

## Primary Surplus Behavior and Risks to Fiscal Sustainability in Emerging Market Countries: A “Fan-Chart” Approach

OYA CELASUN, XAVIER DEBRUN, and JONATHAN D. OSTRY\*

*This paper proposes a probabilistic approach to public debt sustainability analysis (DSA) using “fan charts.” These depict the magnitude of risks—upside and downside—surrounding public debt projections as a result of uncertain economic conditions and policies. We propose a simulation algorithm for the path of public debt under realistic shock configurations, combining pure economic disturbances (to growth, interest rates, and exchange rates), the endogenous policy response to these, and the possible shocks arising from fiscal policy itself. The paper emphasizes the role of fiscal behavior, as well as the structure of disturbances facing the economy and due to fiscal policy, in shaping the risk profile of public debt. Fan charts for debt are derived from the “marriage” between the pattern of shocks on the one hand and the endogenous response of fiscal policy on the other. Applications to Argentina, Brazil, Mexico, South Africa, and Turkey are used to illustrate the approach and its limitations. [JEL H62, H63, C15, E62]*

**I**n the public debt context, a sustainable position is often viewed as one where the government (or public sector) is solvent. To be deemed solvent, a government must be expected to honor current and future financial obligations, including the implicit commitment to continue providing certain public goods, services, and transfers in the future. Solvency thus implies that the present value of government disbursements (including inherited debt amortization, interest payments, and non-interest expenditure) should not

---

\*Oya Celasun and Xavier Debrun are Economists in the IMF’s Research and Fiscal Affairs Departments, respectively. Jonathan D. Ostry is Deputy Director in the Research Department of the IMF.

exceed the present value of revenues or, equivalently, that the present value of future revenues net of non-interest spending (the primary balances) should at least cover the existing public debt. The intertemporal budget constraint and the relationship between the primary balance and the public debt have therefore been at the center of the literature on debt sustainability.

In practice, the notion of sustainability is less straightforward than it appears. First, at a conceptual level, it always implies a judgment as to what constitutes an acceptable strategy for the government to ultimately satisfy its intertemporal budget constraint (Mendoza and Oviedo, 2004). By definition, solvency excludes outright default or forced restructuring.<sup>1</sup> Yet most analysts would also exclude the inflation tax from the set of acceptable strategies, limiting the latter to adjustments in the primary balance. Hence, solvency is only a necessary condition for sustainability, and defining sufficient conditions involves judgment. Second, at a technical level, the forward-looking nature of solvency makes it difficult to assess. No one knows for sure the primary surplus a government will be able (or willing) to generate in 5, 10, or 20 years, nor the future path of interest rates, inflation, and productivity growth over that period. Absent uncertainty, of course, assessing solvency would be a mere arithmetical exercise. In reality, however, it requires making judgments under uncertainty, as well as recognizing that such judgments are subject to risk.

Concrete assessments of debt sustainability—including those performed by IMF country teams—rely on medium-term simulations of the debt-to-GDP ratio given specific macroeconomic forecasts and fiscal policy assumptions. In the absence of reliable sustainability thresholds, however, such estimates do not in themselves allow for determination of the sustainability of a particular debt position. Instead, the expected dynamics of the debt-to-GDP ratio over the medium term (generally 5 to 10 years) are interpreted as a signal of whether underlying policies can be sustained under plausible macroeconomic conditions without endangering solvency. Specifically, a declining trend in the debt ratio signals that government policies are unlikely to jeopardize sustainability, whereas a positive trend or even stabilization at a high level may motivate concerns about sustainability, especially if other factors—such as the fiscal adjustment needed to stabilize or reduce the debt ratio—point to likely difficulties in keeping debt under control.

Uncertainty about future macroeconomic conditions and fiscal policy inevitably weakens the basis for drawing compelling policy conclusions using such analyses. This paper proposes a methodology that improves our understanding of the risks surrounding debt dynamics, and explicitly acknowledges the probabilistic nature of debt sustainability analysis (DSA) exercises. A more nuanced and credible assessment of sustainability results.

Of course, standard DSA does recognize the importance of uncertainty and appraises risks to the baseline debt projection by simulating alternative debt paths under less favorable conditions—so-called bound tests. This approach to risk assessment is entirely deterministic, however, involving a set of scenarios in which one key variable at a time is hit by an adverse shock—including lower growth,

---

<sup>1</sup>Solvency is generally defined as the ability to meet financial obligations on time.

higher interest rates, a lower primary balance, and exogenous debt increases, such as those resulting from exchange rate depreciation or the recognition of off-budget obligations. The calibration of the shocks generally uses a multiple or fraction of the unconditional variance of the underlying series.

The bound-tests approach gives a broad sense of the sensitivity of the sustainability assessment to adverse developments, but significant methodological limitations undermine its credibility and operational relevance. First, both the correlations among shocks and the joint dynamic response of the variables relevant for debt dynamics are ignored. Indeed, simulated deterministic disturbances can realistically be of only two types: large and transitory, or small and permanent.<sup>2</sup> Second, fiscal policy is assumed not to react to the simulated economic developments, as if a deterministic policy process could reasonably prevail in an intrinsically stochastic environment. Third, in an uncertain world, of course, each bound test formally has a near-zero probability of occurrence, making any meaningful quantification of risk impossible.

Measuring risk to debt dynamics requires a stochastic simulation apparatus whereby bound tests covering a range of shock combinations can be generated. With a framework capable of randomly generating a large sample of bound tests, frequency distributions of the debt ratio can be derived for each year of a projection, yielding an explicitly probabilistic assessment of debt sustainability. Specifically, we use “fan charts” to depict confidence bands for varying degrees of uncertainty around the median projection, which are more informative than assessments based solely on the trend in the central projection for debt.

This paper proposes a stochastic DSA algorithm along these lines. The algorithm preserves a certain degree of standardization while allowing for country-specific risk factors to be reflected. To illustrate its versatility, it is applied to five emerging market countries with fairly different risk profiles, namely, Argentina, Brazil, Mexico, South Africa, and Turkey.

The paper also draws on earlier work looking at public debt sustainability in relation to primary surplus behavior. IMF (2003) focuses on determining debt thresholds beyond which sustainability could be considered at risk given average fiscal behavior. The same study introduces the concept of “overborrowing,” defined as the excess of current public debt over the annuity value of future primary surpluses. Using an expanded version of the data set in IMF (2003), Abiad and Ostry (2005) refine the estimations of fiscal reaction functions (including a richer set of political and institutional variables) and of the determinants of overborrowing, and calculate the impact on sustainable debt levels of a variety of fiscal and institutional reforms. The present paper marries the approach to fiscal policy reaction functions in Abiad and Ostry and the stochastic analysis of debt issues in Garcia and Rigobon (2005) and Penalver and Thwaites (2004). These latter papers, by focusing on higher-frequency macroeconomic data, pay insufficient attention to the constraints on the evolution of public debt created by the endogenous response of fiscal policy

---

<sup>2</sup>Recent adjustments to the IMF’s DSA template placed greater emphasis on small and permanent shocks, leaving only exogenous debt increases as large and temporary disturbances.

to debt shocks. An appendix containing technical details of our analysis is available in an earlier version of this paper (Celasun, Debrun, and Ostry, 2006).

## I. Debt Sustainability Analysis and Risk

This section compares deterministic DSA with stochastic approaches, highlighting the value added of the latter over the former. As previously mentioned, a given debt position is sustainable as long as it does not exceed the present value of all future primary surpluses. Because the path of primary surpluses over the indefinite future is essentially unknown, however, implementing this definition is a tremendous challenge, calling for operational alternatives.

### DSA Frameworks

The time horizon for macroeconomic projections rarely extends beyond a few years, so the solvency concept is typically operationalized by asserting that sustainability is not in jeopardy if the expected path of primary surpluses causes the debt-to-GDP ratio to decline over the simulation horizon. Concerns about sustainability arise if the debt ratio trends upward or if it stabilizes at a high level relative to peer countries or relative to its own historical track record; sustainability could also be called into question if the magnitude of fiscal adjustment required to stabilize the debt ratio seems excessive. In the standard DSA setup, the assessment does not relate to the sustainability of a particular debt position but to whether given policies lead to particular trends in the debt-to-GDP ratio, which may in turn motivate calls for policy corrections.

A key issue with the DSA as just described is the omission of uncertainty—for example, about income growth, interest rates, fiscal policies, exchange rates, or even the recognition of contingent liabilities. A natural response is to subject the DSA's baseline projection to a series of shocks (bound tests) that deteriorate the outlook. The standard bound-testing approach is deterministic, however, and limited to either isolated shocks or ad hoc combinations of them. Whereas the unconditional variance of the underlying series determines the magnitude of the simulated disturbances, actual covariances—especially between fiscal and nonfiscal variables—are ignored. This may lead to severely underestimating risks to debt sustainability if adverse shocks—say to growth, interest rates, and exchange rates—combine in an explosive cocktail for debt dynamics.<sup>3</sup>

Calibrating deterministic bound tests to reflect economic and policy patterns observed in a given economy thus constitutes another challenge for the standard DSA framework. One possibility is to devise a small number of standardized scenarios—where isolated shocks are expressed as a proportion of the historical standard deviations of the variables—such that both the shock itself and the resulting debt path appear plausible in probabilistic terms (IMF, 2003). By its nature, this approach lends itself to the construction of bound tests applicable to many different countries, and requires only a fairly parsimonious data set.

