures. Whether sharper effects can be detected in other categories of government spending is an interesting question for future research. To create more effective measures of sectoral fiscal performance and associated outcomes one needs to identify disaggregated categories of spending that are comparable across a wide 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.

Auerbach, A.J. (2006) "Budget Windows, Sunsets, and Fiscal Control," *Journal of Public Economics* 90: 87-100.

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.

Bazzi, S. and M.A. Clemens (2013) "Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth," *American Economic Journal: Macroeconomics* 5(2): 152-86.

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 (2001) "Uganda: The Budget and Medium-Term Expenditure Framework Set in a Wider Context," working paper, World Bank.

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.

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

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.

Bun, M.J.G. and M. de Haan (2010) "Weak Instruments and the First Stage F-statistic in IV Models with a Nonscalar Error Covariance Structure," working paper.

Cameron A.C. and P. Trivedi (2005) *Microeconometrics*, Cambridge U. Press.

Castro, I. and W. Dorotinsky (2008) "Medium-Term Expenditure Frameworks: Demystifying and Unbundling The Concept," working paper.

Diermeier, D. and R. Vlaicu (2011) "Executive Control and Legislative Success," *Review of Economic Studies* 78(3): 846-871.

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, IDB.

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. (2004) "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.

Hayakawa, K. (2009a) "First Difference or Forward Orthogonal Deviation - Which Transformation Should be Used in Dynamic Panel Data Models? A Simulation Study," *Economics Bulletin* 29(3): 2008-2017.

Hayakawa, K. (2009b) "A Simple Efficient Instrumental Variable Estimator in Panel AR(p) Models When Both N and T are Large," *Econometric Theory* 25(3): 873-890.

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," ODI.

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

IMF (2007) *Fiscal Policy Response to Scaled-Up Aid: Macro-Fiscal and Expenditure Policy Challenges*, Washington, DC: IMF.

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*, IMF.

Lane, C. and A. Glassman (2008) "Smooth and Predictable Aid for Health: A Role for Innovative Financing?" Brookings Institution working paper.

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

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.

Stock, J.H. and J.H. Wright (2000) "GMM with Weak Identification," *Econometrica* 68(5): 1055-1096.

Stock, J.H. and M. Yogo (2005) "Testing for Weak Instruments in Linear IV Regression," in D.W.K. Andrews and J.H. Stock (eds.) *Identification and Inference for Econometric*

Models: Essays in Honor of T. Rothenberg, Cambridge, UK: Cambridge University Press.

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.

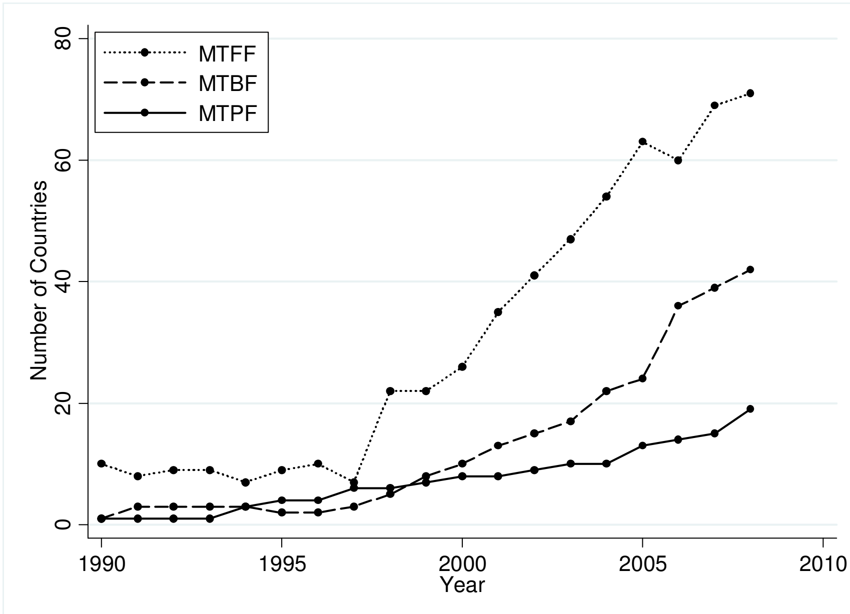
Wooldridge, J. (2002) *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

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*, Washington, DC: World Bank Publications.

