WPS6657

Policy Research Working Paper 6657

The Quality of Budget Execution and Its Correlates

Douglas Addison

The World Bank
East Asia and the Pacific Region
Poverty Reduction and Economic Management Unit
October 2013

Policy Research Working Paper 6657

Abstract

What determines the quality of budget execution around the world, measured in terms of a government's ability to accurately hit its own revenue and expenditure targets? The answers could be relevant to the topics of macroeconomic stability, national development, public service delivery, and political reputation. This paper takes a step toward finding answers through the exploration of a new database of budgets and budget outcomes and potential cross-country correlates of budget execution in levels and in composition. Few countries within the data sample execute their budgets well, in levels or in composition. Expenditure deviations are positively but rather loosely correlated with revenue deviations. Within this broad tendency, there is considerable variation in behavior not only across countries, but also across time within countries. In explaining the cross-country variations, the data confirm traditional drivers for common pool behavior while also supporting constructive roles for political institutions and the technical capacity for public financial management. This is good news for reform minded governments.

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

The Quality of Budget Execution and Its Correlates

Douglas Addison
DAddison@worldbank.org

Key words: agency problem, appropriation, budget, collective-action, common-pool, democracy, economic development, expenditure, fiscal policy, political economy, political institutions, public economics, public financial management, public service delivery, revenue

JEL Classification: D720, E620, H3, H410, H610, H680

Sector Boards: Economic Policy (EP) and Public Sector Governance (PS)

The author is grateful to Tony Bennett, Marek Hanusch, Tobias Haque, Philip Keefer, Stephen Knack, Brandon Lundberg, William McCourt, Gary Reid, Frans Ronsholt and Clay Wescott for their thoughtful comments, suggestions, and guidance.

I. Introduction

There is broad consensus that the credibility of government budgets is important. The ability to set and achieve viable fiscal targets can be important to the attainment of various macroeconomic goals. National objectives for development and public service delivery are easier to achieve when funds are disbursed as allocated. Political parties and political careers may be indirectly dependent upon the quality of budget execution to the extent key goals are achieved, corruption and waste are avoided, inflation and unemployment are tolerable, and debt does not reach unsustainable levels.

This consensus has found expression in the way that government performance is assessed by aid agencies. The World Bank's Country Performance and Institutional Assessment, used to inform IDA allocations to client governments, includes criteria reflective of budgetary credibility. Specifically, there is an assessment of the quality of budgetary and financial management. This measures the extent to which there is: (a) a comprehensive and credible budget, linked to policy priorities; (b) effective financial management systems to ensure that the budget is implemented as intended in a controlled and predictable way; and (c) timely and accurate accounting and fiscal reporting. The African Development Bank and the Asian Development Bank use similar measures in their own assessments.

Similarly, the multi-donor Public Expenditure and Financial Accountability (PEFA) framework, for example, is built around the idea that one of the critical dimensions of an open and orderly PFM system is budgetary credibility.[1] To this end, it includes several indicators to help assess credibility relative to the approved budget. Among these are: *i*) aggregate expenditure deviations; *ii*) compositional expenditure deviations; *iii*) revenue deviations; and *iv*) the management of arrears.

Yet, as will be shown in this paper, many governments do not execute their budgets well. Aggregate and compositional targets are routinely missed even in the best performing countries. In fact, good budget execution is an example of something that is easy to talk about and very difficult to achieve. It requires well set goals that are within reach, financially, technically, and politically. Revenue targets and expenditure allocations that are not consistent with macroeconomic and fiscal constraints are unlikely to be achieved. Expenditure programs that are not well matched to service delivery or investment costs are likely to fail. Programs that do not have political backing will be obstructed. Once feasible goals have been agreed, technical skills are needed to procure inputs with good value for money, political solutions are needed to deal with self-serving politicians and their constituents, administrative solutions are needed to focus bureaucrats on service delivery rather than personal gain and counter-cyclical financing may be needed to offset unexpected shocks.

In essence, governments are somehow making simultaneous decisions about aggregate revenues and expenditures, net financing, and the composition of spending – and doing so while taking into account a variety of macroeconomic, technical and political considerations that may or may not be well aligned. Looked at in this light, it is surprising that any governments execute their budgets well. Yet, some do better than others and it would be useful to know why.

[1] The PEFA Program was founded in December 2001 as a multi-donor partnership between the World Bank, the European Commission, and the UK's Department for International Development, the Swiss State Secretariat for Economic Affairs, the French Ministry of Foreign Affairs, and the Royal Norwegian Ministry of Foreign Affairs, and the International Monetary Fund. For more information, see PEFA Secretariat (2005) and PEFA Secretariat (2011).

Given the importance that can be assigned to the quality of budget execution, one would expect that there would be a solid literature describing and explaining patterns of budgetary outcomes within and across countries. In fact, as reviewed below, this literature is quite uneven. There is a quite deep literature on the accuracy of revenue forecasting. Most of these studies, however, are not cross-country in nature and tend to focus on state and local governments within federal systems in developed nations where data are available for long time periods. Aggregate expenditure outcomes are described in considerable detail in the fiscal economics literature but almost always in the context of actual spending. Deviations from budgetary targets are rarely explored and, when they are, the topic is usually related to fiscal rules.[2] The literature associated with political institutions and budgetary procedures includes empirical tests of some specific hypotheses regarding the quality of fiscal outcomes, notably with regard to deficits and borrowing, but little or nothing is said about budgetary expenditure deviations per se. The public financial management (PFM) literature provides good information about the technical means by which budget execution can be improved[3] but this author has yet to find any systematic cross-country analysis of budgetary deviations or what drives them.

This paper makes three contributions. First, it introduces a new database of budgets and budgetary outcomes that should prove useful to researchers and practitioners in several fields. The database is derived from raw data found in PEFA country reports and includes 159 observations from 45 countries over multiple three year time periods with between 9 and 23 budget heads (line items) per observation for a total of just over three thousand budget heads.

Second, it provides an in-depth analysis of cross-country patterns of deviations from aggregate and compositional budget targets. The analysis here examines the degree to which actual revenues and spending deviate from approved budgetary allocations. This is followed by an analysis of compositional deviations. In general, it seems that very few countries are able to execute their budgets well, regardless of whether governments are prioritizing macroeconomic stability or public service delivery. There are widespread errors in forecasting revenues and within-year counter-cyclical financing does not always compensate for revenue deviations. Expenditure deviations therefore tend to be positively correlated with revenue deviations. Within this broad tendency, there is considerable variation in behavior not only across countries but also across time within countries. Within any specific country and year, it is also common to observe over-spending in some budget heads occurring simultaneously with under-spending in others. Compositional deviations from budgetary expenditure allocations therefore tend to be greater than deviations in aggregate expenditures.

Third, a number of hypotheses are proposed and tested that might help explain why some governments execute their budgets well while others do not. Most of these hypotheses can be grouped into just a few categories: the impact of debt service burdens, common pool behavior, the characteristics of political institutions, and the technical capacity for PFM. While this paper does not focus exclusively on any one variable, the capacity for PFM is of particular importance because investments in PFM capacity and quality offer a way forward for reform minded governments. In particular, it would be useful to learn if the quality of PFM affects the quality of budgetary execution after controlling for other variables.

The results of the hypothesis testing are interesting. Common pool tendencies to spend beyond approved borrowing limits are driven by the existence of various pressure groups that politicians

[2] An example can be found in Section VI.B of Kopits (2001).

[3] Shah (2007) for example includes an entire chapter on the subject.

believe they need to cater to, more so in democracies. Larger populations create such pressure since they are more likely to have more interest groups, including various ethnic factions and aging populations.

Revenue surpluses constitute a second common pool to spend from, in addition to soft borrowing constraints. Revenue deviations are partially the consequence of policy choices: unambitious revenue targets are correlated with unprogrammed revenue gains while overly ambitious revenue targets are correlated with revenue shortfalls. Revenue targets are more likely to be achieved in countries with presidential systems and/or strong capacity for public financial management.

Counter-cyclical financing is often used to at least partially smooth out unanticipated revenue shocks. Countries tend to do this better when they have strong capacity for PFM and high per-capita incomes. Low capacity countries, measured by small expenditure shares for their civil service, small populations, or very young populations tend to underspend in the face of unanticipated revenue gains. In addition, the presence of mature programmatic political parties appears to contribute to compositional accuracy in budget execution. Finally, high debt service burdens are correlated with, and must surely be the consequence of, revenue shortfalls and excessive spending. Yet, countries with high debt service burdens also tend to apply positive revenue shocks to debt reduction while negative revenue shocks induce reduced spending.

These results point to some policy implications. For most countries, reducing revenue deviations is the key to keeping expenditures on track both in level and in composition. To achieve this, governments should manage the ambitiousness of their revenue policy targets and make the kinds of investments that improved the capacity for public financial management. In some situations, capacity building may also require bringing in more skilled workers and ensuring they are adequately paid and motivated. Taking these actions will make it easier to achieve revenue policy targets while also reducing common pool problems. Such investments are surely easier to afford in wealthy countries but governments in low income countries may be able to advance to the extent that external financial and technical assistance is effective.

The remainder of the paper is organized in the following manner. A short literature review is presented in Section II. A new database is introduced and analyzed in Section III. Section IV delves deeper into the literature in order to develop a number of hypotheses that might help explain the patterns found in Section III. Testable hypotheses are presented in Section IV. Data characteristics and the econometric strategy for inference testing are discussed in Section V. The test results and policy implications are discussed in the Section VI. Conclusions and a proposed agenda for future work by interested researchers are found in Section VII.

II. Literature Review

Little has been published about the distribution of budgetary revenue and expenditure deviations around the world. Much more has been written about the distribution of budgetary balances. In the case of revenues, Buettner and Kauder (2010) is a rare example of a cross-country survey of revenue forecasting practices and outcomes within OECD countries. They found a wide range of outcomes within the period 1995 through 2009. All countries displayed a mix of positive and negative deviations from budgetary targets, but some countries showed a wide spread between maximum and minimum deviations while other countries tended to show a more narrow

distribution.[4] Most other studies of revenue forecast performance tend to focus on state and local governments within wealthy federal systems such as state governments. Examples include Mocan and Azad (1995) who cover the United States of America, Couture and Imbeau (2009) who test data from Canadian provinces, and Chatagny and Soguel (2011) who use data from Swiss cantons.

Examples of cross-country work on expenditure deviations are equally rare. Peters (2002) cites three. In one of these, Petrei (1998) documents the magnitude of aggregate expenditure deviations in six Latin American countries from 1991 to 1995. The deviations ranged from -24.65 percent (Brazil 1995) to 81.41 percent (Argentina 1992) with an average of 7.03 percent. In another, Kostopoulos (1999) includes a short description of aggregate budgetary deviations, and sector spending deviations, for 20 Sub-Saharan African countries. He finds that aggregate deviations were moderate while expenditure deviations for specific sectors tended to be much larger, frequently exceeding 30 percent. In the third, the IMF Fiscal Affairs Department (2001a) found that over one third of a sample of 25 highly indebted poor countries had "significant differences" reported between allocations and expenditure commitments.

What can explain the wide variation across countries in the quality of budget execution found in the studies cited above? The multi-disciplinary approach used here follows works by authors such as Alesina *et al* (1999), Fabrizio and Mody (2006) and Dabla-Norriss *et al* (2010), who explore whether various aspects of the budget process are correlated with aggregate fiscal discipline[5] after controlling for economic shocks and variables representing various political institutions. The question they address is an important one: it could be that political institutions are the most relevant factor in explaining fiscal discipline, and the quality of PFM might merely be a reflection of these institutions. If so, the prospects for successful PFM reform programs in poor performing countries would be limited, particularly since political institutions tend to remain in place for decades at a time.

Persson and Tabellini (2000) and Persson (2002) provide convincing evidence that aggregate fiscal outcomes are influenced by political institutions after conditioning for economic shocks. In particular, fiscal outcomes are correlated with electoral rules, rules governing the maintenance of executive tenure, and rules governing the separation of powers. Even so, Alesina *et al* (*ibid*), Fabrizio and Mody (*ibid*), and Dabla-Norriss (*ibid*) and several others find good evidence that budgetary procedures are effective in promoting aggregate fiscal discipline, after controlling for political institutions. If so, then there should be room for PFM reforms to proceed independently of political institutions.

When public service delivery is important, good budget execution will require the ability and desire to defend expenditure targets against unexpected shocks within a given fiscal year. Unexpected revenue shortages, perhaps induced by recession or by falling commodity prices, will need to be met by increased borrowing. Excess revenues would similarly be met by reduced borrowing if wasteful spending is to be avoided. In other words, good budget execution requires the willingness and capacity to deploy effective within-year counter-cyclical financing. Finding correlates for such counter-cyclical financing is one of the major goals pursued in this paper. Inspiration in this regard comes from Dabla-Norris (*ibid*) who explore the interaction between the quality of budgetary institutions and counter-cyclical policy as part of their broader inquiry into fiscal discipline. They

[4] They also found a strong degree of cyclicality in the distribution of forecast errors over time – though they did not explore the topic further.

[5] The primary dependent variables are the fiscal deficit as a share of GDP or the stock of debt as a share of GDP. Fiscal discipline is presumed to be stronger when deficits and debt stocks are smaller.

find that countries with strong budget institutions have lower fiscal deficits and debt and therefore they have more scope for conducting counter-cyclical policies when facing shocks than those with weaker institutions.

The ideas contained in the sources cited here, as well as several others, are used to form a number of hypotheses in Section IV below where the literature review is deepened. The hypotheses put forth in this paper are also informed by an examination of the dependent variables used to measure the quality of budget execution. The distribution of the cross-country data is interesting on its own merits but it also inspires some interesting questions. Thus Section III presents an in-depth analysis of the quality of budget execution across countries.

III. The Quality of Budget Execution

This section of the paper presents several measures of the quality of budget expenditure execution based on the fundamental idea that a well-executed budget should display minimal deviation from approved allocations, in aggregate and in composition.[6] Some care is needed in defining what deviation is to be measured. Many countries approve a budget at the start of a fiscal year and subsequently approve supplemental budgets to make within-year corrections to account for ad-hoc policy changes and various shocks. It would be admirable if all budgetary deviations were the consequence of such approved, within-year corrections. The fact that such corrections are necessary is, however, indicative of weaknesses in macroeconomic forecasting and in planning as well as the usual unanticipated shocks. This paper therefore follows the PEFA methodology which measures only the deviations between actual outcomes and the original budget at the start of each year rather than budgets that may have been subsequently modified.

These measures are derived from a new database of 159 observations from 45 central governments around the world. This is a larger number of countries than what has appeared in the literature to-date. The core of the database is constructed from the raw budgetary data used to calculate scores for PEFA indicators PI-1 and PI-2 for expenditure deviations from approved budgetary allocations in level and composition and PI-3 for revenue deviations.[7] These data exclude project related expenditures financed by donors and debt service obligations.[8] The paper does not make use of the PI-1, 2, and 3 scores themselves.

Of the 45 governments represented in the database, 14 are from Sub-Saharan Africa, 9 are from Europe and Central Asia, 7 are from Latin America and the Caribbean, 6 are from South Asia, 5 are from East Asia and the Pacific, 3 are from the Middle East and North Africa, and 1 is from a European OECD government.[9] Eight of these countries were assessed twice, bringing the total number of country observations to 53. The PEFA assessments include data for three years per country, so that the total number of country observations is 159. The years measured depend upon the period of each PEFA assessment and vary from 2002 to 2010.

---

[6] Those who are interested only in budget execution as a means of achieving public service delivery or public investment goals, without regard to macroeconomic objectives, might prefer a slightly different definition. For this narrower focus, the goal would be on minimizing the gap between *real* (inflation adjusted) allocations and expenditures. This would help ensure that the intended physical volumes of inputs and outputs are obtained.

[7] The PEFA website includes many more countries, allowing for future expansion of the database. Some of the country cases included in the PEFA website were rejected due to lack of published budgetary data, exclusion of domestically funded capital expenditures, less than three years of data, and other problems.

[8] These are outside the direct control of the client governments and are difficult for some governments to track.

[9] The PEFA website is particularly lacking in reports on developing countries from the Middle East and upper income countries in general.

The number of budget expenditure categories reported per country observation varies from a low of 9 to a high of 23, with the majority having 20 categories plus a 21rst containing "all other" budget categories. The total number of budget categories in the database is 3,069. In a few cases, these categories are either exact or close matches to the functional (sector) categories described in the IMF (2001b) Government Financial Statistics Manual. In most cases, however, these categories are actually ministry "budget heads" that may combine several sectors. This distinction is worth noting because the impact of ministerial accountability on performance can be observed more directly when the spending data refer to ministries rather than functional groups that reflect the combined decisions of more than one minister.[10] On the other hand, some of the political economy literature is more oriented around economic categories such as recurrent versus project spending.[11] In some cases, pensions are given a single budget line rather than being distributed across ministries or sectors. The data do not, at present, allow one to differentiate between statutory and discretionary spending nor between recurrent and capital spending.[12] (Many governments use a development budget that contains both recurrent and capital expenditures with the latter financed mainly by external donors.)

A. Budgetary Deviations in Levels

Before looking at the data, it is helpful to set up a simple algebraic framework to put aggregate deviations in context. Define the portion of budgetary expenditures (E) under the direct control of the government as primary expenditures less spending financed from external project support. The counterparts to E, as set out in equation 1 below, are revenues (R), net borrowing less external project support (B), and interest (I).[13] Equation 1 is simply the government's budget constraint with external project support subtracted from both sides of the equation.

$$(1) \qquad E = R + B - I$$

This budget constraint can be rewritten as:

$$(2) \qquad E = R + NF = TNR$$

where NF refers to net financing and TNR refers to total net resources. This can be re-written again in terms of the percentage budgetary deviations where subscripts *a* and *b* refer to actual outcomes versus budgetary targets. This will be helpful when comparing compositional deviations across countries indexed by *i* with different sizes of budgets and with different currencies.

$$(3) \qquad (E_{a,i} - E_{b,i})/E_{b,i} = (R_{a,i} - R_{b,i})/E_{b,i} + (NF_{a,i} - NF_{b,i})/E_{b,i} = (TNR_{a,i} - TNR_{b,i})/E_{b,i}$$

or

[10] It is common for governments to consolidate or split ministries from year to year. This can confound comparisons over time.

[11] Drazen and Eslava (2005) is a good example.

[12] Statutory expenditures on pensions and other transfers may be quite high in wealthy countries which tend to have older populations. With the exception of Norway, such countries do not appear in the sample used here.

[13] Net borrowing is therefore the sum of domestic borrowing, external commercial borrowing and external budget support. Data on planned and actual borrowing for these three categories was not consistently available. The PEFA indicator D-1 for direct budget support proved to be a poor source as it compared actual budget support outcomes with donor projections rather than approved budgetary estimates.

(4) $e_i = r_i + nf_i = tnr_i$

Figure 1 shows that the distribution of deviations in total expenditures *e* is centered almost on zero and includes a range of observations from a low of -45 percent to a high of 44 percent. This is a surprisingly wide range! Moreover, three observations exceed two standard deviations on the negative side of the distribution and another three exceed two standard deviations on the positive side. Outcomes over time are not uniform. Fifteen countries had expenditure deviations that were consistently positive; 13 countries had deviations that were consistently negative; and 17 countries had mixed outcomes. Longer time periods appear to increase the probability of a mixed outcome: of the eight countries with six observations rather than three, five had mixed outcomes.

