Staying the Course: Maintaining Fiscal Control in Developing Countries*

Christopher S. Adam and David L. Bevan
Department of Economics, University of Oxford, UK.

June 2003

Abstract

We use panel data on 83 developing and 25 OECD countries for the period from 1970-2000 to examine variations in the persistence of episodes of fiscal stability. Persistence is defined as the length of time the cyclically adjusted conventional fiscal balance exceeds a specific threshold, where the latter is based on plausible target values for the steady-state public debt-to-GDP ratio. We estimate hazard functions based on a range of alternative deficit thresholds. Four principal results emerge: (i) the fiscal stance and the determinants of fiscal stability differ significantly between OECD and developing countries and between middle-income and low-income countries; (ii) apart from the level of income, conventional structural characteristics of economies play a relatively minor role in explaining the persistence of fiscal stability; (iii) a history of poor fiscal management has a deleterious effect on efforts to maintain a sustainable fiscal stance. For middle-income and OECD countries, but not low-income countries, this legacy depreciates rapidly; (iv) in contrast to comparable work on the OECD we find that revenue reforms rather than expenditure cuts play the major role in underpinning fiscal stability, particularly for low-income countries.

1 Introduction

Sustained fiscal discipline promotes economic stability and underpins growth. There may be some dispute over what constitutes an optimal fiscal stance over the long-run but little, if any, about the importance attached to ensuring that a sustainable stance persists (for example Easterly, *et al*, 1994, Abed *et al,* 1998*,* Adam and Bevan, 2002). In practice, however, weak fiscal discipline as well as occasional catastrophic losses of fiscal control are all too prevalent,

*A first draft of this paper was presented to The Brookings Institution Trade Forum Conference, May 15-16 2003. We are grateful to our discussants, Sanjeev Gupta and Steve Radelet, and to conference participants, for extremely insightful and helpful comments on the first draft of this paper. We also thank Sarah Janzen for excellent research assistance and Miana Plesca for expert advice on manipulating duration data.

especially in developing countries. Fiscal policy debates in these countries therefore naturally converge on two issues: how best to rectify a fiscal stance which is in some sense too lax; and how to maintain this stance once the initial adjustment has been achieved. This paper is concerned with the second of these issues, and in particular with the question: why do some periods of fiscal stability last longer than others? At some level the answer is trivial: maintaining fiscal stability is a simple policy choice and is achievable as a matter of political will. In practice, of course, it is not that straightforward. The sustainability of any fiscal stance may be undermined or bolstered by external factors, while structural characteristics of the economy may influence its vulnerability to collapse. It will also likely be shaped by the country's fiscal history – for example, by its level of public debt and inflation, and the government's track record for fiscal management – and by how the fiscal rectification precipitating a spell of stability was engineered.

Although clearly central to the policy debate, the question of what factors support continued fiscal control has been the subject of remarkably little empirical work in the economics literature. There is, of course, a substantial literature on the politics and economics of fiscal consolidation in the OECD (see the survey paper by Perotti, 1998), but only recently has attention turned to developing countries (for example, Abed *et al* (1998), Adam and Bevan (2001), Gupta *et al* (2002), and Bulir and Moon (2003)). Here, though, the focus has predominantly been on the process of rectification of perceived fiscal imbalances and the impact of these on growth and inflation. This current paper takes a different tack: by using techniques of duration analysis to focus explicitly on the persistence of spells of stability. As we note below, others have used similar techniques to examine the related phenomenon of fiscal consolidation; this paper, we believe, provides the first analysis concerned with the persistence of the fiscal stance *per se* in developing countries.

To give structure to our analysis we organize the empirical part of the paper around four broad themes, each of which draws on different strands in the literature on fiscal policy. The first is concerned with determinism: to what extent do the structural characteristics of economies pre-determine the prospects for sustained fiscal control? For example, are low-income countries at greater risk of losing fiscal control than richer ones? Are natural resource dependent economies, or those more open to trade, particularly vulnerable? The second examines the question of path dependency: how does a history of poor fiscal control influence the durability of today's (stable) fiscal stance? The third is concerned with restraint: can the nominal exchange rate regime or the presence of external agents such as the IMF help lock-in a sustainable fiscal stance? Finally we examine the question of composition: does the form of a fiscal consolidation influence its subsequent duration?

These issues are of particular relevance for 'post-stabilization' economies, those that have achieved a degree of short-term macroeconomic stability following protracted periods of poor economic management, often through draconian fiscal adjustments. For these countries, the overriding macroeconomic

challenge is how to maintain these hard-won gains, and hence an understanding of the policy configurations that are most conducive to continued stability is valuable. For most of these countries, however, macroeconomic and fiscal stability is a relatively recent phenomenon and, as such, their own (post-stabilization) fiscal performance is necessarily censored and can, at best, provide only partial insights into the durability of reforms they have adopted.[1] Only by drawing on a wider sample of countries over a longer period of time can we hope to develop an empirical understanding of what exactly underpins durable fiscal control.

Our duration analysis approach requires us to define what we mean by fiscal consolidation and stability, and then how consolidation relates to sustained reform. Although the fiscal component of stabilization goes by a bewildering variety of names – fiscal adjustment, consolidation, correction, rectification, reform and tightening, to name just the main ones – the analysis usually focuses on some measure of the fiscal deficit. This may be the primary or conventional deficit and in the latter case may be gross or net of external official financing. Given data limitations, shared by most of the literature in this area, we focus on the conventional deficit after interest payments on public debt, expressed as a percent of GDP. What type of behaviour of this ratio is to qualify as a consolidation event? Should it be a reduction in the deficit in excess of some specified amount, or is it more fruitful to specify some threshold for the deficit, with consolidation representing a shift from above to below this threshold? The closely related issue is how to relate consolidation with persistence. This is quite tricky. It would be possible to regard a successful sustained consolidation as involving a one-step jump from an unsustainable to a sustainable configuration and remaining at that level. Persistence of the reform then requires maintenance of the new level, and failure is a relapse from it. We call this the 'level approach'. Alternatively, consolidation might be regarded as an ongoing stepwise climb from the one to the other. It would be regarded as infeasible to make the necessary reform in one jump, so that it would be better regarded as an ongoing process. Persistence of reform would then involve continuous reduction of the deficit at some minimum rate. We call this the 'gradient approach'. Finally, we can combine these two polar cases into a composite where an extended adjustment process moves the deficit to a sustainable level, which is then maintained. A period of continued (gradient) improvement is followed by a period of holding the line at the newly established level. Much of the existing literature focuses on the gradient approach; we believe that the level approach gets closer to the nature of sustainability as generally understood in the fiscal policy debate in developing countries. We therefore concentrate on this approach, leaving for

[1] Both Adam and Bevan (2001) and Gupta *et al* (2002) classify countries as 'post-stabilization' only if they are in a fiscally stable configuration so that the group consists exclusively of countries for which a spell of fiscal stability is ongoing. Hence nothing can be inferred about the likelihood of this stability persisting from their own post-stabilization performance.

subsequent work a more complete 'composite approach'.

The remainder of the paper is structured as follows. Section 2 relates the approach taken in the paper to existing work on fiscal consolidation and its persistence and develops an operational definition of fiscal sustainability. Section 3 describes the data used in the analysis and highlights three key stylized facts on fiscal performance, contrasting the experience of industrialized, middle-income and low-income countries. Based on evidence from 108 countries for the period 1970-2000 we show that countries differ markedly in terms of the volatility and cyclical properties of the deficit, and in the (unconditional) persistence of periods of fiscal stability. Sections 4 and 5 build on these stylized facts by estimating and discussing empirical duration models for the persistence of fiscal control. Section 6 summarizes the main findings of the paper, drawing out lessons for 'post-stabilization' countries, and the appendix describes our sensitivity analysis and provides further discussion of the statistical properties of the duration models used in the paper.

It is worth mentioning at this point what the paper does not do. Our focus on developing countries, combined with the demands placed on the data when dealing with 'multiple failures' – where the same country may experience a number of spells of fiscal (in)stability – means that there is a wide range of important issues we are unable to address in detail at this stage. Chief amongst these are the role of political institutions in shaping fiscal outcomes, an issue which commands centre stage in the OECD literature, and a more detailed analysis of the role of expenditure and revenue compositions. The natural next step in the analysis will be to refine the results presented here as data on both become more readily available.[2]

2 Fiscal Consolidation and Persistence

There is a substantial and rapidly growing body of work that sets out to examine the process of restoring fiscal control. This literature has mostly addressed two groups of issues (see Perotti, 1998). The first is concerned with the initiation and mechanics of the reform process: why it was necessary, what determined when it started, how it was implemented, and why this may have been delayed (for example, Alesina and Drazen, 1991 and Tornell and Lane, 1999). The second is concerned with the consequences of reform, particularly in the macroeconomic arena. These themes, which derived from the post-war experience of OECD countries, are reflected in the emerging literature on fiscal control in developing countries, much of which emanates from the IMF (see, for example, Abed *et al,* 1998, Gupta *et al,* 2002 and Bulir and Moon, 2003).

A recent strand in both branches of the literature is concerned with why some reforms persist while others do not. Perotti (1998), for example, notes that for the OECD countries, consolidations are more likely to persist if they

[2] For example, we hope to be able to exploit the data on political institutions in Africa which have recently been made available by the Africa Research Program at Harvard University (http:\\africa.gov.harvard.edu\\).

are achieved by expenditure reduction than by revenue increases. Von Hagen *et al* (2002) and Illera and Mulas-Granados (2001) use duration analysis to study the persistence of fiscal consolidation for OECD, and the 15 EU member states, respectively. In both cases, and despite somewhat different measures of fiscal consolidation (see below), they find that successful consolidations are, on average, underpinned by expenditure reductions rather than revenue enhancements. They also find that a high debt-to-GDP ratio raises the persistence of consolidation, arguably on the ground that it stiffens the resolve of the reforming government. Gupta *et al*'s (2002) study of 39 low-income countries during the 1990s adds to these findings. They suggest that persistence is greater when the initial fiscal adjustment is large, when there is a reallocation towards capital spending, when growth is faster, and when external financing is lower. More importantly, however, and in contrast to the findings for the OECD, they find that fiscal consolidations are more durable when brought about by revenue enhancements rather than expenditure cuts. It turns out that this last result, which has strong policy implications, holds not just for consolidations (i.e. gradient effects) but also for the persistence of the fiscal stance (level effects).

Measuring persistence. Measures of persistence or fiscal sustainability are necessarily arbitrary and always contestable. For example, Alesina and Perotti, (1997) define persistence, following a 1.5% or greater reduction in the primary deficit-to-GDP ratio, as a reduction in the debt-to-GDP ratio of at least 5% three years later. For von Hagen *et al* (2002) consolidation is an episode in which either the cyclically-adjusted budget balance increases by at least 1.25% of (cyclically adjusted) GDP in two consecutive years, or by at least 1.5% in one year and is positive in both preceding and subsequent years. A consolidation is successful if two years after the initial adjustment the budget balance is at least 75% of its value in the first year of the consolidation episode. Illera and Mulas-Granados (2001) and Gupta *et al* (2002) take a simpler approach, defining consolidation as occurring when the cyclically-adjusted budget balance improves (Illera and Mulas-Granados), or improves by at least 1.5% of GDP (Gupta *et al*), and a failure as when this consolidation does not take place. The duration of consolidation episodes is then defined by the interval between consecutive failures.

While von Hagen *et al's* definition embodies elements of the 'level approach' (is the line being held at some appropriate benchmark?), the latter two papers utilize some variant of what we have called the gradient approach (are things continuing to improve?). This does not seem entirely satisfactory in the present context, where we are interested in the maintenance of fiscal control once it has been achieved. Of course, a country with a dramatically high deficit might indeed require a substantial time in which to rectify its fiscal stance. This would then involve an extended sequence of consecutive periods of incremental consolidation before it should be allowed to rest on its fiscal laurels, and the gradient approach would be an appropriate way to study the persistence of this transition effort. However other countries will be

quite close to fiscal sustainability. For them, a reasonably quick step change followed by consolidation would be appropriate, not a continual process of tightening the screws. In any event, a successful transition will eventually lead the country to what appears to be a satisfactory fiscal stance and the interesting question then becomes how long this condition is likely to continue to hold. For these reasons, we prefer the 'level approach' to consolidation. It does have at least two drawbacks, however. First, although we do examine some aspects of a country's fiscal history, for example whether the country has a sequence of previous failed reforms, and the scale of the initial adjustment required to restore fiscal stability, we do not examine the detailed trajectory of the deficit before the threshold is crossed. Second, the level approach necessarily treats as 'failures' periods where significant adjustment may have taken place but the country nonetheless remains below the threshold. In future work, we aim to develop a consistent integration of these two dimensions of fiscal adjustment.