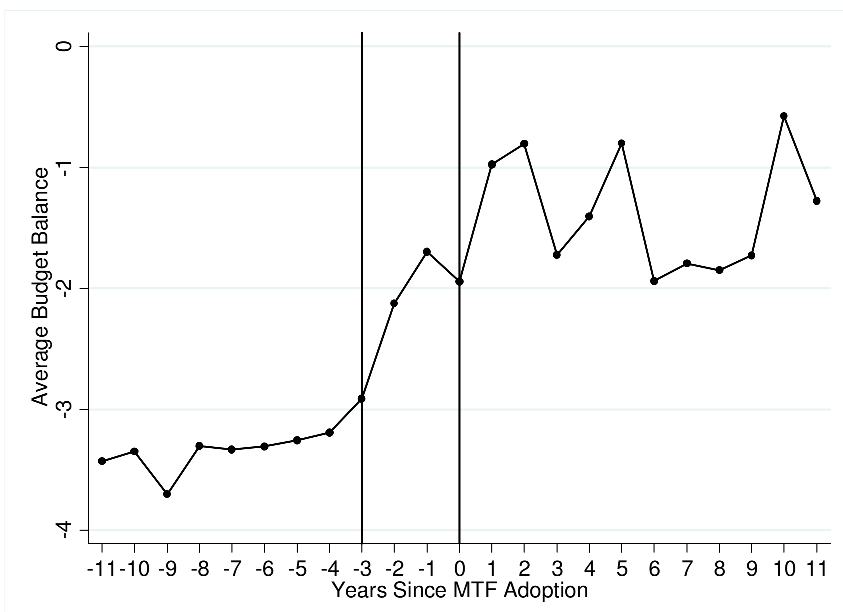
Ylaoutinen, S. (2004) "Fiscal Frameworks in Central and Eastern European Countries," Finnish Ministry of Finance discussion paper 72.

Figure 1: MTF Growth Worldwide, 1990-2008



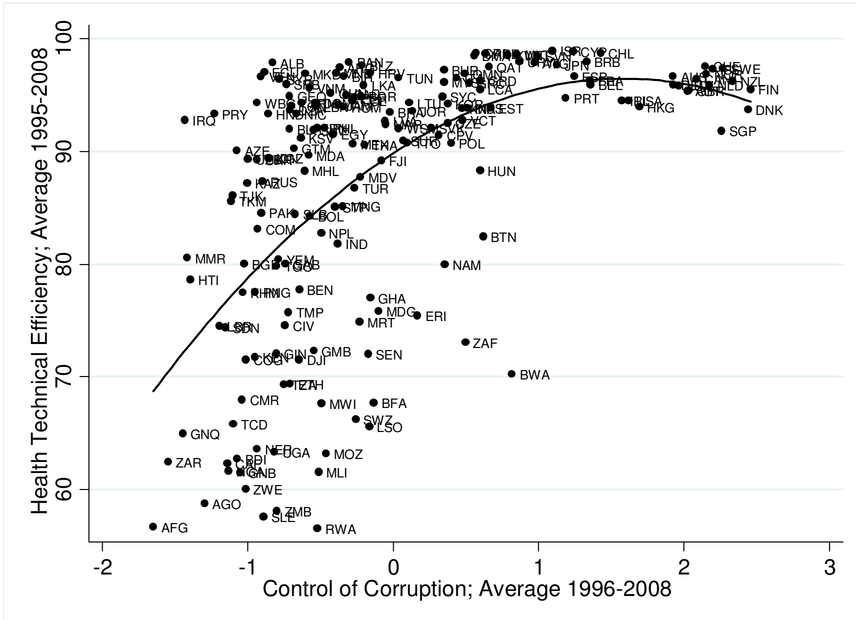
Notes: Authors' calculations based on a sample of 181 countries for 1990-2008. More solid lines indicate more advanced MTF phases.

Figure 2: Budget Balance Before and After MTF Adoption



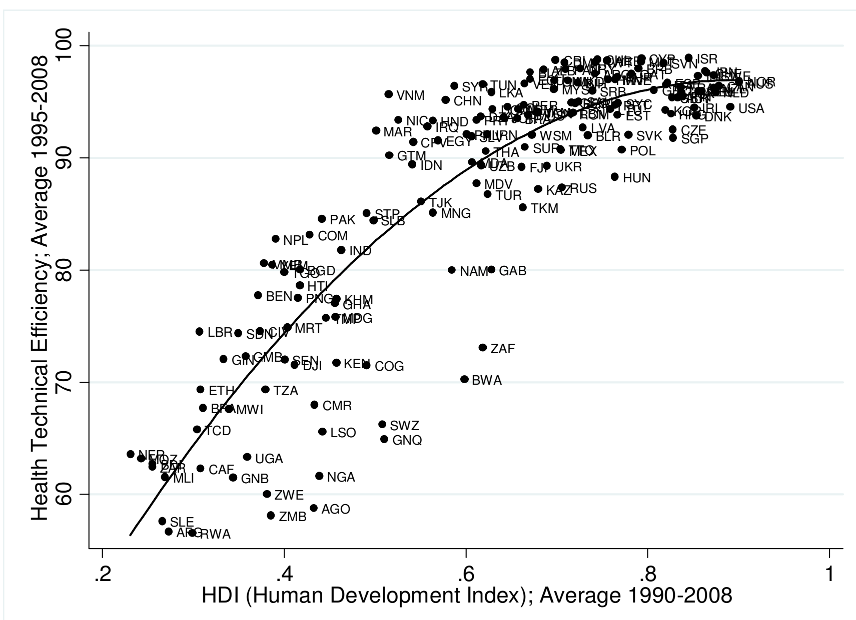
Notes: Authors' calculations based on the 120 countries that adopted an MTF during the sample period 1990-2008.