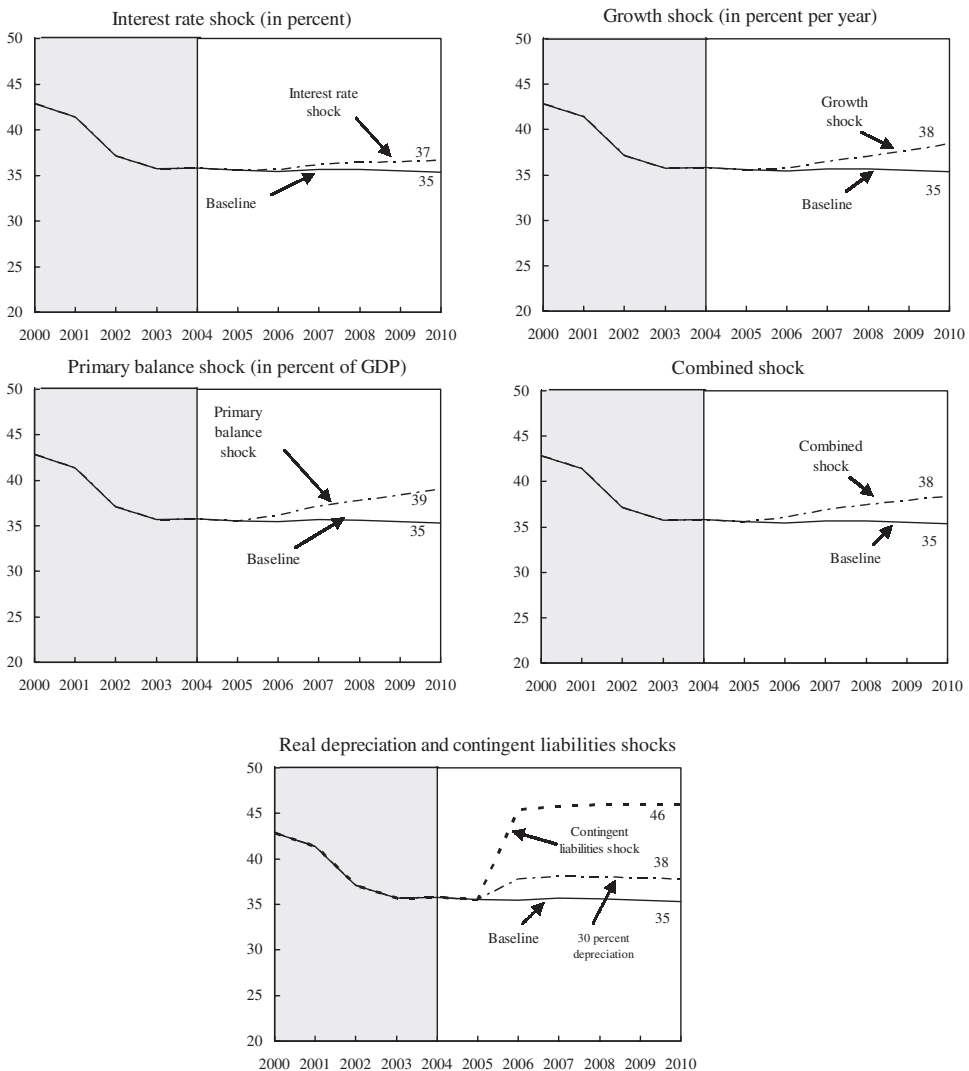
---

<sup>3</sup>That would be the case if a slowdown in economic activity were typically followed by a depreciation of the currency, rising interest rates, and a deteriorating primary balance.

PRIMARY SURPLUS BEHAVIOR AND RISKS TO FISCAL SUSTAINABILITY

For the sake of illustration, the deterministic bound-testing exercise is presented in Figure 1 for South Africa over 2004–10. The template provides debt paths corresponding to several scenarios: the baseline; small but permanent adverse shocks to real GDP growth, the real interest rate, and the primary balance; a combination of these three shocks; and two large temporary disturbances—namely, a 30 percent real depreciation and a shock to the debt stock equivalent to 10 percent of GDP. In line with South Africa’s stable economic environment and low external indebtedness, the selected bound tests suggest fairly low risks to public debt sustainability over the medium term (Debrun, 2005b).

Figure 1. South Africa: IMF Standard Debt Sustainability Analysis, 2004–10



Sources: South African National Treasury; and IMF staff estimates.

## Benefits of an Explicit Risk Assessment

One drawback of the deterministic bound-testing approach is that the underlying scenarios hardly ever follow shock patterns observed in the economy. Specifically, the method ignores co-movements among the determinants of debt dynamics, even though such co-movements are central in the stability of the debt ratio (see Goldfajn, 2005). Furthermore, the standardization of those tests implies fairly different degrees of realism across countries. Because the likelihood of the resulting debt paths cannot be precisely assessed, observers have no choice but to judge the plausibility of these bound tests on the basis of their core assumptions—for example, a growth slowdown without repercussions on interest rates or fiscal policy—rather than on their probabilistic merits in terms of debt outcomes.

A legitimate question is thus to ask whether a diagnostic based on a few stylized scenarios is robust to more realistic constellations of shocks. If a joint distribution of economic disturbances can be estimated for the country under review, stochastic simulations reflecting actual co-movements of shocks in the economy can produce a sample of more realistic bound tests from which country-specific frequency distributions of debt can be derived. These frequency distributions provide a quantitative assessment of the risks to the baseline debt projections that may ultimately help refine fiscal policy recommendations.

Another issue is the extent to which the sustainability diagnostic reflects plausible fiscal policy behavior and accounts for the fact that fiscal policy itself is a source of uncertainty. Commonly used DSA scenarios assume that fiscal policy is invariant to the stylized shocks. This assumption contributes to the policy dialogue by highlighting the consequences of policy inaction, but there are strong a priori grounds as well as ample evidence that the primary balance tends to respond systematically to variations in public debt and to the business cycle, among other factors. Ignoring this feedback may thus distort the risk assessment.<sup>4</sup> In particular, assuming a constant primary surplus in the face of business cycle fluctuations reduces the estimated risk to debt dynamics because slowdowns are generally accompanied by fiscal easing, whereas expansions often fail to improve the primary balance. Also, the estimated residuals of the reaction functions provide information on the stability of fiscal behavior, with erratic policy impulses being another independent source of uncertainty. Conversely, the passive policy assumption fails to capture the stabilizing response of the primary surplus to the debt itself, thereby increasing estimated risks to debt.

In addition to improved reliability of the risk analysis, accounting for systematic features of the policy process should lead policy advice to focus more on fiscal effort (measured as the departure from the estimated reaction function) rather than on the unconditional change in the fiscal balance. Deviations of actual policies from the benchmark provided by the estimated reaction function may prove useful in assessing such effort, and thereby assessing the room for fiscal adjustment (Abiad and Ostry, 2005; and Debrun, 2005a). Table 1 summarizes some key differences between the usual DSA framework and the extended DSA proposed in this paper.

---

<sup>4</sup>Our simulation tool can, however, accommodate any normative policy scenario, including the constant policy assumption.

**Table 1. Debt Sustainability Analysis (DSA) and Risk Assessment**

	Deterministic Bound Testing (commonly used DSA)	Probabilistic Approach (used in this paper)
Basis of diagnostic	A few stylized, isolated shocks; exogenous policies.	Large number of random shock constellations drawn from an estimated joint distribution; endogenous fiscal policy.
Calibration of shocks	Fraction or multiple of historical standard deviations of underlying variables; calibration based on the likelihood of the resulting debt path for a sample of emerging market economies.	Based directly on the estimated joint distribution of disturbances (country specific).
Output	Large temporary shocks provide a probabilistic upper bound to the debt ratio; small permanent shocks delineate interval of most probable outcomes.	Frequency distributions of the debt ratio over time; “fan charts.”
Main advantages	Amenable to standardized bound tests across countries; low data requirement.	Better reflection of country specificity (in terms of shocks and fiscal policy behavior); explicitly probabilistic output.

Sources: IMF and authors' calculations.

## The Simulation Algorithm: Overview and Limitations

The simulation algorithm presented here seeks to satisfy three main objectives: (1) provide a sensible way to account for fiscal behavior and simulate changes to it, (2) provide a tool easily applicable to different emerging market countries while referring to a common benchmark for policy assessment, and (3) keep data requirements close to those in the standard DSA.

The first objective uses the fiscal reaction function as an essential building block. The second points to panel data techniques to estimate average behavior across a group of countries to serve as a “positive” benchmark for each individual member of the group. The third suggests relying mainly on annual data.

These objectives impose a significant departure from existing algorithms.<sup>5</sup> Specifically, there is a need to initially separate the estimation of the fiscal reaction function from that of the other economic relationships before merging them again when simulating the behavior of the public debt ratio. There are (at least) two compelling reasons for doing so. First, the estimation of the variance-covariance

<sup>5</sup>IMF (2003) develops a tractable stochastic simulation tool that shares many features with subsequent research, including ours. However, this algorithm relies on annual data for all relevant variables, placing a premium on the availability of long and stable time series. In contrast, the simulation tools developed by Penalver and Thwaites (2004) and Garcia and Rigobon (2005), among others, require high-frequency data, including for fiscal variables. Such data requirements limit the number of countries to which these algorithms can be applied. To better internalize data constraints, we focus on a shorter time span (during which regime shifts and structural changes are less likely) while keeping annual fiscal data at the center of the analysis. This increases the number of countries to which our algorithm can be applied.



matrix of shocks inevitably relies on time-series techniques (an unrestricted Vector Auto-Regression or VAR model) that demand data more frequent than annual data. However, in contrast to financial variables and GDP, budgetary data at such frequency are often either unavailable or unreliable for the purpose of policy evaluation, exhibiting a very high noise-to-signal ratio.<sup>6</sup> The second reason is that the VAR framework restricts the specification of the reaction function in undesirable ways; for example, the primary surplus responds to contemporaneous variations in the output gap, not lagged ones, as imposed by the VAR framework.