FIGURE 1: DEVIATIONS IN TOTAL EXPENDITURES, e_i

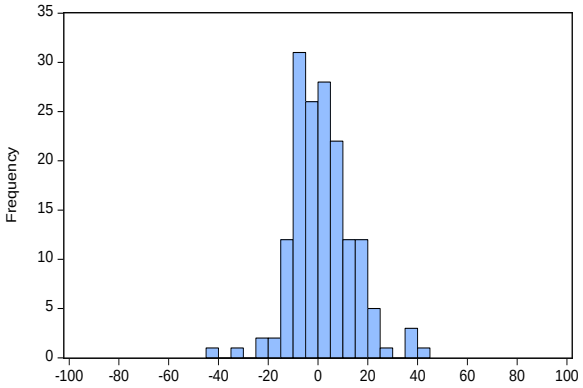
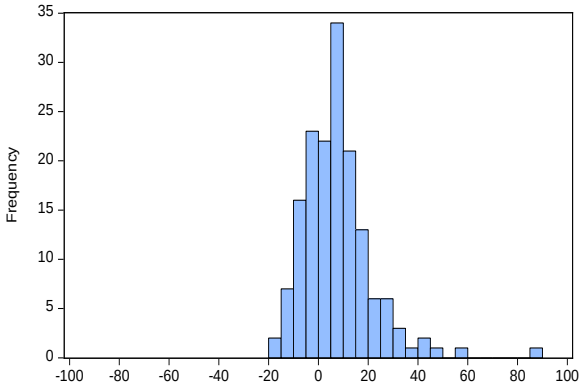


FIGURE 2: DEVIATIONS IN REVENUES, r_i



By contrast, Figure 2 shows that the majority of revenue deviations *r* were surpluses: 89 of 159 observations were in substantial surplus, meaning they were greater than 5 percent of the budget target (the forecast target was under-estimated), 45 were within +/- 5 percent, and only 25 were less than -5 percent of the target (the forecast target was over-estimated). Are some countries consistent over- or under-performers with regard to revenues? Examination of the data shows that 23 countries had observations with revenue deviations that were consistently positive; 2 countries (Grenada 2006-08 and Serbia 2007-09) had deviations that were consistently negative; and 20 countries had mixed outcomes.

B. Correlation between Revenue Deviations and Expenditure Deviations

Talvi and Vegh (2004) found evidence that actual spending in developing countries is highly correlated with actual revenues. This is less true for within-year budgetary deviations and there is evidence that some countries do save unanticipated revenue gains. Table 1 and Figures 3 and 4 portray the within-year correlation between aggregate revenue and aggregate expenditure deviations. In fact, there is a clear positive correlation between revenue deviations and expenditure deviations. Yet, the relationship is quite loose. As Table 1 shows, only 37 out of 89 observations had positive expenditure deviations greater than 5 percent matched with positive revenue deviations greater than 5 percent – while only 10 out 25 observations showed negative expenditure deviations of more than 5 percent matched with negative revenue deviations of more than 5 percent.

Talvi and Vegh (*ibid*) argued this positive correlation may occur because it is politically costly to build surpluses rather than spend. Alternatively, the positive within-year correlation could be appropriate behavior to the extent that governments have prioritized macroeconomic stability and

are trying to defend their deficit and borrowing targets: revenue and expenditure deviations would need to cancel each other out. The data in Figure 3 show that very few observations meet the canceling out criteria: only a few data points are on or close to the 45 degree line.

If public service delivery requires that expenditure targets be met with minimal deviations, then there will be a potential trade-off with macroeconomic stability since within-year counter-cyclical financing will be required to offset unanticipated revenue shocks. This leads to an interesting observation: one third of the 89 observations with unanticipated surplus revenues were paired with minimal deviations from allocated expenditure targets, implying that most of the surpluses were saved. Thus, some countries have shown a form of fiscal prudence. More broadly, there is an apparent tendency towards within-period counter-cyclical financing as implied by the clear negative correlation between revenue deviations and deviations in net financing in Figure 4. The implementation is far from perfect. If counter-cyclical financing had been perfect at all times in all countries, then Figure 3 would consist of a flat line drawn through the origin. In fact, there are many observations where net financing deviations were pro-cyclical (in the upper-right and lower-left quadrants of Figure 4).

TABLE 1—AGGREGATE REVENUE AND EXPENDITURE DEVIATIONS

Revenue Deviations		Expenditure Deviations	
Positive (r > +5%)	89	Negative (e < -5%)	21
		Near Zero (-5% < e < +5%)	31
		Positive (e > +5%)	37
Near Zero (-5% < r < +5%)	45	Negative (e < -5%)	18
		Near Zero (-5% < e < +5%)	17
		Positive (e > +5%)	10
Negative (r < -5%)	25	Negative (e < -5%)	10
		Near Zero (-5% < e < +5%)	6
		Positive (e > +5%)	9

FIGURE 3: DEVIATIONS IN REVENUES AND EXPENDITURES

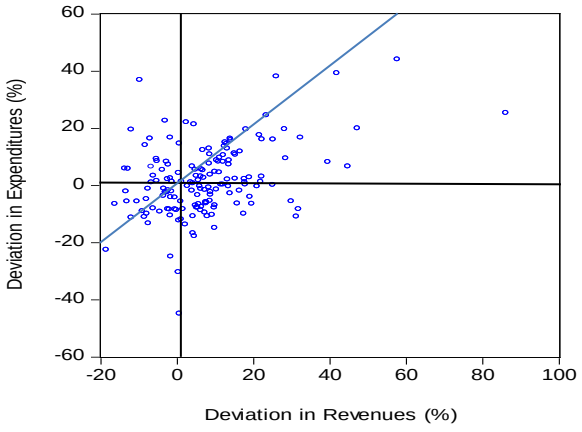
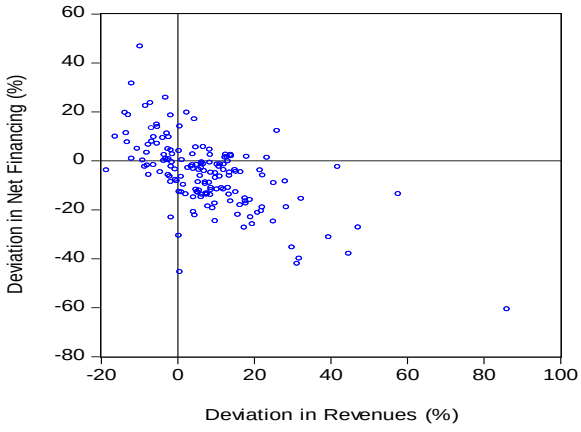


FIGURE 4: DEVIATIONS IN REVENUES AND NET FINANCING



It is also rather striking to see in Table 1 and in the bottom-right quadrant of Figure 3 that there are 21 observations of aggregate *under*-spending *despite the availability of unanticipated surplus revenues*. This may point to issues of administrative spending capacity. The 9 observations in the upper-left quadrant, with excess spending despite revenue shortfalls, may be less surprising due to the common pool and principal-agent problems discussed in Section IV below.

C. Expenditure Deviations in Composition

There are several ways one could measure compositional deviations in expenditure execution. The first is an overall measure of the quality of budget execution. It is focused on the average percentage deviation within the budget.[14] The second is focused on the percent deviation within the over-spent, under-spent, and accurately spent portions of a budget. The third is a measure of the share of budget heads over-spent, under-spent and accurately spent. The fourth is a measure of the average depth of over-spending or under-spending within each group of budget heads.[15]

In order to get a glimpse of the compositional data using the first measure, one can construct a measure of the average absolute compositional deviations within a single budget, $|d|$. Start by calculating the difference between the approved budgetary allocation for each budget head and the actual spending for that head. These deviations can be expressed as percentages of the approved allocations and will be positive or negative. In order to create a measure of overall ability to deliver an approved budget, the absolute values of the percentage deviations can be averaged across the total number of budget heads. The calculation is summarized in equation 5:

$$(5) \qquad |d| = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\left(E_{a,i}-E_{b,i}\right)}{E_{b,i}}\right|$$

where n is the number of budget heads, E is the expenditure assigned to budget head i, and the a and b subscripts refer to actual and budgeted allocations respectively.

There is typically a very wide range of intra-budgetary deviations for any given government and year. This is not surprising: each year brings new macroeconomic shocks and policy changes. For example, the lowest average absolute compositional deviation within the database was 4.5 percent, found in Burundi in 2005. Yet, the absolute value of the intra-budgetary deviations within the Burundi budget ranged from a low of 0.3 percent to a high of 19.8 percent. The worst average absolute compositional deviation within the sample was 435 percent, found in Bolivia in 2006. The absolute value of the intra-budgetary deviations in this case ranged from a low of 3.3 percent to a high of 7,673 percent!

It is important to note that, the good performers in one year are not necessarily equally as good in the next year. For example, the Burundi average worsened from 4.5 percent in 2005 to 24 percent in 2006 and 15 percent in 2007 while Bolivia improved to 135 percent in 2007 and 97 percent in 2008. Moreover, ministries or spending categories with large deviations in one year do not necessarily show similarly large deviations in other years. This may suggest a role for unanticipated shocks or, perhaps, ad hoc policy changes.

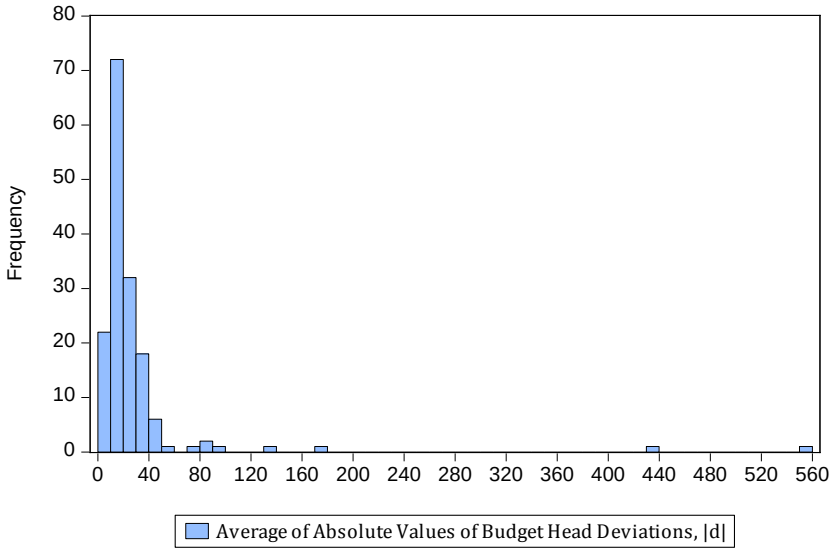
The distribution of the average absolute compositional deviation across all 159 observations is shown in the histogram in Figure 5. The median average deviation is 16 percent. The sample mean of the individual average deviations within a country is 29 percent with a standard deviation of 59

[14] It is similar to, but not the same as indicator PI-2 found in the PEFA Framework.

[15] Another measure, from Brender and Drazen (2010), is focused on changes in composition rather than expenditure deviations per se. If actual spending in each expenditure category within a budget fell short of approved allocations by 40 percent, then the composition will be unchanged: each category will retain its intended share of the budget. In contrast to the other measures in this paper, the Brender-Drazen (*ibid*) measure would imply such an outcome is as desirable as one with no deviations in any category.

percent. The distribution has quite a long tail: the maximum average deviation in the data set is 557 percent.

FIGURE 5: DISTRIBUTION OF COMPOSITIONAL DEVIATIONS



While the measure of absolute compositional deviation clearly illustrates how few countries are able to accurately execute their budgets, it also obscures an interesting second puzzle in the data: under-spending and over-spending almost always occur at the same time within a budget. This is a puzzle because the most obvious reactions to a shortfall in net resources would be to either cut all budget heads by the same amount – or cut all the low priority programs. Without more information about extenuating circumstances, increasing spending on some ministries or programs during a shortfall would seem counterintuitive. Similarly, it would seem odd, at least at first glance, to reduce spending for some budget heads when surplus resources become available.

To further explore compositional deviations, the data for each country observation are divided into three groups: over-spent budget heads, under-spent heads, and those accurately spent within ±5 percent of budgeted allocations. These three groups are tied together by a simple relationship:

(6) $e_i = e_{o,i} + e_{u,i} + e_{b,i}$ where $e_{u,i} < 0$

where e symbolizes the percentage deviation in expenditures within each group, and the subscripts refer to the over-spent group o, the under-spent group u, and the group spent accurately within ±5 percent of budgeted allocations b. The expenditure deviations are expressed as percentages of total budgeted expenditures, E_b. Table 2 provides an example of how the data are generated. Figures 6a through 6c show the distribution of each measure across the sample.

Substantial portions of many budgets within the data sample were over-spent even when total net resources fell below budgeted amounts and vice-versa, with portions under-spent despite the availability of excess resources. In fact, as shown in Table 3, for 56 instances of excess net resources, there were 557 heads that were over-spent by more than 5 percent while 221 heads under-spent by more than 5 percent. Similarly, for the 49 instances of net resource shortfalls, there were 568 heads under-spent by more than 5 percent while 162 heads were over-spent by more than 5 percent. There was also simultaneous over- and under-spending even in the 54 cases when resources were close to their targets.

TABLE 2—EXAMPLE OF CALCULATIONS FOR BHUTAN, 2005-06

	Over-spent group	Under-spent group	Accurately-spent group	Total
Number of budget heads (count)	1	7	4	12
Budgetary allocations (Ngultrum millions)	682	5,478	3,494	9,654
Actual expenditures (Ngultrum millions)	740	4,815	3,425	8,980
Deviations (Ngultrum millions)	58	-663	-69	-674
Deviations w.r.t. to own base (%)	8.50	-12.10	-1.97	-6.98
	e_o	e_u	e_a	e
Deviations w.r.t. to total spending (%)	0.60	-6.87	-0.71	-6.98
	f_o	f_u	f_b	
Fraction of budget heads (%)	8.33	58.33	33.33	100.00
	d_o	d_u	d_b	
Average percent deviation per share of heads (%)	7.21	-11.77	-2.14	

Source: Bhutan Public Financial Management Accountability Assessment, World Bank Report No. 58444-BT.

Even the best performers are not perfect. South Africa between 2005/06 and 2007/08 was the best performer, with the largest share of the budget accurately spent, within the sample. The government managed to keep only 17 of 21 budget heads within 5 percent of approved allocations in 2006/07 when the deviation in net resources was only -0.6 percent. In that year, three heads were under-spent while one was over-spent. In 2005/06, only 14 out of 21 budget heads were within 5 percent of approved allocations when the deviation in net resources was +0.4 percent. Four heads were over-spent and three were under-spent. In 2007/08, only 15 of 21 budget heads were close to approved allocations, four over-spent, and 2 under-spent. One finds simultaneous under- and over-spending even in the best performing governments.

DISTRIBUTIONS FOR MEASURES OF COMPOSITIONAL DEVIATION

Figure 6a. Over-spent group, e_o

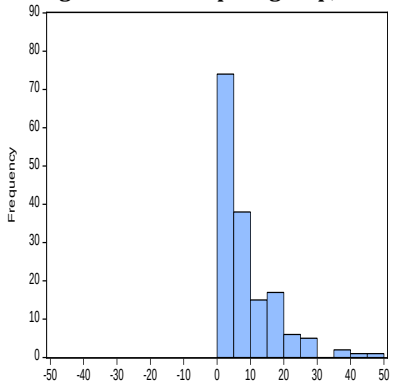


Figure 6b. Under-spent group, e_u

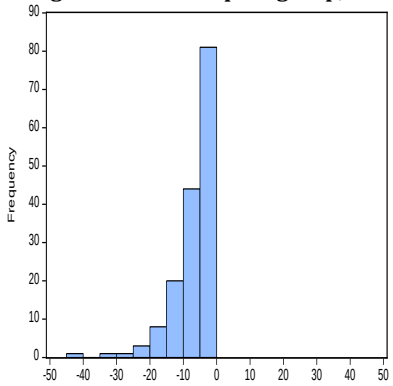
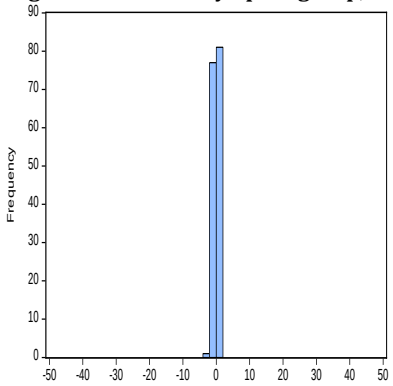


Figure 6c. Accurately-spent group, e_b



Compositional deviations tend to be larger than the deviations in total net resources because of simultaneous over- and under-spending. For example, over-spending some budget heads during a period of an unanticipated resource shortfall necessarily requires that the remaining budget heads be cut beyond what the shortfall would have otherwise required. As shown in Table 3, for the 49 cases of resource shortfalls, the average deviation in net resources was -10.7 percent while the average deviation within under-spent budget heads, e_u, was -13.1 percent, and the average deviation in over-spent heads, e_o, was 2.5 percent. For the 56 cases of surplus resources, the average deviation in net resources was 14.6 percent while the average deviation within over-spent budget heads was 17.4 percent and the average deviation within under-spent heads was -3.0 percent.

TABLE 3—DISTRIBUTION OF COMPOSITIONAL DEVIATIONS

	Net Resource Deviation < -5%	Net Resources Close to Budget	Net Resource Deviation > 5%	Totals
Country Observations	49	54	56	159
Heads Under-spent < -5%				
Number of Heads	568	313	221	1,102
Average Deviation, e_u	-13.1%	-4.4%	-3.0%	-6.6%
Heads within ±5%				
Number of Heads	257	428	261	946
Average Deviation, e_b	-0.2%	0.0%	0.2%	0.0%
Heads Over-spent > 5%				
Number of Heads	162	302	557	1,021
Average Deviation, e_o	2.5%	4.5%	17.4%	8.4%
Total				
Number of Heads	987	1,043	1,112	3,142
Average Deviation, e	-10.7%	0.1%	14.6%	1.8%

D. Compositional Deviations and Net Resource Deviations

Figures 7a through 7c illustrate how over- and under-spending varies with deviations in total net resources ($e = tnr$). Figure 7a is focused on over-spent heads e_o while Figure 7b is focused on under-spent heads e_u. Both figures should and do display positive slopes: the positive deviations in over-spent heads should increase as resource deviations become more positive while the negative deviations in under-spent heads should become less negative as resource deviations become more positive. Figure 7c shows there is no obvious relationship between resource deviations and the depth of over- or under-spending for the well-spent share of budget heads.

PERCENT DEVIATIONS AS A FUNCTION OF RESOURCE DEVIATIONS

Figure 7a. Over-spent Groups, e_o

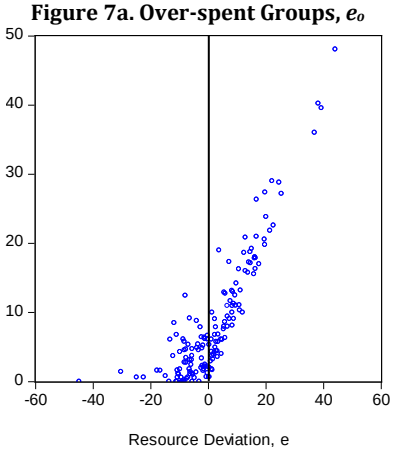


Figure 7b. Under-spent Groups, e_u

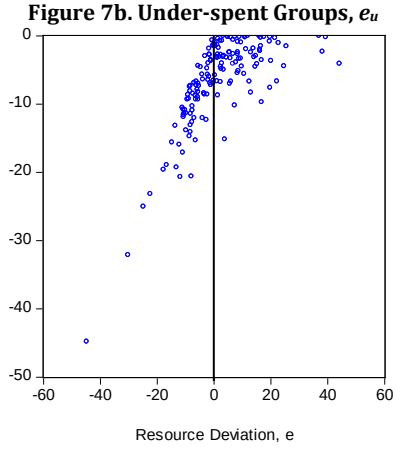
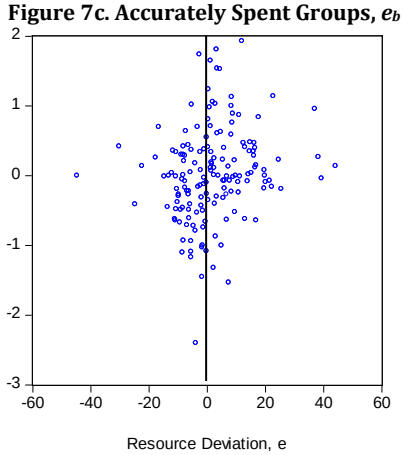


Figure 7c. Accurately Spent Groups, e_b



The correlations in Figures 7a through 7c pose three puzzles. First, why is there over- and under-spending when e is close to zero? Second, a glance at Figure 7a shows over-spending despite unanticipated net resource shortfalls, while Figure 7b shows under-spending despite gains in net resources. Inspection of the data shows that some of the under-spending is paired with over-spending that occurred in excess of gains in net resources. Similarly, some of the over-spending is accompanied by under-spending that occurred in excess of losses in net resources. Yet to be explained, however, is why this sort of thing happens at all. Common pool and agency problems, explored in Section IV below may provide part of the answer.