Figure 3: Health Technical Efficiency and Control of Corruption



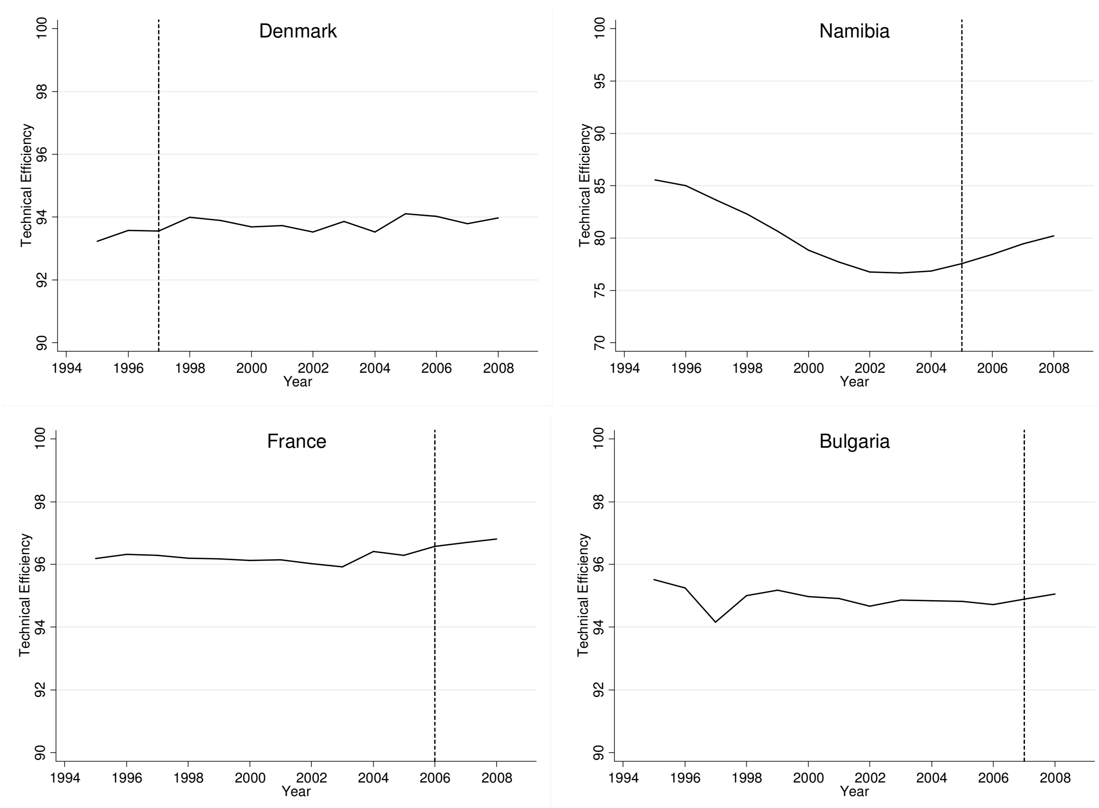
Notes: Authors' calculations based on a sample of 181 countries for 1990-2008. Country codes adjacent to each scatter point. Scatter plot quadratic fit shown.

Figure 4: Health Technical Efficiency and Development



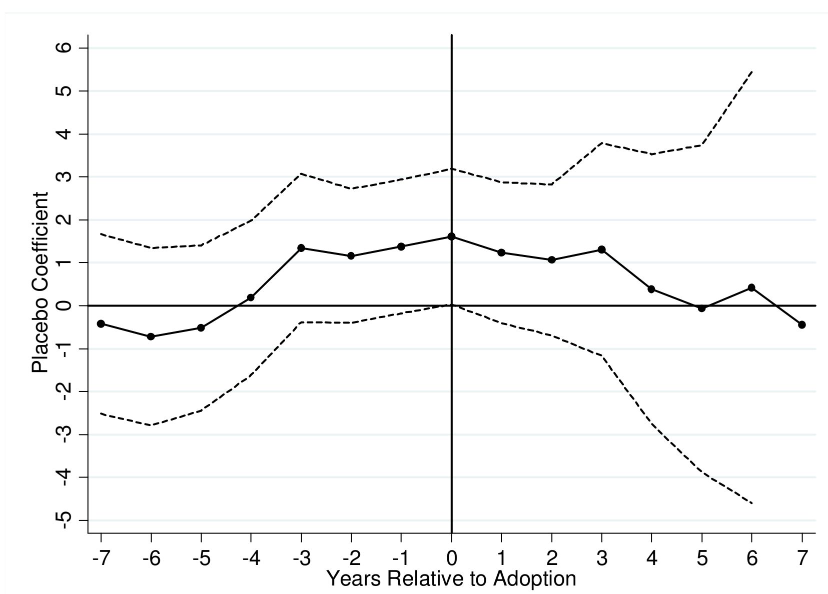
Notes: Authors' calculations based on a sample of 181 countries for 1990-2008. Country codes adjacent to each scatter point. Scatter plot quadratic fit shown.

Figure 5: Health Technical Efficiency, Selected Countries



Notes: The figure depicts the time series for the technical efficiency measure estimated in Table 3 for four countries that have adopted an MTPF during the sample period. In each graph the dotted reference line indicates the year of MTPF reform.