Accordingly, our algorithm comprises three building blocks, as follows:

1. A fiscal reaction function—in the positive sense of a description of average fiscal policy patterns—is estimated for a panel of emerging market economies with annual data. In line with the literature, the specification is given by

$$p_{i,t} = \alpha_0 + \rho d_{i,t-1} + \gamma ygap_{i,t} + X_{i,t}\beta + \eta_i + \varepsilon_{i,t}, \quad t = 1, \dots, T, \quad i = 1, \dots, N, \quad (1)$$

where  $p_{i,t}$  is the ratio of the primary balance to GDP in country  $i$  and year  $t$ ;  $d_{i,t-1}$  is the public debt-to-GDP ratio observed at the end of period  $t - 1$ ;  $ygap_{i,t}$  is the output gap;  $\eta_i$  is an unobserved, country fixed effect; and  $X_{i,t}$  is a vector of control variables.

2. For each country, we estimate an unrestricted VAR model comprising the nonfiscal determinants of public debt dynamics. Formally, we have  $Y_t = \gamma_0 + \sum_{k=1}^p \gamma_k Y_{t-k} + \xi_t$ , where  $Y_t = (r_t^{us}, r_t, g_t, z_t)$ , and  $\gamma_k$  is a vector of coefficients;  $r_t^{us}$  denotes the real foreign interest rate;  $r_t$  the real domestic interest rate;  $g_t$  the real GDP growth rate;  $z_t$  the (log of the) real effective exchange rate; and  $\xi_t$  is a vector of well-behaved error terms:  $\xi_t \sim N(0, \Omega)$ . This model serves two purposes. First, the variance-covariance matrix of residuals  $\Omega$  characterizes the joint statistical properties of the contemporaneous, nonfiscal disturbances affecting debt dynamics. Specifically, the simulations use a sequence of random vectors  $\hat{\xi}_{t+1}, \dots, \hat{\xi}_T$  such that  $\forall \tau \in [t + 1, T]$ ;  $\hat{\xi}_\tau = Wv_\tau$ , where  $v_\tau \sim N(0, 1)$ ; and  $W$  is such that  $\Omega = W'W$  ( $W$  is the Choleski factorization of  $\Omega$ ). Second, the VAR generates forecasts of  $Y$  consistent with the simulated shocks. Because shocks occur each period, the VAR produces joint dynamic responses of all elements in  $Y$ . It should be noted that the method is not sensitive to the ordering of variables in the VAR. Ordering matters only to the extent that one tries to capture causal relationships between innovations and the other variables (for example, for impulse response functions). In the present context, stochastic simulation results are shaped by the variance-covariance matrix of reduced-form errors  $\Omega$ , which is unique (see also Garcia and Rigobon, 2005).
3. For each simulated constellation of shocks, quarterly VAR projections are annualized, and the simulated annual output gap results from the growth differential between predicted GDP growth and the (annualized) steady-state growth rate

<sup>6</sup>Higher-frequency budgetary data (quarterly or monthly) serve cash management purposes. Even though countries with IMF-supported programs or greater leeway to vote for supplementals during the budget year may well undertake policy corrections on a quarterly basis in response to shocks or slippages, the overall policy stance still tends to be reflected in the annual figures. Wyplosz (2005) finds that standard reaction functions fit monthly data for Brazil very poorly.



produced by the VAR (to ensure that shocks to the output gap are zero on average). The corresponding debt path can then be calculated recursively using equation (1) and the conventional stock-flow identity

$$d_t \equiv (1 + g_t)^{-1} \left[ (1 + r_t^{us})(1 + \Delta z_t) d_{t-1}^* + (1 + r_t) \tilde{d}_{t-1} \right] - p_t + s_t,$$

where  $d_t^*$  captures the foreign currency-denominated debt;  $\tilde{d}_t$  designates the public debt in domestic currency; and  $s_t$ , stock-flow adjustments—for instance, due to the recognition of contingent liabilities or the realization of assets. Notice that in each simulation, the primary surplus incorporates a fiscal policy shock  $\phi_{i,t} \sim N(0, \sigma_{\phi_i}^2)$ , where  $\sigma_{\phi_i}^2$  is the country-specific variance of the reaction function's residuals.

This algorithm can produce an arbitrarily large number of debt paths corresponding to different shock constellations. Frequency distributions of the debt ratio can then be obtained for each year of projection, and used to draw fan charts and the probabilistic sustainability indicators discussed in Section III.

Results from using this algorithm are still subject to a number of limitations. First, in many emerging market economies, identifying stable economic relationships is a challenge: the lack of long time series combined with ongoing structural change and shifts in policy regime inevitably reduce the reliability of econometric estimates for predicting future developments. A related issue is that the uncertainty captured by econometric models is not limited to the behavior of residuals, but also finds an expression in the standard errors of estimated coefficients. One way to incorporate this in our algorithm and provide a more encompassing assessment of risk would be to randomize the coefficients of the underlying economic model. However, this raises a number of concerns. One is that randomization could yield a nontrivial number of parameter constellations associated with unstable solutions for the system. Such an approach, moreover, would likely obfuscate rather than illuminate policy discussions. Beyond this, compounding two very different sources of uncertainty—well-identified shocks on key variables on the one hand and doubts about the reliability of the model used to discipline our projections on the other—would considerably reduce the transparency of the simulation procedure and complicate the interpretation of the results. One nevertheless needs to remain aware that the risk assessment emerging from our method is probably a lower bound.

Second, extreme situations—such as crises—are bound to remain low-probability events in this framework because the simulated shocks are drawn from a joint normal distribution. The normality assumption may be at odds with a more frequent occurrence of adverse events (asymmetric distribution of shocks) or the greater likelihood of extreme events (“fat tails”) in some countries or at certain times. However, explicitly dealing with nonnormal disturbances could undermine the simplicity and cross-country comparability of the current approach. Moreover, it would presume that future disturbances are bound to resemble past ones. That said, evidence of nonnormal shock distributions should encourage a discussion of the likely implications on the balance of risk reflected in our fan charts.

Third, the combination of annual and quarterly data shuts off any feedback effect of fiscal policy on economic variables (the causation runs only in the other

direction) and, in particular, interest rates (through credibility effects) and growth (through the quality of fiscal policy). This may be an important issue if fiscal reforms, for example, are likely to produce significant changes in the future course of growth and interest rate spreads. Our simulation tool allows imposing such effects on the results by overriding the steady-state values from the VAR.

The next section turns to the first building block of our model and provides a benchmark fiscal reaction function for a group of emerging market economies.

## II. Debt Dynamics and the Conduct of Fiscal Policy

As shown by Bohn (1998), governments concerned with solvency would be expected to raise the primary balance in response to an increase in the public debt-to-GDP ratio. If all other determinants of fiscal policy are stationary, the positive correlation between debt and the primary surplus is sufficient to guarantee that the debt ratio will revert to some finite steady-state value. This section describes the estimation of fiscal reaction functions for a group of emerging market economies over 1990–2004, with regressors including debt and a range of other economic, policy, and institutional variables of interest.

### Primary Surplus Behavior and Public Debt Sustainability

A growing number of studies have estimated fiscal reaction functions (Mélitz, 1997; Galí and Perotti, 2003; IMF, 2003 and 2004; and Wyplosz, 2005, among others). The aim has been to identify a stable relationship between fiscal policy (measured by the primary balance) and various potential determinants. Such an exercise does not necessarily attempt to establish causality; the idea is rather to extract information on the key considerations that may shape and be correlated with policy decisions. Debt sustainability is expected to be one of those considerations, along with cyclical developments and institutions affecting a government's incentives. Accordingly, a version of equation (1) is commonly estimated.

One difficulty with estimating fiscal rules is the scarcity of relevant budgetary data, which has led researchers to use panel data methods. The most meaningful data from the perspective of policy evaluation are available annually, in line with the budget procedure of most countries. Although fiscal policy adjustments may occur more frequently, these signals are blurred by the intrinsically noisy nature of high-frequency budgetary data.<sup>7</sup>

Panel estimation assumes similar fiscal behavior across countries. To account for possible heterogeneity, we control for a large number of potential determinants of fiscal policy, and also explore the possibility of nonlinear relationships between the primary balance and some of its determinants.

### Estimating Fiscal Reaction Functions

In line with the literature, the reaction function we estimate, given in equation (1), relates the annual primary fiscal balance to the outstanding level of public debt, the

---

<sup>7</sup>One notable exception is in the case of an IMF-supported program, where quarterly reviews of policy implementation take place.

gap between actual and trend output, and a number of potentially important drivers of the primary balance in emerging market economies. These include real oil prices in oil exporters, an index of institutional quality, and two indicator variables accounting for whether the country is in a state of sovereign default and whether it is committed to an IMF-supported program.