Third, why are the data for e_o and e_u distributed as if inside two triangles? The data show that over-spending and under-spending tend to occur simultaneously most frequently when net resource deviations are small. When resource deviations are very large, there is almost no under-spending. Conversely, when the shock is very negative, there is almost no over-spending.

To probe more deeply into this question, two additional compositional measures are introduced. One is the share of over-spent, under-spent, and accurately-spent budget heads (f_o, f_u, and f_b). For example, in Table 2, there were 7 under-spent heads out of a total of 12, so f_u = 58 percent. These shares sum to unity:

(7) $\quad 1 = f_o + f_u + f_b$

The other measure is the depth expenditure deviations within each share above (d_o, d_u, and d_b). Thus one can write:

(8) $\quad e = f_o \cdot d_o + f_u \cdot d_u + f_b \cdot d_b$ where $d_u < 0$

For the majority of countries in the sample, most ministries were able to obtain a share of the largest windfalls and almost every ministry shared the pain of large unexpected losses. This can be seen in Figures 8a and 8b where the over-spent fraction of the budget f_o is positively correlated with net resource shocks while the fraction of the under-spent budget f_u is negatively correlated with such shocks. As these fractions approach unity, the depth of expenditure deviations (d_o, d_u) must necessarily approach the aggregate deviation e.

An unsurprising consequence is that the share of the budget f_b that is accurately spent tends to be higher when unanticipated shocks to net resources are small. This is depicted in Figure 8c. The looseness of the relationship, however, suggests that this broad tendency may be shaped in some countries by other considerations beyond the strength of unanticipated shocks. This will be explored more in the sections below.

BUDGET SHARES AS A FUNCTION OF RESOURCE DEVIATIONS

Figure 8a. Heads Over-spent, f_o

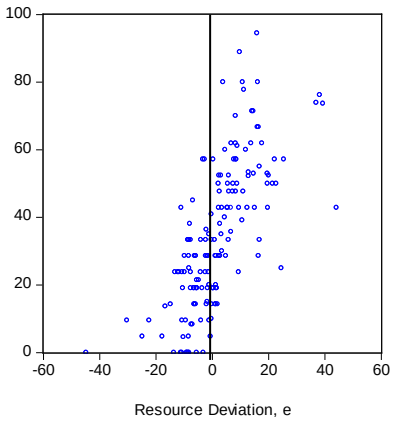


Figure 8b. Heads Under-spent, f_u

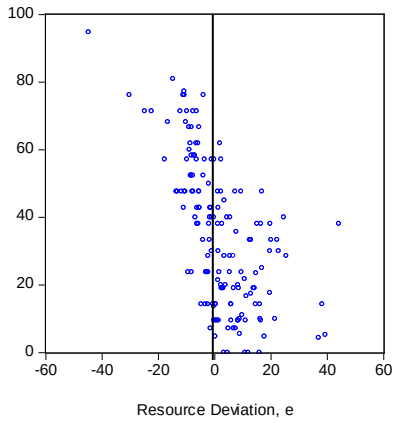
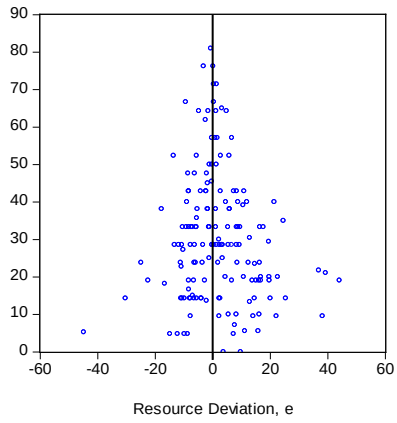


Figure 8c. Heads Within-Budget, f_b



Before concluding this section, it is useful to summarize the three multiple levels of decision making that take place in budget execution: policy makers and stakeholders choose the size of the *aggregate* deviation, *e,* which is equivalent to deciding the amount of revenue deviation, *r*, and within period counter-cyclical financing, *nf*; they choose spending *across* groups of budget heads

(e_o, e_u, e_b); and they decide the trade-off between the fraction of the budget heads *within* each group (f_o, f_w, f_b) and the depth of expenditure deviations (d_o, d_u, d_b). These multiple levels are described by the following accounting relationships:

(4) $e = r + nf$
(6) $e = e_o + e_u + e_b$ where $e_u < 0$
(7) $1 = f_o + f_u + f_b$
(8) $e = f_o \cdot d_o + f_u \cdot d_u + f_b \cdot d_b$ where $d_u < 0$ and

The main conclusion thus far is that the quality of budget execution is rather poor in most countries, regardless of whether governments are prioritizing macroeconomic stability or public service delivery. Revenue targets are usually missed and under-estimated more often than over-estimated. Even so, unanticipated revenue surpluses are saved less frequently than spent: only some governments employ counter-cyclical financing to smooth out revenue shocks. The result is a loose positive correlation between expenditure deviations and revenue deviations. Even in the best performers, it is rare for more than 15 of the 20 largest ministries to spend within 5 percent of their budgeted allocations. Moreover, the best performing countries in one year are not necessarily the best in the next year. Over-spending is common, even in the face of revenue short-falls. Yet, there are also more than a few observations of aggregate under-spending despite positive revenue deviations. This is rather surprising. Within the composition of the aggregate total, over-spending and under-spending often occur simultaneously, at least when resource shocks are not large. This simultaneity causes compositional deviations to be larger than the deviations in total net resources. Compositional over-spending and under-spending appears to be positively correlated with net resource deviations on a cross-country basis. In addition, there appears to be a tendency for a greater number of ministries to share in the impact of an unanticipated shock as the size of shocks increase, though this tendency is far from uniform within the sample.

The observations above could motivate a large number of hypotheses. There is a lot to explore. The remainder of this paper is devoted to the exploration of only a few topics: what are the correlates for aggregate revenue and expenditure deviations, r and e, and what are the correlates for compositional accuracy, defined as the share of the budget accurately spent, f_b?

IV. Hypotheses

This section presents a number of hypotheses that might explain the cross-country distribution of the budgetary deviations described above. These hypotheses are drawn from three strands of literature: political institutions, fiscal economics, and public financial management. The discussion below proceeds in stages, from general concepts, to a framework for hypothesis testing, and concludes with a list of specific, testable hypotheses.

A. General Concepts

The now standard point of departure for discussions of budgeting and budget execution begins with a discussion of the theories of the common pool-problem and the principal-agent problem. Both problems create pressures for spending above socially optimal levels. As discussed in von Hagen (2002), the common pool problem arises when politicians can concentrate their spending on various targeted groups while taxing all members of the larger population.[16] This creates a gap

[16] Weingast, Shepsle, and Johnsen (1981) put the common pool problem at the heart of their model of geographically targeted "pork barrel" projects.

between the net social benefits perceived by the targeted groups and the net social benefits for society. In consequence, each of the targeted groups and their politicians will demand more spending than what is socially optimal for society as a whole. The principal-agent problem in budgeting is created when spending decisions are delegated by voters to politicians who they cannot fully control. Likewise, politicians must delegate to bureaucrats who they cannot fully control. This lack of control allows politicians and bureaucrats to pursue their own interests, often by finding ways to extract rents through the budgetary process. As a result, public expenditure is increased beyond the socially optimal level.

The focus of the logic described above is on the gap between actual and socially optimal spending. By contrast, this paper is focused instead on the gap between actual spending and approved budgetary allocations. Even so, the logic can be adapted. For example, once a budget is formed, politicians and bureaucrats may not believe that the budget constraint is binding and, thus, that additional borrowing will be allowed.[17] This potential for borrowing creates the equivalent of a common-pool resource from which additional spending above the approved allocations can be financed. The cost of poor budget execution will be manifested in less public service delivery or less public investment but that loss will be spread over the entire population whereas, in contrast, members of favored groups will enjoy targeted transfers or other spending.

Unanticipated revenue shocks, *r*, could be interpreted as a second common pool. Politicians and bureaucrats may expect more revenues than forecast, knowing the Finance Ministers often under-estimate revenues. Indeed over the course of a fiscal year they may come to see that extra revenues are forthcoming. Tornell and Lane (1999) argue that weak governance allows competing groups to exploit such a common pool for purposes of redistribution more than proportionately as resources increase (the voracity effect). They focus on shocks outside the control of the government. The idea advanced here is that some finance ministers (23 out of 45 in the sample used here) may create a common pool of surplus revenue by deliberately under-estimating annual revenue targets.

Why would a government behave in such a manner? One possibility advanced by van der Ploeg (2007) is that the Finance Minister wishes to insure against the possibility of revenue shortfalls due to forecast errors, believing that any surplus could be applied to debt reduction rather than wasted on unplanned spending. Another possibility is that political leaders may want to create opportunities for discretionary expenditures in order to dispense patronage to needed allies.

In summary, there are two common pools for politicians and bureaucrats to draw from. The first is associated with soft budget constraints and the ability to *borrow* beyond budgeted amounts. The second arises from within-period *revenue* shocks, *r*, of which some portion is truly unexpected due to uncertainty or weak technical capacity. In some countries, another portion may be anticipated, at least when and where finance ministers are known to routinely under-estimate their revenue targets. Either way, common pool behavior becomes a possibility once it is known that extra revenues are available. Yet, a capable government that has prioritized service delivery and effective investment over other goals will want to contain common pool behavior in order to maintain a predictable flow of resources to its ministries by diverting unanticipated revenue shocks to net financing.

[17] Rather than assuming a government is too weak to stop any over-spending, it might be more realistic to assume a government is strong enough only to reduce spending by politically weak ministers (or technically weak ministries) in order to compensate for over-spending by others.

These ideas form the foundation of three over-arching hypotheses. The first is that there may be a common set of underlying variables driving deviations in both revenues and expenditures, at least for some governments. The second hypothesis is that common-pool and agency (CP&A) problems associated with soft budget constraints are modulated across countries by political institutions or rules and by their capacity for public financial management.[18] Good rules and strong capacity for financial management could minimize if not eliminate these problems altogether. Third, CP&A problems arising from unanticipated revenue shocks can be reduced through appropriate rules and adequate capacity for within-period counter-cyclical financing. This is a fairly direct parallel to the argument from Dabla-Norris *et al* (2010) that fiscal policies tend to be less pro-cyclical in countries where budgetary institutions are strong.

B. Framework for Hypothesis Testing

These over-arching hypotheses can inform the super-structure of a framework for the empirical testing of more specific hypotheses about the correlates of aggregate revenue and expenditure deviations, *r* and *e*, and the correlates of the share of the budget accurately spent, f_b.

Equation 9, shown in Table 4, sets out the hypothesized model for aggregate revenue deviations, *r*. In the case of revenue deviations, *r*, it is assumed that budgetary deviations will be a function of revenue policy, *Rp*, revenue volatility, *Rv*, the will and capacity to achieve policy goals, and several variables likely to encourage or discourage common-pool behavior.

The variable designated *Rp* for revenue policy is constructed to measure the impact of administrative and policy changes relative to actual outcomes in the previous year. It is expected that unambitious policies will be associated with under-estimated revenue targets and thus actual revenue will exceed the target, leading to a positive revenue deviation. Overly ambitious policies and revenue shortfalls are also possible. This variable can therefore take on positive or negative values. It is constructed as the ratio of what revenues would have been, without any revenue policy changes, to budgeted revenues, less one.[19] Thus, an ambitious revenue target would generate a negative value for *Rp*, an unambitious target would lead to a positive value for *Rp*, and one would expect a positive correlation between *Rp* and *r*.

Higher levels of revenue volatility *Rv* are expected to make it more difficult for governments to accurately forecast what revenues will be.[20] This variable is calculated as the standard deviation in the growth rate of nominal revenues for three years prior to any current year. It thus captures volatility in both the real tax base and in prices. This variable is always positive. It is assumed, following the prudence hypothesis from van de Ploeg (2007), that governments will choose to insure against difficulties in making revenue forecasts by under-estimating their revenue targets in order to generate surplus revenues $r > 0$, with this tendency growing as *Rv* grows.

---

[18] It is assumed here that the modulating variables have an equal impact on CP&A problems arising from unprogrammed borrowing and from unanticipated revenues.

[19] A crude proxy for what revenues would have been, without any revenue policy changes, is calculated by inflating the actual revenue outcome from the previous year by actual real GDP growth and actual price inflation, e.g. by GDP_t/GDP_{t-1}. In other words, it is a proxy for what revenues would have been *with* actual changes in the tax base and *without* the expected impact of changes in administration or policy. A more sophisticated proxy would deal with changes in trade and other aspects.

[20] Wildavsky (1986) offers a related but different line of thought. He argued that wealthy countries with stable tax bases would tend to display incremental budgets, with only small changes from year to year, while the least developed countries with unstable tax bases would be forced to use "repetitive" budgeting, adjusting their expenditures frequently within each year in accord with cash availability. In fact, tax base stability could be an inherent feature of wealthy countries: Ramey and Ramey (1995) observed, wealthier countries tend to have greater stability in output growth.

As indicated above, the revenue model also allows for the possibility that the variables X_j could promote CP&A problems directly while variables Y_k could reduce them. Specific hypotheses for these variables appear in the next section.

It is assumed here that some aspects of institutional quality or capacity will make it easier for some governments to achieve their revenue targets than others. This is modeled by including an interactive term $\rho_i \cdot Rp \cdot \Sigma Z(Y_i)$ where Y_i is transformed in the manner shown here so that $Z(Y_i) = 0$ at maximum beneficial value of Y_i:

$$Z(Y_i) = \frac{Y_{max} - Y_i}{Y_{max} - Y_{min}}$$

As an example of the above, perhaps deviations coming from revenue policy are minimized when the capacity for public financial management is maximized.

Macroeconomic and fiscal forces must also have a role. Habitual revenue shortfalls and over-spending contribute to high interest obligations. Thus, there should be a negative correlation between *r* and *PCT_INT* and a positive correlation between *e* and *PCT_INT* where *PCT_INT* is defined as ratio of interest payments made to actual revenues received. The direction of causality could run in the other directions as well, with high debt service burdens inducing policy responses. Couture and Imbeau (2009) find that the presence of anti-deficit laws in Canadian provinces tends to reduce revenue forecast errors. van der Ploeg (2012) argues that, when debt burdens are high, prudent governments would want to under-estimate their revenue targets in order to generate a surplus that could be applied to debt reduction. This could be modeled by including an interactive term $Rp \cdot Z(PCT\_INT) > 0$ so that deviations from revenue policy are reduced when interest obligations are high.

The assumptions for expenditure deviations, *e*, should be similar: budgetary deviations should be a function of several variables that encourage or discourage borrowing-based common-pool behavior, unanticipated revenue shocks and the will and capacity to use within period counter-cyclical financing to smooth out revenue shocks.

Aggregate expenditure deviations, *e,* are expected to conform to equation 10. The interactive terms reflect the relationship shown in Figure 3, that revenue deviations and expenditure deviations are positively correlated. It also allows for the possibility that some governments might have the will and capacity to smooth out revenue deviations by using within period counter-cyclical financing. This is modeled through an interactive term $\varepsilon_i \cdot r \cdot \Sigma Z(Y_i)$ so that Y_i will vary between 1 for its' least beneficial value to zero for its' most beneficial value. In this way, the impact of unanticipated revenue deviations can be assumed completely eliminated in countries with the greatest degree of will and capacity for counter-cyclical financing. Countries lacking sufficient will or capacity would pass on the full impact of any revenue deviations, multiplied by the coefficients ε_i.

The aggregate expenditure model also allows for the possibility that the variables X_j could worsen CP&A problems directly while variables Y_k could reduce them. If there is always some over-spending, per the CP&A problems, then $\varepsilon_0 + \varepsilon_i \cdot \Sigma X_i - \varepsilon_j \cdot \Sigma Y_j + \varepsilon_k \cdot r \cdot \Sigma Z(Y_k)$ would always be positive. It is possible, however, that some countries might display a tendency towards under-spending due to administrative or technical capacity constraints. If so, then the equation could generate negative values. This might help explain the puzzling observations where aggregate under-spending occurred despite surplus net resources.

It is also assumed that a heavily indebted government could reduce spending below approved allocations in order to slow or reverse the rate of borrowing. This suggests that if interest obligations are a large share of actual revenues, then a larger share of any unanticipated revenue surpluses should be saved to reduce the stock of debt rather than spent on various ministries.[21] This implies there should be a positive correlation between aggregate expenditure deviations and $r \cdot Z(PCT\_INT)$ when $r > 0$. Conversely, expenditures should be cut in response to unanticipated revenue shortfalls.

Once the size and direction of the aggregate deviation in expenditures is known, one can ask how the composition of expenditure will be affected. It is assumed here that the fraction of budget heads that are accurately spent, f_b, will decrease as the absolute value of aggregate budget deviations e increases, modulated by variables hypothesized to impact on the CP&A problems.

The model does not capture everything. For example, when faced with an unanticipated shock from very high inflation, many governments will seek to cut spending (and/or increase taxation) in order to reduce the need for monetary financing. When faced with a large output gap, many governments will expand spending (and/or reduce taxation) in order to stimulate real demand and job creation. Large relative price changes originating from sources such as exchange rate movements or fluctuations in fuel prices can motivate compositional changes in budget execution. While the impacts from each of these forces are implicitly included in the data for aggregate revenue deviations, it would be better if they could be studied explicitly. Unfortunately, this is not possible without an additional investment in data gathering: few governments publish the necessary data that allow a distinction between actual and expected outcomes for general inflation, fuel price inflation, exchange rates, or output gaps.

TABLE 4 – EMPIRICAL MODELS FOR HYPOTHESIS TESTING

		Equation 9 r	*Equation 10* e	*Equation 11* f_b
	Dependent Variables			
1	Aggregate Revenue Deviation, r	1		
2	Aggregate Expenditure Deviation, e		1	$-\phi_1 \cdot \lvert e \rvert$
3	Compositional Expenditure Deviation, f_b			1
4	**Constants**	ρ_0	ε_0	ϕ_0
	Variables Maintained as Exogenous			
5	Revenue Policy	$+ \rho_1 \cdot Rp$		
6	Revenue Volatility	$+ \rho_2 \cdot Rv$		
7	Debt Service Burden	$- \rho_3 \cdot PCT\_INT$	$+ \varepsilon_1 \cdot PCT\_INT$	$- \phi_2 \cdot PCT\_INT$
8		$+ \rho_4 \cdot Rp \cdot Z(PCT\_INT)$	$+ \varepsilon_2 \cdot r_{pos} \cdot Z(PCT\_INT)$	
9	Contributors to Common Pool Problems	$+\rho_i \cdot \Sigma X_i$	$+\varepsilon_i \cdot \Sigma X_i$	$+\phi_i \cdot \Sigma X_i$
10	Will and Capacity to Limit C.P. Problems	$-\rho_j \cdot \Sigma Y_j$	$-\varepsilon_j \cdot \Sigma Y_j$	$+\phi_j \cdot \Sigma Y_j$
11	Revenue Policy · Will and Capacity	$- \rho_k \cdot Rp \cdot \Sigma Z(Y_k)$		
12	Revenue Deviation · Will and Capacity		$+\varepsilon_k \cdot r \cdot \Sigma Z(Y_k)$	
13	Error terms	λ	μ	ν

Before concluding this section, it can be noted that the equations 9, 10 and 11 are cascading in the sense that the outcome of revenue deviations in equation 9 drive the outcome of expenditure deviations in equation 10 which, in turn, drive the outcome of compositional accuracy in equation 11. In theory, it is possible that one equation could be a linear combination of another if hypothesis

[21] No portion of an unanticipated surplus may be used for spending in Belgium, as noted by Shah (2007).

testing rejects all variables except those in rows 7, 9 and 10. In practice, this does not happen, as documented in Section VI below. All equations are well identified. In addition, there should be no risk of simultaneity bias provided that each of the explanatory variables are exogenous as maintained here.[22]

C. Key Hypotheses

Short descriptions of key hypotheses concerning the role of political institutions and the capacity for public financial management are provided below. These are put into three broad categories: (*i*) variables expected to worsen common pool behavior; (*ii*) variables related to various aspects of political institutions; and (*iii*) the technical capacity for public financial management.[23] The variables in the first group correspond to ΣX_i above while variables in the second and third groups correspond to variables ΣY_i. Each will ultimately be summarized in the context of equations 9, 10 and 11 along with the expected signs for each variable in Table 6 at the conclusion of this section.