Figure 6: Placebo Coefficients



Notes: The graph plots D-GMM estimates of the effect of MTF on Budget Balance. The reference specification is in Table 8 Column (5). Dotted lines are 95% confidence bounds.

Table 1: Summary Statistics

Variable	Obs.	Mean	Std. Deviation Across	Std. Deviation Within	Min	Max
Total Budget Balance	2,991	–2.24	11.87	9.59	–151.33	384.15
Primary Budget Balance	2,544	0.19	7.23	6.48	–147.51	121.61
Health Spending Volatility	2,282	6.69	7.62	6.35	0.002	83.19
Health Technical Efficiency	2,359	86.48	12.11	1.67	39.21	99.00
MTF	3,378	0.29	0.45	0.35	0	1
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
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
Human Development Index	2,659	0.61	0.18	0.03	0.193	0.942
Control of Corruption	2,313	–0.06	1.00	0.21	–2.06	2.59
Openness	3,069	85.28	48.73	16.89	0.31	456.65
Conflict	3,439	0.05	0.22	0.17	0	1
Balanced Budget Rule	3,439	0.21	0.41	0.27	0	1
Debt Rule	3,439	0.18	0.39	0.26	0	1

Notes: Statistics are based on a sample of 181 countries for 1990-2008. See Section 3 in the main text and the Online Appendix for data sources, units of measurement, and variable construction details. The differences in number of observations across variables reflect data availability in the different data sources.

Table 2: MTF 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 MTF	11	132	125	0	–4

Notes: The summary statistics are based on a sample of 181 countries for 1990-2008. 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: Stochastic Frontier Model for Public Health Delivery

Dependent Var.:	$\log(Life\_Exp_{i,t})$ (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 summarized in Table 1. 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. See Section 3 in the main text. ***, **, * indicate statistical significance at 1%, 5%, 10%.

Table 4: Potential Exclusion Restriction Threats

	$x_{i,t-1}$ on $MTF_{i,t-1}$				$x_{i,t-1}$ on $MTF_{i,t-2}$				$x^{*}_{i,t-1}$ on $MTF^{GMM}_{i,t-2}$, covariates			
x :	Coef.	S.E.	Cntrs.	Obs.	Coef.	S.E.	Cntrs.	Obs.	Coef.	S.E.	Cntrs.	Obs.
GDP Growth	0.671	0.491	179	3,250	0.311	0.479	179	3,087	0.784***	0.185	175	2,768
Debt-to-GDP	3.946	6.933	130	1,987	2.173	6.056	130	1,927	–0.395	1.989	127	1,644
Unemployment	0.331	0.452	163	1,737	–0.001	0.492	159	1,673	0.429**	0.212	137	1,411
Inflation	7.252	20.61	177	3,197	28.21	20.28	177	3,051	–18.78**	7.721	173	2,716
Interest Ext. Debt	0.111	0.255	126	2,231	–0.387	0.398	126	2,132	0.379	0.301	124	1,928
Openness	5.475**	2.234	175	3,069	5.816***	2.169	175	2,907	–	–	–	–
Foreign Direct Inv.	0.527	0.458	180	3,170	0.516	0.499	180	3,029	–0.217	0.320	173	2,663
Conflict	0.010	0.015	181	3,378	0.024*	0.014	181	3,197	–	–	–	–
Election Year	0.009	0.013	167	3,013	–0.003	0.015	167	2,960	0.007	0.010	165	2,645
Checks&Balances	–0.072	0.127	167	3,013	–0.078	0.144	167	2,868	0.012	0.067	165	2,552
Majority	0.985	2.108	165	2,388	1.249	2.046	163	2,244	–0.167	0.964	160	1,990
Gov. Fragm.	–0.010	0.024	164	2,803	–0.006	0.025	164	2,678	–0.012	0.012	160	2,378
Bal. Budget Rule	0.085**	0.037	181	3,378	0.094**	0.039	181	3,197	–	–	–	–
Debt Rule	0.064*	0.038	181	3,378	0.072*	0.039	181	3,197	–	–	–	–
Expenditure Rule	0.040	0.030	181	3,378	0.054	0.033	181	3,197	–0.038	0.025	175	2,797
Revenue Rule	–0.009	0.016	181	3,378	–0.010	0.018	181	3,197	–0.029	0.020	175	2,797
Indebted Poor	0.027	0.022	181	3,378	–0.005	0.024	181	3,197	0.049***	0.013	175	2,797
Aid-to-GDP	0.198	0.387	178	3,229	–0.300	0.378	178	3,076	–0.160	0.149	173	2,741
IMF Program	–0.006	0.035	181	3,378	–0.029	0.035	181	3,197	–0.027	0.017	175	2,797
IMF Missions	0.286	0.188	181	3,378	0.151	0.243	181	3,197	0.277*	0.160	175	2,797

Notes: Estimates based on a sample of 181 countries for 1990-2008. Number of observations reflects data availability. All regressions include year and country fixed effects. Standard errors clustered at the country level. Variables with asterisk (*) indicate transformation using orthogonal deviations; see Section 3 in the main text. In panel 3 the covariates are $Openness_{i,t-1}$, $Conflict_{i,t-1}$, $BalBudget_{i,t-1}$, and $DebtRule_{i,t-1}$. ***, **, * next to a coefficient indicate statistical significance at 1%, 5%, 10%.