The estimation of equation (1) needs to take into account three sources of endogeneity bias. The first is simply the endogeneity of the output gap with respect to contemporaneous fiscal policy shocks,  $\epsilon_{i,t}$ . The remaining two sources stem from the dependence of lagged debt on past values of the primary surplus. As to the second of the three sources, clearly the lagged debt level,  $d_{i,t-1}$ , will necessarily be correlated with the country-specific and time-invariant determinants of primary surpluses,  $\eta_i$ ; countries able to generate higher surpluses on average—captured by higher values of the fixed effects  $\eta_i$ —will tend to have a lower level of public debt, and if this is not properly accounted for, the negative correlation between debt levels and the unobserved country fixed effects would exert a downward bias on the estimated primary surplus response to debt. As to the third source of endogeneity, to the extent that there is persistence in the idiosyncratic error term,  $\epsilon_{i,t}$ , the dependence of lagged debt on past surpluses will render lagged debt in equation (1) endogenous.<sup>8</sup>

Ideally, one could address these potential endogeneity problems with adequate instrumentation. To tackle the first source of endogeneity, the output gap needs to be instrumented with variables that are exogenous to the idiosyncratic primary surplus shocks. The second source of endogeneity—the endogeneity of debt to the primary surplus fixed effects—can in principle be addressed by the inclusion of country indicator variables among the regressors, but this method would still be subject to the third endogeneity problem if there is strong serial correlation in the idiosyncratic errors.<sup>9</sup> Moreover, the use of country dummies can potentially introduce an additional bias, commonly referred to as the small-sample bias of the fixed-effects estimator.<sup>10</sup> By contrast, instrumenting both the output gap and lagged debt (with variables that are exogenous to the primary surplus fixed effects and to the idiosyncratic errors) would simultaneously address all three potential endogeneity problems. That said, reliable instrumental variable (IV)-based estimations require the use of suitable exogenous instruments that are strongly correlated with the endogenous regressors. Such ideal instruments are difficult to find in our context. Moreover, IV-based regressions are generally not very efficient, yielding estimates with relatively large standard errors.

<sup>8</sup>For instance, a positive shock to the surplus in period  $t - 1$ , that is, a positive realization of  $\epsilon_{i,t-1}$ , would reduce the debt stock,  $d_{i,t-1}$ . Thus, persistence in the idiosyncratic policy shocks (serial correlation between  $\epsilon_{i,t-1}$  and  $\epsilon_{i,t}$ ) would result in a negative correlation between  $d_{i,t-1}$  and  $\epsilon_{i,t}$ .

<sup>9</sup>The inclusion of country indicator variables addresses the endogeneity of debt to  $\eta_i$  by transforming the equation to eliminate  $\eta_i$ . Specifically, including country dummies is equivalent to performing the regression on the variables in deviations from country means.

<sup>10</sup>A large body of literature analyzes the bias of the least squares with dummy variables estimator in dynamic models that include the lagged dependent variable as a regressor. The bias of this estimator decreases with the time dimension of the sample and the variance of the lagged dependent variable that is attributable to factors other than the disturbance terms (see Kiviet, 1995; or Judson and Owen, 1999). Our model falls into the category of dynamic panel models, given that lagged debt is among the regressors.

Against this background, our strategy is to estimate five possible specifications. The first two, a Limited Information Maximum Likelihood (LIML) regression and a system Generalized Method of Moments (system GMM) specification, respectively, instrument for the output gap and lagged debt, and exclude country dummies. A third uses instruments for the output gap only and includes country dummies to account for the fixed effects. This specification eliminates the first two sources of endogeneity, but not the endogeneity from the persistence in idiosyncratic policy shocks; it should yield results similar to the first two methods if the serial correlation in the errors is weak and if the small-sample bias associated with the use of country dummies is small.<sup>11</sup> Beyond this, columns 4–5 in Table 2, discussed below, incorporate nonlinearities to better capture heterogeneities in fiscal behavior across countries and circumstances.

### Estimation Results

The panel presented here comprises 34 countries and a maximum of 15 years (1990–2004); data on primary balances and public debt levels were obtained from IMF country desk economists for the largest available coverage of the fiscal sector.<sup>12</sup> For the linear reaction function, we present in columns 1–3 of Table 2 three specifications. The first eliminates the country effects by using first differences and instruments for the lagged change in debt and contemporaneous change in the output gap, using as instruments lags of one-year U.S. bond rates, changes in real oil prices, lagged fiscal costs of banking crises, and import demand in industrial-country trading partners.<sup>13</sup> We run a LIML regression, which is preferable to GMM if the instruments are not very strong. In the second regression, we implement Blundell-Bond (1998) system-GMM (SGMM), which jointly estimates the level and differenced forms of the equation, using lagged differences and levels of the endogenous regressors as instruments in addition to the exogenous instruments used in the LIML regression. Third, we estimate a version with country dummies, instrumenting only the output gap with import demand in industrialized trade partners (GMM-DV).

All three estimations suggest a positive response of primary surpluses to the debt. With a positive coefficient on the output gap, primary balances are estimated to be countercyclical. However, this effect is driven mostly by the worsening of the balance during recessions rather than improvements during booms, as discussed below in the context of the nonlinear specification. IMF-supported programs encourage countries to run higher surpluses. And countries in default run larger primary balances, reflecting their restricted market access.

The estimated coefficient of the index of institutional quality is mostly negative, but significant only in the GMM-DV regression. An interesting observation is that the estimated country fixed effects in the GMM-DV regression are positively correlated with the average institutional quality over the sample period (Figure 2), suggesting that countries with stronger institutions run larger primary balances on

<sup>11</sup>Monte Carlo simulations suggest that the magnitude of the (positive) bias on  $\rho$  is less than 20–30 percent of the true coefficient (Celasun and Kang, 2006).

<sup>12</sup>See Celasun, Debrun, and Ostry (2006) for a description of the data set.

<sup>13</sup>The fiscal costs of banking crises typically take the form of below-the-line expenditures, thereby increasing the public debt burden without affecting recorded primary surpluses.

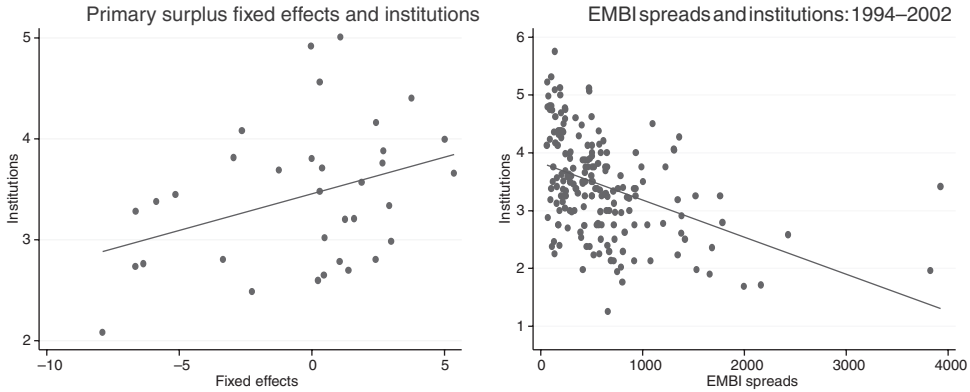
Table 2. Estimates of the Fiscal Reaction Function, 1990–2004

	Dependent Variable: Level or Difference of the Primary Fiscal Balance				
	(1) LIML (Difference)	(2) System- GMM	(3) GMM with DV	(4) LIML (Difference)	(5) GMM with DV
Lagged debt	0.039 [0.032]	0.030*** [0.007]	0.046*** [0.008]	0.121 [0.172]	0.097*** [0.036]
Output gap	0.104 [0.109]	0.217*** [0.072]	0.328*** [0.113]		
Real oil price	0.481*** [0.072]	0.084** [0.030]	0.354*** [0.082]	0.487*** [0.112]	0.361*** [0.086]
Institutions	0.374 [0.484]	-0.219 [0.322]	-0.675*** [0.258]	0.463 [0.445]	-0.380 [0.256]
IMF program	0.765** [0.347]	1.121 [0.689]	1.108*** [0.328]	0.777** [0.344]	0.939** [0.328]
Default	0.870** [0.351]	0.884 [0.813]	1.190*** [0.401]	0.749*** [0.297]	1.077*** [0.368]
Debt spline (50 percent)				-0.108 [0.194]	-0.062* [0.037]
Positive output gap				-0.092 [0.358]	0.181 [0.631]
Negative output gap				0.258 [0.246]	0.268 [0.225]
Constant		-0.684 [1.479]	-0.963 [1.138]		-3.628 [2.892]
Country dummies	No	No	Yes	No	Yes
Observations	349	399	418	368	418
Hansen test ( <i>P</i> -value)	0.84	1.00	0.45	—	0.03
AR(1) test ( <i>P</i> -value)		0.05			
AR(2) test ( <i>P</i> -value)		0.09			
Cragg-Donald statistic	7.23		19.63	1.96	

Source: Authors' calculations.

Notes: System-GMM = system Generalized Method of Moments; LIML = Limited Information Maximum Likelihood. Standard errors are in brackets. \* denotes significance at 10 percent; \*\* at 5 percent; \*\*\* at 1 percent. *P*-values of the test statistics are reported for the tests of overidentifying restrictions and the tests of serial correlation in the residuals of the difference equation in the system-GMM regressions (the AR tests). In the LIML regression in the first column, the second and third lags of U.S. one-year bond rates, second and third lags of the changes in real oil prices, lagged fiscal banking crisis costs, and the contemporaneous value of trade partners' import demand were used as instruments for lagged debt and the output gap. The Blundell and Bond (1998) system-GMM regression in column 2 uses the second lags of the output gap and debt, in addition to the banking fiscal cost measure and the trade partners' import demand. The third column presents a GMM regression with country dummies, where the output gap is instrumented with the contemporaneous and lagged values of trade partners' import demand. The equation in column 4 is exactly identified, hence there is no test of overidentifying restrictions. The instruments include lagged fiscal banking crisis costs, the contemporaneous value of trade partners' import demand, the interaction of these variables with a dummy that indicates whether debt exceeds 50 percent, and a dummy that indicates whether the output gap is positive. The estimation in column 5 instruments only for the positive and negative output gap terms, using the interactions of the trade partners' import demand measure.