1. Common Pool Behavior

H01: Fractionalization. As noted at the beginning of this section, the common-pool problem arises when politicians can make targeted expenditures to various groups of potential voters. It follows then that the common pool problem should be more evident when society is divided into various factions.[24] Thus, ethnic, linguistic or religious fractionalization could be expected increase aggregate expenditure deviations *r* and *e* and thus reduce compositional accuracy f_b. To test these ideas, we use the three fractionalization variables from Alesina *et al* (2003), assuming that the data do not change from year to year.

H02: Dependency ratios. The same logic can be used to generate a set of hypotheses based on demographics. Tonizzo (2008) argued that a large share of the population over 65 years of age would increase the redistributive demand for social services, thereby increasing total expenditures. Perhaps, if the share of the elderly is very high, politicians might feel pressure to over-spend on hospitals, pensions, and other schemes. Similarly, if the share of youth in the population is high, perhaps parents would pressure politicians for more spending on pre-schools and day-care. To test these ideas, dependency ratios for those below age 15 and above age 65 are included as potential explanatory variables. These data are sourced from the UN Population Division.

H03: Civil Service Employment. Most countries in the sample used for this paper are developing countries that finance most of their capital investments from external aid. Conversely, rich countries such as Norway tend to have small capital budgets since the private sector takes on a greater role in providing infrastructure and utility services. Either way, the portion of expenditures assigned to wages and salaries in this sample is often substantial. Government hiring can be a form of patronage, and thus a likely contributor to the common-pool problem. If so, there should be a positive correlation between the share of spending on civil service employment, *PCT_CSE*, and

[22] The author tested for the possibility that there might be endogeneity between PEFA scores and aggregate deviations in revenues or expenditures: perhaps it is easier for countries with low budgetary deviations to earn high PEFA scores. This possibility was rejected by the data. Some of variables could be deemed endogenous over longer time periods. For example, youth dependency ratios could be positively correlated with expenditure deviations if such deviations led to higher birth rates through poorly executed health and education policies.

[23] Some of these variables, which could be assigned to more than one group, will be re-grouped in Section VI after hypothesis testing.

[24] There is an implicit assumption here that fractionalized groups are always and everywhere organized enough to warrant attention from one or more politicians. This is almost certainly not true.

aggregate deviations, *r* and *e*. Alternatively, civil service employment could be viewed as an essential input to public service delivery and, as such, something to be protected. Such spending should be well insulated from revenue shortfalls, if only because politicians want to avoid unrest and/or they know it is difficult to replace skilled workers lost to attrition over poor pay conditions. This protection would be achieved through counter-cyclical financing and the avoidance of over-spending on non-wage related activities.

In this context, it is worth noting that there could be an alternative interpretation for the youth dependency ratio in the set of hypotheses H02 above. It may instead be related to capacity for PFM: very young populations tend to be less educated and therefore there is a smaller pool of talent for a government civil service to draw upon.

2. Political Institutions

Von Hagen (2002) lists three options for bringing spending levels closer to the social optimum. These include the adoption of electoral rules that encourage political accountability and competition; the adoption of fiscal rules, such as a prohibition against borrowing; and the adoption of better decision making procedures. He notes that there is an accumulation of strong empirical evidence that governments that incorporate one or more of these elements of cost centralization tend to have lower spending, lower fiscal deficits and less debt that those lacking them. These ideas can also be adapted to help explain cross country patterns of budget deviations.

The next several hypotheses therefore focus on the impact of political institutions and rules governing elections. These are grouped according to the following typology: autocracies versus anocracies versus democracies[25], and, within the democracies, executives elected by voters versus those elected by parliaments and plurality voting versus proportional voting for seats within parliaments or legislatures. Table 5 shows how the observations within the sample are classified according to this typology.

TABLE 5—TYPOLOGY OF POLITICAL INSTITUTIONS

	Proportional	Plurality	Mixed	No Data	Total
Presidential	**21**	**39**	**41**	**0**	**101**
Autocracy (Polity IV score -10 to -6)	0	4	5	0	9
Anocracy (Polity IV score -5 to +5)	8	34	23	0	65
Democracy (Polity IV score =6 to +10)	11	1	12	0	24
No data	2	0	1	0	3
Other Systems a/	**12**	**25**	**9**	**0**	**46**
Autocracy (Polity IV score -10 to -6)	0	0	0	0	0
Anocracy (Polity IV score -5 to +5)	0	3	4	0	7
Democracy (Polity IV score =6 to +10)	12	22	5	0	39
No Data	**0**	**0**	**0**	**12**	**12**
Total	**33**	**64**	**50**	**12**	**159**
Autocracy (Polity IV score -10 to -6)	0	4	5	0	9
Anocracy (Polity IV score -5 to +5)	8	37	27	0	72
Democracy (Polity IV score =6 to +10)	23	23	17	0	63
No data	2	0	1	12	15

a. Non-presidential systems include parliamentary and mixed systems.
Sources: Polity IV database (Marshall, Jaggers, and Gurr) and the Database of Political Institutions (Beck).

[25] Based on Polity2 scores from the Polity IV database provided by Marshall, Jaggers, and Gurr (2010) at http://www.systemicpeace.org/polity/polity4.htm. These scores range from -10 for a hereditary monarchy (pure autocracy) to +10 for a consolidated democracy. Anocracies are those countries with scores in the middle. The empirical tests use the full 21 point scale.

H04: Democracy versus Autocracy. Several authors have written on the impact of democracy and autocracy on economic outcomes such as real growth or the composition of spending.[26] Tornell and Lane (1999) argue that the pro-cyclical voracity effect should be reduced in democracies, to the extent that entrenched interests become less powerful and power becomes more diffuse. Lake and Baum (2001) argue that the political competition[27] inherent in democracies tends to reduce opportunities for collecting rents at the expense of providing public goods. They interpreted a positive correlation between democracy and the quality of health and education outcomes as evidence in support of their hypothesis. Tonizzo (2008) found that increasing the degree of political competition tends to reduce government size, measured as a share of GDP, while fiscal policies tend to become more redistributive.

While these authors do not address budgetary deviations directly, it is possible to build on their ideas. In particular, autocracies might be more likely to over-estimate revenues or over-spend, since they can tolerate the inflationary consequences of domestic borrowing more easily than a democracy could.[28] The higher degree of accountability in democracies should motivate better budget execution for both revenues and expenditures. Moreover, following Tornell and Lane (*ibid*), democracies should exhibit less pro-cyclical behavior.

There is, however, a tension between the impact of political competition and the impact of enlarging the share of voters (starting from zero in a dictatorship). Meltzer and Richard (1981) argue that increasing the number of voters will typically mean that more poor people become empowered, simply because the poor almost always outnumber the rich. They will then demand more services, thus driving up the level of government spending. Mueller and Stratmann (2003) and Fabrizio and Mody (2006) find empirical evidence that public spending increases with voter participation. If a substantial portion of the increased demand is for targeted services, then the common-pool problem should be exacerbated by an enlarged franchise.[29]

This set of hypotheses can be tested using the variable *Polity* from Marshall, Jaggers, and Gurr (2010). The variable *Polity* ranges from -10 for autocracy to +10 for full democracy. If the competitive effect dominates the franchise effect, then high *Polity* scores should reduce CP&A tendencies in revenues and expenditures and promote stability in expenditures despite unanticipated revenue shocks. Alternatively, if the franchise effect dominates, then high polity scores should have the opposite effect, effectively worsening common pool problems.

H05: Mature Programmatic Political Parties. Cruz and Keefer (2010) and Keefer (2011) offer an important counter-point: democracy *per se* may not be as important to good economic development outcomes as the nature of the political parties operating in a country. Those that are programmatic (e.g. those with an identifiable economic agenda) are more likely to get good results, possibly in non-democratic countries as well. In essence, programmatic parties have solved the

[26] See for example and McGuire and Olson (1996) or Deacon (2003).

[27] Political competition in this context refers to competition between parties and candidates rather than between voters and is also distinct from the concept of voter turn-out.

[28] On the other hand, one could argue that the choices of autocrats will depend on their particular policy preferences – and their preferences might be independent of their status as autocrats. For example, the late Chilean dictator, Augusto Pinochet has been portrayed as an autocrat dedicated to fiscal sustainability.

[29] A more direct test would be to use a variable that measures voter participation directly. An interesting and somewhat related test would be to introduce an interaction term to modify the impact of democracy by the impact of differing degrees of spatial segregation. Alesina and Zhuravskaya (2008) find that more segregated societies have a lower quality of government.

collective action problem of managing self-interested politicians and bureaucrats who might otherwise succumb to CP&A temptations.[30]

Keefer provides two measures for the quality of political parties: one is the fraction of parties that are programmatic and the other is the number of years the ruling party has existed less the number of years the current executive has been in office. The latter captures the fact that many countries are ruled by parties that are more oriented around a single personality than an agenda. Thus, a ruling party that was started by one man who has been the executive for the duration of the party would receive a score of zero.[31] By contrast a decades old mature party with leaders holding office for no more than 8 years at a time would receive a very high score. The highest score in the sample is 118 years for Norway. The set of hypotheses tested here are that mature programmatic ruling parties, denoted here by "*MPP*", would be more likely to minimize aggregate and compositional deviations directly and through counter-cyclical behavior.

H06: Presidential systems versus parliamentary systems. Table 5 shows that presidential systems found in 101 of the 159 country observations in the sample. Persson, Roland, and Tabellini (1997) argue that the presidential system with its separation of powers is likely to display smaller government and less corruption. In particular, the concentration of power found in parliamentary systems tempts abuse and is thus empirically correlated with higher degree of corruption. Cheibub (2006) also finds favor with presidential systems, providing evidence that they generate smaller budget deficits than parliamentary systems because voters know it is the executive branch that is responsible for budget execution and can vote out the president (or his/her party in congress) if they are not satisfied. He also finds evidence that the relatively weaker performance of parliamentary systems cannot be traced to the costs of maintaining coalition governments as seen in some countries.[32]

These ideas can be applied to budget execution. If presidential systems are more likely to have smaller budget deficits and are less likely to be corrupt, then perhaps they would also display a comparatively better quality of budget execution directly and through counter-cyclical financing. To test these hypotheses, a dummy (denoted as "*System*") can be employed, based on information taken from the World Bank Database of Political Institutions.[33] The dummy is set to one for presidential systems, so that a zero signifies parliamentary systems and assembly-led electoral systems.[34]

Gerring, Thacker, and Moreno (2005) provide a counter-argument. They maintain that parliamentary systems should exhibit a better quality of governance than presidential systems because political control is more concentrated, and a wider range of views (within the dominant party or coalition) can be reconciled with lower transaction costs.

---

[30] For this reason, hypotheses H04 and H05 for democracy and programmatic parties could also correctly be assigned to the common-pool hypotheses H01 – H03 in the sense that they describe how the common-pool tendency is ameliorated.

[31] The definition of the variable departs from Keefer (2011) in one small way: the number of years the party *affiliated with the chief executive* has existed is used rather than the number of years that the *largest* political party has existed. In most cases, this leads to the same results. Countries also scored as zero include those without political parties, with executives who are not affiliated with political parties, or with executives who have been in power longer than the party they are affiliated with.

[32] Roubini and Sachs (1989) for example, found a tendency towards larger deficits in governments characterized by short average tenures and coalition governments with a large number of political parties.

[33] See Beck *et al* (2001).

[34] Note: Since "*System*" is a dummy variable, *Z*(*System*) is the equivalent of 1–*System*.

H07: Plurality versus proportional elections. Table 5 shows there are 33 observations in the sample with proportional elections for legislative or parliamentary seats. Proportional voting rules assign seats to candidates in proportion to the votes they earned, so that very few if any votes are lost. In these systems, broad social programs become useful to political parties and party coalitions because they need to persuade at least half of all the voters in order to gain dominance. By contrast, in plurality elections, multiple candidates may contest a district seat but only one person, who gains the most votes, will win. This implies that a party in a plurality system needs to convince only half of the voters in half of the districts in order to dominate parliament. Votes for losing candidates are thus essentially lost. Lizzeri and Persico (2001) argue that politicians in plurality systems have incentives to cater to the preferences of those who can help them to get enough votes to win. They will attempt to do so by promising targeted spending to citizens most likely to vote for them. The tactic of targeted spending has been explicitly linked to the common-pool problem. Thus, it seems reasonable to formulate a set of hypotheses that proportional voting reduces the common-pool problem.[35] The need to maintain credibility with a broad range of voters should also motivate more counter-cyclical financing in order to protect spending targets. Proportional voting data are taken from the Database of Political Institutions dummy variable *PR*.[36]

Counter-arguments favoring pluralities exist. As Duverger (1954) observed, most plurality systems tend to gravitate towards just two parties while proportional systems tend to have several parties. Proportional systems could therefore require higher spending on side payments and patronage to hold coalitions together. Persson and Tabellini (2000) found empirical evidence that countries with plurality elections have smaller expenditures and less corruption.

3. Technical Capacity for Public Financial Management

As noted in the introduction to this paper, appropriate techniques of public financial management are needed if a government is to successfully pursue its objectives within prevailing political and economic realities. Von Hagen (2007) suggests that CP&A problems can be reduced by introducing budgetary rules and procedures backed by centralized mechanisms that impose costs on politicians who do not display adequate restraint. Premchand (1999) offers a good overview of the technical side of how to do these things. It turns out that public financial management is a highly technical field of study and practice.[37] Topics include the roles of the executive and the legislature, policy formulation, budget formulation, budget structure options, internal control options, payment systems, accounting, financial reporting, and the importance of external audit to accountability.

The PEFA framework provides a means of exploring many of these topics empirically.[38] The framework provides 28 indicators of public financial management distributed across 6 key dimensions. These include 1) budgetary credibility; 2) comprehensiveness and transparency; 3) policy-based budgeting; 4) predictability and control in budget execution; 5) accounting, recording and reporting; and 6) external scrutiny and audit. Each indicator seeks to measure performance of a key PFM element against well-defined criteria using a four point ordinal scale from A to D.

[35] As Duverger (1954) observed, most plurality systems tend to gravitate towards just two parties while proportional systems tend to have several parties. Proportional systems could therefore require higher spending on side payments and patronage to hold coalitions together. Persson and Tabellini (2000) found empirical evidence that countries with plurality elections have smaller expenditures and less corruption.

[36] Note: Since "*PR*" is a dummy variable, *Z*(*PR*) is the equivalent of 1–*PR*.

[37] Books such as Premchand (1984) or Shah (2007) offer deeper guidance.

[38] Dabla-Norris et al (2010) offer a competing index. The author did not have access to this data at the time of writing.

PFM capacity in each of these key dimensions is critical to good budget execution. Producing a comprehensive budget reduces the risk that public spending outside the budget could drain resources from the approved budget. Ensuring budgetary transparency makes the common pool problem and the agency problem less likely by increasing the degree of accountability felt by public officials. Policy-based budgeting helps ensure that public expenditures are well aligned with public goals, thus increasing the efficiency of spending. When goals are clearly identified, however, there is also more felt accountability to reach those goals, particularly when budgets are made available for scrutiny by the public, by parliamentarians, and by external audit agencies. Such accountability must surely motivate a better quality of budgetary execution. There are technical aspects as well. The capacity for minimizing budgetary deviations is stronger when governments are able to predict well what revenues and net financing are expected, when they can limit spending to what was approved, and when accurate accounting reports throughout each year are made available to track progress.

***H08: Index of PFM Capacity*.** With the preceding logic in mind, an index of PFM capacity was constructed from the average of the 24 PEFA indicators in dimensions 2 through 6. The scores D to A are converted to values 1 through 4. The indicators from the first dimension are excluded since the raw data underlying these indicators is already employed in r, e and f_b. Each indicator is weighted equally: to-date there have been no studies that suggest such an approach is not appropriate.

Not all country reports have scores assigned to each of the 24 indicators. As recorded by the PEFA Secretariat (2009), there are several reasons why a score may not be assigned. It is assumed here that the missing score(s) should take the same value as the average of all the other scores in the case of insufficient information or when the scope of the assessment was limited so that the indicator was not assessed.[39] When the indicator is not applicable in the specific country context, then the index is also the average of all the other scores. Within the sample, the index ranges from a low of 1.6 to a high of 3.6 with an average of 2.6 and a standard deviation of 0.4.

It should be noted that the PEFA country reports do not provide annual scores. The index is therefore assigned for each of the three annual observations within each PEFA country report even though they are generally most pertinent to the last year in the report.

The set of specific hypotheses advanced here are that a higher PEFA score should reduce CP&A problems, produce more accurate revenue targets despite any volatility, insulate expenditures from unanticipated revenue shocks via counter-cyclical financing and improve compositional outcomes.

H09: Population Size. de Renzio (2009) proposed that there could be economies of scale in operating budgetary control systems in larger countries. This would imply that CP&A problems should be easier to contain in countries with larger populations and, in addition, compositional deviations should also be lower in high population countries. Economies of scale in budget management systems might also make it easier to implement counter-cyclical financing when required and improve compositional outcomes. Alternatively, large populations could exacerbate common pool problems if they are more likely to hold within them many disparate special interest groups for politicians to cater to. The unit of measure is the natural logarithm of population size.

H10: Income per Capita. de Renzio (2007) conjectured that governments in wealthier countries can pay for better talent and better systems of control than other governments. If so, then one

[39] A more sophisticated approach would be to use imputation by chained equations.

could argue for a set of hypotheses that CP&A problems should be reduced by income per capita and that access to better talent and systems might also make it easier to implement counter-cyclical financing and accurate compositional outcomes. The unit of measure is the natural logarithm of gross national income per capita, denominated in purchasing power parity adjusted US dollars.