Table 5a: MTF and Fiscal Discipline: Static vs. Dynamic Model

Dependent Var.:	$Balance_{i,t}$ (Total Budget Balance, as % of GDP)					
Model:	OLS (1)	FE (2)	D-GMM (3)	OLS (4)	FE (5)	D-GMM (6)
$Balance_{i,t\text{-}1}$	–	–	–	0.501***	0.392***	0.391***
				(0.040)	(0.032)	(0.031)
$Balance_{i,t\text{-}2}$	–	–	–	0.165***	0.106***	0.108***
				(0.041)	(0.030)	(0.030)
$MTF_{i,t}$	0.945	–0.261	2.569**	0.215	0.058	2.047**
	(0.641)	(0.597)	(1.275)	(0.249)	(0.314)	(0.853)
$Openness_{i,t\text{-}1}$	0.006	–0.014	–0.020	0.002	0.006	0.000
	(0.004)	(0.017)	(0.018)	(0.002)	(0.008)	(0.008)
$Conflict_{i,t\text{-}1}$	–3.246***	–2.805**	–2.803**	–0.605	–1.024*	–1.021*
	(1.250)	(1.214)	(1.196)	(0.458)	(0.548)	(0.541)
$BalBudget_{i,t\text{-}1}$	1.932**	0.355	0.277	0.572*	–0.066	–0.121
	(0.779)	(0.851)	(0.870)	(0.314)	(0.518)	(0.529)
$DebtRule_{i,t\text{-}1}$	–2.287***	–0.977	–1.130	–0.710**	–0.124	–0.215
	(0.870)	(0.999)	(1.002)	(0.347)	(0.541)	(0.529)
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
Instruments	–	–	–/1	–	–	–/1
AR(1) Test p-val.	–	–	0.031	–	–	0.000
AR(2) Test p-val.	–	–	0.476	–	–	0.618
Sample Period	1990–2008	1990–2008	1990–2008	1990–2008	1990–2008	1990–2008
Countries	163	163	163	162	162	161
Observations	2,724	2,724	2,561	2,546	2,546	2,384

Notes: The unit of observation is a country-year from the sample summarized in Table 1. Columns (1),(2),(4),(5) report OLS standard errors clustered at the country level. Columns (3),(6) report GMM two-step estimates and standard errors with the Windmeijer correction. D-GMM estimates based on the orthogonal deviations transform. In columns (3),(6) the instrument for the transformed $MTF_{i,t}$ is the collapsed second lag of $MTF_{i,t}$. In Column (6) the lagged dependent variables are treated as exogenous. ***, **, * indicate statistical significance at 1%, 5%, 10%.

Table 5b: Instrument Strength and Reduced Form

Dependent Var.:	$MTF^{*}_{i,t}$	$Balance^{*}_{i,t}$	$Balance^{*}_{i,t}$	$MTF^{*}_{i,t}$	$Balance^{*}_{i,t}$	$Balance^{*}_{i,t}$
Model:	OLS	2SLS	OLS	OLS	2SLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
$Balance^{*}_{i,t-1}$	–	–	–	–0.000	0.391***	0.390***
				(0.001)	(0.031)	(0.031)
$Balance^{*}_{i,t-2}$	–	–	–	–0.001	0.108***	0.106***
				(0.001)	(0.030)	(0.029)
$MTF^{*}_{i,t}$	–	2.304*	–	–	1.987**	–
		(1.251)			(0.836)	
$Openness^{*}_{i,t-1}$	0.002***	–0.021	–0.017	0.002***	0.000	0.005
	(0.001)	(0.018)	(0.017)	(0.001)	(0.008)	(0.007)
$Conflict^{*}_{i,t-1}$	0.024	–2.793**	–2.738**	0.022	–1.023**	–0.980*
	(0.037)	(1.187)	(1.164)	(0.038)	(0.541)	(0.527)
$BalBudget^{*}_{i,t-1}$	0.040	0.217	0.309	0.039	–0.135	–0.057
	(0.053)	(0.868)	(0.835)	(0.054)	(0.534)	(0.509)
$DebtRule^{*}_{i,t-1}$	0.055	–1.067	–0.941	0.039	–0.202	–0.126
	(0.063)	(0.978)	(0.963)	(0.065)	(0.529)	(0.522)
$MTF^{GMM}_{i,t-2}$	0.271***	–	0.624*	0.270***	–	0.536**
	(0.021)		(0.334)	(0.022)		(0.217)
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
Underid. p-val.†	0.000	0.000	–	0.000	0.000	–
Weak Id. F ‡	159.70	159.70	–	152.96	152.96	–
Sample Period	1990–2008	1990–2008	1990–2008	1990–2008	1990–2008	1990–2008
Countries	163	163	163	161	161	161
Observations	2,561	2,561	2,561	2,384	2,384	2,384

Notes: The unit of observation is a country-year from the sample summarized in Table 1. Columns (2),(5) present the 2SLS counterparts of the D-GMM estimates in columns (3),(6) of Table 5a. Columns (1),(4) are the first stages of columns (2),(5), respectively. Columns (3),(6) are the reduced-forms of columns (2),(5), respectively. Variables with asterisk (*) indicate transformation using first orthogonal deviations; see Section 3 in the main text. Standard errors clustered at the country level in parentheses. The underidentification tests' null is that the excluded instruments are uncorrelated with the endogenous regressor(s). †Corresponding to Angrist-Pischke Chi-sq. test for first-stage equation. ‡Kleibergen-Paap F test for the second-stage equation. ***, **, * next to a number indicate statistical significance at 1%, 5%, 10%.