**Figure 2. Institutional Quality, Primary Surplus Behavior, and Emerging Market Bond Index (EMBI) Spreads**



Sources: Authors’ calculations; International Country Risk Guide; and Bloomberg.

Notes: The left panel plots the estimated country fixed effects in the GMM-DV regression against average institutional quality over the sample period. The right panel plots EMBI spreads against institutional quality over the period 1994–2002.

average. The negative estimated impact of institutions may stem from the fact that improvements in institutional quality are typically associated with decreases in borrowing costs, implying that countries would need to run smaller primary balances to service a given level of debt as their institutions improve.

Columns 4 and 5 in Table 2 present estimates of the nonlinear fiscal reaction function, which allows for a “kinked” response to debt at 50 percent of GDP, and a different response to the output gap depending on the latter’s sign.<sup>14</sup> The specification takes the form

$$\begin{aligned}
 p_{i,t} = & \alpha_0 + \rho d_{i,t-1} + \bar{\rho} D_{i,t-1} (d_{i,t-1} - 50) + \gamma^p P_{i,t} ygap_{i,t} + \gamma^n N_{i,t} ygap_{i,t} \\
 & + X_{i,t} \beta + \eta_i + \varepsilon_{i,t},
 \end{aligned}
 \tag{2}$$

where  $P_{i,t}$  and  $N_{i,t}$  are dummy variables whether the output gap is positive or negative, respectively.

The larger number of parameters and required instruments limits the choice of estimation methods.<sup>15</sup> Thus, for this specification, we present a LIML regression that excludes country dummies and uses instruments for lagged debt and the output gap,<sup>16</sup> and a GMM regression that includes country dummies and instruments only for the (negative and positive) gap terms.

<sup>14</sup>Previous work suggests a structural shift in the primary surplus equation when debt reaches 50 percent of GDP and a different response of the surplus depending on the sign of the output gap (Abiad and Ostry, 2005).

<sup>15</sup>Given the interaction terms in the specification, the number of excluded instruments needed for the nonlinear specification is double that needed for the linear specification.

<sup>16</sup>In the LIML estimation in column 4, the banking-crisis fiscal cost measure and its interaction with  $D_{i,t-1}$  were used to instrument for lagged debt and splined debt. The interactions of the import demand measure with  $P_{i,t}$  and  $N_{i,t}$  were used as instruments for the gap measures. This equation was thus exactly identified.



Although the findings are less precise than those obtained for equation (1), they qualitatively confirm that the response to debt is stronger when debt is less than 50 percent of GDP, and that the response to booms and recessions is asymmetric. In particular, the worsening in primary balances during recessions exceeds the improvements attained during economic booms.

For our baseline calibration exercise in the next section, we use the GMM-DV parameter estimates (column 3), which are more precise. Estimated country fixed effects are also useful for gauging heterogeneity across countries. SGMM estimates (column 2) are also proposed as a robustness check. For simulations using the non-linear specification, we use the GMM-DV estimates in column 5.

### III. Risks to Debt Sustainability in Five Emerging Market Economies

This section proposes various prospective and retrospective risk analyses of public debt in five emerging market economies with fairly different risk profiles: Argentina, Brazil, Mexico, South Africa, and Turkey. After a discussion of the calibration, we apply the simulation algorithm outlined in Section I to generate a sample of 1,000 simulations from which we derive frequency distributions of public debt. We then discuss possible tools of analysis—fan charts and a summary indicator of debt sustainability. Finally, we examine the sensitivity of the risk analysis to variations in the underlying assumptions.

#### Calibrating the Simulations

For a given country, all simulations assume the same joint distribution of disturbances and co-movements among the variables.

The fiscal reaction function combines both standardized and country-specific features as follows—an upper “hat” designates parameter estimates obtained in Section II as well as the corresponding predictions:

$$\hat{p}_{i,t+\tau} = \Lambda_{i,t+\tau} + \hat{\rho}d_{i,t+\tau-1} + \hat{\gamma} ygap_{i,t+\tau} + \varphi_{i,t+\tau}, \text{ for } \tau = 1, \dots, 5, \quad (3)$$

with  $\Lambda_{i,t+\tau} = \hat{p}_{i,t} - \hat{\rho}d_{i,t-1} - \hat{\gamma} ygap_{i,t} + \kappa_{i,t+\tau}$ ; and  $\varphi_{i,t+\tau}$ , a policy shock drawn from a mean-zero normal distribution with variance equal to the country-specific variance of residuals.<sup>17</sup>

Equation (3) splits fiscal policy into an automatic, a predetermined, and a random part. The automatic component follows the average response of the primary balance to the public debt and to the output gap, and is considered identical for all countries. The predetermined part, summarized by  $\Lambda_{i,t+\tau}$ , captures the impact of all other determinants of the primary surplus, including institutional quality, the existence of an IMF-supported economic program, a default/restructuring option, and, if relevant, the budgetary effect of oil-price fluctuations. By default,  $\kappa_{i,t+\tau} = 0$  so that  $\Lambda_{i,t+\tau}$  is a country-specific constant providing an anchor to simulated primary balance paths. We also allow for non-zero, time-varying values for  $\kappa_{i,t+\tau}$  to

<sup>17</sup>Recall that the fiscal disturbance is assumed to be orthogonal to economic developments.



account for specific information about future policy changes, such as those related to institutional reforms likely to affect fiscal performance, the adoption of an IMF-supported adjustment program, or, alternatively, discretionary impulses envisaged in medium-term budget plans. The estimated reaction functions discussed in Section II may provide some guidance for the calibration of  $\kappa_{i,t+\tau}$ .

Although this paper emphasizes the usefulness of estimated fiscal reaction functions, the simulation framework can accommodate a range of policy behavior, including normative scenarios (for example, constant primary surplus or program targets). Where long data series are available, parameters corresponding to country-specific estimates of the fiscal reaction function can also be used. Finally, it is worth noting that deterministic stress tests representing shocks ignored in the empirical model, such as the materialization of contingent liabilities or other “fat tail” events, can be combined with the stochastic analysis (see Debrun, 2005b).

### Baseline Scenarios

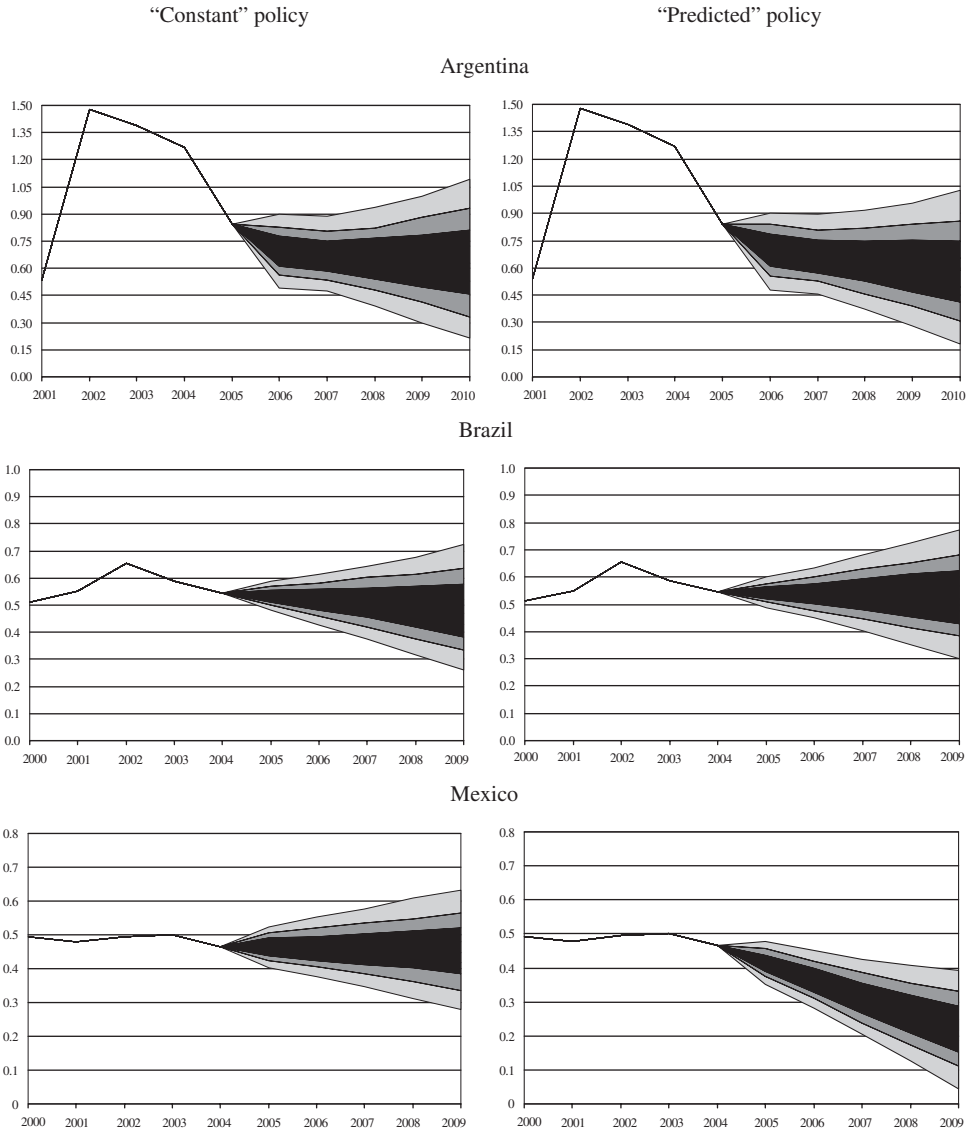
Our baseline scenarios allow only for automatic responses of the primary balance to output shocks and public debt developments. One issue, however, is that  $\Lambda_{i,t+\tau}$  incorporates the residual of the reaction function in year  $t$ . The discrepancy between the predicted surplus and actual behavior may result from discretionary slippages or adjustment efforts. Depending on whether that deviation is expected to be permanent or temporary, two baseline scenarios can be envisaged.