TABLE 6—SUMMARY OF HYPOTHESES

	Equation 9 **Aggregate Deviation, *r***	**Equation 10** **Aggregate Deviation, *e***	**Equation 11** **Compositional Accuracy, f_b**
r	**1**	..	..
e	..	**1**	$-\phi_1 \cdot \lvert e \rvert$
f_b	..	..	**1**
Constants	$+ \rho_0$	$+ \varepsilon_0$	$+ \phi_0$
Rp	$+ \rho_1 \cdot Rp$	..	..
Rv	$+ \rho_2 \cdot Rv$	..	..
PCT_INT	$- \rho_3 \cdot PCT\_INT$	$+ \varepsilon_1 \cdot PCT\_INT$	$- \phi_2 \cdot PCT\_INT$
PCT_INT	$+ \rho_4 \cdot Rp \cdot Z(PCT\_INT)$	$+ \varepsilon_2 \cdot r_{pos} \cdot Z(PCT\_INT)$	..
Common Pool Variables, *Y*			
H01 FCTN_ETHN	$+ \rho_5 \cdot FCTN\_ETHN$	$+ \varepsilon_3 \cdot FCTN\_ETHN$	$- \phi_3 \cdot FCTN\_ETHN$
H01 FCTN_LANG	$+ \rho_6 \cdot FCTN\_LANG$	$+ \varepsilon_4 \cdot FCTN\_LANG$	$- \phi_4 \cdot FCTN\_LANG$
H01 FCTN_RELIG	$+ \rho_7 \cdot FCTN\_RELIG$	$+ \varepsilon_5 \cdot FCTN\_RELIG$	$- \phi_5 \cdot FCTN\_RELIG$
H02 DRY 1/	$+ \rho_8 \cdot DRY$	$+ \varepsilon_6 \cdot DRY$	$- \phi_6 \cdot DRY$
H02 DRO	$+ \rho_9 \cdot DRO$	$+ \varepsilon_7 \cdot DRO$	$- \phi_7 \cdot DRO$
H03 PCT_CSE 1/	$+ \rho_{10} \cdot \%CSE$	$+ \varepsilon_8 \cdot \%CSE$	$- \phi_8 \cdot \%CSE$
Political Institutions, X			
H04 POLITY 2/	$- \rho_{11} \cdot POLITY$	$- \varepsilon_9 \cdot POLITY$	$+ \phi_9 \cdot POLITY$
H05 MPP	$- \rho_{12} \cdot MPP$	$- \varepsilon_{10} \cdot MPP$	$+ \phi_{10} \cdot MPP$
H06 SYSTEM	$- \rho_{13} \cdot SYSTEM$	$- \varepsilon_{11} \cdot SYSTEM$	$+ \phi_{11} \cdot SYSTEM$
H07 PR	$- \rho_{14} \cdot PR$	$- \varepsilon_{12} \cdot PR$	$+ \phi_{12} \cdot PR$
Capacity for PFM, X			
H08 PEFA	$- \rho_{15} \cdot PEFA$	$- \varepsilon_{13} \cdot PEFA$	$+ \phi_{13} \cdot PEFA$
H09 ln(POP) 2/	$- \rho_{16} \cdot POP$	$- \varepsilon_{14} \cdot POP$	$+ \phi_{14} \cdot POP$
H10 ln(GNIPC)	$- \rho_{17} \cdot GNIPC$	$- \varepsilon_{15} \cdot GNIPC$	$+ \phi_{15} \cdot GNIPC$
H11 ODAGDP	$- \rho_{18} \cdot ODAGDP$	$- \varepsilon_{16} \cdot ODAGDP$	$+ \phi_{16} \cdot ODAGDP$
Interactive Variables			
H03 PCT_CSE 1/	$+ \rho_{19} \cdot Rp \cdot Z(PCT\_CSE)$	$+ \varepsilon_{17} \cdot r \cdot Z(PCT\_CSE)$	..
H04 POLITY 2/	$+ \rho_{20} \cdot Rp \cdot Z(POLITY)$	$+ \varepsilon_{18} \cdot r \cdot Z(POLITY)$	..
H05 MPP	$+ \rho_{21} \cdot Rp \cdot Z(MPP)$	$+ \varepsilon_{19} \cdot r \cdot Z(MPP)$	..
H06 SYSTEM	$+ \rho_{22} \cdot Rp \cdot Z(SYSTEM)$	$+ \varepsilon_{20} \cdot r \cdot Z(SYSTEM)$	..
H07 PR	$+ \rho_{23} \cdot Rp \cdot Z(PR)$	$+ \varepsilon_{21} \cdot r \cdot Z(PR)$	..
H08 PEFA	$+ \rho_{24} \cdot Rp \cdot Z(PEFA)$	$+ \varepsilon_{22} \cdot r \cdot Z(PEFA)$	..
H09 ln(POP)	$+ \rho_{25} \cdot Rp \cdot Z(POP)$	$+ \varepsilon_{23} \cdot r \cdot Z(POP)$	..
H10 ln(GNIPC)	$+ \rho_{26} \cdot Rp \cdot Z(GNIPC)$	$+ \varepsilon_{24} \cdot r \cdot Z(GNIPC)$	..
H11 ODAGDP	$+ \rho_{27} \cdot Rp \cdot Z(ODAGDP)$	$+ \varepsilon_{25} \cdot r \cdot Z(ODAGDP)$	..
H12 BATTLE	$+ \rho_{28} \cdot BATTLE$	$+ \varepsilon_{26} \cdot BATTLE$	$- \phi_{17} \cdot BATTLE$

1. If this variable takes the opposite sign from what is predicted, then it may be more relevant to PFM capacity.
2. If this variable takes the opposite sign from what is predicted, then it may be more relevant as a common pool variable.

H11: External assistance. Many governments are supported financially by external donors for a wide variety of purposes. The impact could be positive in some ways, increasing the pace of

development and improving the capacity for PFM. de Renzio *et al* (2011) found a positive but small correlation between PEFA scores and external technical assistance for building PFM capacity. Aid flows might also dampen some aspects of PFM capacity. For example, Mocan and Azad (1995) find a slight decrease in revenue forecast accuracy by US states when the share of federal grant increases. Some authors, such as Heller (1975), have observed that higher levels of external financial assistance are correlated with lower revenue efforts. This may be because government expectations that low revenue efforts will be compensated by aid are often met. Perhaps similar logic obtains for within-year behavior: aid dependent governments might be less likely to prioritize revenue capacity building and more likely to tolerate unanticipated revenue shortfalls when they believe external aid will help fill the gap. The hypothesis tested here is that revenue deviations will be negatively correlated with official development assistance as a share of GDP (*ODAGDP*) while expenditure deviations are also negatively correlated with *ODAGDP* and compositional accuracy, f_b, is positively correlated with *ODAGDP*.

Conflict. As an additional control, unrelated to any of the three categories above, it may be useful to include a variable measuring the impact of conflict. Several countries in the data sample suffered from substantial conflict during the periods under measurement. de Renzio *et al* (2011) found that countries in conflict and post-conflict situations tended to have lower than average PEFA scores. Since battle outcomes and battle needs are often difficult to predict, it seems reasonable to propose that conflict could increase aggregate deviations in *r* and *e* and thus reduce compositional accuracy f_b as well.[40] To test for any impact, a measure of battle deaths as a share of the population is constructed.

V. Data and Econometric Strategy

The empirical tests below will focus on three dependent variables: deviations in total revenues *r*, deviations in total spending *e*, and the accurately spent fraction of the budget, f_b. The hypotheses advanced in section IV require 18 potential explanatory variables. More information about each can be found in Annex 1 while the statistical properties of the independent and dependent variables are summarized in Annex 2. Not all of the 159 observations for compositional deviations were accompanied by data for each of the explanatory variables proposed above. Some variables had missing observations for Kosovo, Montenegro, Serbia, and Timor-Leste. This reduced the sample to 144 observations with full data, representing 41 countries of which 7 are repeated twice. All 159 observations for each of the independent variables are provided in Annex 3 for other researchers to use.

Model selection is the main econometric challenge in this paper. The independent variables can enter the regression equations independently or in interaction with revenue policy or revenue deviations. In the case of equation 9 for aggregate revenue deviations, there are 28 potential correlates including interaction terms. This generates 2^{28}-1 possible models to choose from, approximately 268 million! Within these, there may be many models where the variables of interest take the expected signs and are significantly different from zero. Unfortunately, due to the frequent collinearity between variables, standard errors will be larger and some variables may be incorrectly rejected. The econometrician will also often observe that a regressor that is positive in one model becomes negative or insignificant in other models. One could also easily miss the fact that an estimated coefficient and its variance are not stable over the majority of models. Worse, one may succumb to temptation to report only the favorable models.

[40] An alternative hypothesis is that the initiation and end of battle are uncertain and generate spending shocks.

In order to expeditiously deal with these challenges, a technique referred to as Bayesian Averaging of Classical Estimates (BACE) is employed. As explained by Sala-i-Martin (1997) and Doppelhofer et al. (2000), rather than assuming there is a single correct model, as in a simplification search, they assign weights corresponding to a "level of confidence" for each model. This then allows them to calculate a weighted OLS average for each of the potential variables. The only prior assumption required is the expected model size, e.g. out of K variables proposed, perhaps only k would be expected to be significant.[41]

The main statistics generated by BACE for each regressor include the posterior inclusion probability (PIP), the posterior mean coefficient, and the posterior mean standard deviation. The latter two are calculated unconditionally and conditional on inclusion. The PIP indicates the probability that a regressor drawn from set K should be included in a robust model of size k. The posterior means are the weighted averages of all the coefficients and their standard deviations for all models. The posterior means that are conditional upon inclusion use weights based on the assumption that each regressor *must* be included in any model of size k. For this purpose, the weights are revised to sum to one only for those models containing the regressor. The ratio of the posterior mean coefficients to the posterior mean standard deviations is analogous to the t-statistic. Doppelhofer *et al.* (2000) caution that this ratio is *not* the same as a t-statistic but they go on to suggest that it can be used in much the same way. A ratio with an absolute value of 2 or higher indicates that approximately 95% of the probability density associated with the posterior mean for a regression coefficient will exclude zero. In addition, in the language of Doppelhofer *et al.* (2000), a variable is referred to as being "robust" when PIP > k/K *and* it is well estimated so that the absolute value of the mixed t-ratio is greater than two.

The author's implementation of BACE will allow testing of up to 19 variables without exceeding the memory and capacity of a standard laptop computer.[42] This is not a problem for the regression equations for f_b which requires only 16 variables. The regression equation for aggregate deviations e requires 26 variables while the regression for r requires 28 variables. For these two cases, a stepwise procedure will be employed to eliminate the least significant variables before the remaining 19 variables are tested using the BACE procedure. The variable with the highest probability of being rejected as insignificant is removed. The variable with the next highest p-value, given the removal of the first variable, is also removed. Next both of the removed variables are checked for re-inclusion into the model: any variable whose p-value is lower than 5 percent is added back in to the model. Once the adding back step has been performed, the next variable is removed. This process is repeated until 19 variables remain. In order to cast the net fairly wide, the expected model size for r and e is set at 16 out of 19 variables. The expected model size for f_b is set at 14 out of 17 variables.

It should be noted that the procedure above rests on the assumption that the error term in a regression equation of interest is normally distributed. This is a fundamental requirement for hypothesis testing. Therefore, before any variables are eliminated through the procedures above, the distribution of the regression residual is examined using the Jarque-Bera test. If the null hypothesis of a normal distribution is rejected, then dummy variables may be added to obtain a normally distributed error term. The dummy variables will be selected from the largest positive and negative residuals in the error term. The dummy variables can be retained as useful explanatory variables if they can be linked to relevant circumstances.

[41] Magnus *et al* (2010) propose an alternative technique that is less computationally demanding. The author did not have access to software using this methodology.
[42] Using EVIEWS software.

Following the procedure outlined above, a preferred OLS equation for each independent variable will be constructed on the basis of the robust BASE variables. This final equation will be subjected to common tests for the distribution of the error term, heteroskedasticity in the error term, model mis-specification, and influential points.

VI. Test Results and Interpretations

A. Aggregate Revenue Deviations

The results of the analysis for aggregate revenue deviations are shown in Table 7. Column 1 shows the hypothesized signs for each independent variable in equation 9. Columns 2 and 3 show the results for an encompassing equation 9a that includes 28 hypothesized variables. Prior to testing, several dummy variables were added to equation 9a in order to obtain a normally distributed error term. The dummy variables were selected from the largest positive and negative residuals in the error term. Further investigation suggests that these should be retained as explanatory variables because each represents an unusual set of circumstances. These are summarized below:

- **Georgia 2004:** The government had implemented tax policy reform that far exceeded expectations. According to the PEFA report for Georgia, the authorities had set cautious targets until they knew reform would work. They expected a gain of 1-2 percent of GDP but instead gained 4.1 percentage points, from 9.4 percent of GDP in 2003 to 13.5 percent of in 2004. By contrast, 2005 and 2006 saw adverse shocks from drought and floods and energy price shock.
- **Ghana 2006:** There is a discontinuity in the nominal GDP data from WDI. The authorities revised the national accounts from 2006 onwards and the WDI reflect this but the data prior to 2006 were not adjusted to compensate.
- **Ghana 2008:** The combined impact of tax administration reforms and unanticipated rapid growth in the tax base led to a revenue surplus. According the PEFA report for Ghana, there were "significant improvements made with respect to taxpayers' access to information on tax liabilities and administrative procedures and on the collection of tax arrears." Nominal GDP growth accelerated from 20 percent per annum in 2006-07 to 30 percent in 2008.
- **Peru 2006:** According to the PEFA report for Peru, a substantial portion of tax revenues is derived from commodity exports, particularly from copper. Copper prices increased by 83 percent in 2006.
- **Solomon Islands 2004:** In July 2003, an Australian-led peace-keeping force landed in the Solomon Islands, putting an end to several years of conflict and economic mis-management including a large build-up of payments in arrears. The Regional Assistance Mission to Solomon Islands (RAMSI) provided substantial technical assistance to the Ministry of Finance and Treasury. With their help, actual revenues in 2004 were 73 percent higher than budgeted revenues due to a combination of deliberate under-targeting in order to create a surplus to pay down the stock of arrears and a failure to anticipate how rapidly the economy would rebound following the end of conflict.

Columns 4 and 5 show the results of the Bayesian averaging after selecting the 19 most significant variables from equation 9a. Columns 6 and 7 show the author's initial preferred equation 9b, based on the Bayesian outcomes. Four variables are indicated for inclusion in equation 9b on the basis of their PIP scores. The variables PCT_INT, DRY, *rp·Z(SYSTEM)* and *rp·Z(GNIPC)* have PIP scores above

the 16/19 threshold. Of these, only *rp·Z(SYSTEM)* is robustly significant (PIP > 10/19 *and* |unconditional mixed t-statistics| > 2).

As an additional test of robustness, all outlier observations were investigated to see if any acted as influential points. These are shown in Table 8. One of these, Jamaica 2006/06, caused *PCT_INT* to become statistically insignificant at the 10 percent level. All other variables remained significant with or without the other influential points.

TABLE 7: AGGREGATE REVENUE DEVIATIONS, *r*

	Hypothesized Signs	Encompassing Equation 9a		Bayesian k/K = 16/19		Preliminary Equation 9b	
		Coeff.	p-value	PIP	Mixed-t	Coeff.	p-value
	1	2	3	4	5	6	7
Constant		-2.819	0.894	..	..	14.974	**0.000**
Rp	+	146.725	**0.081**	0.569	0.721	..	**..**
Rv	+	-1.438	0.882	..	..	..	**..**
PCT_INT	–	-0.186	**0.030**	**0.855**	-1.547	-0.121	**0.063**
Rp·Z(PCT_INT)	+	-34.324	0.513	..	..	..	**..**
Common Pool							
H01 FACTION_ETHN	+	0.068	0.162	0.551	0.718	..	**..**
H01 FACTION_LANG	+	-0.029	0.492	0.345	-0.143	..	**..**
H02 FACTION_RELIG	+	0.028	0.446	0.339	0.102	..	**..**
H02 DRO	+	0.525	**0.059**	0.705	1.050	..	**..**
H02 DRY	+	-0.161	**0.016**	**0.864**	-1.580	-0.141	**0.000**
H03 PCT_CSE	+	0.135	**0.054**	0.558	0.738	..	**..**
Political Institutions		..	..	..	..	..	**..**
H04 POLITY	–	-0.198	0.369	0.484	-0.578	..	**..**
H05 MPP	–	0.094	**0.088**	0.559	0.717	..	**..**
H06 SYSTEM	–	0.051	0.982	..	..	..	**..**
H07 PR	–	-3.430	0.119	0.422	-0.452	..	**..**
Capacity for PFM							
H08 PEFA	–	2.386	0.332	0.555	0.738	..	**..**
H09 ln(POP)	–	0.796	0.182	0.693	1.021	..	**..**
H10 ln(GNIPC)	–	-1.522	0.436	0.404	0.023	..	**..**
H11 ODAGDP	–	19.252	**0.097**	0.693	1.017	..	**..**
Interactive Terms							
H03 Rp·Z(PCT_CSE)	+	-30.461	0.438	..	..	..	**..**
H04 Rp·Z(POLITY)	+	25.598	0.511	..	..	..	**..**
H05 Rp·Z(MPP)	+	-125.702	0.115	0.526	-0.635	..	**..**
H06 Rp·Z(SYSTEM)	+	37.413	**0.021**	**0.968**	**2.466**	41.230	**0.000**
H07 Rp·Z(PR)	+	-1.242	0.954	..	..	..	**..**
H08 Rp·Z(PEFA)	+	-10.548	0.831	..	..	..	**..**
H09 Rp·Z(lnPOP)	+	-29.473	0.513	0.421	-0.418	..	**..**
H10 Rp·Z(lnGNIPC)	+	134.577	**0.025**	**0.879**	1.673	85.990	**0.000**
BATTLE	+	0.071	0.867	..	..	..	..
Georgia 2004		14.717	0.129			22.632	**0.004**
Ghana 2006		-66.382	**0.000**			-50.867	**0.000**
Ghana 2008		21.522	**0.006**			23.056	**0.002**
Peru 2007		37.628	**0.000**			33.171	**0.000**
Solomon Islands 2004		50.697	**0.000**			45.449	**0.000**
Observations		144				144	
R-squared		0.778				0.725	
Adjusted R-squared		0.713				0.707	
S.E. of regression		7.342				7.429	
Sum squared resid		5,984				7,394.733	
Log likelihood		-473				-487.914	
F-statistic		12.123	**0.000**			39.264	**0.000**
Normal error term 2/		0.841				0.556	
Heteroskedasticity 3/		0.307				0.654	
Omitted Variables 4/		0.146				0.480	

1. Unconditional inclusion. 2. Jarque-Bera probability.
3. White $n \cdot R^2$ probability (without cross terms). 4. Ramsey RESET F-statistic probability (1 term).

TABLE 8: INFLUENTIAL POINTS FOR AGGREGATE REVENUE DEVIATIONS

	Outliers 1/	Influential Points 2/
PCT_INT	Afghanistan 2004/05 and 2006/07, Jamaica 2003/04 and 2005/06, Sierra Leone 2004	Jamaica 2005/06
H02 DRY	None	None
H06 Rp · Z(SYSTEM)	Solomon Island 2004, Tonga 2008/09, Yemen 2005 and 2006	None
H10 Rp · Z(GNIPC)	Afghanistan 2004/05, Ghana 2006, Maldives 2005, Yemen 2005 and 2006	None

1. Observations outside of the mean plus or minus 2 standard deviations or, for highly skewed distributions, the most extreme observations.
2. Single observations that, when excluded, make a coefficient insignificant at the 10 percent level.

Before concluding that equation 9b is suitable, three more diagnostic tests are made. The Jarque-Bera test fails to reject the null hypothesis of a normally distributed error term. The null hypothesis of no heteroskedasticity is rejected by White's test. In order to facilitate hypothesis testing, White's heteroskedasticity-consistent standard errors & covariance are employed. The Ramsey RESET test can detect mis-specification from incorrect functional form, simultaneity, and omitted variables. It fails to reject the null of no mis-specification using squared fitted values of the dependent variable. The adjusted r-squared is 0.707 with the dummies and 0.503 without them. Even though these results are very encouraging, some caution must be maintained since it is already known from the discussion in Section IV above that there could be some missing explanatory variables such as unexpected shocks to relative prices. The bottom line, however, is that equation 9b holds up well on its econometric merits.

While many of the proposed correlates for aggregate revenue deviations were rejected, those that were not tell an interesting story. The constant is highly significant and positive, creating a bias of 15 percentage points towards revenue surplus – consistent with Figure 2, Section III. As expected, revenue deviations are highly correlated with the degree of ambition in revenue policy: unanticipated surpluses are most common when revenue targets are not ambitious. The two interactive terms modify this tendency: governments are more likely to hit their revenue targets when they have presidential systems and when PFM capacity, measured by *ln*(GNIPC), is high. As expected, debt service takes a negative sign: under-performing revenue is one of the causes of indebtedness.

The results indicate a high youth dependency ratio pushes revenue deviations in a negative direction, contrary to H02. It may be, as conjectured in Section IV, that *DRY* is instead an indicator of the size of the talent pool a government may draw upon.[43] Populations with a large volume of youth below age 15 will hold fewer educated, experienced adults who can serve productively as civil servants. In other words, *DRY* might better serve as an indicator of capacity for PFM.