Table 6: Reverse Causality: Simultaneous Equations Model

Dependent Var.:	$Balance_{i,t}$	$MTF_{i,t}$	$Balance_{i,t}$	$MTF_{i,t}$	$Balance_{i,t}$	$MTF_{i,t}$
Model:	OLS (1)	OLS (2)	2SLS (3)	2SLS (4)	OLS (5)	OLS (6)
$Balance_{i,t}$	–	–0.002* (0.001)	–	–0.048** (0.020)	–	–
$MTF_{i,t}$	–0.434 (0.302)	–	4.942* (2.851)	–	–	–
$Openness_{i,t}$	–0.007 (0.006)	–	–0.020** (0.009)	–	–0.009 (0.006)	0.002*** (0.000)
$Conflict_{i,t}$	–1.803*** (0.494)	–	–1.999*** (0.535)	–	–1.839*** (0.494)	0.032 (0.033)
$BalBudget_{i,t}$	–	0.102*** (0.031)	–	0.147*** (0.043)	1.027** (0.464)	0.086*** (0.031)
$DebtRule_{i,t}$	–	0.034 (0.032)	–	0.008 (0.041)	–0.573 (0.480)	0.047 (0.032)
$Checks\&Balances_{i,t}$	–	–0.012** (0.005)	–	–0.016** (0.007)	–0.117 (0.082)	–0.011** (0.005)
$Growth_{i,t}$	0.142*** (0.019)	0.003** (0.001)	0.130*** (0.021)	0.010*** (0.003)	0.144*** (0.019)	0.002** (0.001)
Country Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.450	0.663	0.433	0.512	0.501	0.667
Sample Period	1990–2008	1990–2008	1990–2008	1990–2008	1990–2008	1990–2008
Countries	154	154	154	154	154	154
Observations	2,553	2,553	2,553	2,553	2,553	2,553

Notes: The unit of observation is a country-year from the sample summarized in Table 1. All variables are in levels and contemporaneous to each other. Columns (1),(2) present OLS estimates of the two-equations system. Columns (3),(4) present 2SLS estimates of the two-equations system. Variables other than the dependent variables $Balance_{i,t}$ and $MTF_{i,t}$ are treated as exogenous in their respective equations. Columns (5),(6) are the reduced-forms of columns (3),(4), respectively. System standard errors in parentheses. ***, **, * indicate statistical significance at 1%, 5%, 10%.

Table 7: MTF and Fiscal Discipline: Alternative Specifications

Dependent Var.:	$Balance_{i,t}$ (Total Budget Balance, as % of GDP)					
Model:	D-GMM (1)	D-GMM (2)	D-GMM (3)	D-GMM (4)	D-GMM (5)	S-GMM (6)
$Balance_{i,t\text{-}1}$	0.417***	0.417***	0.382***	0.463***	0.418***	0.438***
	(0.034)	(0.034)	(0.085)	(0.056)	(0.034)	(0.039)
$Balance_{i,t\text{-}2}$	0.111***	0.111***	0.011	0.125***	0.111**	0.120***
	(0.036)	(0.035)	(0.064)	(0.044)	(0.036)	(0.042)
$MTF_{i,t}$	2.007**	1.983	2.009*	1.936**	2.096**	1.841**
	(0.833)	(1.235)	(1.206)	(0.815)	(0.845)	(0.727)
Specification:	Instr. for $y^*_{i,t-1}, y^*_{i,t-2}$	Use Instr. $MTF^{GMM}_{i,t-6}$	Three-Yr. Windows	Macro Factors	Intl. Organiz.	Add Level Equation
Year Effects	Yes	Yes	Yes[§]	Yes	Yes	Yes
Instruments	2/1	2/1	2/1	2/1	2/1	3/2
Underid. p-val.[†]	0.000	0.000	–	–	–	–
Weak Id. F [‡]	51.411	41.706	–	–	–	–
AR(1) p-val.	0.000	0.000	0.000	0.000	0.000	0.000
AR(2) p-val.	0.730	0.735	0.615	0.594	0.786	0.725
Sample Period	1990–2008	1990–2008	W1-W7	1990–2008	1990–2008	1990–2008
Countries	161	161	158	161	161	162
Observations	2,384	2,384	591	2,363	2,384	2,546