In the first one (the “constant” policy scenario), any deviation from the predicted primary surplus in  $t$  is assumed to persist over the entire simulation horizon, as if the most recent stance signaled a sustained departure from past primary surplus behavior (in that case,  $\kappa_{i,t+\tau} = 0$  over the entire simulation horizon). That scenario is reminiscent of the constant policy scenario commonly found in deterministic DSAs, although our framework freezes only the (non-debt-related) discretionary part of policy. Alternatively (the “predicted” policy scenario below), recent deviations from the predicted primary surplus may result from purely temporary factors, in which case it is more appropriate to assume that the simulated surplus paths follow those predicted by the reaction function (so that  $\kappa_{i,t+\tau} = -\hat{\epsilon}_{i,t}$  over the simulation horizon). In both cases, surplus behavior follows a common reaction function (see Section II). Simulations are performed for  $t = 2004$  and  $\tau = 5$ , except for Argentina, for which we used  $t = 2005$  to account for the debt exchange operation. In line with the DSA analysis presented in IMF (2005), for Argentina we used federal-level fiscal data rather than the more comprehensive data used for the other countries in the sample.

Fan charts summarize risks to debt dynamics by representing the frequency distribution of a large sample of debt paths generated by means of stochastic simulations (Figure 3). Different colors delineate deciles in the distributions of debt ratios, with the zone in dark grey representing a 20 percent confidence interval around the median projection, and the overall colored cone, a confidence interval of 80 percent. A number of results flow from the charts.

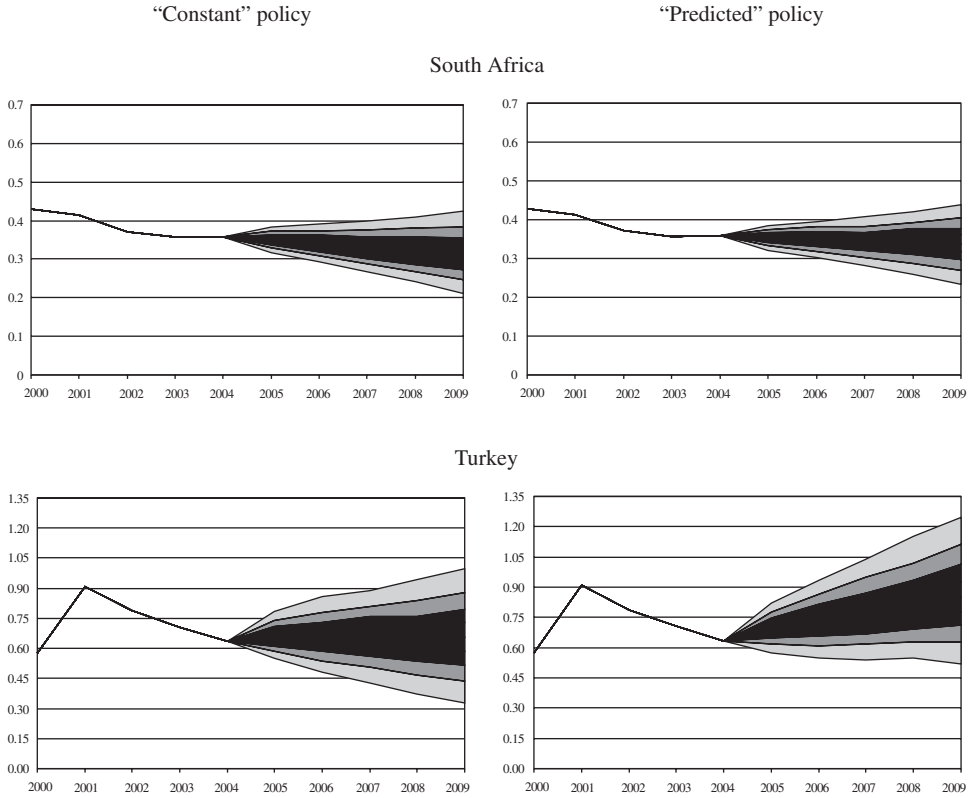
First, compared with the outcome of simple bound tests (reported in Figure 1 in the case of South Africa), it appears that these fall within the 40 percent confidence

Figure 3. Fan Charts for Public Debt-to-GDP Ratios in Five Emerging Market Economies, 2005-10



interval only, confirming that deterministic bound tests imperfectly account for the overall magnitude of risks to public debt sustainability. Second, in the five countries, the fiscal reaction function generally appears sufficiently responsive to the public debt to ensure that the median debt path is sustainable (in the sense that the debt ratio is stable or declining over the simulation horizon). However, in most instances, the response proves too weak to prevent growing debt ratios in the two upper bands of the charts, representing the second and third deciles of the debt ratio. Hence, in those

Figure 3 (concluded)



Source: Authors’ calculations.

cases, there is at least a 30 percent chance that combinations of adverse economic and policy shocks may lead to concerns over debt sustainability.

Third, the overall risk profile for the different countries reflects the intrinsic volatility of their respective economies, with less-volatile economies—South Africa, for example—exhibiting narrower confidence intervals than others. Of course, our assessment is contingent on the relatively short period of time over which the VARs are estimated. In particular, wide confidence intervals inevitably reflect past crises, and may thus overestimate the true magnitude of risks. This is especially evident in the simulations for Argentina and Turkey.

Fourth, the fan charts for Mexico and Turkey illustrate the issue of large residuals in the last year of observation. In Mexico’s case, a highly negative residual (close to 5 percent of GDP) was observed in 2004, reflecting the lesser sensitivity of Mexico’s primary balance to oil prices in comparison with other oil producers in the panel. The large increase in real oil prices in 2004 thus showed up as a residual. A more detailed analysis of Mexico could discuss alternative scenarios, including a reaction function excluding oil prices as a determinant of the primary balance. In Turkey, large positive residuals were observed in 2003–04, reflecting significant

fiscal adjustment efforts (Debrun, 2005a). Here, the charts clearly show that sustaining such high surpluses (under the constant policy scenario) is key to containing the risks to debt sustainability.

Finally, fan charts reflect the intrinsic asymmetry of risks to the debt ratio because its level magnifies the impact of growth rate and interest rate shocks. In Argentina and Turkey, for instance, the distribution of debt ratios is clearly skewed toward the upside, suggesting that the standard response of the primary balance to debt may prove insufficiently stabilizing in unfavorable times.

### Sustainability Assessments

So far, the value added in our approach has been to produce probability distributions of debt at different horizons, rather than a deterministic path for debt.<sup>18</sup> But what is the policymaker to make of all these distributions? Presumably, recognizing that the statement, “There is a 30 percent chance that debt will be below a given ratio to GDP in three years’ time, and a 20 percent chance that it will be above another threshold over the same horizon,” contains more information than the simple assertion, “Under present policies, the debt ratio will decline over time or rise over time,” the policymaker is ultimately likely to be interested in whether the country’s debt profile and distribution is a problem or not. This is the issue we try to get at in this section, recognizing that what is an acceptable risk for one policymaker may be an unacceptable risk for another.

At a basic level, the prospect of a downward trend in the debt ratio in a deterministic setup corresponds to the probability that the debt ratio will fall below its initial value in our stochastic setting. A probability is clearly more informative from the point of view of generating useful sustainability indicators than what is available to the policymaker as a result of a deterministic exercise. But policymakers may be interested in more than simply an assessment of whether debt is likely to decline or not, particularly if they are very concerned about the possibility that the debt ratio could rise (that is, are very averse to upside risk).

Clearly, two countries with the same probability of a declining debt ratio may face quite different upside risks for debt. To account for the overall risk profile, one can posit a sustainability indicator that combines the probability that debt will decline over time with the risk that debt will not rise beyond some specific (or tolerable) amount over the same horizon. A plausible indicator could be calculated as  $Pr(d_{t+\tau} < d_t) \times [1 - Pr(d_{t+\tau} > (d_t + x))]$ , where  $x$  is a positive markup over the initial public debt  $d_t$ . The value of the sustainability indicator increases with the likelihood of desirable outcomes; that is, a nonincreasing trend, and well-contained upside risks. That value has no intrinsic meaning, and it is up to the policymaker to set a critical threshold below which the debt situation would be a cause for concern. Such a threshold essentially depends on the degree of risk aversion and on the perceived need for a reduction in debt. These factors are reflected in the markup  $x$ , the level

<sup>18</sup>Although not undertaken as an explicit exercise in this paper (but see Abiad and Ostry, 2005), the approach can also be used to assess how the distribution of debt shifts in response to changes in policy fundamentals that are captured in the estimated reaction function (such as institutional reform, an IMF-supported program, etc.).

of upside risk deemed acceptable (that is, how small should  $Pr(d_{t+\tau} > (d_t + x))$  be in order to consider that upside risks are well contained), and the desired probability that debt will decline in the future.