B. Aggregate Expenditure Deviations

The results of the analysis for aggregate expenditure deviations are shown in Table 9. Column 1 shows the hypothesized signs for each independent variable in equation 10. Columns 2 and 3 show the results for an encompassing equation 10a that includes all 26 hypothesized variables. Columns 4 and 5 show the results of the Bayesian averaging after removing the 7 least significant variables. Columns 6 and 7 show the author's preliminary results (equation 10b), based on the Bayesian

[43] This could be difficult to test more deeply: educational attainment is highly correlated not only with dependency ratios (young and old) but also with income per capita and PEFA scores, with a high probability of two-way causality between some or all of these variables.

outcomes. Nine variables are indicated for inclusion on the basis of their PIP scores. Of these, three are robustly significant (PIP > 16/19 *and* |unconditional mixed t-statistics| > 2).

TABLE 9: AGGREGATE EXPENDITURE DEVIATION, *e*

	Hypothesized Signs	Encompassing Equation 10a		Bayesian Outcomes k/K=16/19		Preliminary Equation 10b	
		Coeff.	Prob. 2/	PIP	Mixed-t 1/	Coeff.	Prob.
	1	2	3	4	5	6	7
Constant	0	-13.193	0.509	..	..	-1.237	0.877
PCT_INT	+	0.211	**0.040**	**0.918**	1.907	0.119	0.139
*Z_PCT_INT*Rpos*	+	0.595	**0.032**	**0.908**	1.825	0.389	**0.040**
Common Pool							
H01 FCTN_ETHN	+	-0.074	0.127	0.614	-0.841	..	**..**
H01 FCTN_LANG	+	0.008	0.858	..	..	..	**..**
H01 FCTN_RELIG	+	0.060	0.134	0.579	0.776	..	**..**
H02 DRO	+	0.495	**0.043**	**0.866**	1.590	0.429	**0.055**
H02 DRY	+	-0.184	**0.001**	**0.951**	**-2.249**	-0.154	**0.003**
H03 PCT_CSE	+	-0.045	0.604	0.340	-0.203	..	**..**
Political Institutions							
H04 POLITY	−	0.634	**0.060**	**0.924**	1.868	0.276	**0.032**
H05 MPP	−	-0.026	0.739	..	..	..	..
H06 SYSTEM	−	4.639	0.150	0.646	0.902	..	..
H07 PR	−	-3.549	0.145	0.776	-1.247	..	..
Capacity for PFM							
H08 PEFA	−	-1.545	0.616	0.510	-0.646	..	..
H09 ln(POP)	−	1.347	**0.053**	**0.940**	**2.032**	0.292	0.461
H10 ln(GNIPC)	−	-0.023	0.990	..	..	..	..
H11 ODAGDP	−	-9.490	0.509	0.483	-0.591	..	..
Interactive Terms							
H03 R·Z(PCT_CSE)	+	-0.808	**0.042**	**0.985**	**-2.816**	-0.525	**0.023**
H04 R·Z(POLITY)	+	0.348	0.415	0.489	0.587	..	**..**
H05 R·Z(MPP)	+	-0.065	0.901	..	..	..	**..**
H06 R·Z(SYSTEM)	+	0.003	0.985	..	..	..	**..**
H07 R·Z(PR)	+	-0.127	0.534	0.350	-0.139	..	**..**
H08 R·Z(PEFA)	+	0.829	0.184	0.763	1.192	..	**..**
H09 R·Z(lnPOP)	+	-0.802	0.107	**0.858**	-1.513	-0.560	**0.011**
H10 R·Z(lnGNIPC)	+	0.571	0.401	**0.860**	1.576	1.117	**0.000**
H11 R·Z(ODAGDP)	+	0.055	0.875	..	..	..	..
BATTLE	+	0.143	0.777	..	..	..	..
Ghana 2008						22.667	**0.005**
India 2007/08						31.405	**0.000**
Pakistan 2007/08						21.691	**0.007**
Sierra Leone 2007						25.172	**0.001**
Observations		144		144		144	
R-squared		0.544				0.572	
Adjusted R-squared		0.443				0.529	
S.E. of regression		8.182				7.523	
Sum squared resid.		7,833				7,358	
Log likelihood		-492				-488	
F-statistic		5.374	**0.000**			13.357	**0.000**
Normal error term 3/		0.750				0.487	
Heteroskedasticity 4/		0.013				0.325	
Omitted variables 5/		0.008				0.246	

1. Unconditional inclusion.
2. White heteroskedasticity-consistent standard errors & covariance.
3. Jarque-Bera probability.
4. White $n \cdot R^2$ probability (without cross terms).
5. Ramsey RESET F-statistic probability (1 term).

Four dummies were added to ensure the error term is normally distributed. Each can be considered a legitimate explanatory variable:[44]

- **Ghana 2008:** The coincidence of an election year and rapidly escalating food and fuel prices created strong pressure to spend above the approved budget. The 2004 elections also created spending pressures but without additional commodity price movements.
- **India 2008/09:** Revenues fell below target due to the global recession while expenditures where sharply expanded within-year for stimulus purposes via a supplemental budget.
- **Pakistan 2007/08:** The budget was adversely affected by a combination of slower than expected real GDP growth, increased subsidies to partially insulate consumers from rising fuel and food prices, and a loss of fiscal discipline during a year of major political tumult.
- **Sierra Leone 2007**: In 2007, donors withheld almost half the expected level of grant financing due to fiduciary concerns. This led to the imposition of daily cash rationing in the short run and helped motivate successful, deeper investments in PFM capacity in the medium term.

As an additional test of robustness, all outlier observations were investigated to see if any acted as influential points. Several variables were affected by moderate changes in their slopes but only one, $Z(PCT\_INT)\cdot r_{pos}$, was made insignificant by two influential points (see Table 10). There are no obvious reasons to remove those two points.

TABLE 10: INFLUENTIAL POINTS FOR AGGREGATE EXPENDITURE DEVIATIONS

	Outliers 1/	Influential Points 2/
$Z(PCT\_INT)\cdot r_{pos}$	Solomon Islands (2004)	Jamaica (2003/04), Yemen (2005)
H02 DRO	Norway (2004 - 2006)	
H02 DRY	Uganda (2005/06 – 2007/08)	
H03 r·Z(PCT_CSE)	Sierra Leone (2007), Solomon Islands (2004)	
H04 POLITY	Bhutan (2005/06 - 2007/08)	
H09 lnPOP	India (2005/06 - 2007/08)	
H09 r·Z(lnPOP)	Solomon Islands (2004)	

1. Observations outside of the mean plus or minus 2 standard deviations or, for highly skewed distributions, the most extreme observations.
2. Observations that, when excluded, make a coefficient insignificant at the 10 percent level.

The final equation is not perfect but it is sufficient to allow simple inference testing of the possibility that the hypothesized correlations in Section IV above are or are not significantly different from zero. The Jarque-Bera test fails to reject the null hypothesis of a normally distributed error term. White's test fails to reject the null hypothesis of no heteroskedasticity. The Ramsey RESET test fails to reject the null of no mis-specification using squared fitted values of the dependent variable. The adjusted r-square is 0.57 with dummies and 0.41 without them. As before, even though this is very encouraging, some caution must be maintained since it is already known from the discussion in Section IV above that there *is* at least one missing explanatory variable: unexpected shocks to external grants and net financing.

The results provide mixed support for the hypotheses advanced in section IV above. The constant is not statistically different from zero in equation 10b. This is borne out by Figure 1 which shows that *e* is distributed symmetrically around zero. The interactive terms define the capacity to

[44] Three of the four dummies overlap with an acceleration of the global price of petroleum an done overlaps with the global depression. The author tried adding variables for petroleum price inflation and OECD growth, on the (admittedly weak) supposition that policy makers expect next year's outcome to be the same as the previous year's outcome. These variables proved to be statistically insignificant with and without the dummies. A better test, if data were available, would use each government's own forecast versus the actual outcome.

maintain expenditure targets through counter-cyclical financing. Four of the hypothesized interactive terms are not rejected. These are Z(*PCT_INT*)·r_{pos}, *Z(PCT_CSE)·r*, *Z(lnPOP)·r*, and *Z(lnGNIPC)·r.*

Expenditure deviations are positively correlated with debt service burdens in the vast majority of all potential models (PIP = 0.876) but not in the specific model depicted by equation 10b. Aggregate expenditure deviations are positively correlated with Z(*PCT_INT*)·r_{pos}, implying governments tend to avoid over-spending when the debt service burden is high.

As expected, aggregate expenditure deviations are positively correlated with Z(*GNIPC*)·*r*. This means that unanticipated revenue deviations and expenditure deviations are positively correlated when the capacity for PFM, measured by *GNIPC*, is small. Conversely, expenditures are more likely to remain on target despite unanticipated revenue shocks when the capacity for PFM is high.

Contrary to expectations, expenditure deviations are negatively correlated with *Z(PCT_CSE)·r* and *Z(lnPOP)·r*. This result is hard to understand at first glance since it appears to convert unanticipated positive revenue shocks into negative expenditure deviations when *PCT_CSE* and *lnPOP* are small. The negative correlation makes only partial sense: low population countries and/or those with a small civil service could lack the capacity to spend even when revenues are in excess – but this should also be true when revenues fall short. In fact, when revenue deviations are confined to shortfalls, the coefficient for *Z(PCT_CSE)·r* becomes statistically insignificant.

When looked at as part of a bigger picture one could conclude that within-period counter-cyclical financing is promoted when the cumulative impacts from *lnGNIPC* and *PCT_INT* just balance those coming from *PCT_CSE* and *lnPOP* so that the cumulative multiplier $\Sigma\varepsilon_i \cdot Z(X)$ is close to zero. Within the sample, the cumulative multiplier ranges from a low of -0.38 to a high of +1.03 with an average of 0.34. Only 19 of the observations fall within ±0.10. Another 8 observations fall between -0.38 and -0.10 and 117 observations fall between +0.10 and +1.03. This suggests that the majority of governments in the sample prefer to spend their unanticipated revenue gains and minimize the need for unprogrammed increases in borrowing when revenues fall short of targets. The overall will or capacity for achieving good budget execution through counter-cyclical financing is quite rare, at least within the sample of countries that is used here, mainly developing countries.

Two variables expected to worsen common pool problems were not rejected but only one took the expected sign: aggregate expenditure deviations are positively correlated with higher old age dependency ratios, but negatively correlated with higher youth dependency ratios. The sign for the youth dependency ratio is opposite of expectations: it may be instead that *DRY* is capturing weak PFM capacity to spend despite unanticipated revenues. Such an outcome is possible, and easily observed in reality, when government procurement units are inexperienced or burdened with unhelpful procedures.

Only one variable expected to improve the capacity to deal with common pool problems, *POLITY*, was not rejected but it did not take the expected negative sign. In other words, *e* is increased by democracy (suggesting the franchise effect dominates the competition effect). Although the variable *lnPOP* is not significant in the model for equation 10b, it is positively correlated with expenditure deviations in the vast majority of potential models (PIP=0.953), suggesting that expenditure deviations are increased by large populations which might include more groups for

politicians to spend on. Both variables *POLITY* and *lnPOP* can therefore be reclassified as contributors to the common pool problem.[45]

TABLE 11—FRACTION OF BUDGET ACCURATELY SPENT, f_b

	Hypothesized Signs	Encompassing Equation 11a		Bayesian Outcomes k/K = 14/17		Preliminary Equation 11b	
		Coeff.	Prob.	PIP	Mixed-t 1/	Coeff.	Prob.
	1	2	3	4	5	6	7
Constant	..	33.136	0.319	..	..	20.1439	**0.007**
Aggregate Deviations, |*e*|	−	-0.717	**0.000**	**1.000**	**-4.355**	-0.7408	**0.000**
PCT_INT	−	0.047	0.733	0.321	0.278	..	..
Common Pool							
H01 FCTN_ETHN	−	-0.306	**0.000**	**1.000**	**-4.560**	-0.2241	**0.000**
H01 FCTN_LANG	−	-0.043	0.535	0.348	-0.293	..	..
H01 FCTN_RELIG	−	0.123	**0.048**	**0.856**	1.562	..	..
H02 DRY	−	0.163	0.151	0.573	0.799	..	..
H02 DRO	−	0.283	0.510	0.335	0.266	..	..
H03 PCT_CSE	−	-0.007	0.951	0.330	-0.312	..	..
Political Institutions							
H04 POLITY	+	-0.715	**0.043**	0.778	-1.278	..	..
H05 MPP	+	0.330	**0.000**	**0.995**	**2.953**	0.1892	**0.000**
H06 SYSTEM	+	-0.382	0.914	0.315	0.197	..	..
H07 PR	+	0.950	0.781	0.357	0.359	..	..
Capacity for PFM							
H08 PEFA	+	11.488	**0.005**	**0.947**	**2.200**	9.2921	**0.001**
H09 ln(POP)	+	0.405	0.689	0.377	0.380	..	..
H10 ln(GNIPC)	+	-3.907	0.179	0.620	-0.876	..	..
H11 ODAGDP	+	-29.687	**0.074**	0.601	-0.843	..	..
ln(BATTLE)	−	0.851	0.301	**0.853**	1.575	..	..
Burundi 2005		40.369	0.007	..	..	49.740	**0.000**
Uganda 2006-07		41.068	0.004	..	..	51.093	**0.000**
Observations		144		144		144	
R-squared		0.537				0.479	
Adjusted R-squared		0.466				0.456	
S.E. of regression		12.990				13.110	
Sum squared resid		20,923				23,547	
Log likelihood		-563				-571	
F-statistic		7.557	**0.000**			20.949	**0.000**
Normal error term 2/		0.600				**0.752**	
Heteroskedasticity Test 3/		0.409				**0.849**	
Omitted Variable Test 4/		0.019				0.001	

1. Unconditional inclusion.
2. Jarque-Bera probability.
3. White $n \cdot R^2$ probability (without cross terms).
4. Ramsey RESET F-statistic probability (1 term).

C. Compositional Expenditure Deviations

The results of the analysis for aggregate expenditure deviations are shown in Table 11. Column 1 shows the hypothesized signs for each independent variable in equation 9. Columns 2 and 3 show the results for an encompassing equation 11a that includes all 17 hypothesized variables. Columns 4 and 5 show the results of the Bayesian averaging. Columns 6 and 7 show the author's preferred equation 11b, based on the Bayesian outcomes. Four variables are indicated for inclusion on the basis of their PIP scores. Each are robustly significant (PIP > 14/17 *and* |unconditional mixed t-statistics| > 2). As an additional test of robustness, all outlier observations were investigated to see if any acted as influential points: none proved to do so. Two of these, however, do skew the

[45] This does not need negate the scale effect found by de Renzio (2007) where PEFA scores are positively correlated with population size. It may be that this remains true but the impact of the common pool problem overwhelms the scale effect.

distribution of the residual term. These were therefore added to the equation as dummy variables. It is not clear whether these can be justified as useful explanatory variables or not: neither the PEFA reports nor IMF reports describe policies or events in the context of compositional expenditure deviations.

The Jarque-Bera test fails to reject the null hypothesis of a normally distributed error term when the two dummy variables are included. The null hypothesis of no heteroskedasticity is not rejected by White's test. The Ramsey RESET test, however, does reject the null of no mis-specification using squared fitted values of the dependent variable.

One possible solution to the RESET test outcome is to substitute a non-linear relationship between f_b and $|e|$ for the hypothesized linear relationship found in 11b. This is motivated by the idea that it might become progressively more difficult to increase the share of ministries absorbing a shock, even though the demand to do so would increase as the strength of the shock increased. This might occur because political and/or administrative frictions increase as the number of ministries involved rises. To test this idea, a number of alternatives were tested: e^2, $|e|^{0.5}$ and $ln|e|$. All were found to be acceptable substitutes in terms of statistical significance but only the latter was found to cure the RESET problem. The result is equation 11c in Table 12.

Table 12—Fraction of Budget Accurately Spent, f_b

	Equation 11b		Equation 11c		Equation 11d	
	Coeff.	Coeff.	Coeff.	Prob.	Coeff.	Prob.
	6	7	8	9	10	11
Constant	20.144	**0.007**	25.978	**0.001**	-2.453	0.741
Aggregate Deviations, $\|e\|$	-0.741	**0.000**	..	**..**	..	..
Aggregate Deviations, $ln\|e\|$	..	**..**	-6.062	**0.000**	..	..
Aggregate Deviations, $ln\|e\|$ when $e < 0$	..	**..**	..	**..**	-5.365	**0.000**
Aggregate Deviations, $ln\|e\|$ when $e > 0$	..	**..**	..	**..**	-5.883	**0.000**
H01 FCTN_ETHN	-0.224	**0.000**	-0.202	**0.000**	-0.215	**0.000**
H05 MPP	0.189	**0.000**	0.138	**0.010**	0.142	**0.008**
H08 PEFA	9.292	**0.001**	8.471	**0.002**	9.495	**0.001**
Burundi 2005	49.740	**0.000**	49.233	**0.000**	48.476	**0.000**
Uganda 2006-07	51.093	**0.000**	51.891	**0.000**	50.319	**0.000**
Observations	144		144		144	
R-squared	0.479		0.505		0.512	
Adjusted R-squared	0.456		0.484		0.487	
S.E. of regression	13.110		12.769		12.725	
Sum squared resid	23,55		22,336		22,023	
Log likelihood	-571		-567		-566	
F-statistic	20.949	**0.000**	23.323	**0.000**	20.404	**0.000**
Normal error term 1/	**0.752**		**0.518**		**0.730**	
Heteroskedasticity Test 2/	**0.849**		**0.636**		**0.429**	
Omitted Variable Test 3/	0.001		**0.426**		**0.593**	

1. Jarque-Bera probability.
2. White $n \cdot R^2$ probability (without cross terms).
3. Ramsey RESET F-statistic probability (1 term).

The results reject most of the hypothesized relationships but those included in equation 11c are all robust and conform to expectations. Overall, the fraction of budget heads that are accurately spent f_b tends to fall as aggregate expenditure deviations move away from zero in either direction. The rate of decline slows as aggregate deviations move further away from zero. Equation 11d shows there is little to be gained by dividing aggregate expenditure deviations into negative and positive observations: the coefficients suggest the response is essentially symmetrical. This, of course, may not be the case within specific countries across time.

The general tendencies above are modified by two variables. Compositional accuracy is reduced by ethnic fractionalization. The good news is that f_b is increased by the presence of mature political ruling parties, *MPP*, and by the quality of public financial management, *PEFA*.

D. Discussion and Policy Implications

The results above lend support to the over-arching hypotheses that common pool behavioral tendencies tend to worsen the quality of budgetary execution while political institutions can help reduce such tendencies directly and through the use of counter-cyclical financing. The qualification to this broad statement is that common pool drivers are absent from the revenue deviation story which is instead driven by debt burdens, revenue policy, political systems, and the capacity for PFM. By contrast, expenditure deviations are driven directly or indirectly by variables from all categories including drivers of common pool behavior.

High debt service burdens are correlated with revenue shortfalls and excessive spending, although the spending impact is not included in Table 13 since it is only significant at just above the 10 percent level. Countries with high debt service burdens also tend to save positive revenue shocks via the interactive term *rpos·Z(PCT_INT)* while negative revenue shocks induce reduced spending.