Notes: The unit of observation is a country-year from the sample summarized in Table 1, except in Column (3) where data is averaged within three-year non-overlapping windows. Each numbered column reports the estimates of the indicated change from the baseline specification in Table 5a Column (6). See Section 4 in the main text. All models include $Openness_{i,t\text{-}1}$, $Conflict_{i,t\text{-}1}$, $BalBudget_{i,t\text{-}1}$, and $DebtRule_{i,t\text{-}1}$ as covariates. Column (4) adds $Growth_{i,t\text{-}1}$, $Inflation_{i,t\text{-}1}$ as covariates. Column (5) adds $IndebtedPoor_{i,t\text{-}1}$, $IMF\_Missions_{i,t\text{-}1}$ as covariates. The instruments for the two transformed lags of the dependent variable are its collapsed second and third lag. The instrument for the transformed $MTF_{i,t}$ is its collapsed second lag. [§]Window fixed effects replace year effects. The first window is reduced to the single year 1990. [†]Corresponding to Angrist-Pischke Chi-sq. test for first-stage equation of the 2SLS analog of the model. [‡]Kleibergen-Paap F test for the second-stage equation of the 2SLS analog of the model. ***, **, * indicate statistical significance at 1%, 5%, 10%.

Table 8: Instrument Validity: Overidentifying Restrictions Tests

Dependent Var.:	$Balance_{i,t}$ (Total Budget Balance, as % of GDP)				
Model:	D-GMM (1)	D-GMM (2)	D-GMM (3)	D-GMM (4)	D-GMM (5)
$Balance_{i,t-1}$	0.459***	0.576	0.419***	0.418***	0.468***
	(0.048)	(0.485)	(0.034)	(0.034)	(0.047)
$Balance_{i,t-2}$	0.076**	0.040	0.110***	0.110**	0.075*
	(0.038)	(0.294)	(0.036)	(0.036)	(0.040)
$MTF_{i,t}$	1.766**	1.781**	1.956**	1.902**	1.608**
	(0.837)	(0.861)	(0.801)	(0.885)	(0.806)
Year Effects	Yes	Yes	Yes	Yes	Yes
Instruments	5/1	(5–2)/1	2/3	2/(3–1)	4/4
Underid. p-val.†	0.000	–	0.000	–	0.000
Weak Id. $F^{\ddagger}$	27.905	–	32.745	–	23.275
AR(1) p-val.	0.000	0.243	0.000	0.000	0.000
AR(2) p-val.	0.641	0.787	0.748	0.747	0.620
Hansen p-val.	0.148	0.023	0.933	0.729	0.267
Diff-in-Hansen p-val.	–/–	–/–	–/–	–/–	0.216/0.229
Sample Period	1990–2008	1990–2008	1990–2008	1990–2008	1990–2008
Countries	161	161	161	161	161
Observations	2,384	2,384	2,384	2,384	2,384

Notes: The unit of observation is a country-year from the sample summarized in Table 1. The instruments for the lags of the dependent and for MTF start at lag two. In column (2) lags two and three are removed. In column (4) lag two is removed. The null hypothesis for the Hansen test is that the full set of instruments in valid. The null hypothesis for the difference-in-Hansen test is that the excluded set of instruments is valid; the first test excludes the four balance instruments, the second test excludes the four MTF instruments. † and ‡ See Table 7 notes. ***, **, * indicate statistical significance at 1%, 5%, 10%.

Table 9: MTF Phases and Fiscal Discipline

Dep. Var.:	$Balance_{i,t}$ (Total Budget Balance, as % of GDP)								
Model:	OLS (1)	FE (2)	D-GMM (3)	OLS (4)	FE (5)	D-GMM (6)	D-GMM (7)	D-GMM (8)	S-GMM (9)
$Balance_{i,t\text{-}1}$	–	–	–	0.498***	0.391***	0.405***	0.404***	0.355***	0.452***
				(0.039)	(0.032)	(0.025)	(0.024)	(0.088)	(0.040)
$Balance_{i,t\text{-}2}$	–	–	–	0.163***	0.106***	0.106***	0.109***	0.025	0.129***
				(0.041)	(0.030)	(0.027)	(0.030)	(0.065)	(0.043)
$MTFF_{i,t}$	0.864	–0.155	2.791**	0.207	0.079	1.717*	1.664**	2.755***	0.826*
	(0.647)	(0.602)	(1.401)	(0.252)	(0.329)	(0.937)	(0.851)	(0.926)	(0.432)
$MTBF_{i,t}$	–0.048	–0.229	4.204**	–0.199	0.141	2.028*	2.048*	2.891**	0.963
	(0.753)	(0.815)	(1.646)	(0.307)	(0.416)	(1.217)	(1.146)	(1.382)	(1.399)
$MTPF_{i,t}$	3.372***	1.822	8.172***	1.138***	1.122*	4.401**	4.303**	6.107**	2.196***
	(0.863)	(1.324)	(3.048)	(0.387)	(0.659)	(2.097)	(1.906)	(2.397)	(0.802)
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes[§]	Yes
Instruments	–/–	–/–	–/51	–/–	–/–	–/48	2/48	2/23	3/51
Underid. p-val.[†]	–	–	0.001	–	–	0.003	0.004	–	–
Weak Id. F [‡]	–	–	10.560	–	–	11.907	11.203	–	–
AR(1) p-val.	–	–	0.037	–	–	0.000	0.000	0.002	0.000
AR(2) p-val.	–	–	0.428	–	–	0.801	0.754	0.489	0.740
Hansen p-val.	–	–	0.940	–	–	0.697	0.722	0.259	0.615
Sample Period	1990–2008	1990–2008	1990–2008	1990–2008	1990–2008	1990–2008	1990–2008	W1-W7	1990–2008
Countries/Obs.	163/2,724	163/2,724	163/2,561	162/2,546	162/2,546	161/2,384	161/2,384	158/591	162/2,546