How then do we locate the threshold? In most economic applications of duration analysis, defining alternative states and hence the 'failure event' is simple: death is a well-defined state, as is unemployment, job tenure, legal bankruptcy, conflict, and so on. Matters are much harder when we are dealing with forward-looking concepts such as fiscal sustainability. Any threshold which we propose is therefore going to be arbitrary but the challenge is to define one that is plausible, operationally useful, and testable, or at least amenable to sensitivity analysis. Two issues need to be confronted: the first is the level of the threshold itself, and the second is whether this level should vary across countries and over time. On the first issue, one approach would be to rely on actual practice to define the target sustainable deficit. One obvious reference value would be the 3% of GDP conventional deficit target embodied in the 1997 Growth and Stability Pact of the European Union. Others would include the balanced-budget rules employed at various times by countries such as Indonesia, Zambia, Tanzania and Uganda. A second approach would draw on the literature on growth and fiscal deficits to ask 'over the long-run, what level of the (conventional) budget deficit maximizes growth?' In recent work based on a panel of 45 developing countries, we derive a point estimate for this of 1.5% of GDP with a 95 percent confidence interval ranging from around 4.5% to just short of a balanced budget (Adam and Bevan, 2002). We could therefore select as the threshold alternative points from this distribution. A third, and perhaps the most natural, approach would use the simple arithmetic of debt sustainability to ask 'over the long-run, and given assumptions about the long-run rate of growth and interest rates, what level of the (conventional) budget deficit is consistent with a sustainable level of public indebtedness?' Approaching the issue from this perspective makes clear the instrumental role of the fiscal deficit and emphasises that notions of fiscal sustainability rest ultimately with concerns about the public debt ratio, although it also shifts by one remove the open question of what constitutes a sustainable steady-state level for the latter.

On the second issue, a debt-sustainability perspective naturally implies the threshold will change over time in response to changing long-run growth and interest rates. Specifically, with the slowdown in world growth rates since the mid-1970s and the corresponding rise in real interest rates, at least until the late 1990s, we would anticipate a gradual tightening of the threshold over our sample period. It is not, however, sensible to handle this evolution by simply subdividing the sample. Sample-splitting increases the proportion of observations in the sample that are left- and right-censored, in other words periods of stability which either begin before the start of the sub-period (left-censoring) or terminate after its conclusion (right-censoring). As we discuss in the appendix, the higher the proportion of censored spells of stability in the sample, the less reliable will be our econometric estimates. Moreover, allowing for discontinuities in the threshold at essentially arbitrary points in time raises a serious problem of observational equivalence. Around such points we will be unable to discriminate between episodes of fiscal stability that come to an end because of a genuine loss of fiscal control and those, otherwise stable spells, that end because of arbitrary adjustment of the threshold; statistically both are equivalent but economically they are rather different events.

Nonetheless, debt-sustainability conditions have changed over time and it would be inappropriate not to reflect this in our analysis. In addition to using a fixed threshold for the entire sample, we define a second class of threshold which evolves slowly between two fixed levels of the deficit along a smooth transition path defined by

$$y_t = \alpha_0 D_0 + \alpha_1 D_1 + \beta_0 (1 - D_0 - D_1) \left(\frac{e^{-\beta_1 (t - T_0)}}{1 + e^{-\beta_1 (t - T_0)}} \right) \qquad (1)$$

where where T_o denotes the mid-point of the transition (taken to be 1983), D_0 and D_1 denote the start- and end-points of the transition (which we select as 1975 and 1991 respectively) and α_0 and α_1 denote the pre- and post-transition thresholds.[3]

In settling on an average value for the threshold we confront a trade-off between specifying a threshold which is too loose – in the strict sense that it is inconsistent with what may be regarded as a sustainable public debt stock – and one that is so tight so that it eliminates too many spells of 'reasonable' fiscal performance, leaving as exemplars of sustainable fiscal performance only a small number of (relatively short) spells of stability drawn from a limited number of (disproportionately) successful countries. The core results presented below are based on a fixed threshold, corresponding to a conventional deficit of 3% of GDP, and a smooth-transition threshold which converges on this value from an initial level of 4.5% of GDP in the early 1970s. Both are shown in Figure 1along with two other lines. The first is a tighter threshold which follows the same smooth transition path but this

[3]The parameters of the transition segment, β_0 and β_1, governing the curvature of the transition, are tied down by smooth-pasting requirements between the three components of the curve.

time from a deficit level of 3% to one of 1.5% of GDP. This threshold is used in the sensitivity presented in the appendix. The second, disountinuous, line is not used in the analysis but it helps locate our thresholds in the debt-sustainability arithmetic. It traces an implied set of long-run 'sustainable' deficits for developing countries based on actual real interest rates, proxied by the US long-bond rate, and actual five-year average GDP growth rates, given a steady-state target debt to GDP ratio of 70 percent in present value terms.[4]

*** Figure 1 here ***

Though we feel that our central thresholds are broadly plausible, we recognize that they may be viewed as being too relaxed; both the e 3% fixed deficit threshold and the 70% present-value debt to GDP are at the high end of the range. For example, the target for public indebtedness under the Maastricht Treaty is 60% of GDP, and for countries eligible to debt-relief under the HIPC Initiative, the target present value debt stock is 150% of exports. With exports rarely exceeding 25%-30% of GDP in developing countries this suggests a debt-to-GDP target of around 40%-50% for external debt. Even adding on an additional 10-15% for domestic debt would imply a target little in excess of 60%. We have intentionally chosen to err on the side of leniency here for a very practical reason. For developing countries as a whole, the last 30 years of fiscal history is littered with episodes of poor fiscal discipline so that the unconditional 'raw failure rate' in these data is already rather high.[5] For example, using a fixed threshold deficit of 3% of GDP our data implies a raw failure rate of almost exactly 50% (60% for low income countries). Assuming a present-value debt-to-GDP ratio of 50 percent would cut the sustainable fiscal deficit to around 1.5% GDP for the 1990s (and 3% in the 1970s) and imply a raw failure rate of around 70% percent (and 82% for low income countries). In order to generate a distribution of durations sufficient for the empirical models to gain some purchase on the underlying processes of interest we are obliged to set the threshold rather looser than strict debt-sustainabilty arithmetic might demand. A key element in defending this approach is, of course, the sensitivity of our results to the threshold. As we show in the appendix, where we vary the threshold across country groups, the qualitative import of our results remains intact to reasonable variations in the threshold.

3 Stylized facts on fiscal performance

Our data cover the period from 1970 to 2000 and are drawn from four sources: the Penn World Tables (Version 6.1); the World Bank's World Development

[4] Notice that by using a present value basis for the debt ratio we have accounted for the concessionality in new borrowing. Approximately the same outcome would emerge if we established a debt ratio target in nominal terms and used instead an average cost of capital measure that reflected the concessional element in external borrowing for developing countries.

[5] The raw failure rate is defined as the proportion of time spent below the threshold (see also Table 3).

Indicators (2001); and the IMF's Government Finance Statistics (GFS) and International Financial Statistics (IFS). After excluding the countries of the former Soviet Union and other centrally planned economies, and those for which we have insufficient data we are left with a usable sample of 108 countries, 25 of which are members of the OECD.[6] Of the remaining 83 developing countries, 37 are low-income countries. The sample is not, however, fully representative since the GFS in particular suffers from systematic censoring. Our sample is under-represented in chronically performing low-income countries for whom reporting to the GFS is often an early casualty of economic crisis. It is also under-represented in those who have managed to turn around their fiscal performance but still do not report to the GFS.[7] A full list of the countries used in the statistical analysis which follows is provided in the data appendix.

Before embarking on our duration analysis we briefly descrive three key features of comparative fiscal performance between OECD and developing countries.

First, average fiscal deficits are comparable but fiscal outcomes are much more volatile in developing countries. Table 1 shows the average dispersion of the conventional budget balance by country groups and over time.

*** Table 1 here ***

Fiscal deficits in developing countries have converged towards those in the OECD and were lower on average in the second half of the period. This process of convergence reflects a significant fiscal adjustment over time by middle-income developing countries whose average budget deficit fell by over 2.5 percentage points of GDP between 1970-84 and 1985-2000. Despite this convergence in means, the average within-country standard deviation of the fiscal balance amongst developing countries has remained approximately one full percentage point per annum higher than in the OECD (a difference of almost 50 percent). Average fiscal deficits in low-income countries have shown no tendency to converge and have remained more or less constant at around 5.25% of GDP, significantly higher (and more variable) than those in middle income countries.[8]

Second, fiscal performance is more asymmetric in developing countries. In Table 2 we examine the cyclicality of the fiscal stance in response

[6] Our definition of the OECD excludes the following current members: the Czech Republic, Hungary, Poland, the Slovak Republic, Korea and Mexico but does include Malta.

[7] One such example would be Uganda whose economic recovery has been underway for over 10 years but the authorities still do not report their fiscal data to the GFS. As it happens, we have been able to augment the GFS data from domestic data sources for Uganda and a small number of other countries.

[8] Notice, however, that these figures do not take into account the concessional element in external financing to low-income countries which might otherwise argue for a higher structural fiscal deficit. We examine this possibility in the appendix.

to underlying economic conditions. We follow the approach suggested by Bayoumi and Eichengreen (1995) and utilized by Gavin and Perotti (1997) in their study of fiscal policy in Latin America. We regress the annual change in the fiscal balance on the rate of growth of real GDP, the lagged fiscal balance, and a set of country dummy variables. The coefficient on the growth variable is then taken as a measure of the cyclical covariation of fiscal performance with economic performance; a positive coefficient therefore implies a counter-cyclical fiscal response (i.e. the fiscal stance tighthens in good times and eases as bad times).

*** Table 2 ***

A comparison of columns [1] and [3] highlights the first major difference between the OECD and developing countries. In the OECD, a one percentage point increase in the rate of output growth is associated on average with a statistically significant counter-cyclical fiscal contraction of 0.20 percentage points (column [3]). By contrast the fiscal response in developing countries is almost ten times smaller and is statistically significant at the 10% level. This non-cyclical response is only present amongst the middle-income developing countries(column[5]); for low-income countries, the fiscal response is counter-cyclical but the semi-elasticity is only one third of the size of that for the OECD. The second distinction is that in both the OECD and developing countries, the fiscal response is asymmetric in the face of favourable and adverse macroeconomic developments, but this assymetry is much more marked in developing countries. Splitting real GDP growth between positive and negative outcomes (column[2]), the counter-cyclical fiscal response in the OECD is approximately twice as strong when growth is negative as when it is positive (a one percent decline in output is accompanied by an increase in the fiscal deficit of around 0.32 percent of GDP as opposed to a 0.15 percentage point response to an equivalent positive growth in GDP).[9] For the developing country group we observe significant counter-cyclicality in bad times (although not as strong as for the OECD), but for middle-income countries the response in good times is siginficantly pro-cyclical.[10] Splitting the data in this manner unpicks the finding in column [7] that for the OECD the fiscal response was counter-cyclical on average; the reality is that this number is driven by highly significant counter-cyclicality in bad times (of which there are many) and insignificant counter-cyclicality in good times.

Third, and underscoring the previous point, fiscal stability is much less persistent in developing countries. Table 3 summarizes the data on

[9] These results are invariant to whether we distinguish between positive and negative growth or if we use the Gavin and Perotti (1997) distinction between "good" and "bad" outcomes where the latter denotes growth more than one standard deviation below the mean.

[10] Although estimated over a different sample (of countries and time periods) these results accord with those reported by Gavin and Perotti (1997) for a sample of industrialized and (Latin American) middle-income countries.

episodes of fiscal stability, where stability in this case corresponds to a period where the conventional fiscal balance exceed of the alternative threshold levels defined in Section 2.

*** Table 3 ***

The upper panel reports the average incidence rate (i.e. the proportion of time spent *below* the threshold), the number of episodes of fiscal stability (of all durations), and the mean duration of these episodes at each threshold. The lower panel reports the corresponding (non-parametric) Kaplan-Meier survivor probabilities.[11] At the three percent deficit cut-off countries have, on average, spent about as much time above the threshold as below it. There is no significant difference between the OECD and developing countries as a whole, although within the latter group, low income countries are much less likely to be in a stable configuration. The average length of a spell of fiscal stability is approximately 5 years for the OECD, and almost two years less for developing countries taken as a whole (3.2 years).[12] This large gap appears to be influenced partly by the relatively short mean duration ofepisodes of fiscal control amongst natural resource dependent countries, most of whom are middle-income developing countries (the mean duration is 2.15 years as opposed to 3.6 years for non resource-dependent economies). Conditional on being in a stable configuration, there is no significant difference between low and middle income in the mean duration of stable episodes. The most striking aspect of the table, however, is seen in the Kaplan-Meier survivor functions in the lower panel. At the three percent deficit cutoff, the unconditional probability of an episode of of fiscal stability lasting 3 year is around 0.70 but less than half this for developing countries, while the probability of a spell lasting more than 5 years is less than 0.10.

Changing the thresholds obviously changes the summary statistics – the more permissive threshold in column [2] corresponds to a lower overall incidence rate and longer mean durations and *vice versa* for the tighter threshold – but it does not, in fact, alter the underlying patterns across country groups. However, as we argued earlier, the tight threshold has a devastating effect on the distribution of durations in the sample, especially for developing countries.

[11] The survivor function is estimated non-parametrically as

$$\hat{S}(m) = \prod_{l=1}^{m} \frac{n_l - e_l}{n_l} \qquad m = 1, .., l, .., M-1$$

where n_l is the number of countries 'at risk' over the interval defined by l and e_l the number whose cyclically adjusted fiscal balance falls below the threshold in that interval. The survivor function therefore measures the probability of a spell of fiscal stability persisting to m years under the assumption that multiple spells of stability by a country are independent. This latter assumption is relaxed in the econometric analysis in Section 4.

[12] In both these cases the mean is biased downwards by the presence of right-censored observations on countries that were still in a stable configuration at the end of the sample (or the end of our data for them).

The mean survival duration is just over two years, but as the Kaplan-Meier survivor functions show the distribution is highly skewed: the unconditional probability of a spell surviving beyond three years is only 16 percent for middle income countries and only 12 percent for low income countries. Given that our data are annual, this rapid fall off in the unconditional survival probability seriously undermines the discriminatory power of the duration models we use in the next section.