Because the sustainability indicator is the product of these two probabilities—the probability of a declining debt ratio times the probability that the debt ratio will not rise by more than the markup—once these probabilities are set, a critical threshold is established. The analyst or the policymaker can then focus on problematic cases where the indicator’s value falls below the threshold value of the sustainability indicator, and not worry the rest of the time.

In Table 3, we calculate the sustainability indicator for  $x = 0.1$  (10 percent of GDP). Which entries in the table are worrisome, and which are benign? To illustrate, we pick a threshold value of 0.4. Many combinations of  $Pr(d_{t+\tau} < d_t)$  and  $[1 - Pr(d_{t+\tau} > (d_t + x))]$  can deliver such a value. For example, the level of comfort associated with that threshold corresponds to a probability of at least 50 percent that the debt ratio declines and a probability of less than 20 percent that debt rises by more than 10 percent of GDP:  $0.40 [= 0.50 \times (1 - 0.20)]$ ; alternatively, of course, probabilities of decline of less (more) than 50 percent could be offset by lower (higher) probabilities that the debt ratio will rise by more than 10 percent of GDP.

In Table 3, entries where the sustainability index is less than 0.4 are italicized.<sup>19</sup> The indicator underscores vulnerability, particularly in the case of Turkey, which stems mainly from the upside risks to debt driven by the shocks—notably to domestic real interest rates and the exchange rate—identified through the VAR. Mitigating this vulnerability requires, as shown by a comparison of the constant policy and predicted policy fan charts, persisting with Turkey’s strong fiscal effort of recent years.

Probabilistic indicators of debt sustainability can prove useful for policymakers if they convey a credible signal that fiscal policy needs to be changed in order to reduce the likelihood of adverse outcomes to acceptable levels. A test of our approach to sustainability is thus whether it gives appropriate warnings of trouble on the eve of a crisis. Given data constraints, we investigate this issue—namely, the track record of our sustainability indicator—in three “eve-of-crisis” cases: Argentina and Brazil at end-2000 and Turkey at end-1999, using only information available to policymakers at the time.<sup>20</sup> As a control, we also introduce a retrospective analysis of South Africa (at end-2000) to check whether our model would have properly differentiated the risks faced by these countries.

The fan charts in Figure 4 point to upside risks in all three crisis countries, with Argentina looking particularly vulnerable to explosive outcomes. Turkey’s debt dynamics also looked unmanageable under most circumstances, and Brazil’s situation appeared under control, albeit with significant upside risks.

<sup>19</sup>More cells would be italicized if the policymaker wanted to reduce the probability of upside risk to, say, 10 percent from 20 percent: in that case the problematic entries would be all those where the sustainability indicator is lower than 0.45. At the extreme, if the policymaker were extremely averse to upside risk, all cells below 0.50 would be italicized, signaling more worrisome cases than those shaded in Table 3.

<sup>20</sup>The dating of crises is problematic for Brazil and Turkey, because neither country defaulted. Although it would be straightforward to extend the exercise to other dates, the dates chosen correspond to instances when sovereign spreads rose to very high levels (above, say, 1,000 basis points) at some point during the following year.

**Table 3. Probabilistic Debt Sustainability Assessment**  
(*“Constant” policy scenario*)

	$t + 1$	$t + 2$	$t + 3$	$t + 4$	$t + 5$
Debt ratio lower than in $t$					
Argentina ( $t = 2005$ )	0.83	0.85	0.82	0.76	0.73
Brazil ( $t = 2004$ )	0.59	0.63	0.64	0.63	0.64
Mexico ( $t = 2004$ )	0.51	0.54	0.53	0.55	0.56
South Africa ( $t = 2004$ )	0.60	0.64	0.69	0.70	0.71
Turkey ( $t = 2004$ )	0.40	0.45	0.46	0.47	0.47
Debt ratio more than 10 percent of GDP higher than in $t$					
Argentina ( $t = 2005$ )	0.06	0.06	0.10	0.15	0.20
Brazil ( $t = 2004$ )	0.00	0.05	0.10	0.15	0.19
Mexico ( $t = 2004$ )	0.03	0.08	0.11	0.16	0.20
South Africa ( $t = 2004$ )	0.00	0.00	0.00	0.03	0.04
Turkey ( $t = 2004$ )	0.21	0.28	0.34	0.35	0.38
Sustainability index					
Argentina ( $t = 2005$ )	0.78	0.79	0.74	0.65	0.59
Brazil ( $t = 2004$ )	0.59	0.60	0.57	0.54	0.52
Mexico ( $t = 2004$ )	0.50	0.50	0.47	0.46	0.45
South Africa ( $t = 2004$ )	0.60	0.64	0.69	0.68	0.68
Turkey ( $t = 2004$ )	0.32	0.32	0.30	0.30	0.29

Source: IMF staff calculations.

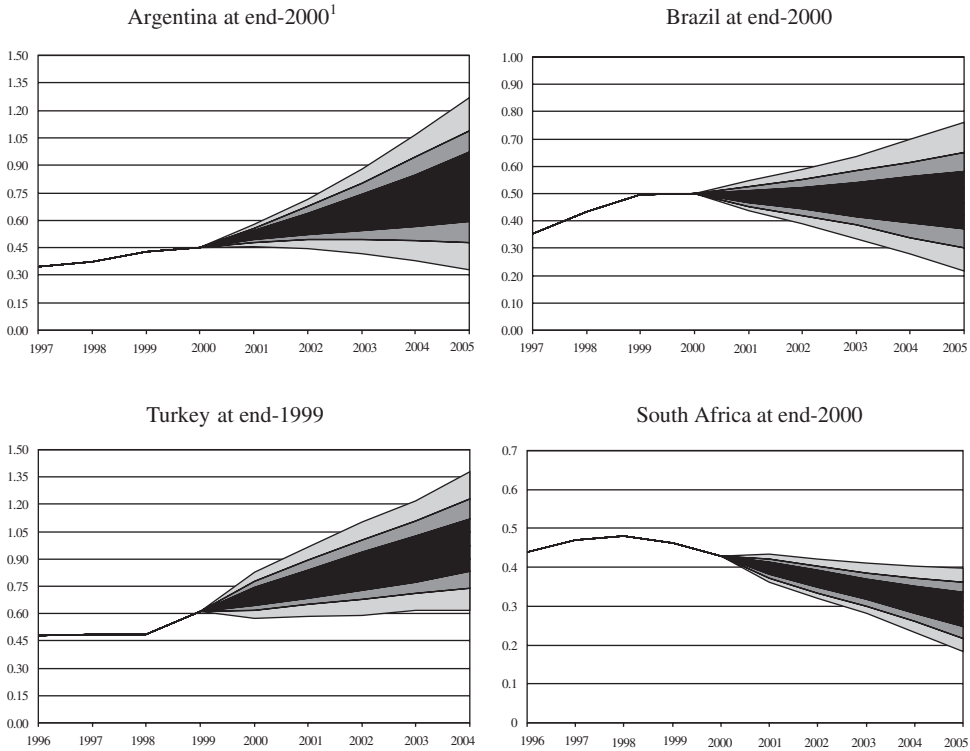
This indicates that the endogenous debt-stabilizing response typically observed in emerging market economies was insufficient to prevent explosive debt dynamics, calling for significant fiscal adjustment efforts in Argentina and Turkey, and prudent fiscal management in Brazil. In contrast, our model nicely predicts the declining debt path that was effectively observed in South Africa.

This impression is confirmed by our sustainability indicator (Table 4), which exhibits low and declining values for Argentina and Turkey over the simulation horizon. The situation in Brazil at end-2000 looked less alarming than that in Argentina and Turkey, although the dramatic decline in the indicator over the simulation horizon signaled a rapid deterioration in the sustainability outlook. In sum, however, the proposed sustainability indicator would have been providing strong signals of impending troubles for Argentina and Turkey in 2000 and 1999, respectively, and signaling potential problems in the case of Brazil in 2000.<sup>21</sup> As expected, South Africa exhibits a very high value of the sustainability indicator.

The outcome of simulation exercises is potentially highly sensitive to initial assumptions. The impact of plausible variations in our baseline assumptions on the sustainability analysis is reported in Celasun, Debrun, and Ostry (2006) for Argentina,

<sup>21</sup>As already mentioned, it would be worth repeating the exercise for Brazil for 1998 and 2001, which are arguably more plausible dates ahead of crises in the following years.

Figure 4. Fan Charts on the Eve of Troubled Times  
 (“Constant” policy scenario)



Source: Authors' calculations.

<sup>1</sup>This uses the “U.S. junk bond” rate as the relevant foreign interest rate. Other specifications of the VAR proved unstable.

Brazil, and Turkey. Robustness checks include a change in the specification of the fiscal reaction function along the lines suggested in Section II, the use of parameters from the SGMM estimates of the linear reaction function, a deterministic primary surplus path, and the use of a risky foreign interest rate (in lieu of the U.S. rate). The first two tests did not materially affect the risk analysis performed above, but the latter two suggest that (1) the deterministic policy assumption leads to underestimation of risks, and (2) the use of risky foreign interest rates significantly widens fan charts.