Table 13—Surviving Correlates

	Revenue Deviations, *r*	Expenditure Deviations, *e*	Compositional Deviations, f_b
Expenditure Deviations, *e*	..	..	*−ln|e|*
Debt Burden			
PCT_INT	*-PCT_INT*	*+rpos·Z(PCT_INT)*	..
Common Pool			
H02 DRO	..	*+DRO*	..
H09 *ln*POP	..	*+lnPOP*	..
H04 POLITY	..	*+POLITY*	..
H01 FCTN_ETHN	..	..	*−FCTN_ETHN*
Political Institutions			
H06 SYSTEM	*+Rp·Z(SYSTEM)*	..	..
H05 MPP	..	..	*+MPP*
PFM Capacity			
H10 *ln*GNIPC	*+Rp·Z(lnGNIPC)*	*+r·Z(lnGNIPC)*	..
H02 DRY	*-DRY*	*-DRY*	..
H09 *ln*POP	..	*−r·Z(lnPOP)*	..
H03 PCT_CSE	..	*−r·Z(PCT_CSE)*	..
H08 PEFA	..	..	*+PEFA*

Common pool tendencies are driven by the existence of various pressure groups that politicians believe they need to cater to. Surviving correlates associated here with common pool behavior include ethnic factions and aging populations. Population is reassigned as a common pool correlate on the grounds that larger populations would be more likely to have more interest groups to cater to. The degree of democratization is also reassigned as a common pool correlate: politicians in democracies probably feel a greater pressure to respond to pressure groups than those in autocracies.

Unambitious revenue targets, measured by *Rp*, are correlated with unprogrammed revenue gains while overly ambitious revenue targets are correlated with revenue shortfalls. This tendency is reduced through interactive terms for countries with presidential systems and/or high income per capita. In addition, mature programmatic political parties appear to be critical to compositional accuracy in budget execution.

Counter-cyclical financing can be used to smooth out unanticipated revenue shocks. Countries tend to do this better when they have high per-capita incomes. Low capacity countries, measured by small expenditure shares for their civil service, small populations, or very young populations tend to underspend in the face of unanticipated revenue gains.

Table 14—Summary of Preferred Regression Equations

	Preferred Equation 9c		Preferred Equation 10c		Preferred Equation 11c	
	Coeff.	**Prob.**	**Coeff.**	**Prob.**	**Coeff.**	**Prob.**
Constant	15.602	**0.000**	..	**..**	25.978	**0.001**
Expenditure Deviations, *ln|e|*	..	**..**	..	**..**	-6.062	**0.000**
Debt Service Burden						
PCT_INT	-0.119	**0.076**	0.135	**0.052**	..	**..**
$r_{pos} \cdot Z(\%INT)$	..	**..**	0.564	**0.000**	..	**..**
Common Pool						
H02 *DRO*	..	**..**	0.642	**0.000**	..	**..**
H09 lnPOP	..	**..**	..	**..**	..	**..**
H04 POLITY	..	**..**	0.219	**0.072**	..	**..**
H01 *FCTN_ETHN*	..	**..**	..	**..**	-0.202	**0.000**
Political Institutions						
H06 *Rp* · Z(*SYSTEM*)	43.099	**0.000**	..	**..**	..	**..**
H05 *MPP*	..	**..**	..	**..**	0.138	**0.010**
PFM Capacity						
H08 *Rp* · Z(*PEFA*)	90.765	**0.000**	..	**..**	..	**..**
H08 *r · Z(PEFA)*	..	**..**	1.241	**0.000**	..	**..**
H02 *DRY*	-0.150	**0.000**	-0.133	**0.000**	..	**..**
H09 *r · Z(lnPOP)*	..	**..**	-1.431	**0.000**	..	**..**
H03 *r · Z(PCT_CSE)*	..	**..**	..	**..**	..	**..**
H08 *PEFA*	..	**..**	..	**..**	8.471	**0.002**
Burundi 2005	..	**..**	..	**..**	49.233	**0.000**
Georgia 2004	19.441	**0.000**	..	**..**	..	**..**
Ghana 2006	-44.308	**0.000**	..	**..**	..	**..**
Ghana 2008	22.851	**0.003**	25.782	**0.001**	..	**..**
India 2007/08	..	**..**	37.554	**0.000**	..	**..**
Pakistan 2007/08	..	**..**	23.563	**0.003**	..	**..**
Peru 2006	36.004	**0.000**	..	**..**	..	**..**
Solomon Isl. 2004	35.226	**0.000**	..	**..**	..	**..**
Uganda 2006/07	..	**..**	..	**..**	51.891	**0.000**
Zambia 2003	..	**..**	25.721	**0.001**	..	**..**
Observations	144		144		144	
R-squared	0.712		0.581		0.505	
Adjusted R-squared	0.692		0.550		0.484	
S.E. of regression	7.607		7.358		12.769	
Sum squared resid	7,755		7,200		22,336	
Log likelihood	-491		-486		-567	
F-statistic	36.749	**0.000**	..	**..**	23.323	**0.000**
Normal error term 1/	0.438		0.834		0.518	
Heteroskedasticity 2/	0.624		0.971		0.636	
Omitted Variables 3/	0.571		0.531		0.426	

1. Jarque-Bera probability.
2. White $n \cdot R^2$ probability (without cross terms).
3. Ramsey RESET F-statistic probability (1 term).

There is a positive role for PFM capacity measured by the PEFA index in compositional accuracy. While this is good news, the econometric procedures outlined in Sections VI.B and VI.C above rejected any role for the PEFA index in minimizing aggregate revenue and expenditure deviations, favoring instead *lnGNIPC*. Yet, it is obvious that wealthier governments can procure better capacity if they want to – and it seems they often do. de Renzio (2007) found that PFM capacity, as measured by the PEFA index, is highly correlated with the natural log of income per capita.[46] This

[46] In the author's recreation of de Renzio's results, income per-capita had a t-statistic of 7.5. The next highest t-statistic was 4.2 for population.

suggests it should be possible to substitute PEFA for *ln*(GNIPC) in equations 9 and 10. The results, shown in Table 14, are positive and hold up well econometrically with only minimal degradation in explanatory power relative to the original results above.[47] Moreover, in the case of equation 10, the variable *Z(PEFA)·r* appears in 14 of the 15 best (highest r-square) models.

The empirical results do point to some policy implications. For most countries, minimizing revenue deviations is the key to keeping expenditures on track both in level and in composition. To achieve this, the model suggests three options within the immediate control of governments. One is to manage the ambitiousness of revenue policy targets. Another option is to make the kinds of investments that generate higher PEFA scores. Doing so will make it easier to achieve revenue policy targets while also reducing common pool problems. Equations 9b and 10b suggest this is more likely in wealthy countries. It is worth noting, however, that de Renzio *et al* (2011) found a positive but small correlation between PEFA scores and external technical assistance for building PFM capacity. This suggests that interested and motivated governments in low income countries may be able to advance despite their low income environments.

Finally, some governments suffer from under-spending even in the face of a revenue surplus. As noted above, some of these governments often have a smaller than average share of expenditures devoted to civil service employment. The problem could be traced to low numbers of employees, low numbers of skilled employees, or high numbers of underpaid and unmotivated employees. Governments in these situations could consider bringing in more skilled workers and ensuring they are adequately paid and motivated.

VII. Conclusions

This paper introduced and exploited a new database with several measures of the quality of budget execution, based on deviations from approved budgetary targets for revenues and expenditures. It was discovered that very few countries are able to execute their budgets well, regardless of whether they have prioritized macroeconomic stability or public service delivery. There are widespread errors in forecasting revenues and within-year counter-cyclical financing does not fully compensate for revenue deviations. Expenditure deviations therefore tend to be positively correlated with revenue deviations. Within this broad tendency, there is considerable variation in behavior not only across countries but also within countries over time. Compositional accuracy is rare. It is common to observe over- and under-spending occurring simultaneously, even when net resource deviations are small.

For the majority of governments in the sample, the key to keeping expenditures on track is keeping revenues on target. Most governments tend to set their targets too low relative to what is likely to be achieved, thus generating unprogrammed surpluses. These revenue surpluses constitute a second common pool to spend from, in addition to the common pool of excess borrowing. Compositional deviations from budget targets tend to increase as aggregate deviations increase – although the rate of increase appears to slow as more and more ministries become involved. In addition, compositional deviations tend to be greater than deviations in total net resources because upward deviations by some ministries will require downward deviations in others, and vice versa.

The paper introduced and tested a number of hypotheses to explain the cross-country distribution of budget execution quality. These are inferred from the literature on political institutions and

[47] Two variables dropped out of equation 10 when *PEFA* was substituted for *lnGNIPC*. These were the constant, already insignificant in equation 10b, and *r·Z(PCT_CSR)*.

public financial management. The data support the common pool hypothesis and the idea that the tendency towards common pool behavior can be ameliorated through political institutions and the capacity for public financial management.

There is good news in the data for governments seeking policy options for better budget execution. Since expenditure deviations are strongly correlated with revenue deviations, one option for many governments is to set revenue targets that are more in line with what is most likely to be achieved: most governments set targets that are too low in this regard. Countries can increase their chances of success by investing in PFM capacity and by ensuring they have enough well trained and motivated civil servants. To achieve these goals, they may draw either on their national wealth or on external assistance from the international donor community.

There are many opportunities for researchers who may want to challenge or expand upon the work begun in this paper. The data sample could be expanded to include more countries, ideally with more representation from high income countries, and with more years per country.[48] This might eventually allow a panel approach to explore within-country correlates across time.

Future research could explore additional independent variables such as how many ministries or programs are to be over/under-spent or the trade-off between the number of ministries involved and the average depth of over/under-spending.

It may be possible to expand the list of hypotheses. The list of proposed hypotheses is a first attempt with lots of room for improvement. For example, it was noted in Section IV.C that it would be possible to test for the impact of unanticipated recessions, inflation, relative price changes, political events, and natural disasters – if comprehensive data could be found to allow it.

It would also be useful to see if the PEFA index could be disaggregated into various components to see if any stand out as particularly important or unimportant to good budget execution, paralleling earlier work by Dabla-Norriss (2010) and de Renzio *et al* (2011).

Another possibility would be to extend the political economics theory of election cycles to the issue of budgetary deviations. Do election cycles have an impact on aggregate or compositional deviations? If so, under what circumstances? Are some sectors routinely protected more than others? The work by Drazen and Eslava (2005) is intriguing: it provides a behavioral model that points in the right direction. They find that the composition of spending tends to shift away from recurrent spending and towards projects immediately before elections at the local level. Having ordered such a shift, would politicians also push their bureaucracy to ensure the shift is well implemented or would the promise be sufficient so that subsequent budgetary deviations are politically costless?

Finally, future researchers might find ways to use budgetary deviations as explanatory variables rather than independent variables. It would be especially exciting to see the assertions about the correlation between budgetary credibility and the quality of public service delivery or investment given a more empirical grounding.

[48] Future researchers should be aware that portions of the PEFA methodology changed in January 2011.

Bibliography

Alesina, Alberto, Ricardo Hausmann, Rudolf Hommes, and Ernesto Stein, 1999. Budget Institutions and Fiscal Performance in Latin America, *Journal of Development Economics*, Vol. 59, pp. 253–73.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg, 2003. Fractionalization, *Journal of Economic Growth* Vol. 8, No. 2.

Alesina, Alberto and Ekaterina Zhuravskaya, 2008. Segregation and the Quality of Government in a Cross-Section of Countries, *Working Paper 14316*, Cambridge MA: National Bureau of Economic Research.

Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer and Patrick Walsh, 2001. New Tools in Comparative Political Economy: The Database of Political Institutions, *World Bank Economic Review*, 15(1).

Brender, Adi and Allen Drazen, 2012. Elections, Leaders, and the Composition of Government Spending, *Journal of Public Economics*, Vol. 96.

Buettner, Thiess and Bjoern Kauder, 2010. Revenue Forecasting Practices: Differences across Countries and Consequences for Forecasting Performance, *Fiscal Studies*, Vol. 31, No. 3.

Chatagny, Florian, and Nils C. Soguel, 2011. SSES 2011 Rationality and Determinants of Tax Revenue Forecasting Errors: Evidence from Swiss cantons.

Cheibub, Jose, 2006. Presidentialism, Electoral Identifiability, and Budget Balances in Democratic Systems, *American Political Science Review*, Vol. 100, No. 3.

Couture, Jerome, & Louis M. Imbeau, 2009. Do Governments Manipulate Their Revenue Forecasts? Budget Speech and Budget Outcomes in the Canadian Provinces, in *Do They Walk Like They Talk?* (pp. 155-166), Springer, New York.

Cruz, Cesi, and Philip Keefer, 2010. Programmatic Political Parties and Public Sector Reform, APSA 2010 Annual Meeting Paper.

Dabla-Norris, Era, Richard Allen, Luis-Felipe Zanna, Tej Prakash, Eteri Kvintradze, Victor Lledo, Irene Yackovlev, and Sophia Gollwitzer, 2010. Budget Institutions and Fiscal Performance in Low-Income Countries, *IMF Working Paper 10/80*.

Deacon, Robert., 2003. Dictatorship, Democracy, and the Provision of Public Goods, *Departmental Working Papers*, Department of Economics, UCSB, UC Santa Barbara.

de Renzio Paolo, 2009. Taking Stock: What do PEFA Assessments Tell Us About PFM Systems Across Countries?, *Working Paper 302*, Overseas Development Institute.

de Renzio, Paolo, Matt Andrews, and Zac Mills, 2011., Does Donor Support to Public Financial Management Reforms in Developing Countries Work? An Analytical Study of Quantitative Cross-Country Evidence, *Working Paper* 329, Overseas Development Institute.

Doppelhofer, G., X. Sala-i-Martin, and R. Miller, 2000. 'Determinants of Long-run Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach', *Working Paper 7750*, Cambridge MA: National Bureau of Economic Research.

Drazen, Allan and Marcela Eslava, 2005. Electoral Manipulation via Expenditure Composition: Theory and Evidence, *Working Paper 11085*, NBER Working Paper Series.

Duverger, Maurice., 1954. *Political Parties: Their Organization and Activity in the Modern State*, Wiley, New York.

Fabrizio, Stefania and Ashoka Mody, 2006. Can Budget Institutions Counteract Political Indiscipline?, *Economic Policy,* Vol. 21, No. 48.

Gerring, John; Strom Thacker, and Carola Moreno, 2005. Are Parliamentary Systems Better?, *Comparative Political Studies*, Vol. 42, No. 3.

Heller, Peter, 1975. A Model of Public Fiscal Behavior in Developing Countries: Aid, Investment, and Taxation, *The American Economic Review*, Vol. 65(3).

IMF 2001a. *Tracking of Poverty-Reducing Public Spending in HIPCs*, Washington, D.C.

IMF, 2001b. Government Financial Statistics Manual, Washington, DC.

Keefer, Phil, 2011. Collective Action, Political Parties and Pro-Development Public Policy, *World Bank Working Paper Series,* No. 5676.

Kopits, George, 2001. Fiscal Rules: Useful Policy Framework or Unnecessary Ornament?, *IMF Working Paper* 01/145.

Kostopoulos, Christos, 1999. Progress in Public Expenditure Management in Africa: Evidence from World Bank Surveys, Africa Region Working Paper Series – 1, The World Bank.

Lake, David and Matthew Baum, 2001. The Invisible Hand of Democracy: Political Control and the Provision of Public Services, *Comparative Political Studies*, Vol. 34, No. 6.

Lewis, M. Paul (ed.), 2009. *Ethnologue: Languages of the World*, Sixteenth edition. Dallas, Tex.: SIL International. Online version: http://www.ethnologue.com/.

Lizzeri, Allesando and Nicola Persico, 2004. Why Did The Elites Extend the Suffrage? Democracy and the Scope of Government, With an Application to Britain's "Age of Reform," *The Quarterly Journal of Economics*, Vol. 119(2).

Marshall, Monty; Jaggers, Keith; and Ted Gurr, 2010. *Polity IV Project, Political Regime Characteristics and Transitions, 1800-2009*, Center for International Development and Conflict Management, University of Maryland, College Park, MD.
http://www.systemicpeace.org/polity/polity4.htm

McGuire, Martin and Mancur Olson, Jr., 1996. The Economics of Autocracy and Majority Rule: The Invisible Hand and the Use of Force, *Journal of Economic Literature*, Vol. 24, No. 1.

Meltzer, Allan H., and Richard, Scott F, 1978. "Why Government Grows (and Grows) in a Democracy." *Public Interest*, Vol. 52.

Mocan, H. Naci and Sam Azad, 1995. Accuracy and Rationality of State General Fund Revenue Forecasts: Evidence from Panel Data, *International Journal of Forecasting*, Vol. 11.

Mueller, Dennis, and Thomas Stratmann, 2003. "The Economic Effects of Democratic Participation," *Journal of Public Economics*, 87:2129–2155.

PEFA Secretariat, 2005. Public Financial Management Performance Measurement Framework, World Bank, Washington, D.C.

PEFA Secretariat, 2011. Public Financial Management Performance Measurement Framework, World Bank, Washington, D.C.

PEFA Secretariat, 2009. Issues in Comparison and Aggregation of PEFA Assessment Results Over Time and Across Countries, World Bank, Washington, D.C.

Persson, Torsten, 2002. Do Political Institutions Shape Economic Policy?, *Econometrica*, Vol. 70(3).

Persson, Torsten, Gerard Roland and Guido Tabellini, 1997., Separation of Powers and Political Accountability, *Quarterly Journal of Economics*, 112, 310–327.

Persson, Torsten and Guido Tabellini, 2000. Political Institutions and Economic Policy Outcomes: What Are the Stylized Facts?, *Mimeo*, Institute for International Economic Studies.

Peters, Lucian, 2002. Sound Budget Execution for Poverty Reduction, *A Background Paper for the World Bank Institute Program on "Fiscal Policy for Poverty Alleviation" Module on Budget Execution*, World Bank, Washington, DC.

Petrei, A. Humberto, and Romeo E. Petrei. 1998. *Budget and Control: Reforming the Public Sector in Latin America*, Inter-American Development Bank, Washington, D.C.

Premchand, Agrigapudi, 1984. *Government Budgeting and Expenditure Controls: Theory and Practice*, International Monetary Fund.

Premchand, Agrigapudi, 1999. Public Financial Management: Getting the Basics Right, in *Governance, Corruption and Public Financial Management*, Asian Development Bank.

Ramey G. and V. Ramey, 1995. Cross-Country Evidence on the Link between Volatility and Growth, *American Economic Review*, Vol. 85, No. 5, 1138-1150.

Roubini N. and J. Sachs, 1989. Political and Economic Determinants of Budget Deficits in the Industrial Democracies," *European Economic Review*, 33, 903-938.

Sala-i-Martin, X., 1997. 'I Just Ran 2 Million Regressions', *American Economic Review*, 87(2): 178-183.

Shah, Anwar, editor, 2007. *Budgeting and Budgetary Institutions*, The World Bank.

Talvi, Ernesto and Carlos Vegh, 2004. Tax Base Variability and Procyclical Policy, *Working Paper 7499*, National Bureau of Economic Research, Washington, DC.

Tonizzo, Martina., 2008. Political Institutions, Size of Government and Redistribution: An Empirical Investigation, *Development Studies Institute Working Paper Series, No. 08-89.*

Tornell, Aaron and Philip Lane, 1999. The Voracity Effect, *The American Economic Review*, Vol. 89(1).

UN Population Division, World Population Prospects, online database. http://esa.un.org/unpd/wpp/index.htm

van der Ploeg, Frederick, 2007. Prudent Budgetary Policy: Political Economy of Precautionary Taxation, *CESifo Working Paper*, No. 1973.

Von Hagen, Jurgen, 2002. Fiscal Rules, Fiscal Institutions, and Fiscal Performance, *The Economic and Social Review*, Vol. 33(3).

Von Hagen, Jurgen, 2007. Budgeting Institutions for Better Fiscal Performance, in *Budgeting and Budgetary Institutions*, Anwar Shah, editor, The World Bank.

Wildavsky, Aaron, 1986. *Budgeting: a Comparative Theory of Budgetary Processes*, Transaction Publishers, New Jersey.

Weingast, B., Shepsle, K., and C. Johnsen, 1981. The Political Economy of Benefits and Costs: A Neoclassical Approach to Distributive Politics, *The Journal of Political Economy*, Vol. 89, No. 4, pp. 642-664.