The key implication we take from these three stylized facts is that while broadly plausible for the OECD, the assumption of a cyclically-determined but broadly stationary fiscal stance is the exception rather than the rule for developing countries as a whole over the last three decades. For them, on average, fiscal policy does not follow a standard cyclical pattern. Rather it is highly assymetric, vulnerable in the face of adverse output shocks, and not particularly buoyant in the face of positive shocks with the consequence that extended periods during which a broadly sustainable fiscal stance is maintained are comparatively rare. In the next section we employ the methods of duration analysis to further understand these patterns in duration, both between and within country groups. We start with a brief description of our empirical approach.

4 Duration analysis

A natural approach to analyzing the duration of spells of fiscal stability in circumstances where countries typically experience multiple periods of stability is to specify and estimate models of the hazard function.[13] Define the duration (or survivor) function, i.e the probability that the current spell of fiscal stability, denoted T, exceeds some duration t, as $S(t) = \Pr(T \geq t)$. Then, defining its counterpart, the failure function, as $F(t) = 1 - S(t)$, the density of which is $f(t) = dF(t)/dt$, we derive the hazard as

$$\lambda(t) = f(t)/S(t) \tag{2}$$

which measures the rate at which spells of fiscal stability terminate in the interval $t + h$ given that they have survived as long as t. In the language of duration analysis, this termination is referred to as a 'failure'. The hazard can then be parameterized using the class of proportional hazard models of the form:

$$\lambda(it) = \lambda(t; \mathbf{x}(it), \mu(i)) = \mu(i)\gamma[\mathbf{x}(it)]\lambda_0(t) \tag{3}$$

where $\lambda_0(t)$ is the baseline hazard (assumed to be common across countries), $\mathbf{x}$ is a vector of covariates, some of which vary over time, and $\mu(i)$ is an unobservable random country-specific effect.[14] The basic intuition underpinning

[13] If we were analyzing only single spells of fiscal stability (or could assume that multiple spells for a country were independent), and if the explanatory variables were all time-invariant, we could equivalently model the duration of spells directly, using censored regression or Tobit methods.

[14] Consistent estimation in the presence of time-varying covariates requires that there is no feedback from the duration of the spell itself to the conditioning variables. For most

the proportional hazard model is that the vector of covariates serves to modify, or shift, the baseline hazard function. Factors that shift the hazard in a positive direction are associated with a higher risk of (instantaneous) failure and hence a shorter duration for the spell of fiscal stability. Following standard practice, we assume that the factor of proportionality takes the specific functional form $\gamma[\mathbf{x}(it)] = \exp[\mathbf{x}(it)\boldsymbol{\beta}]$, where the β coefficients can be interpreted as the semi-elasticities of the hazard.

Three features of our empirical strategy set it apart from the existing literature in this area. The first is the form of the underlying baseline hazard, $\lambda_0(t)$. A basic assumption is that the baseline hazard is constant, so that $\lambda_0(t) = \lambda$ and the instantaneous probability of failure is independent of how long the spell of stability has already lasted. This assumption can be tested as a restriction on a specific functional form on the hazard, with the Weibull, log-normal or gamma distributions being the most common, depending on assumptions about the likely shape of the hazard.[15] Von Hagen *et al* (2002), for example, assume a (monotonic) Weibull hazard of the form $\lambda_0(t) = \alpha t^{\alpha-1}$ which allows them to directly test the hypothesis of 'consolidation fatigue', the idea that, controlling for other factors, maintaining a fiscally prudent regime gets harder or easier with time. Fatigue in this context simply corresponds to case where the baseline hazard increases with the duration of the current spell of stability ($\alpha > 1$) where as decreasing duration dependence ($\alpha < 1$) is consistent with notions of stability becoming embedded over time.

Here we use the more flexible piecewise-constant hazard (Lancaster, 1990, Chapter 8). By dividing the total duration time into M intervals defined by the cut-off points $c_1, c_2, ..c_{M-1}.$, and constructing a corresponding set of dummy variables

$$d_m(t) = \left\{ \begin{matrix} 1 & \text{if } c_{m-1} \leq t < c_m \\ 0 & \text{otherwise} \end{matrix} \right\} \qquad m = 1, 2..M$$

so that $d_m = 1$ if t lies in the m^{th}interval, the piecewise-constant baseline hazard can be defined as

$$\lambda_0(t) = \exp\left\{ \sum_{m=1}^{M} \lambda_m d_m(t) \right\}. \tag{4}$$

This specification thus assumes the hazard to be constant (of magnitude λ_m) within each interval of time defined by m, but variable across intervals. Since there is no restriction on the length of the intervals (since in the limit $m \to \infty$

of our covariates any potential feedback is avoided since we take start of period values of the variables. Within our set of time-varying covariates, the measure of conflict and the terms of trade can reasonably be assumed to satisfy this exogeneity condition. For the rate of growth, however, the condition may be less likely to be satisfied, although the contemporaneous feedback from the fiscal stability to growth is likely to be weak.

[15] In circumstances where the form of the baseline hazard itself is not of specific interest, Cox's proportional hazard model provides an alternative estimation method (for example Gupta *et al* (2002)).

as the hazard tends to the constant exponential form), nor any requirement that the intervals to be equal, this flexible form nests any of the parametric hazards, whether they be monotonic as in the case of the Weibull, or non-monotonic as in the case of the log-normal or gamma parametric distributions. In the models reported below, we define the intervals as spells of two years starting from $t = 0$, concatenating all durations greater than $t = 10$ into a single category. This form of piecewise-constant hazard marginally dominates alternative parametric characterizations of the hazard, although it can readily be restricted to a standard one-parameter Weibull hazard. However, as we show below, by using the piecewise-constant formulation and exploiting this equivalence we are able to develop a clearer understanding of the role of past fiscal failures on the likelihood of current stability.

Second, we may wish to allow the parameters on the covariates themselves to vary with time. Of particular interest in this case is the parameter on our measure of fiscal history. We measure a country's fiscal history as its past propensity to failure (i.e the number of times it experienced a loss of fiscal control over the time under observation) up to the beginning of the current spell of stability. A natural question is whether the effect of this legacy is constant over the duration of a spell or whether its effects can be gradually erased as a result of an evolving track record of stability. To investigate this we make use of the piecewise-constant structure used to define the baseline hazard. Letting $h(i)$ denote the fiscal legacy for country i at the beginning of a spell (and the remaining covariates $\mathbf{z}(it)$), we can re-express the factor of proportionality in (3) as

$$\gamma[\mathbf{x}(it)] = \exp[\mathbf{z}(it)\boldsymbol{\beta} + \sum_{m=1}^{M} h(i)\delta_m d_m] \tag{5}$$

where the indicator variables, d_m are defined as above, and δ_j represents the parameter on the effect the fiscal legacy in the j^{th} interval. It is then a simple matter to test the restriction that $\delta_j = \delta \ \forall j \in 1..M$. In principle, of course, this time-varying structure can be applied to any or all of the covariates. Although not reported, we can readily accept the null hypothesis that the effects of the other key time-varying covariates are time-invariant.

The final feature of our model concerns unobservable heterogeneity. Given the limitations in our data and the necessary incompleteness of the underlying model, unobserved heterogeneity is bound to be an important feature of our estimated models. For example, as noted earlier, we do not control for measures of political structure across our countries. We might reasonably expect, as the literature on fiscal consolidation in the OECD suggests, that different political structures are more or less able to 'lock-in' fiscal reforms. Failure to account for this, and any other heterogeneity across countries, may lead to biased estimates of both the covariates and the estimated duration dependence (see Lancaster 1990). To handle this (3) also includes the latent variable $\mu(i)$ to capture time-invariant unobserved heterogeneity, which

is assumed to enter the hazard function multiplicatively. Along with the additional assumption that $\mu(i)$ is common across all failures for country i and independent of the covariates, the heterogeneity is handled analogously to a random effect in standard panel data models. Estimation of (3) requires a specific distribution for $g[\mu(i)]$. Here we again follow standard practice and assume the heterogeneity follows a Gamma distribution. Normalizing the expected value of the heterogeneity such that $E(\mu(i)) = 1$ this implies that the average hazard in the presence of gamma-distributed heterogeneity is $\lambda(it) = \lambda(t; \mathbf{x}(it), \mu(i)) = \gamma[\mathbf{x}(it)]\lambda_0(t)$ and we have only one parameter, the variance of the distribution of $\mu(i)$, to estimate. Denoting this by θ, the presence of heterogeneity can then be tested directly as the restriction $H_0 : \theta = 0$.[16]

5 Explaining the persistence of fiscal control

We now turn to the evidence, starting with the role played by structural characteristics. From there we explore how the history of previous fiscal failures and initial conditions prevailing at the beginning of a spell of fiscal stability shape its duration before finally focussing on the question of how the composition of adjustment influences its duration. The full set of conditioning variables is described in the data appendix and the vector of significant variables is summarized in Table 4, split according to whether the fiscal balance exceeds or falls short of the 3 percent of GDP fiscal deficit threshold. With the exception of the log of real income, all the variables are either dummy variables, index numbers, or raw percentages so that the coefficients measure the semi-elasticity of the hazard.[17] [18]

*** Table 4 ***

In the following tables we concentrate on the results for the full set of developing countries, contrasting these results with the sample of OECD sample and distinguishing between middle- and low-income developing countries as required. Further results detailing the sensitivity of the main findings to alternative definitions of the threshold are relegated to the appendix.

5.1 Economic structure and persistence

Our first model tests the role of the level and strucutre of economic activity. The vector of covariates consists of a dummy variable for natural resource

[16] An alternative non-parametric and less restrictive strategy for estimating the unobserved heterogeneity is suggested by Heckman and Singer (1984)

[17] We maintain the convention throughout that a value of one percent is recorded as 1.00.

[18] The Weibull and piecewise exponential hazard models estimated here have an equivalent accelerated failure-time (AFT) representation in which the coefficients would then represent the marginal effect of the covariate on the (log of the) survival time. As reported in Table 3 above, the mean survival duration for all developing countries is approximately 3 years. Re-estimating the models in Tables 5-7 under the AFT representation, suggests that a unit change in a covariate with an estimated coefficient of 0.10 corresponds to an increased duration of around 4 months.

dependency, the agricultural share in GDP, the openness of the economy to international trade, the overall size of government, and income inequality, plus the level of per capita income. For these variables we use their average value over the period leading up to the (current) spell of fiscal stability. For countries with repeated spells of stability, these covariates will change between spells but will be constant over the duration of each spell. We also include as time-varying covariates a dummy variable for the presence of conflict and the terms of trade.

*** Table 5 ***

Columns [1] and [2] of Table 5 report the results from this first model under the assumption that the baseline hazard follows a Weibull distribution; in column [2] we also control for potential unobserved country-specific heterogeneity. It is worth commenting briefly on three statistical features of the model before turning to the results themselves. First, as indicated by the likelihood ratio test for the restriction $\theta = 0$, there is, as we expected, evidence of highly significant unobserved heterogeneity. This is present regardless of the subset of countries over which we estimate the hazard functions. Henceforth, therefore, we only report results where we have controlled for unobservable heterogeneity. The second feature, which again is present across all sub-samples, is that the data imply positive duration dependence, which is consistent with the notion of consolidation fatigue described by von Hagan *et al* (2002). The Weibull α parameter of 1.14 implies that after 5 years the risk of failure, other things equal, is around 43 percent higher than after one year. Looking ahead to the model estimated under the time-varying threshold (column[6]), this point estimate rises to 1.21 which corresponds to a 70 percent risk elevation after five years. Even more striking are the results for the low income countries (from column [8]) and the OECD (column [9]) which imply the presence of extremely strong consolidation fatigue (in the case of the latter, the autonomous risk of failure after 5 year is almost 25 times higher than at the beginning of the spell!). As we shall show shortly, it would appear that this result actually reflects an incomplete specification of the model; once we take into consideration the role of initial conditions, the positive duration dependence disappears, at least for the developing countries. Finally, even though in most cases the vector of covariates contribute to a statistically significant shift in the hazard, the low pseudo-R^2measures indicate that a large amount of the variation in individual durations remains unexplained. This is a common feature of duration models; unfortunately, since none of the papers reviewed above actually report goodness of fit measures it is not possible to tell whether the fit we find here is in line with the literature.

Turning next to the role of the covariates, we consider first the developing country sample (column [2]). For this group of countries, and conditional on the baseline hazard, most of the structural characteristics (mineral

dependency, openness to trade, the agricultural share in GDP and the measure of inequality) are both statistically insignificant and have an extremely weak effect. For example, the coefficient on the openness of the economy ($\beta = 0.0016$) implies that a country whose openness to external trade is 10 percentage points of GDP higher than the average country, faces a risk that a spell of fiscal stability will collapse over the following year that is only 1.6 percent higher. Similarly, even quite large movements in the Gini coefficient matter very little for the hazard. The only structural characteristics that do shift the baseline hazard to any measurable degree are the average size of government (which increases the hazard) and the level of real income (which decreases it). For a country with a government which is, say, 10 percentage points larger than the mean, a difference of approximately one standard deviation in this sample, the hazard is roughly 13 percent higher, other things equal, while a doubling of real per capita income would reduce the risk that a spell of stability ends by as much as 37 percent.

The two time-varying covariates are also moderately important, at least for developing countries taken together. The outbreak of conflict (at a time of fiscal stability) raises the risk of fiscal collapse by around 24 percent, while a ten point improvement in the terms of trade (roughly a 10 percent increase) reduces the risk of collapse by approximately 2 percent.