### Implications of the Analysis

The analysis in this section shows that our simulation algorithm could be easily applied to various emerging market economies in order to help develop a “risk-management” approach to public debt sustainability. The central idea was to provide policymakers with graphic representations (in the form of fan charts) that summarize the distribution of public debt, as well as numerical sustainability indicators that take into account characteristics of this distribution, notably the probability that debt will decline over some horizon and the probability that upside risks to the debt ratio will be limited.



**Table 4. Probabilistic Debt Sustainability Assessment:  
Looking Back at the Eve of Troubled Times  
(“Constant” policy scenario)**

	2000	2001	2002	2003	2004	2005
Debt ratio lower than in 2000 (1999 for Turkey)						
Argentina	...	0.07	0.11	0.13	0.17	0.18
Brazil	...	0.71	0.66	0.61	0.59	0.58
South Africa	...	0.87	0.93	0.96	0.95	0.95
Turkey	0.18	0.14	0.12	0.09	0.10	...
Debt ratio more than 10 percent of GDP higher than in 2000 (1999 for Turkey)						
Argentina	...	0.24	0.60	0.68	0.72	0.75
Brazil	...	0.01	0.06	0.14	0.20	0.25
South Africa	...	0.00	0.00	0.00	0.00	0.00
Turkey	0.19	0.64	0.73	0.80	0.83	...
Sustainability index						
Argentina	...	0.06	0.04	0.04	0.05	0.05
Brazil	...	0.71	0.63	0.53	0.47	0.44
South Africa	...	0.87	0.93	0.96	0.95	0.94
Turkey	0.14	0.05	0.03	0.02	0.02	...

Source: Authors' calculations.

The algorithm we proposed was also used to generate fan charts and numerical sustainability indicators on the eve of crises in order to check whether our approach would have provided credible early warning to policymakers: the indicator we developed indeed was flashing bright red on the eve of crises, and the fan charts give an equally supportive visual impression of the prospect of troubled times ahead. The analysis also illustrates the extent to which fiscal policy—summarized in the estimated fiscal reaction function—provides adequate insurance to stabilize the debt ratio in the face of shocks (on the eve of the above crises, it clearly did not). The algorithm is also flexible enough to study alternative scenarios and assess their implications for risks to debt dynamics.

#### IV. Conclusions

This paper has developed an algorithm for generating an explicit risk analysis of debt dynamics and applied it to five emerging market economies. The method builds on the standard approach to debt sustainability used at the IMF and elsewhere by accounting for country-specific risk factors regarding the economy and the fiscal policy process. Overall, we obtain a more complete, objective, and realistic assessment of risks than is possible with nonstochastic DSA templates, which rely on a few deterministic and standardized bound tests.

Our approach introduces greater realism in three critical dimensions. First, it uses estimates of joint probability distributions of economic shocks to construct a large number of scenarios that capture covariances among disturbances as well as the dynamic response of the economy. Second, it allows for fiscal policy to adjust to these shocks (to debt and growth, for example) according to the pattern commonly observed in emerging market economies, as given by our estimated fiscal reaction functions; these functions also appropriately shift the focus of policy analysis from adjustments in the nominal primary balance to measures of fiscal effort (defined as the difference between the predicted and actual or projected surplus). Third, we allow for fiscal policy itself to be a source of risk.

The debt sustainability assessment proposed here is explicitly probabilistic and can prove useful for policymakers in a variety of ways. First, the method offers a flexible tool allowing policymakers to capture country-specific features relevant for debt dynamics, and to have clearer signals of the risks involved in delaying fiscal adjustment or undertaking fiscal expansions. Second, more complete information on the debt risk profile should in turn improve medium-term budgetary planning. Indeed, one would expect greater awareness of the risks to public debt to promote caution in the conduct of fiscal policy. For example, this could imply less reliance on debt to finance new expenditure programs, thereby reducing the likelihood of the dynamics spinning out of control as a result of macroeconomic risks such as lower growth or higher interest rates. In addition, an explicit quantification of risks could help in the design of consolidation strategies: governments could evaluate the merits of alternative adjustment plans not only in terms of their impact on future trends of debt, but also on the upside risks to the debt path itself. More generally, governments with low credibility and facing a volatile environment could better internalize the costs of policies implying higher debt ratios, whereas governments with greater credibility and facing a more stable environment could avoid taking excessive comfort in a benign baseline outlook.

Applications to five emerging market economies with different risk profiles illustrate the merits of our approach. In particular, we show how the fan charts and sustainability indicators can be used to guide policymakers in making judgments about whether the present course of fiscal policy will or will not lead to problems (a crisis) down the road. Just as important, the estimated fiscal reaction functions—which connect fiscal behavior to various economic and institutional fundamentals—should also be of use in guiding policymakers on how to forestall problems by pursuing reforms that will shift the distribution of public debt paths that the economy faces. On both scores, the approach advocated here yields significant value added relative to deterministic approaches that are still widely used to assess debt sustainability issues in emerging market countries.

## REFERENCES

- Abiad, Abdul, and Jonathan D. Ostry, 2005, “Primary Surpluses and Sustainable Debt Levels in Emerging Market Countries,” IMF Policy Discussion Paper No. 05/6 (Washington: International Monetary Fund).

- Blundell, Richard, and Stephen Bond, 1998, "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, Vol. 87 (August), pp. 115–43.
- Bohn, Henning, 1998, "The Behavior of U.S. Public Debt and Deficits," *Quarterly Journal of Economics*, Vol. 113 (August), pp. 949–63.
- Celasun, Oya, Xavier Debrun, and Jonathan D. Ostry, 2006, "Primary Surplus Behavior and Risks to Fiscal Sustainability in Emerging Market Countries: A 'Fan-Chart' Approach," IMF Working Paper 06/67 (Washington: International Monetary Fund).
- Celasun, Oya, and Joong Shik Kang, 2006, "On the Properties of Various Estimators for Fiscal Reaction Functions," IMF Working Paper 06/182 (Washington: International Monetary Fund).
- Debrun, Xavier, 2005a, "Lessons from an Empirical Model of Fiscal Policy in Emerging Markets" in *Turkey at the Crossroads—From Crisis Resolution to EU Accession*, by Reza Moghadam and others, IMF Occasional Paper No. 242 (Washington: International Monetary Fund).
- , 2005b, "Public Debt in South Africa: A Risk Analysis," in *South Africa: Selected Issues*, IMF Country Report No. 05/345 (Washington: International Monetary Fund).
- Galí, Jordi, and Roberto Perotti, 2003, "Fiscal Policy and Monetary Integration in Europe," *Economic Policy*, Vol. 37 (October), pp. 533–72.
- Garcia, Márcio, and Roberto Rigobon, 2005, "A Risk Management Approach to Emerging Markets' Sovereign Debt Sustainability with an Application to Brazilian Data," in *Inflation Targeting, Debt, and the Brazilian Experience, 1999 to 2003*, ed. by Francesco Giavazzi, Ilan Goldfajn, and Santiago Herrera (Cambridge, Massachusetts: MIT Press).
- Goldfajn, Ilan, 2005, "Comment on Chapter 5," in *Inflation Targeting, Debt, and the Brazilian Experience, 1999 to 2003*, ed. by Francesco Giavazzi, Ilan Goldfajn, and Santiago Herrera (Cambridge, Massachusetts: MIT Press).
- International Monetary Fund, 2003, "Public Debt in Emerging Markets: Is It Too High?" Chapter III in *World Economic Outlook, September 2003: Public Debt in Emerging Markets*, World Economic and Financial Surveys (Washington: International Monetary Fund).
- , 2004, "Has Fiscal Behavior Changed Under the European Economic and Monetary Union?" Chapter II in *World Economic Outlook, September 2004: The Global Demographic Transition*, World Economic and Financial Surveys (Washington: International Monetary Fund).
- , 2005, *Argentina: 2005 Article IV Consultation—Staff Report; Staff Supplement; Public Information Notice on the Executive Board Discussion; and Statement by the Executive Director for Argentina*, IMF Country Report No. 05/236 (Washington: International Monetary Fund).
- Judson, Ruth A., and Ann L. Owen, 1999, "Estimating Dynamic Panel Data Models: A Guide for Macroeconomists," *Economics Letters*, Vol. 65 (October), pp. 9–15.
- Kiviet, Jan F., 1995, "On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models," *Journal of Econometrics*, Vol. 68 (July), pp. 53–78.
- Méltiz, Jacques, 1997, "Some Cross-Country Evidence About Debts, Deficits and the Behaviour of Monetary and Fiscal Authorities," CEPR Discussion Paper No. 1653 (London: Centre for Economic Policy Research).
- Mendoza, Enrique G., and P. Marcelo Oviedo, 2004, "Public Debt, Fiscal Solvency and Macroeconomic Uncertainty in Latin America: The Cases of Brazil, Colombia, Costa Rica, and Mexico," NBER Working Paper No. 10637 (Cambridge, Massachusetts: National Bureau of Economic Research).
- Penalver, Adrian, and Gregory Thwaites, 2004, "Analysing Sovereign Debt Sustainability: A New Probabilistic Approach" (unpublished; London: Bank of England).
- Wyplosz, Charles, 2005, "Institutions for Debt Sustainability in Brazil," in *Inflation Targeting, Debt, and the Brazilian Experience, 1999 to 2003*, ed. by Francesco Giavazzi, Ilan Goldfajn, and Santiago Herrera (Cambridge, Massachusetts: MIT Press).