ANNEX 1—KEY VARIABLES, DEFINITIONS AND SOURCES

	Definitions	Sources	Notes
Revenue deviation, *r*	Deviation in revenue, as a percent of budgeted total expenditure.	Assorted PEFA Reports and author's calculations.	See Section III for methodology. See Annex 3 for actual data.
Expenditure deviation, *e*	Deviation in total expenditure, as a percent of budgeted total expenditure.	Assorted PEFA Reports and author's calculations	See Section III for methodology. See Annex 3 for actual data.
Accurate spending, f_b	Fraction of total budget heads spent within ±5 percent of their original budgetary allocations, expressed as a percent of budgeted total expenditure.	Assorted PEFA Reports and author's calculations	See Section III for methodology. See Annex 3 for actual data.
rp	Revenue policy.	Author's calculations	See Section IV for methodology.
rv	Revenue volatility .	Author's calculations. Primary sources include the World Development Indicators and reports from the International Monetary Fund.	Values for Afghanistan 2004/05 and 2005/06 were set equal to values for 2006/07 due to lack of data.
PCT_INT	Interest paid as a share of actual revenue receipts.	Assorted PEFA reports and reports from the International Monetary Fund.	
H01 FCTN_ETHN	Ethnic fractionalization.	Alesina et al (2003).	Data for Maldives, Samoa, and Yemen were estimated by the author using published population estimates.
H01 FCTN_LANG	Linguistic fractionalization.	Alesina et al (2003) and Ethnologue (2009).	Data for Cape Verde, El Salvador, Grenada, Maldives, and Samoa were estimated by the author using data from Ethnologue and other sources.
H01 FCTN_RELIG	Religious fractionalization.	Alesina et al (2003).	Data for Maldives and Samoa were calculated by the author using data from published population estimates.
H02 DRY	Dependency ratio, population below age 15.	World Development Indicators.	
H02 DRO	Dependency ratio, population above age 65.	World Development Indicators.	
H03 PCT_CSE	Civil service wages and salaries paid as a share of actual total expenditures.	World Development Indicators and assorted PEFA reports.	Data for Nepal 2004/-5 and 2005/06 assumed to be same as observed in 2003/04.
H04 POLITY	Index of autocracy and democracy	Marshall, Jaggers and Gurr, 2010.	Ten year average, current year and nine preceding years.
H05 MPP 1/	Index of mature programmatic political parties, defined as the age of the executive's political party less the number of years the executive has been in office.	The Database of Political Institutions, 2010, and author's calculations.	
H06 SYSTEM 1/	Dummy indicating a country is using a presidential system.	The Database of Political Institutions, 2010.	Data for Tonga were estimated by the author.
H07 PR 1/	Dummy indicating a country is using proportional voting.	The Database of Political Institutions, 2010.	Data for Tonga were estimated by the author.
H08 PEFA	Index of the quality of public financial management.	Assorted PEFA Reports and author's calculations.	See Section IV for details.
H09 ln(POP)	Natural log of total population.	World Development Indicators.	
H10 ln(GNIPC)	Natural log of gross national income per capita, measured at purchasing power parity.	World Development Indicators.	
H11 ODAGDP	Official development assistance as a share of GDP.	World Development Indicators.	
H12 BATTLE	Battle related deaths per 100,000 people.	World Development Indicators.	

1. All DPI variables for Senegal in 2010 are assumed to be the same as what was observed in 2009.

ANNEX 2—DESCRIPTIVE STATISTICS

	Bounds	Type	Mean	Max.	Min.	Std. Dev.	Skewness	Kurtosis
Dependent Variables								
Deviation in expenditures, *r*	None	Continuous	7.388	86.069	-18.549	13.714	1.855	10.671
Deviation in expenditures, *e*	None	Continuous	3.047	44.128	-22.417	10.963	0.829	4.261
Compositional accuracy, f_b	0 to 100	Continuous	30.868	80.950	0.000	17.769	0.676	2.992
Independent Variables								
Rp	None	Continuous	0.032	0.663	-0.293	0.127	1.277	7.328
Rv	0 to ∞	Continuous	0.091	0.505	0.006	0.081	2.277	9.436
PCT_INT	0 to 100	Continuous	9.865	46.978	0.000	9.919	1.884	6.045
H01 FCTN_ETHN	0 to 100	Continuous	47.248	93.020	3.000	26.557	-0.255	1.756
H01 FCTN_LANG	0 to 100	Continuous	43.285	92.270	0.800	29.292	-0.107	1.670
H01 FCTN_RELIG	0 to 100	Continuous	46.583	86.030	0.230	25.619	-0.309	1.885
H02 DRY	0 to ∞	Continuous	60.207	99.939	21.350	19.673	-0.160	2.206
H02 DRO	0 to ∞	Continuous	8.797	22.604	3.428	4.584	1.471	4.652
H03 PCT_CSE	0 to 100	Continuous	26.563	55.844	3.559	11.991	0.344	2.299
H04 POLITY	-10 to 10	Continuous	3.452	10.000	-9.400	5.333	-0.623	2.152
H05 MPP	0 to ∞	Continuous	16.007	118.000	0.000	22.387	1.950	7.207
H06 SYSTEM	Dummy	0 or 1	0.681	1.000	0.000	0.468	-0.774	1.600
H07 PR	Dummy	0 or 1	0.215	1.000	0.000	0.412	1.385	2.919
H08 PEFA	1 to 4	Continuous	2.570	3.609	1.641	0.433	0.246	2.496
H09 ln(POP)	0 to ∞	Continuous	15.382	20.854	11.525	1.961	-0.162	3.256
H10 ln(GNIPC)	0 to ∞	Continuous	7.971	10.884	5.799	0.962	0.353	3.300
H11 ODAGDP	0 to ∞	0 to 100	0.092	0.495	0.000	0.107	1.939	6.588
H12 BATTLE	0 to 100	Continuous	0.443	16.759	0.000	1.909	5.956	43.721

1. Sample includes 144 observations. Sample omits Kosovo (6 observations), Montenegro (3 observations), Serbia (3 observations), and Timor-Leste (3 observations).
Source: Author's calculations.

ANNEX 3—QUALITY OF BUDGET EXECUTION

Obs	Country	Year Start	Heads	Deviation in Revenue, *r*	Deviation in Total Expenditure, *e*	Share of Heads Accurately Spent, f_b
1	Afghanistan	21-Mar-04	21	-8.5449%	-10.9706%	14.29%
2	Afghanistan	21-Mar-05	21	13.8613%	-2.7156%	19.05%
3	Afghanistan	21-Mar-06	21	8.7509%	-2.2483%	42.86%
4	Armenia	1-Jan-05	14	3.9997%	6.7138%	57.14%
5	Armenia	1-Jan-06	14	9.8849%	4.8592%	64.29%
6	Armenia	1-Jan-07	14	13.5045%	7.3889%	42.86%
7	Belarus	1-Jan-05	10	15.2949%	10.8258%	20.00%
8	Belarus	1-Jan-06	10	11.0213%	4.6233%	40.00%
9	Belarus	1-Jan-07	10	16.5107%	11.9148%	40.00%
10	Benin	1-Jan-04	21	-7.9199%	-9.8251%	4.76%
11	Benin	1-Jan-05	21	5.6226%	-6.5115%	19.05%
12	Benin	1-Jan-06	21	0.4473%	-12.1987%	4.76%
13	Bhutan	1-Jul-05	12	4.7959%	-6.9816%	33.33%
14	Bhutan	1-Jul-06	12	6.4163%	-7.4074%	33.33%
15	Bhutan	1-Jul-07	12	1.5108%	-8.2532%	16.67%
16	Bolivia	1-Jan-06	21	19.0809%	-3.9260%	14.29%
17	Bolivia	1-Jan-07	21	5.3875%	-7.6379%	9.52%
18	Bolivia	1-Jan-08	21	12.5993%	15.1158%	19.05%
19	Botswana	1-Apr-05	21	6.1757%	-8.6341%	4.76%
20	Botswana	1-Apr-06	21	7.9175%	-10.6730%	14.29%
21	Botswana	1-Apr-07	21	9.8132%	-14.7999%	4.76%
22	Burundi	1-Jan-05	21	8.7542%	-3.1394%	76.19%
23	Burundi	1-Jan-06	21	-0.5541%	-8.2855%	19.05%
24	Burundi	1-Jan-07	21	-4.5256%	-9.1272%	33.33%
25	Cape Verde	1-Jan-05	20	-6.2617%	3.4440%	25.00%
26	Cape Verde	1-Jan-06	20	-9.0756%	-8.9268%	40.00%
27	Cape Verde	1-Jan-07	20	16.3554%	-1.7562%	45.00%
28	Dominican Republic	1-Jan-04	20	23.3133%	24.5947%	35.00%
29	Dominican Republic	1-Jan-05	20	4.4284%	21.4595%	40.00%
30	Dominican Republic	1-Jan-06	20	-1.3369%	1.3467%	50.00%
31	Dominican Republic	1-Jan-07	15	13.0965%	12.9482%	13.33%
32	Dominican Republic	1-Jan-08	17	0.6058%	14.7073%	23.53%
33	Dominican Republic	1-Jan-09	17	-12.0160%	19.5738%	29.41%
34	El Salvadore	1-Jan-06	21	12.0617%	8.3128%	28.57%
35	El Salvadore	1-Jan-07	21	13.9163%	16.4089%	23.81%
36	El Salvadore	1-Jan-08	21	8.5491%	10.9592%	42.86%
37	Georgia	1-Jan-04	21	25.1200%	16.1033%	19.05%
38	Georgia	1-Jan-05	20	32.3112%	16.7825%	20.00%
39	Georgia	1-Jan-06	21	21.5420%	17.6730%	33.33%
40	Ghana	1-Jan-03	21	-2.5145%	2.2885%	9.52%
41	Ghana	1-Jan-04	21	6.7711%	12.4373%	23.81%
42	Ghana	1-Jan-05	21	-6.2941%	-7.9572%	14.29%
43	Ghana	1-Jan-06	21	-5.2116%	8.5736%	23.81%
44	Ghana	1-Jan-07	21	-8.3159%	14.1293%	9.52%
45	Ghana	1-Jan-08	21	25.9725%	38.1913%	9.52%
46	Grenada	1-Jan-06	21	-7.1371%	16.4645%	33.33%
47	Grenada	1-Jan-07	21	-5.4371%	9.3761%	28.57%
48	Grenada	1-Jan-08	21	-1.8000%	16.7615%	19.05%
49	India	1-Apr-05	23	8.3519%	12.9474%	30.43%
50	India	1-Apr-06	23	12.1087%	10.6212%	39.13%
51	India	1-Apr-07	23	-9.7790%	36.9478%	21.74%
52	Jamaica	1-Apr-03	21	21.8527%	1.3474%	28.57%
53	Jamaica	1-Apr-04	21	-2.7992%	8.3760%	42.86%
54	Jamaica	1-Apr-05	21	-12.8114%	5.9163%	38.10%
55	Kosovo	1-Jan-03	21	0.3001%	-30.2760%	14.29%
56	Kosovo	1-Jan-04	21	-1.6795%	-24.8395%	23.81%
57	Kosovo	1-Jan-05	21	-1.7903%	-10.4377%	33.33%
58	Kosovo	1-Jan-06	22	9.1547%	-10.2080%	27.27%
59	Kosovo	1-Jan-07	22	31.2562%	-10.8066%	22.73%
60	Kosovo	1-Jan-08	22	4.1960%	-16.6543%	18.18%
61	Kyrgystan	1-Jan-06	9	11.1339%	9.7451%	0.00%
62	Kyrgystan	1-Jan-07	10	22.2288%	16.2190%	10.00%
63	Kyrgystan	1-Jan-08	10	8.6569%	3.8115%	0.00%
64	Lesotho	1-Apr-02	21	-3.0943%	-2.3352%	42.86%
65	Lesotho	1-Apr-03	21	7.5204%	-6.2246%	47.62%

Obs	Country	Year Start	Heads	Deviation in Revenue, r	Deviation in Total Expenditure, e	Share of Heads Accurately Spent, f_b
66	Lesotho	1-Apr-04	21	-2.0182%	-8.3771%	42.86%
67	Malawi	1-Jul-04	21	9.7322%	-7.6935%	14.29%
68	Malawi	1-Jul-05	21	17.4109%	-9.8360%	14.29%
69	Malawi	1-Jul-06	21	10.2878%	-1.3260%	33.33%
70	Maldives	1-Jan-05	21	-13.6947%	5.9499%	38.10%
71	Maldives	1-Jan-06	21	7.5146%	-5.7587%	33.33%
72	Maldives	1-Jan-07	21	-7.5049%	-13.2128%	28.57%
73	Mauritania	1-Jan-05	21	4.5276%	-17.7377%	38.10%
74	Mauritania	1-Jan-06	21	8.6158%	-5.3333%	38.10%
75	Mauritania	1-Jan-07	21	-1.5150%	-2.0182%	47.62%
76	Moldova	1-Jan-02	20	0.3745%	4.4257%	20.00%
77	Moldova	1-Jan-03	21	18.8727%	2.8488%	52.38%
78	Moldova	1-Jan-04	21	6.7336%	5.3559%	33.33%
79	Moldova	1-Jan-05	18	14.9759%	11.2556%	5.56%
80	Moldova	1-Jan-06	18	10.3781%	8.9075%	33.33%
81	Moldova	1-Jan-07	18	13.9810%	15.8871%	5.56%
82	Montenegro	1-Jan-05	14	29.8989%	-5.4825%	35.71%
83	Montenegro	1-Jan-06	10	20.9531%	-0.2796%	50.00%
84	Montenegro	1-Jan-07	10	39.4820%	8.2664%	10.00%
85	Morocco	1-Jan-05	21	10.8159%	8.3593%	33.33%
86	Morocco	1-Jan-06	21	15.2093%	2.4100%	28.57%
87	Morocco	1-Jan-07	21	22.1301%	3.1611%	28.57%
88	Nepal	16-Jul-03	21	7.7330%	-5.8782%	23.81%
89	Nepal	16-Jul-04	21	2.7028%	-0.2008%	28.57%
90	Nepal	16-Jul-05	21	-13.1608%	-5.5558%	14.29%
91	Norway	1-Jan-04	21	17.8332%	0.3837%	71.43%
92	Norway	1-Jan-05	21	13.4570%	-0.2912%	57.14%
93	Norway	1-Jan-06	21	25.0153%	0.2067%	76.19%
94	Pakistan	1-Jul-05	10	18.0004%	19.6850%	20.00%
95	Pakistan	1-Jul-06	10	13.1887%	2.1703%	30.00%
96	Pakistan	1-Jul-07	10	-3.1196%	22.6833%	20.00%
97	Peru	1-Jan-05	21	11.9822%	0.3500%	28.57%
98	Peru	1-Jan-06	21	44.6715%	6.7429%	19.05%
99	Peru	1-Jan-07	21	31.7365%	-8.2292%	42.86%
100	Samoa	1-Jul-03	20	-3.4907%	-1.0723%	50.00%
101	Samoa	1-Jul-04	20	6.2018%	3.2427%	65.00%
102	Samoa	1-Jul-05	21	6.9013%	2.7717%	42.86%
103	Samoa	1-Jul-06	21	5.5462%	1.8869%	23.81%
104	Samoa	1-Jul-07	21	-7.6503%	-1.0969%	33.33%
105	Samoa	1-Jul-08	21	-5.3647%	1.6284%	57.14%
106	Senegal	1-Jan-04	20	6.2030%	-1.1414%	25.00%
107	Senegal	1-Jan-05	20	9.9132%	-6.8989%	15.00%
108	Senegal	1-Jan-06	20	4.7373%	5.5439%	10.00%
109	Senegal	1-Jan-08	21	-1.6500%	-4.0426%	14.29%
110	Senegal	1-Jan-09	21	-16.3130%	-6.4270%	14.29%
111	Senegal	1-Jan-10	21	-1.7102%	2.5950%	14.29%
112	Serbia	1-Jan-07	21	-0.2319%	-8.5501%	47.62%
113	Serbia	1-Jan-08	21	-2.4445%	-8.2382%	33.33%
114	Serbia	1-Jan-09	21	-10.5031%	-5.5607%	52.38%
115	Solomon Islands	1-Jan-04	14	86.0689%	25.4458%	14.29%
116	Solomon Islands	1-Jan-05	14	15.7388%	-6.2623%	28.57%
117	Solomon Islands	1-Jan-06	14	17.6320%	2.3280%	14.29%
118	South Africa	1-Apr-05	21	11.4831%	0.3529%	66.67%
119	South Africa	1-Apr-06	21	8.2817%	-0.6354%	80.95%
120	South Africa	1-Apr-07	21	1.1480%	1.4230%	71.43%
121	Tajikistan	1-Jan-04	14	12.9336%	14.5518%	14.29%
122	Tajikistan	1-Jan-05	14	5.0405%	3.4036%	28.57%
123	Tajikistan	1-Jan-06	14	8.4354%	7.7230%	7.14%
124	Timor-Leste (Half Year)	1-Jan-07	19	0.6023%	-44.7840%	5.26%
125	Timor-Leste	1-Jan-08	19	41.8400%	39.3187%	21.05%
126	Timor-Leste	1-Jan-09	21	-12.0434%	-11.1379%	23.81%
127	Tonga	1-Jul-03	21	-2.6875%	-2.4254%	61.90%
128	Tonga	1-Jul-04	21	5.3697%	-3.3360%	28.57%
129	Tonga	1-Jul-05	21	12.4148%	13.8230%	19.05%
130	Tonga	1-Jul-06	21	-2.4033%	-1.6142%	38.10%
131	Tonga	1-Jul-07	22	5.9171%	-0.2618%	45.45%
132	Tonga	1-Jul-08	22	-13.4207%	-2.0663%	13.64%

Obs	Country	Year Start	Heads	Deviation in Revenue, r	Deviation in Total Expenditure, e	Share of Heads Accurately Spent, f_b
133	Trinidad And Tobago	1-Oct-04	21	28.4483%	9.5236%	33.33%
134	Trinidad And Tobago	1-Oct-05	21	19.5306%	-6.3874%	23.81%
135	Trinidad And Tobago	1-Oct-06	21	17.7144%	1.1695%	38.10%
136	Tunisia	1-Jan-06	21	4.0871%	0.8465%	57.14%
137	Tunisia	1-Jan-07	21	6.3133%	5.7487%	52.38%
138	Tunisia	1-Jan-08	20	13.5745%	8.8037%	40.00%
139	Vanuatu	1-Jan-03	14	-8.0589%	-4.7540%	64.29%
140	Vanuatu	1-Jan-04	14	3.5766%	1.1915%	64.29%
141	Vanuatu	1-Jan-05	14	7.2757%	-1.5231%	64.29%
142	Yemen	1-Jan-04	21	28.1281%	19.7597%	19.05%
143	Yemen	1-Jan-05	21	57.6420%	44.1279%	19.05%
144	Yemen	1-Jan-06	21	47.2160%	20.0638%	14.29%
145	Zambia	1-Jan-02	21	-2.3344%	7.2833%	4.76%
146	Zambia	1-Jan-03	21	2.4168%	22.1821%	9.52%
147	Zambia	1-Jan-04	21	1.0517%	-11.8215%	28.57%
148	Kenya	1-Jul-04	21	4.1733%	-10.7196%	28.57%
149	Kenya	1-Jul-05	21	0.9434%	-5.5796%	33.33%
150	Kenya	1-Jul-06	21	-0.6392%	-4.1465%	42.86%
151	Uganda	1-Jul-05	21	3.8661%	-1.7755%	38.10%
152	Uganda	1-Jul-06	21	7.2474%	-9.3834%	66.67%
153	Uganda	1-Jul-07	21	2.1424%	-13.5954%	52.38%
154	Sierra Leone	1-Jan-04	21	5.0791%	-7.6400%	28.57%
155	Sierra Leone	1-Jan-05	21	-3.8348%	5.5014%	28.57%
156	Sierra Leone	1-Jan-06	21	-6.7430%	1.1190%	33.33%
157	Sierra Leone	1-Jan-07	21	-18.5484%	-22.4171%	19.05%
158	Sierra Leone	1-Jan-08	21	-3.5767%	-3.5404%	23.81%
159	Sierra Leone	1-Jan-09	21	-6.7996%	6.5959%	28.57%