One surprising result from columns [1] and [2] is the counter-intuitive sign and lack of significance associated with the dummy variable for natural resource dependency, regardless of whether we control for individual country heterogeneity. This may reflect the manner in which we have defined this characteristic, although using alternative measures including continuous variables for the oil/mineral share in production or exports does not alter this result. To check whether this arises from a more general missspecification we re-estimate the model for developing countries splitting the sample between resouce-dependent and non-resource dependent economies. As the pooling LR test suggests, the two country groups are distinct, but it would appear that this mainly reflects large difference in the effects of conflict and the terms of trade. Continued fiscal stability in resource-dependent economies is more than twice as vulnerable to adverse terms of trade movements and to the outbreak of conflict compared to the non-resource dependent group. These findings make good sense: in resource dependent economies fiscal structures are often highly dependent on royalties or export taxes and hence particularly vulnerable to movements in the terms of trade or to the outbreak of conflict which, itself often centres on competition for control of the fiscal 'prize' (see, for example, Collier and Hoeffler (2002)). Given the relatively small size of the sub-samples we re-aggregate the countries but use the results in columns [3] and [4] to interact the natural resource dependent dummy variable with the terms of trade and conflict variables and carry these through the remainder of the analysis.

The next step is to address the question of whether these results are driven solely by the fixed deficit threshold. In columns [5] and [6] we re-estimate the

model, including the new interactive terms for conflict and the terms of trade, for the fixed 3% deficit threshold and our first smooth transition threshold. Since this threshold is lower throughout the early part of the sample period we naturally observe a lower raw failure rate (47% as opposed to 54%) and a slightly higher pseudo-R^2 but there is no significant difference in the results. To keep matter manageable, we restrict the discussion in the remainder of the main text paper to the results based on this threshold, relegating to the appendix a further discussion the sensitivity of the results to the choice of threshold.

Finally, we split the developing country sample between low- and middle income countries and also report the same model estimated over the OECD sub-sample (columns [7] to [9]). The results in column [8] for the low-income countries are particularly striking; with the exception of the terms of trade, neither of the structural characteristics, nor the level of per capita income, can explain the variation in the duration of spells of fiscal stability within this group of countries. As indicated by the LR test of $\beta = 0$, the model has no statistical significance beyond the baseline hazard (which also explains the fact that the across-group pooling restiction is accepted in this case). Part of the reason for the failure of the model over this sub-sample is that there are very few spells of any significant duration, but part would appear to be due to the fact that there appears to be much less variation in economic structure across low-income countries.[19] Once we control for policy choices, however, we begin to get a clearer picture of what determines fiscal sustainability in these countries.

Structural factors explain marginally more of the variation in the duration of fiscal stability in the OECD.[20] We note that amongst these countries the hazard falls proportionally with income and with the openness of the economy to trade (each percentage point increase in the trade share lowering the risk by one percent) although, as expected, changes in the terms of trade are much less important than in developing countries. And in contrast to developing countries, it would appear that spells of fiscal stability are easier to sustain the lower the degree of income inequality. Again, this is a potentially interesting result, in its own right (although see Perotti (1998) for a discussion of how partial is the current understanding of the relationship between forms of social fragmentation and the political economy of fiscal consolidation), and in terms of the contrast with the developing country sample. Unfortunately, however, concerns over the accuracy of the data on income inequality in developing countries mean that we cannot rule out the possibility that identification of

[19] Notice also that there is much less unobserved heterogeneity in this sample; although still statistically significant, the variance of the individual heterogeneity (denoted by θ) is only one third of that for the middle-income countries.

[20] We must be careful, however, to check the scale of the effects. For example, the coefficient on the role of the agricultural share is around three times larger than for the middle-income developing countries. But agriculture is a similar order of magnitude less important in the OECD so that the two effects more or less cancel out. A similar argument applies in reverse for the coefficient on real per capita income.

the true effect is thwarted by problems of measurement error.

5.2 Initial conditions, fiscal history, and external arrangements

We now broaden our analysis to examine how economic performance and the policy environment influence the likely duration of periods of stability. These results are reported in Table 6 and are built up as follows.

*** Table 6 ***

First, we retain the significant covariates from Table 5 (the size of government, the level of real per capita income, the terms of trade and the dummy variable for the presence of civil or inter-state conflict). Their contribution to the hazard does not change significantly in the extended model. Hence the presence of conflict, a large government, and adverse terms of trade movements all increase the risk of failure, while richer countries face a significantly lower risk of failure, given that they are already in a stable configuration. To these we add three groups of covariates. The first, which describes dimensions of country's fiscal history, includes: the level of public indebtedness, which for developing countries we partition between domestic and external public debt; a measure of the country's recent inflation history, for which we employ a dummy variable which takes the value of 1 if the country's inflation exceeded 40 percent (20 percent for OECD countries) at any point in the five years preceding the current spell of fiscal stability;[21] and the number of times a country has suffered a loss of fiscal control prior to the start of the current spell (of stability).[22] This third measure could take a number of interpretations, the most obvious being that it picks up some of the unobserved heterogeneity otherwise captured by $\mu(i)$, including variations across countries in the degree of politcal commitment to fiscal reform, or differential checks and balances on the fiscal authorities' autonomy. The final element of this group is a measure of the initial fiscal adjustment precipitating the current spell of adjustment. The second group of variables reflect the external environment in which a country operates. This is characterized by the nominal exchange rate regime, where we use the Reinhart and Rogoff (2002) index of exchange rate arrangements, and whether the country has an IMF adjustment programme in place (and hence operates under some externally monitored policy conditionality). The final covariate included at this stage is the contemporaneous rate of growth of per capita GDP.

The legacy of previous failures It makes sense to start once more with the baseline hazard. The first point to notice in column [1] is that as a result of adding this vector of initial conditions and policy variables the strong

[21] The 40 and 20 percent cutoffs were chosen after considering both the average inflation rate itself as well as alternative thresholds.

[22] This variable is normalized by the 'time at risk' and is thus bounded between zero and one. Similar qualitative results obtain if we define this variable simply as the number of past failures.

positive duration dependence observed in Table 5 is reversed so that over time the hazard declines monotonically; in other words good fiscal performance is self-reinforcing. If true, this is an important and encouraging result, but one that requires unpicking. It would appear that this result is directly linked with our measure of the history of previous fiscal failures. As we noted above, it is natural to assume that a country's fiscal history will matter for the likely duration of the current spell of stability. As the results in column[1] indicate this is borne out by the data; a country with a fiscal history one standard deviation worse than the mean will face a 40 percent higher probability *at every point in time* that the current spell of stability will come to an end in the next period given that it has survived to this point, other things equal.[23] As we argued above, however, it is not obvious this history should have a 'permanent' effect on duration. A more plausible view is that ongoing fiscal stability should amortize the legacy of history as governments develop reputations for fiscal management. We can investigate this possibility by allowing the semi-elasticity of the past fiscal history be time-varying and test whether there is evidence of decay. In column [2] we re-estimate the model according to (5) where we allow for both a piecewise constant baseline hazard and for fiscal history to have a time-varying semi-elasticity. The semi-elasticities reported as $history[m1]$ to $history[m5]$ display a statistically significant monotonic decline with the duration of the period of stability.[24] Hence, the adverse effects of a poor fiscal history diminish over time (by around 25 percent after four years of stability and by 60 percent if stability remains intact for a decade[25] But having allowed for this effect, the previous negative duration dependence disappears; the likelihood ratio test against the constancy of the baseline hazard indicates that the null of zero duration dependence is now accepted [Pr=0.789].

This pattern does not repeat for our OECD and low-income sub-sample, although for rather different reasons. For low-income countries, we can again restrict the baseline hazard to a constant form, but while there is some evidence of decay in the effect of fiscal history it is much weaker and no longer statistically significant. The effect of a poor fiscal history decays by less than 10 percent over the first four years and by only 30 percent if a stable environment is maintained for a decade or more.[26] For the OECD countries, we observe a strongly declining, but less significant, effect of the legacy of past failures. But this time we can no longer assume a constant baseline hazard, which is rejected in favour of a significantly positive hazard. As von Hagan *et al* (2002) also found, the OECD countries appear to exhibit a degree of fiscal fatigue. But it is important to note that in our case this result may have

[23] The mean value for fiscal history is 0.33 with a standard deviation of 0.32. The marginal effect of a one standard deviation deterioration in fiscal history is then 1.236*0.32= 0.395.

[24] The likelihood ratio test against the null that the individual piecewise coefficients for fiscal history are equal is decisively rejected [Pr=0.038].

[25] Computed as 1-(1.069/3.283).

[26] The absence of a coefficient at $m2$ reflects the fact that the estimated semi-elasticities at $m1(0-2yrs)$ and $m2(2-4yrs)$ were statistically indistinguishable.

less to do with the political economy explanations advanced by von Hagan *et al* than with the countercyclical pattern of fiscal outcomes in the OECD we noted in Section 3. In such circumstances, for a levels threshold that lies within the normal amplitude of the fiscal cycle hazard models of this type will necessarily tend to exhibit positive duration dependency.[27]

Initial conditions and growth. The legacy of past failures aside, the conditions under which a country embarks on a period of fiscal control matter strongly for its likely duration. The two factors that matter most are a country's recent inflation history, which makes fiscal control harder to sustain across all countries but especially middle-income countries,[28] and the size of the initial fiscal adjustment, measured as the (absolute) distance between the threshold and the average fiscal balance in the three years prior to the start of the current spell of stability. As with Gupta *et al* (2002), we find that the larger the initial adjustment, the more durable is the subsequent spell. This effect is more pronounced in low-income as opposed to middle-income developing countries although even then the quantitative effect is relatively weak; each additional one percentage point of initial disequilibrium translates into a reduction in the hazard of only 10 percent. In the OECD, by contrast, the effect is very strong, with each percentage point of initial adjustment reducing the hazard by closer to 60 percent. Part of this difference in the size of semi-elasticities is due to the fact that the average disequilibrium (relative to the threshold) is much smaller in the OECD than in other countries. Nonetheless the result is important and is consistent with Perrotti (1998) claim that the bigger political cost associated with radical fiscal adjustment means that if such adjustments are observed they are likely to be more successful.

The more surprising result is that public indebtedness appears not to shift the hazard at all for developing countries, even though public debt might be thought of as encapsulating a country's entire fiscal history. Fiscal stability in developing countries is certainly harder to sustain the higher the level of domestic public debt but this effect is very weak, and not statistically significant. Each percentage point increase in domestic indebtedness raises the risk of termination by only one quarter of one percent. External indebtedness has a similarly weak effect quantitatively and is negatively signed for the developing country sample as a whole. Only in the case of OECD countries do we find a statistically significant effect of higher domestic debt in raising the risk of failure. These results on public debt are rather puzzling. Not only are they weak in general but for the OECD they stand in contrast to the finding of von Hagen *et al* (2002) that the likelihood of consolidation persistence was *increasing* in the debt level, although not significantly so. One possible explanation is that the fiscal history variable, and indeed the inflation history,

[27] Though not reported here, a simple demonstration of this argument is that as we relax (tighten) the deficit threshold for the OECD countries the degree of positive duration dependence weakens (increases).

[28] With very few such episodes the 40 percent inflation cutoff is not significant for the OECD, although the 20 percent cutoff is.

measures the same thing as the debt stock, namely the absence of fiscal control in the past. This possibility is not borne out by the data; eliminating either or both leads to a marginal increase in the effect of the debt stocks but not significantly so. The same result may, of course, arise from offsetting forces; high debt stocks raise the fiscal stakes and make current fiscal control both more important and also more difficult.

As expected, we find that, in general, higher contemporaneous GDP growth forestalls the risk of fiscal failure, but that this effect is strongly increasing in the level of per capita GDP. The beneficial effects of growth are large in the OECD, with a one percentage point per annum increase in the rate of growth lowering the hazard by around 60 percent at each point in time during a period of fiscal stability. The same growth improvement eases the hazard by only 14 percent in middle-income countries, and by less than half this amount in low-income countries. At first sight these effects seem surprisingly weak for developing countries – one might have expected growth to play a much stronger role in underpinning good fiscal performance – but a moment's reflection reveals that these results are entirely consistent with the evidence from Section 3 which suggested that on average the fiscal deficit in developing countries was much less sensitive to changes in GDP compared to the OECD, especially in good times.

External conditions and the exchange rate regime. Finally, in this section, we turn to the role of the external policy environment. Here our first finding is that when an IMF adjustment program is in place at the beginning of spell of fiscal consolidation this has a weakly beneficial effect on the duration. Holding other factors constant, the presence of a Fund program reduces the hazard by between 10 and 20 percent for low- and middle-income countries respectively. The effect is, however, statistically insignificant. This result is rather robust, in the narrow sense that the modest size and lack of statistical significance appears to be true regardless of how we measure the effect of Fund programs. For example, the same conclusion emerges if we differentiate programs by type, or measure the number of years a Fund program had been in place, or control for whether a program terminated in the course of a spell of stability. This lack of precision is extremely common in empirical work aimed as assessing the the impact of IMF-supported programs (see for example the review by Bulir and Moon (2003)) Directly testing for the role of Fund programs on fiscal deficits for a sample of 112 countries during the 1990s they find that their measures of IMF program performance were statistically insignificant "although the signs of these parameters are intuitive [and]...tell a relatively consistent story; more structural conditions were associated with below-average performance, while their implementation was associated with better-than-average performance [Bulir and Moon (2003), p23].