Notes: The unit of observation is a country-year from the sample summarized in Table 1. All models include $Openness_{i,t\text{-}1}$, $Conflict_{i,t\text{-}1}$, $BalBudget_{i,t\text{-}1}$, and $DebtRule_{i,t\text{-}1}$ as covariates. OLS and FE columns report standard errors clustered at the country level. D-GMM columns report GMM two-step estimates and standard errors with the Windmeijer correction. The instruments for the transformed lags of the dependent variable are the collapsed second and third lags of $Balance_{i,t}$. The instruments for the transformed $MTFF_{i,t}$, $MTBF_{i,t}$, $MTPF_{i,t}$ are their uncollapsed second lags; column (8) also uses the third lag. [§]Window fixed effects replace year effects. The first window is reduced to the single year 1990. [†] and [‡] See Table 7 notes. ***, **, * indicate statistical significance at 1%, 5%, 10%.

Table 10: Alternative Fiscal Performance Measures

Dep. Var. ($y_{i,t}$):	*Primary_ Balance*$_{i,t}$				*Health_ Spending_ Volatility*$_{i,t}$				*Health_ Technical_ Efficiency*$_{i,t}$			
Model:	OLS (1)	FE (2)	D-GMM (3)	D-GMM (4)	OLS (5)	FE (6)	D-GMM (7)	D-GMM (8)	OLS (9)	FE (10)	D-GMM (11)	D-GMM (12)
$y_{i,t\text{-}1}$	0.48***	0.37***	0.38***	0.40***	0.26***	–0.02	–0.04	0.07	0.99***	0.84***	0.85***	0.95***
	(0.04)	(0.04)	(0.03)	(0.04)	(0.06)	(0.05)	(0.05)	(0.08)	(0.00)	(0.02)	(0.01)	(0.09)
$y_{i,t\text{-}2}$	0.15***	0.09***	0.10***	0.12***	–	–	–	–	–	–	–	–
	(0.05)	(0.03)	(0.03)	(0.04)								
$MTFF_{i,t}$	–0.17	–0.29	3.05***	2.91***	0.48	0.40	–0.78	–1.77	–0.06	–0.09*	0.02	0.01
	(0.29)	(1.96)	(1.09)	(1.10)	(0.53)	(0.62)	(2.16)	(1.96)	(0.05)	(0.05)	(0.13)	(0.21)
$MTBF_{i,t}$	–0.52	–0.41	3.17**	3.08*	0.93	0.90	–7.57**	–6.09*	–0.04	–0.03	–0.04	0.00
	(0.34)	(0.49)	(1.61)	(1.64)	(0.59)	(1.14)	(3.42)	(3.49)	(0.08)	(0.08)	(0.09)	(0.36)
$MTPF_{i,t}$	0.44	0.28	6.27**	5.99**	–0.71	–2.79	–12.7*	–14.8**	–0.08	–0.03	1.11	0.85
	(0.41)	(0.77)	(2.48)	(2.43)	(1.05)	(2.01)	(7.29)	(7.35)	(0.09)	(0.21)	(0.84)	(1.33)
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instruments	–/–	–/–	–/48	2/48	–/–	–/–	–/33	2/33	–/–	–/–	–/33	2/33
Underid. p-val.†	–	–	0.082	0.082	–	–	0.248	0.290	–	–	0.728	0.310
Weak Id. F ‡	–	–	7.91	7.11	–	–	1.62	1.44	–	–	1.29	2.75
AR(1) p-val.	–	–	0.000	0.000	–	–	0.000	0.000	–	–	0.042	0.041
AR(2) p-val.	–	–	0.506	0.397	–	–	0.526	0.799	–	–	0.362	0.316
Hansen J p-val.	–	–	0.945	0.934	–	–	0.602	0.560	–	–	0.306	0.450
Sample Period	1990-2008		1990-2008		1996-2008		1996-2008		1996-2008		1996-2008	
Countries/Obs.	140/2,176		140/2,036		172/1,942		171/1,770		169/1,886		169/1,717	

Notes: The unit of observation is a country-year from the sample summarized in Table 1. All models include $Openness_{i,t\text{-}1}$, $Conflict_{i,t\text{-}1}$, $BalBudget_{i,t\text{-}1}$, and $DebtRule_{i,t\text{-}1}$ as covariates. OLS and FE columns report standard errors clustered at the country level. D-GMM columns report two-step estimates and standard errors with the Windmeijer correction. The instruments for the transformed lags of the dependent variable are its collapsed second and third lags. The instruments for the transformed $MTFF_{i,t}$, $MTBF_{i,t}$, $MTPF_{i,t}$ are their uncollapsed second lag. † and ‡ See Table 7 notes. ***, **, * indicate statistical significance at 1%, 5%, 10%.