A more precise set of results emerge when we examine the role of the exchange rate regime. The traditional argument advanced in favour of a fixed exchange rate regime is that the need to honour the exhange rate peg induces greater fiscal discipline. Tornell and Velasco (1995) develop a con-

trary position which argues that the delayed inflation consequences of lax fiscal policies are delayed under a fixed exchange rate regime creates an incentive for higher fiscal spending. Gavin and Perotti (1997) find support for this view for a sample Latin American economies over the period from 1970-95, but found no significant link between the exchange rate regime and fiscal performance for their sample of industrialized economies. Our findings are consistent with theirs. Using the Reinhart and Rogoff (2002) index of exchange rate arrangements which runs from 1 (denoting a fixed peg) to 6 (a free float) with intermediate values denoting various degrees of managed float,[29] we find that middle-income developing countries embarking on spells of fiscal stability with a more flexible exchange rate regime will tend to face a lower risk of failure. The mean value for our sample is just under 2 (corresponding to a managed crawl). Compared to this a country with a fully floating exchange rate regime, denoted by a 4 on the Reinhart-Rogoff scale, would enjoy a 15-20 percent reduction in the hazard. By contrast, the effect is not significant either for the OECD (again consistent with Gavin and Perotti) or for the low-income countries.

5.3 The impact of composition on the persistence of fiscal consolidation

An improvement in the conventional fiscal balance can be brought about through expenditure reduction, enhanced domestic revenue mobilization, an increase in grant financing, or some combination of all three. In Table 7 we report the impact of alternative forms of fiscal adjustment precipitating in a period of fiscal stability on the duration of the subsequent spell, for the hazard function defined in Table 6.

*** Table 7 ***

To isolate the effects of the composition of adjustment we retain the same piecewise constant hazard and time-varying fiscal history plus the full vector of covariates, including the controls for the terms of trade and the overall size of government. Fiscal adjustments are measured as percentage points of GDP so that the semi-elasticities are directly comparable across groups and between forms of adjustment. For convenience, and since the covariates on the other coefficients do not change significantly between the two models, we report only the coefficients corresponding to the fiscal adjustment. Increased grant financing reduces the risk of collapse but the effect is weak and statistically insignificant in general; a one percentage point increase in grant finance reduces the risk of failure by around one percent for all developing countries.[30] The semi-elasticity is much higher for middle-income countries

[29] To be precise, we aggregated this index further. Categories 4,5 and 6 in the Reinhart-Rogoff classification correspond to "freely floating" (4), "freely falling"(5) and "dual market in which parallel market data are missing"(6). We aggregate these into a single group.

[30] The recording of grants-in-aid in the GFS is often rather arbitrary and we cannot discount the possibility that these weak results also reflect measurement errors.

but this must be set against the fact that for these countries grant financing averages less than 1 percent of GDP compared to 3.6 percent of GDP in low income countries.

What really matters for developing countries, and especially for low-income countries, are fiscal adjustments brought about by improved domestic revenue mobilization. These significantly prolong the duration, both in their own right and, for low-income countries, compared to equivalent reductions in government expenditure.[31] For middle-income countries embarking on a period of fiscal stability, every additional percentage point of GDP either added to domestic revenue or cut from total expenditure reduces the probability of failure by around 15 percent. For low-income countries, on the other hand, a one percent improvement in revenue mobilization reduces the hazard by almost 60 percent, while a corresponding expenditure reduction contributes to a decline in the hazard of only 20 percent. Part of the reason for the large differences in semi-elasiticities is again to do with differences in scale –given the higher average level of grants, domestic revenue is around two percentage points lower in low-income countries. Bur even accounting for this, these results reflect deeper processes at work. First, the fact that fiscal adjustments engineered through (aggregate) expenditure cuts do not persist is entirely consistent with the literature on adjustment failures and repeated adjustments. We know that expenditure reductions tend to be easily reversed and those that are sustained tend to be concentrated in areas such as maintenance and investment expenditure which may well leave public finances less rather than more sustainable in the future. By contrast, sustained improvements in domestic revenue mobilization often requires substantial institutional reform and political committment, both of which may arguably be less easily reversed; what these results show is that the return to this effort is significantly more stable fiscal performance. For our OECD countries the pattern is rather different; as von Hagan et al (2002) find, it is expenditure reductions rather than revenue enhancement that contribute most to sustained fiscal balance and this confirms the findings of other work on OECD countries.

6 Conclusions

The results presented in this paper provide considerable food for thought. Although a substantial amount of the variance in the persistence of spells of fiscal stability remains unexplained, and critical dimensions of the problem – principally the political and institutional determinants of fiscal control in developing countries – are yet to be explored, a number of important empirical results do emerge.

From a methodological perspective we show that flexible-hazard duration models estimated over multiple-failure data offer an empirically useful method for addressing the question of fiscal sustainability. We derive a criterion for

[31] Because of limitations with the fiscal data we have not yet been able to examine in more detail the specifics of fiscal adjustment (i.e. between different types of taxes and expenditure) as is done by von Hagen *et al* (2002) and Gupta *et al* (2002).

sustainability that is rooted in conventional debt sustainability arithmetic and show that the qualitative insights this approach delivers are reasonably robust to plausible variations in this threshold.

From an empirical perspective three key results stand out, all of which have an important implications for 'post stabilization' and deserve further investigation. First, the persistence of periods of fiscal stability and their determinants differ between OECD and developing countries, but much more markedly between middle- and low-income countries. Little of this difference can, apparently, be explained simply in terms of structural features of the economies, although richer countries generally face better prospects, while fiscal stability in natural resource dependent economies is significantly more vulnerable to terms of trade movements and the adverse effects of conflict than elsewhere. But the types of structural feature that often emerge as important explanatory variables in empirical work (the composition of GDP, openness to trade, and income equality) appear to have little impact on persistence. An exception is size of government, as measured by the ratio of expenditure to GDP, which is inversely related to persistence. The evidence also gives some support to the view of Tornell and Velasco (1995) that fiscal discipline is better supported by flexible rather than fixed exchange rates, although on this evidence these effects are relatively weak across all developing countries: they do not suggest that arguments about choice of exchange rate regime should hinge on different regimes' impacts on fiscal discipline.

Second, and in contrast, fiscal prospects depend quite heavily on aspects of a country's fiscal history. A recent history of inflation (but not public debt) and a track record of poor fiscal management act as a drag on fiscal performance. As important, however, is the finding that, for OECD and middle-income countries, as stability endures the adverse effects of a poor track record depreciate; good fiscal policy is, in effect, self re-inforcing. Though traces of the same effect are present for low-income countries, the half-life of a poor fiscal history is significantly longer for this group; radical discontinuities in economic management aside, goverments in low-income countries would appear to have to hold their feet to the fires for much longer to erase the legacy of past fiscal indiscipline.

Third, persistence is more easily attained in a context of growing incomes but only weakly so for developing countries and virtually not at all for low-income countries. Finally, the composition of the initial fiscal consolidation, is an important determinant of how sustained the period of stability is likely to be. Whereas expenditure reductions are central to fiscal consolidation in the OECD, the durability of the fiscal stance in developing countries rests much more heavily on revenue increases, particularly in low-income countries.

Appendix: Statistical properties of the duration models

In this section we examine aspects of robustness of the empirical results presented in the main text. We focus on three issues: (i) the statistical properties of the estimated model; (ii) the sensitivity of the results to our

definition of the threshold; and (iii) the problem of left-censoring.

i) Residual diagnostics A standard methods for assessing the statistical validity of durations models is to examine the Cox-Snell residuals, defined as the change in the integrated hazard for country j at time t.[32] With multiple-spell data it is simpler to transform these to a form which is more readily amenable to visual inspection. We first define a new measure

$$m_{jt} = f_{jt} - c_{jt}$$

where f_{jt} is an indicator variable taking the value one if country j fails at time t and c_{jt} is the corresponding Cox-Snell residual. Hence m_{jt} measures the difference at each point in time between actual failures in the data and predicted failures based on the model. The Cox-Snell residuals are strictly non-negative; hence the m_{jt} residuals are bounded above by 1 (when the model predicts no change in the hazard but there is a failure ($f_{jt} = 1$ but $c_{jt} = 0$) and $-\infty$ (where the model predicts a change in the hazard,$c_{jt} > 0$). A final re-scaling generates a set of residuals that are symmetric and unbounded around zero. The rescaled residuals, denoted D_{jt} and known as 'deviance residuals' are defined as

$$D_{jt} = sign(m_{jt})[-2(m_{jt} + f_{jt}\ln(f_{jt} - m_{jt}))]$$

These can then be interpreted in a manner analogous to the residuals from a linear regression. Positive values of D correspond to cases where the model under-predicts a country's likelihood of failure (given their characteristics) and *vice versa* for negative values. We plot the deviance residuals from the core model (Table 6, column[1]) below, first against time to investigate whether there are any fundamental shifts in the performance of the model over time (Figure A1), and then by country (Figure A2) in order to identify outliers.[33] A number of features are immediately apparent. First, there is no evidence of systematic change in the distribution of the residuals over time suggesting that the core model's predictive power does not significantly improve or deteriorate for more recent episodes. Second, outliers exist on both sides of the distribution; there is no obvious bias in the excess under- or over-prediction of actual failures. Third, as Figure A2 indicates, the systematic over-prediction of failure is concentrated in a very small number of countries. The negative outliers (where fiscal performance was better than the model predicted) correspond to Chad (1973-83), Congo-DRC (1980-83), Yemen (1999) and Uganda (1997-99). The large positive outliers, where the model under-predicts fiscal failures correspond to Kuwait in 1996, Colombia 1996-1997, Peru 1977 and Mauritius in 1973. Taken together, these residual

[32] The integrated hazard is $\hat{H}_{jt} = -\ln(\hat{S}_{jt})$ where S denotes the survivor function. Hence the Cox-Snell residual is $c_{jt} = d\hat{H}_{jt}/dt$.

[33] The country codes are derived from the IFS system as follows: 200-399 denote countries in Latin America and the Carribbean; 400-499 the Middle East and North Africa; 500-599 South and East Asia; 600-799 Sub-Saharan Africa; and 800-899 South Pacific and Oceania.

plots give some supporting evidence to the claim that the model is applicable across the aggregate developing country sample.

*** Figure A2 and Figure A3 ***

ii) Alternative thresholds Table A1 reports reports the results from re-estimating our core model under differing assumptions concering the threshold. To keep this manageable we restrict our attention to a small number of permutations of the model reported in Table 6 (the relevant comparator column is reported at the top of the table). Specifically we examine: (i) the differences between using a fixed and time-varying threshold; (ii) the effect of setting a more demanding threshold, specifically a time-vary threshold which moves from a deficit of 3% of GDP to 1.5% of GDP; and (iii) the effect of using a less demanding threshold for low-income countries, reflecting the fact that external debt financing of low-income budgets often contains a significant grant element.

*** Table A1 ***

Three features emerge from this table. First, and most importantly, the basic qualitative characteristics of the results we report in the main text are robust to changes in the thresholds, at least for the significant covariates. This is particularly so in the comparision between the time-varying and fixed threshold case (column [1]). Second, however, there is a marked deterioration in the overall goodness of fit and individual significance of the covariates at the tighter threshold level (columns [2] and [3]). This is true for both developing and OECD countries and reflects the higher failure rate and correspondingly shorter duration. Third, the results for the model estimated for low-income countries do not vary substantially between the 4.5% to 3% deficit threshold case (Table 6) and the 5% to 3.5% deficit threshold case (Table A1). Individual coefficient are very similar across the two models and in both cases the vector of covariants is statistically insignificant, suggesting that the poor performance of the model for low-income countries was not due to the problem of confronting them with too strict a definition of stability.

iii) Left-censoring The multiple-failure structure of our sample and the fixed observation window leads to two forms of censoring. The first is the conventional right censoring problem which arises if there are uncompleted spells of fiscal stability at the end of our observation period. Suppose for the moment all countries start their final spell of stability within the sample window (i.e. after 1970), and that the true duration of this spell is t_i^*. The observed duration will be $t_i = \min(t_i^*, c)$ where $c = (2001 - t_0)$ and t_0 denotes the start of the most recent spell. Hence final spells will be systematically understated. Under the assumption that the distribution of the true duration of spells is independent of c, the log-likelihood can be written as

$$L = \sum_{i=1}^{N} \{d_i \ln f(t_i \mid \mathbf{x}_i; \beta) + (1 - d_i) \ln[1 - F(t_i \mid \mathbf{x}_i; \beta)]\} \qquad \text{(A1)}$$

where the first term defines the contribution of completed spells to the likelihood, the second the contribution of censored spells (this term being the survivor function), and d_i is the indicator variable taking a value of 1 if a spell has been completed. Given a functional form for the hazard, this likelihood can be maximzed with respect to (which includes the ancillary parameters defining the baseline hazard). The maximum likelihood estimator in this case exhibits the standard consistency properties (see, for example, Wooldridge 2002). The results reported in Tables 5-7 are based on this maximization of this form of likelihood, modified accordingly for the choice of baseline hazard and for the presence of unobserved country-specific heterogeneity.[34]

We also face left-censoring problem in which some countries *initial* spells of fiscal stability are censored. Specifically, our first sample observation on a number of countries is one where it is in a stable configuration. Again assuming the true duration to be t_i^* and the start date to be a_i, then for first spells which start prior to our first observation (denoted t_0) we have

$$t_i = t_i^* + a_i - t_0$$

In these circumstances the log-likelihood (A1) must be augmented by an additional term is

$$L = \sum_{i=1}^{N} \begin{array}{l} \{d_i \ln f(t_i \mid \mathbf{x}_i; \beta) + (1 - d_i) \ln[1 - F(t_i \mid \mathbf{x}_i; \beta)] \\ - \ln \left[\int_0^{t_0} [1 - F(t_0 - a_i \mid \mathbf{x}_i; \beta)] k(a_i \mid \mathbf{x}_i; \eta) da_i \right] \} \end{array} \tag{A2}$$

where $k(a_i \mid x_i; \eta)$ denotes the density of a_i, given the vector of covariates. In the current version of the paper we do not estimate this model. However, using a slightly different sampling technique we can investigate whether this problem of left censoring is serious in our sample. Accordingly, we eliminate from our sample any spells of fiscal control which were underway at the point of first observation on a country. This converts our sample into a flow sample which includes countries in our risk pool only as they entered a period of stability. The final column of Table A1 reports the results of re-estimating our core model under this alternative sample. As can be seen, despite the change in sample and the corresponding rise in the sample failure rate, the fundamental quality of our results does not change.The model displays somewhat greater negative duration dependence and the effect of the fiscal history is weakened. Thus although this re-estimation does not substitute for the estimation of the model under both left- and right-censoring, it does give some support to the robustness of the results reported in the text.

[34] This result depends on the assumption that the distribution of the true duration is independent of the start date, a and the censoring point, c. In circumstances where there are unobservable (or unobserved) time-varying factors such as greater learning about fiscal management techniques or shifting preferences concerning fiscal deficits, we may expect that the true distribution of spells is a function of time. For example, we might reasonably expect, other things equal, spells of control triggered by reforms in the 1990s may have a different distribution to those introduced in the 1970s.

References

[1] Abed, George, Liam Ebrill, Sanjeev Gupta, Benedict.Clements, Ronald McMorran, Anthony Pellechio, Jerald Schiff and Marijn Verhoeven, "Fiscal reforms in low-income countries." *IMF Occasional Paper* No. 160 (1998), Washington DC, International Monetary Fund.

[2] Adam, Christopher and David Bevan, "Fiscal policy design in low-income countries." *UNU/WIDER Discussion Paper* No. 2001/67 (2001), Helsinki, WIDER.

[3] Adam, Christopher and David Bevan, "Fiscal deficits and growth in developing countries." Oxford University Department of Economics Working Paper No. 120 (2002), forthcoming *Journal of Public Economics*.

[4] Alesina, Alberto and Allen Drazen, "Why are stabilizations delayed?" *American Economic Review* vol 81 (1991) pp 1170-1188.

[5] Alesina, Alberto and Roberto Perotti, "Fiscal adjustments in OECD countries: composition and macroeconomic effects" *IMF Staff Papers* vol 44 (1997) pp210-248.

[6] Bulir, Ales and Soojin Moon, "Do IMF-supported programs help make fiscal adjustments more durable?" *IMF Working Paper* 03/38 (2003).

[7] Bayoumi, Tamim and Barry Eichengreen, "Restraining yourself: fiscal rules and stabilization" *IMF Staff Papers* vol 42 (1995) pp 32-48.

[8] Collier, Paul and Anke Hoeffler, "Greed and grievance in civil war" *Centre for the Study of African Economies, Univerisity of Oxford Working Paper* No. 2002/01 (2002).

[9] Easterly, William, Carlo Rodriguez, and Klaus Schmidt-Hebbel, *Public Sector Deficits and Macroeconomic Performance,* (1994) Oxford,World Bank and OUP.

[10] Gavin, Michael, and Roberto Perotti, "Fiscal policy in Latin America", *NBER Macroeconomics Annual* (1997) pp11-61.

[11] Gupta, Sanjeev, Benedict Clements, Emanuele Baldacci and Carlos Mulas-Granados, "Expenditure composition, fiscal adjustment, and growth in low-income countries", *IMF Working Paper* 02/77 (2002).

[12] Heckman, James and Burton Singer, "A method for minimizing the impact of distributional assumptions in econometric models for duration data", *Econometrica*, vol 47, (1984) pp 247-283.

[13] Lancaster, Tony, *The Econometric Analysis of Transition Data,* (1990) Cambridge: CUP.

[14] Loyaza, Norman, Klaus Schmidt-Hebbel and Luis Servin, "The World Savings Data Base" World Bank Mimeo (January 1998) (data from www.worldbank.org/research/projects/savings/data.htm).

[15] Illera, Reyes Mallota and Carlos Mulas-Granados, "Duration of fiscal consolidation in the European Union", *Fundacion de Estudios de Economia Aplicada Working Paper* 2001/19 (2001).

[16] Perotti, Roberto, "The political economy of fiscal consolidations", *Scandinavian Journal of Economics*, vol 100, (1998) pp 367-394.

[17] Pritchett, Lant, "Understanding patterns of economic growth: searching for hills among plateaus, mountains and plains" *World Bank Economic Review* vol 14, (2000) pp 221-50.

[18] Reinhart, Carmen and Kenneth Rogoff "The modern history of exchange rate arrangements: a reinterpretation" *NBER Working Paper* 8963 (2002).

[19] Sun,Yan. "Do fixed exchange rates induce more fiscal discipline" *IMF Working Paper* 03/78 (2003).

[20] Tornell, Aaron and Philip Lane, "The Voracity Effect", *American Economic Review* , vol 89 (1999), pp22-46.

[21] Tornell, Aaron and Andres Velasco, "Fiscal discipline and the choice of exchange rate regime" *European Economic Review,* vol 39, (1995),pp 759-770.

[22] Von Hagan, Jurgen, Andrew Hughes Hallett and Rolf Strauch, "Budgetary consolidation in Europe: quality, economic conditions and persistence" *Journal of the Japanese and International Economies* vol 15, (2002) pp 512-535.

[23] Wooldridge, Jeffrey *Econometric Analysis of Cross Section and Panel Data,* (2002), Cambridge MA: MIT Press.

Table 1: Conventional Budget Balance as % of GDP (unweighted group averages) [a]

		1970-2000 Mean	std dev[b]	t-test	1970-1984 Mean	std dev	t-test	1985-2000 Mean	std dev	t-test
OECD	[n=25]	-3.77%	2.74%		-3.72%	2.78%		-3.83%	2.70%	
All Developing Countries	[n=102]	-4.03%	3.58%	1.18 [c]	-4.84%	3.62%	3.87	-3.14%	3.54%	2.10
Low-Income	[n=37]	-5.31%	3.93%	7.24 [d]	-5.38%	4.09%	2.24	-5.21%	3.76%	8.18
Middle-Income	[n=46]	-3.15%	3.54%	2.60 [c]	-4.44%	3.42%	5.68	-1.85%	3.35%	2.28

Notes
[a] Defined for the consolidated central government and computed after grants [GFS Code 80hh].
[b] Average within-country standard deviation of the budget-balance.
[c] t-tests for difference in means relative to OECD.
[d] t-tests for difference in means between low-income and middle-income developing countries.
Source: IMF Government Finance Statistics. See text for country classifications.

Table 2: Cyclical Properties of the Fiscal Balance
Dependent variable = change in conventional fiscal balance (after grants)
Sample 1970-2000 [Annual data]

OLS Estimation with Fixed Effects									
Country Group		**OECD**		**All Developing**		**Middle-Income**		**Low-Income**	
		[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Real GDP Growth[a]	**Average**	0.200 [6.27]		0.027 [1.71]		-0.008 [0.41]		0.061 [2.35]	
	postive		0.149 [3.27]		-0.035 [1.28]		-0.088 [2.69]		0.024 [0.55]
	negative		0.316 [4.81]		0.101 [3.32]		0.105 [2.48]		0.098 [2.15]
N		25	25	83	83	46	46	37	37
T		23	23	19	19	20	20	17	17
R-square (within)		0.192	0.195	0.229	0.234	0.198	0.207	0.267	0.268
Pooling F-test [p-value]		[0.008]	[0.012]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Symmetry F-test [p-value] [b]			[0.049]		[0.005]		[0.002]		[0.321]

Notes:
[a] GDP growth measured in constant price US dollars per capita [PWT6.1]
t statistics in parentheses
[b] Symmetry test of null that b[positive]=b[negative]
All models include as additional regressors a vector of country-specific intercepts and the lagged fiscal balance.

Table 3 The Persistence of Fiscal Stability (multiple spell data)

Deficit Threshold [a]	**[1] 3 % fixed**			**[2] Smooth Transition [4.5% to 3%]**			**[3] Smooth Transition [3% to 1.5%]**		
Countries	**Incidence Rate** [c]	**Number of Episodes** [d]	**Mean Duration** [e]	**Incidence Rate**	**Number of Episodes**	**Mean Duration**	**Incidence Rate**	**Number of Episodes**	**Mean Duration**
All	54%	105	3.47	46%	115	4.14	66%	104	2.62
OECD	53%	30	4.94	45%	31	5.65	67%	30	3.45
Developing Countries	54%	75	2.94	47%	84	3.60	65%	74	2.16
NR=1 [b]	50%	24	2.15	44%	27	3.18	62%	25	1.71
NR=0	56%	51	3.60	48%	57	3.83	67%	49	2.52
Middle Income	47%	45	2.82	39%	51	3.78	57%	74	2.12
Low Income	65%	30	3.22	58%	33	3.32	77%	27	2.26

Kaplan-Meier Survival Probabilities [f]	**Years**	**OECD**	**Middle Income**	**Low Income**	**OECD**	**Middle Income**	**Low Income**	**OECD**	**Middle Income**	**Low Income**
	1	0.93	0.78	1.00	0.93	0.89	1.00	0.86	0.56	0.75
	2	0.87	0.48	0.57	0.87	0.68	0.57	0.63	0.30	0.32
	3	0.70	0.24	0.32	0.78	0.51	0.32	0.45	0.16	0.12
	5	0.35	0.08	0.08	0.50	0.18	0.12	0.16	0.03	0.02
	8	0.05	0.01	0.01	0.16	0.03	0.01	0.01	0.00	0.00
	10	0.01	0.00	0.00	0.04	0.01	0.00	0.00	0.00	0.00

Notes:
[a] See text for description of thresholds
[b] NR=1 denotes countries whose oil or mineral production accounts for more than 50% of exports.
[c] Proportion of time spent below threshold
[d] Number of new episodes of fiscal stability (all durations)
[e] In years. Survival durations are censored from above implying underestimate of true survival duration.
[f] Proportion of episodes (defined in [d]) surviving longer than 1,2,3…etc years.

Table 4
Summary statistics for covariates by groups and failures [a] [b] [c]

	[1] OECD	[2] All Developing Countries	[3] Middle Income	[4] Low Income
Fiscal Balance [% GDP]	-3.77	-4.03	-3.15	-5.31
Total Expenditure [% GDP]	34.00	23.24	24.12	24.29
Total Revenue (excl grants) [% GDP]	31.61	22.49	21.66	19.87
Natural resource dependency [prevalence]	0.08	0.33	0.38	0.23
Openness to Trade [% GDP]	48.33	74.17	74.04	75.70
Agricultural Share [% GDP]	7.27	25.06	17.74	36.41
Gini coefficient	34.11	45.21	45.96	44.15
Log real GDP per capita	9.67	7.96	8.51	7.12
Terms of Trade	111.48	114.32	113.06	116.18
Annual growth in per capita GDP [%]	2.35	1.40	2.02	0.30
Conflict (prevalence)	-	0.06	0.04	0.09
Number of 'fails'	6.15	5.99	5.85	5.87
Fiscal History (incidence)	0.33	0.33	0.32	0.34
Inflation History [prevalence]	0.03	0.14	0.17	0.17
Exchange Rate Regime	1.94	1.88	1.86	1.79
Domestic Debt [% of GDP]	32.45	11.87	12.84	10.43
External Debt [% of GDP]	-	29.71	27.58	31.02

Notes:
[a] Fail = if deficit is greater than 3 percent of GDP.
[b] see data appendix for complete description of variables
[c] Public debt measured in present value terms
** denotes difference in means significant at 1% level; * significant at 5% level.

Table 5: Persistence of Fiscal Stability and Structural Characteristics
Sample 1970-2000 (multiple-spell data) [a] [b]

Country Grouping	[1] Developing	[2] Developing	[3] Developing	[4] Developing	[5] Developing	[6] Developing	[7] Middle Income	[8] Low Income	[9] OECD
Control for Heterogeneity	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Threshold [c]	Fixed 3%	Fixed 3%	Fixed 3% NR=1	Fixed 3% NR=0	Fixed 3%	Time-varying 4.5% to 3%	Time-varying 4.5% to 3%	Time-varying 4.5% to 3%	Time-varying 4.5% to 3%
Covariates [d] [e]									
Natural resource dependency [NR]	-0.135 [0.71]	-0.133 [0.54]	-	-	0.246 [0.61]	0.21 [0.48]	1.177 [1.75]	-0.694 [1.24]	-1.08 [0.61]
Agriculture [% of GDP]	-0.005 [0.74]	-0.012 [1.30]	-0.017 [0.59]	-0.007 [0.67]	-0.009 [0.99]	-0.009 [0.83]	-0.042 [1.82]	0.015 [1.47]	-0.135 [4.00]
Trade [% of GDP]	0.002 [1.83]	0.002 [0.86]	0.001 [0.39]	0.002 [0.59]	0.002 [0.91]	0.002 [0.70]	-0.006 [1.25]	0.006 [2.94]	-0.011 [1.88]
Government Expenditure [% GDP]	0.009 [1.85]	0.013 [2.76]	0.020 [2.09]	0.007 [2.24]	0.012 [2.38]	0.015 [2.80]	0.037 [3.35]	0.012 [1.93]	0.010 [0.55]
Income Inequality	-0.007 [0.89]	-0.002 [0.27]	-0.015 [0.80]	-0.014 [0.52]	0.001 [0.13]	0.007 [0.98]	0.003 [0.35]	0.021 [2.08]	-0.039 [2.02]
Log real income (per capita)	-0.262 [2.56]	-0.375 [2.89]	-0.480 [1.94]	-0.310 [1.78]	-0.360 [2.56]	-0.433 [2.78]	-1.165 [4.64]	0.011 [0.05]	-0.983 [8.15]
Conflict dummy [tv]	0.173 [0.87]	0.238 [1.55]	0.483 [2.39]	0.17 [0.96]	0.142 [0.80]	0.023 [0.12]	0.335 [1.18]	-0.027 [0.10]	-
NR*conflict [tv]	-	-	-	-	0.269 [2.71]	0.624 [2.54]	0.373 [0.67]	0.108 [0.16]	-
Terms of Trade [tv]	-0.002 [2.34]	-0.002 [2.00]	-0.005 [2.86]	-0.002 [1.14]	-0.002 [1.05]	-0.001 [0.90]	0.004 [0.23]	-0.003 [2.35]	0.000 [0.15]
NR*terms of trade [tv]	-	-	-	-	-0.004 [2.31]	-0.003 [1.89]	-0.008 [1.80]	0.001 [0.15]	-0.006 [0.44]
Weibull α	1.140 [2.07]	1.140 [3.05]	1.090 [1.16]	1.247 [4.02]	1.178 [3.63]	1.209 [3.93]	1.230 [3.24]	1.391 [4.91]	2.480 [17.90]
Heterogeneity θ		0.745	1.126	0.741	0.842	0.974	1.550	0.306	0.766
LR [θ = 0]		235.21	75.47	164.22	245.58	262.20	184.54	32.72	46.84
Pr		[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
LL	943.03	1060.6	243.81	753.69	994.62	781.43	329.69	453.36	328.61
Pseudo R-square	0.011	0.011	0.031	0.006	0.012	0.017	0.057	0.017	0.067
LR [β = 0]	20.93 [0.000]	22.54 [0.000]	15.14 [0.0343]	8.71 [0.000]	24.36 [0.000]	26.31 [0.000]	37.45 [0.000]	15.33 [0.469]	43.76 [0.000]
LR-pooling test				126.2 [0.000]				3.24 [0.932]	
No. countries	83	83	22	61	83	83	37	46	25
Raw failure rate	53.4%	53.4%	50.1%	56.1%	54.3%	46.6%	38.5%	57.7%	44.9%

Notes
[a] z-scores in parentheses
[b] β estimates
[c] See text for description of thresholds
[d] See data appendix for description of covariates
[e] tv denotes time-varying covariate

Table 6: Persistence of Fiscal Stability with Initial Conditions and Fiscal History
Weibull and Piecewise Constant Hazard Models with adjustment for unobserved heterogeneity
All Models estimated with Time-Varying Threshold (4.5% to 3% deficit)
Sample 1970-2001 (multiple-spell data) [a]

Country Grouping	[1] Developing All	[2] Developing All	[3] Developing Middle Income	[4] Developing Low Income	[5] OECD
Hazard	Weibull	Piecewise	Piecewise	Piecewise	Piecewise
Covariates					
Natural resource	0.337	0.262	1.167	-0.493	-0.800
	[0.88]	[0.67]	[1.90]	[0.82]	[0.48]
Government Expenditure [% GDP]	0.015	0.012	0.034	0.011	0.013
	[3.15]	[2.62]	[3.04]	[1.61]	[1.04]
Log real income (per capita)	-0.296	-0.278	-0.668	0.138	-0.574
	[3.20]	[2.91]	[2.45]	[0.55]	[2.44]
Conflict dummy [tv]	0.152	0.203	0.685	-0.006	-
	[0.92]	[1.22]	[2.42]	[0.02]	
Terms of Trade [tv]	-0.001	-0.001	0.000	-0.001	-0.001
	[0.57]	[1.01]	[.09]	[0.43]	[0.05]
NR*ToT [tv]	-0.004	-0.003	-0.009	0.001	-0.002
	[2.16]	[1.95]	[1.92]	[0.19]	[0.15]
Inflation history	0.361	0.334	0.545	0.126	-0.341
	[2.90]	[2.63]	[2.36]	[1.70]	[0.18]
Inflation history [OECD]	-	-	-	-	0.253
					[1.86]
Domestic debt [% of GDP]	0.002	0.003	0.002	0.018	0.010
	[0.52]	[0.48]	[0.38]	[1.02]	[2.32]
External debt [% of GDP]	-0.005	-0.005	-0.015	-0.002	-
	[0.80]	[0.76]	[1.66]	[0.41]	
Real GDP growth [tv]	-0.075	-0.066	-0.142	-0.056	-0.613
	[2.26]	[2.11]	[2.54]	[0.39]	[2.78]
Initial Adjustment	-0.085	-0.078	-0.047	-0.092	-0.602
	[2.58]	[2.35]	[1.60]	[2.10]	[6.25]
Exchange rate regime	-0.056	-0.047	-0.083	-0.026	-0.027
	[1.53]	[1.27]	[1.95]	[0.43]	[0.33]
IMF	-0.003	0.087	-0.206	-0.093	-
	[0.02]	[0.62]	[0.40]	[0.51]	
Fiscal history [tv]	1.236	-	-	-	-
	[5.81]				
Piecewise constant coefficients [b]					
history[m1]	-	3.28	3.80	3.12	5.11
		[3.53]	[3.57]	[2.48]	[1.79]
history[m2]	-	2.61	2.47	-	1.74
		[4.40]	[2.56]		[1.37]
history[m3]	-	2.47	2.69	2.92	1.42
		3.97]	[3.53]	[3.05]	[1.48]
history[m4]	-	2.01	2.27	2.05	0.88
		[2.74]	[3.02]	[2.56]	[1.04]
history[m5]	-	1.07	1.51	1.95	0.75
		[3.56]	[3.50]	[2.03]	[1.03]
LR test [history[mi]=history]	-	4.31	8.51	2.13	2.93
		[0.038]	[0.036]	[0.142]	[0.086]
Weibull α	0.911	-	-	-	-
t [alpha=1]	[1.48]				
LR test [constant baseline]		3.15	4.5	4.41	12.76
		[0.789]	[0.609]	[0.621]	[0.047]
Heterogeneity θ	0.404	0.431	0.918	0.629	0.034
LR [θ = 0]	38.76	46.63	45.77	18.32	0.94
Pr	[0.000]	[0.000]	[0.000]	[0.000]	[0.166]
LL	790.12	801.75	343.58	389.78	345.01
Pseudo R-square	0.032	0.122	0.182	0.063	0.257
LR [β = 0]	49.36	174.31	105.63	46.48	141.17
Pr	[0.000]	[0.000]	[0.000]	[0.002]	[0.000]
LR-pooling test				136.78	
				[0.000]	
No. countries	83	83	46	37	25
Raw Failure Rate	46.53%	46.53%	38.46%	52.05%	44.95%

Notes:
[a] See notes to Table 5
[b] Time-intervals are defined in years as follows: m1=[0,2), m2=[2,4) …m5=[8,inf)

Table 7: The Effect of composition on persistence [a] [b] [c]

Country Grouping	[1] Developing All	[2] Developing Middle-Income	[3] Developing Low-Income	[4] OECD
Fiscal Correction [β coefficients]				
Domestic Revenue	-0.103 [3.11]	-0.145 [2.78]	-0.578 [4.04]	-0.434 [2.81]
Grants	-0.009 [0.33]	-0.126 [1.49]	-0.011 [0.28]	
Total Expenditure	0.020 [0.49	0.149 [2.20]	0.186 [1.67]	1.090 [2.99]
Pseudo R-square	0.042	0.293	0.039	0.134
LR [β = 0] Pr	69.34 [0.000]	166.28 [0.000]	31.94 [0.000]	89.34 [0.000]

Notes:
[a] See notes to Tables 5 and 6
[b] Results conditional on covariates included in Table 6. These are not reported here.
[c] coefficients measure the effect on the hazard of a one percentage point increase in the relevant fiscal item.

Table A1: Sensitivity Analysis [a] [b]

	[1]	[2]	[3]	[4]	[5]
Country Grouping	Developing	Developing	OECD	Low-Income	Developing
Threshold	Fixed 3%	Time-Varying 3% to 1.5%	Time-Varying 3% to 1.5%	Time-Varying 5% to 3.5%	Fixed 3% Adjusted for Left-Censoring
Comparator [c]	Table 6 col[1]	Table 6 col[1]	Table 6 col[5]	Table 6 col[4]	
Covariates					
Natural resource	0.258	0.058	0.300	-0.279	0.168
	[0.78]	[0.21]	[0.33]	[0.47]	[0.57]
Government Expenditure [% GDP	0.010	0.005	0.013	0.013	0.010
	[2.83]	[1.84]	[1.44]	[2.01]	[0.27]
Log real income (per capita)	-0.179	-0.135	-0.300	0.208	-0.151
	[2.91]	[2.58]	[1.87]	[0.84]	[2.62]
Conflict dummy [tv]	0.206	0.086	-	-0.077	0.253
	[1.48]	[0.68]		[0.30]	[1.60]
Terms of Trade [tv]	-0.001	-0.001	0.001	-0.001	-0.002
	[1.04]	[0.53]	[0.61]	[0.58]	[1.62]
NR*ToT [tv]	-0.004	-0.002	-0.003	-0.003	-0.003
	[2.46]	[1.86]	[0.32]	[0.70]	[1.86]
Inflation history	0.338	0.167	0.047	0.195	0.323
	[3.20]	[1.72]	[0.14]	[1.09]	[2.43]
Domestic debt [% of GDP]	0.002	0.002	0.005	0.038	0.007
	[0.53]	[0.44]	[1.12]	[1.65]	[2.36]
External debt [% of GDP]	-0.0004	-0.0001	-	-0.007	-0.0005
	[0.69]	[0.51]		[1.11]	[1.08]
Real GDP growth [tv]	-0.151	-0.084	-0.089	0.016	-0.127
	[2.83]	[1.87]	[1.46]	[0.19]	[1.97]
Initial Adjustment	-0.06	-0.132	-0.505	-0.066	-0.075
	[2.87]	[3.87]	[4.18]	[0.87]	[3.15]
Exchange rate regime	-0.071	-0.019	-0.054	-0.054	-0.031
	[2.46]	[0.77]	[0.97]	[0.91]	[1.11]
IMF	0.032	-0.043	-	-0.173	-0.080
	[0.26]	[0.38]		[0.98]	[1.05]
Fiscal history [tv]	1.001	0.26	0.433	0.754	0.447
	[6.05]	[1.86]	[1.49]	[1.92]	[2.39]
Weibull α	0.872	1.05	1.43	1.19	0.80
t [alpha=1]	[2.35]	[1.06]	[4.80]	[2.13]	[2.69]
Heterogeneity θ	0.104	0.064	0	0.659	0.019
LR [θ = 0]	16.57	12.7	0	15.05	0.93
Pr	[0.000]	[0.000]	[1.000]	[0.000]	[0.167]
LL	1042.67	1414.19	610.18	385.04	759.82
LL(0)	979.585	1372.085	515.555	379.76	743.47
Pseudo R-square	0.064	0.031	0.184	0.014	0.022
LR [β = 0]	126.17	84.21	189.25	10.56	32.70
Pr	[0.000]	[0.000]	[0.000]	[0.5667]	[0.000]
No. countries	83	83	25	37	83
Raw Failure Rate	54.20%	67.34%	67.86%	52.05%	75.31%

Notes:
[a] See notes to Tables 5 and 6

Data Appendix

The vector of covariates consists of two types of variable. The first consists of those that are fixed in perpetuity and those that vary over time. For each spell the latter enter the analysis at their start-of-period values but these will vary across separate spells of fiscal stability. In the list that follows, these variables are denoted [s].Those that do vary over the duration of a spell itself are denotes [tv].

Sources: The source for each variable is denoted in the following list. The abbreviations are as follows:

GFS	Government Finance Statistics, International Monetary Fund, Washington DC (January 2003)
IFS	International Finance Statistics, International Monetary Fund, Washington DC (January 2003)
WDI	World Development Indicators, World Bank, Washington DC (2002)
PWT	Penn World Tables Version 6.1. A.Heston, R.Summers and B.Aten, Centre for International Comparisons, University of Pennsylvania, October 2002.
WIDER	UNU/WIDER-UNDP World Income Inequality Database (September 2000) (www.wider.unu.edu/wiid/wiid.htm)

Variable and Source

Natural resource dependency	=1 if mineral production exceed 50% of exports. Source: WDI [s]
Agriculture	Agricultural share in GDP. Source: WDI [s].
Trade	Trade share in GDP. Source: PWT [s]
Government Expenditure	Government Expenditure Government share in GDP (at constant PPP prices). Source: PWT [s]
Inequality	Most recent country specific household and income based gini coefficient. Source: WIDER [s]
Conflict	=1 if country engaged in civil or inter-state conflict. Source: International Peace Research Institute Oslo (January 2003) www.prio.no [tv]
Real income	(log) Average per capita GDP, PPP prices, Laspeyres index Source: PWT [s].
Growth	Annual growth in per capita GDP. Source: PWT [tv]
Terms of Trade	Income terms of trade (1970=100). PWT [tv]
Fiscal history	Number of previous fiscal failures divided by time at risk Source: see text [s].
Inflation history	=1 if inflation in excess of 40% per annum in last 5 year period before current spell. Source: IFS [s].
Domestic / External Debt	Domestic and External public debt (in present value terms) Source: Loyaza *et al* (1998).
Fiscal correction	Change in government expenditure (GFS Code 82Z), change in domestic revenue (GFS Code 81Y), and grants (GFS Code 81Z) initiating return to fiscal stability [s]. Source: GFS.
Exchange rate regime	Reinhart and Rogoff (2002) 'coarse' exchange rate classification. [s].
IMF	IMF =1 if current period of stability started with IMF support (eg SAF/ESAF or PRGF). [s].
Initial Adjustment	Country specific difference between fiscal threshold and average fiscal deficit in the three years prior to current spell of fiscal stability. Source: see text [s]

Countries included in analysis

OECD: Austraila, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Malta Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States.

Developing Countries (Middle Income)
Argentina, Barbados, Boliva , Botswana, Brazil , Chile, Colombia, Costa Rica , Cyprus, Dominica, Dominican Republic, Ecuador, Egypt, Fiji, Gabon, Guatamala, Honduras, India, Iran, Israel, Jamaica, Jordan, Korea, Lebanon, Malaysia, Mauritius, Mexico, Morocco, Namibia, Panama, Papua New Guinea, Paraguay, Peru, Phillipines, Seychelles, Singapore, South Africa, Sri Lanka, St Kitts, St Vincent, Syria, Thailand, Trinidad and Tobago, Tunisia, Uruguay, Venezeula.

Developing Countries (Low Income)
Bangladesh, Benin, Burkina Faso, Burundi, Cameroon, Chad, Comoros, Congo, Congo-DRC (Zaire) , Cote d'Ivoire, Ethiopia, Gambia, Ghana, Guinea Guinea-Bissau, Guyana, Haiti, Indonesia, Kenya, Lesotho, Madagascar, Malawi, Mali, Mauritania, Nepal, Nicaragua, Niger, Nigeria, Pakistan, Rwanda, Senegal, Sierra Leone, Togo, Uganda, Yemen, Zambia , Zimbabwe.

Figure 1
Alternative Thresholds for Fiscal Sustainability

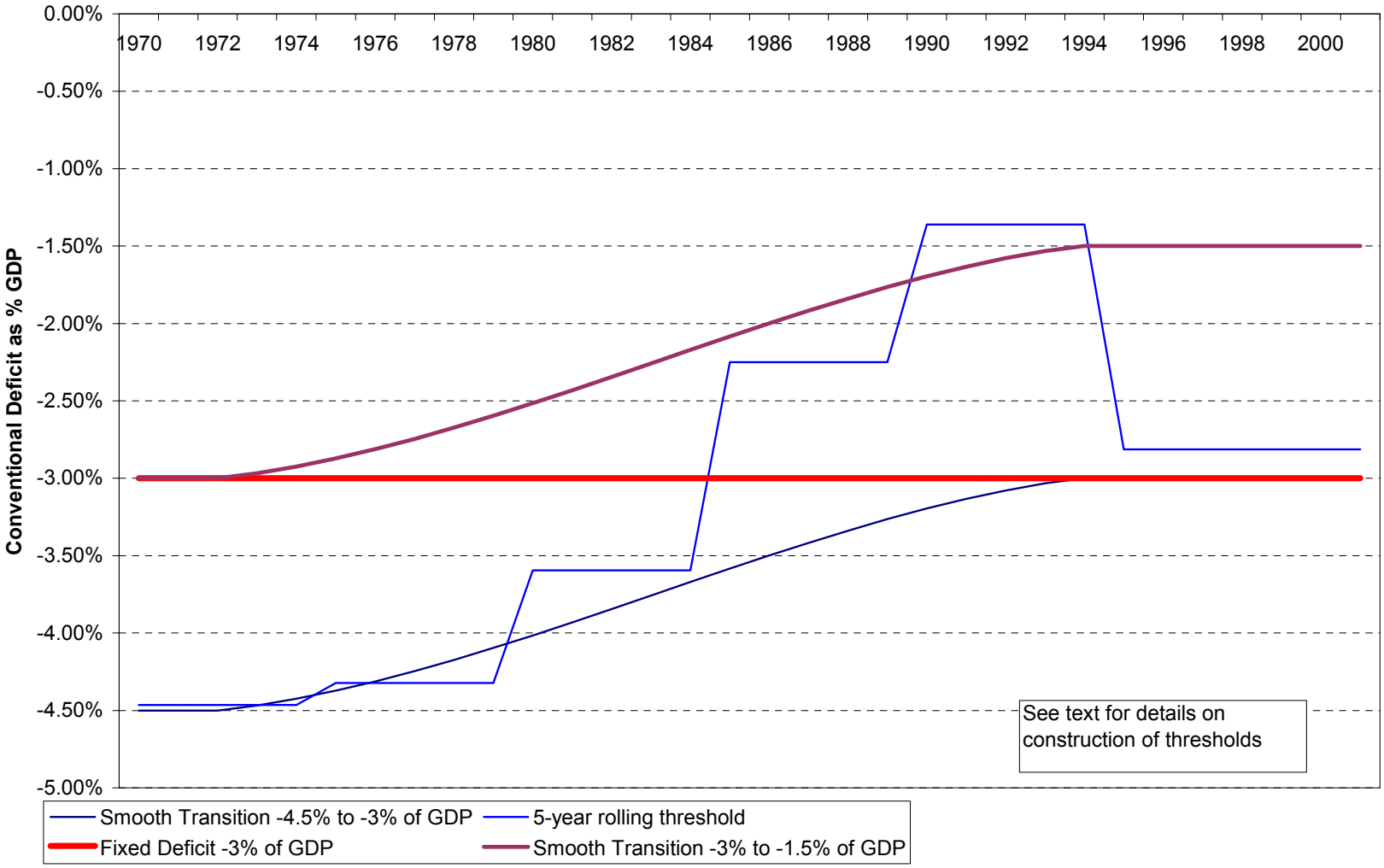


Figure A1

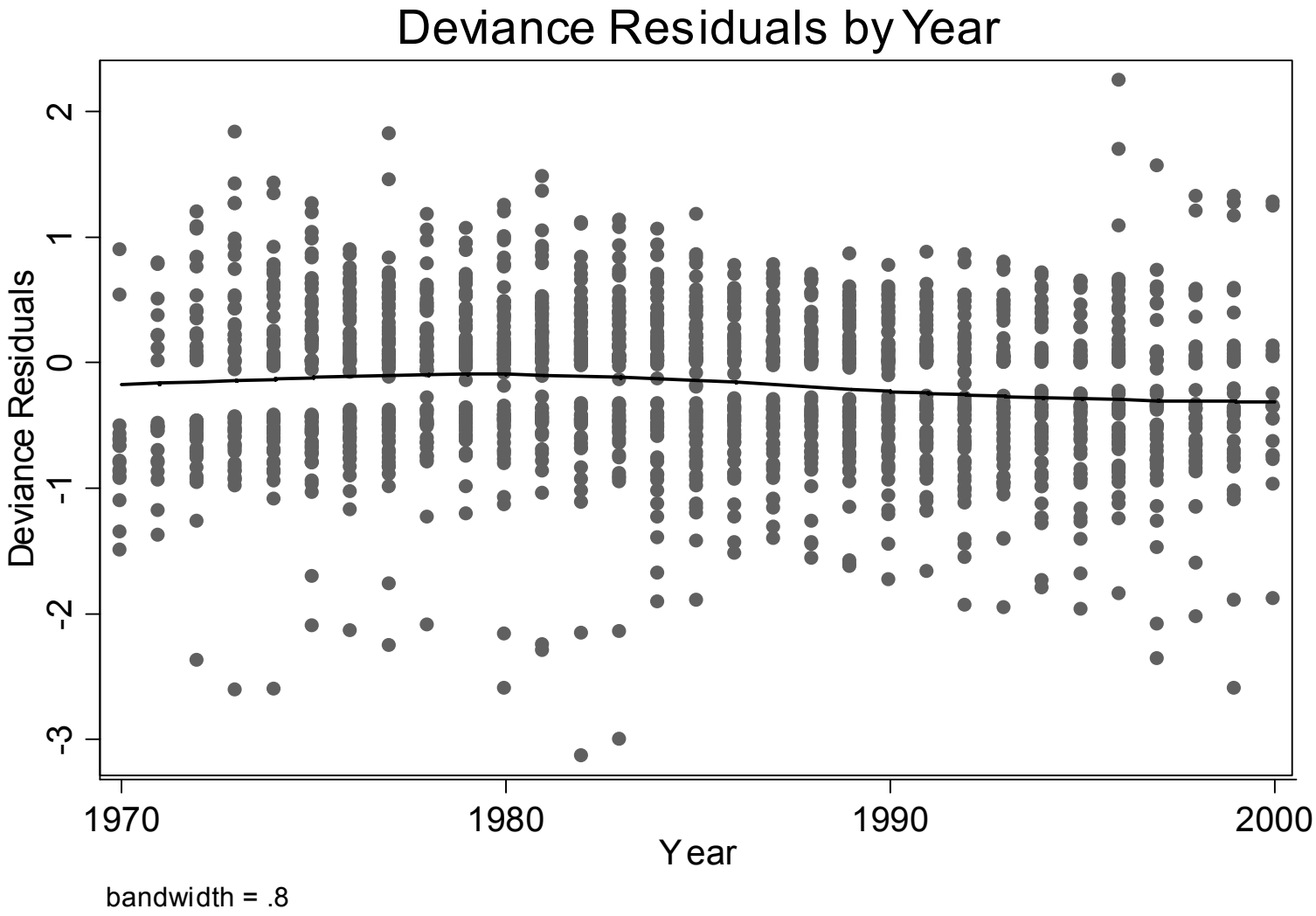
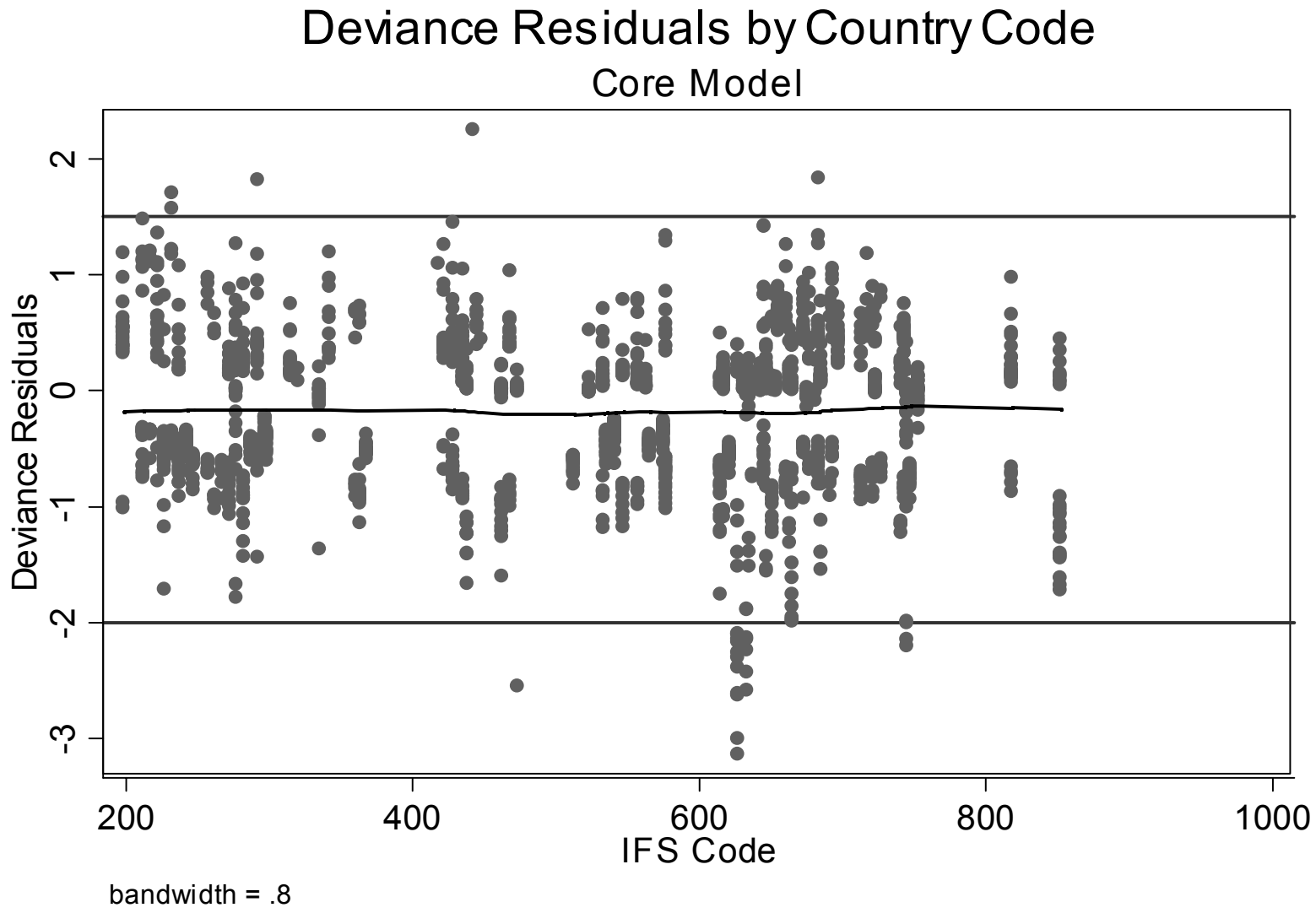

Deviance Residuals by Year
Deviance Residuals
2
1
0
-1
-2
-3
1970
1980
1990
2000
Year
bandwidth = .8

Figure A2


Deviance Residuals by Country Code
Core Model
Deviance Residuals
2
1
0
-1
-2
-3
200
400
600
800
1000
IFS Code
bandwidth = .8