Fiscal Policy and Tax Incidence

Fiscal forecasting: quality and motivation in Mozambique

Sam Jones

Discussion papers

No. 39E, Chapter 10
April 2007

National Directorate of Studies and Policy Analysis

Ministry of Planning and Development

Republic of Mozambique

The intent of the discussion paper series is to stimulate discussion and exchange ideas on issues pertinent to the economic and social development of Mozambique. A multiplicity of views exists on how to best foment economic and social development. The discussion paper series aims to reflect this diversity.
As a result, the ideas presented in the discussion papers are those of the authors. The content of the papers do not necessarily reflect the views of the Ministry of Planning and Development or any other institution within the Government of Mozambique.

The Logo was kindly provided by the Mozambican artist Nlodzy.

Contact:
Sam Jones

National Directorate of Studies and Policy Analysis (DNEAP)

Ministry of Planning and Development
Av. Ahmed Sekou Touré nº 21, 7º andar

Maputo, Mozambique

Tel: (+258) 2 1 499442

Fax: (+258) 2 1 492625

Web: www.mpd.gov.mz

Email: esamjones@gmail.com

Chapter 10

Fiscal forecasting: quality and motivation in Mozambique

Sam Jones

Abstract

Based on a unique dataset of outcome and forecast variables for Mozambique (1995-2005), this paper analyses the quality of public macroeconomic and revenue forecasts. Applying a new error decomposition technique, it is shown that forecasts are persistently optimistic and have deteriorated for major variables over the period to date. In turn, it is demonstrated that past forecasts have not outperformed a simple, naïve forecasting rule. Considering the sources of these errors, the normative influences on forecasting behaviour are discussed. Observed changes in taxation effort appear to be consistent with a principal-agent problem in which the principal attempts to induce higher collection effort by setting forecasts-as-targets. However, the evidence indicates this is at best partially effective and does not mitigate considerable strategic behaviour on the part of the agent.

1 Introduction

High quality fiscal projections can make a considerable contribution to public financial management and the achievement of social goals. Where budget forecasts are consistently misleading the overall efficiency and stability of government expenditures can be jeopardized and the credibility of public financial management systems may decline. Nevertheless, the practice of fiscal forecasting in low income developing countries has received relatively little research attention. This paper breaks with this trend and reviews the quality of macroeconomic and revenue forecasts in Mozambique for the period 1995-2005. As such it provides an example of how forecast evaluations can be undertaken and illuminates some of the key factors which affect the forecasting environment in Mozambique.

Based on a unique dataset forecasts and outcomes for both revenue and macroeconomic variables, the following questions are explored:

- what are the characteristics of the forecast errors;
- how have the forecast errors changed over time;
- how does the forecast model compare to a naïve alternative forecast rule; and
- what do the forecast errors tells us about the nature of the tax system?

The structure of this chapter is as follows. In the next section the relevant literature on fiscal forecasts in developing countries is reviewed. Section 3 presents the dataset and develops the analytical framework. Section 4 presents the results, including a simple description of the forecast errors (§4.1), their decomposition into component parts (§4.3), a formal review of their unbiasedness and efficiency (§4.2) and an analysis of their accuracy compared to a naïve alternative forecasting model (§4.4). In Section 5 explanations for the observed trends in forecasts errors are explored, considering in particular the use of forecasts as normative targets. Section 6 concludes.

2 Background

The literature on forecast evaluation is extremely large and beyond the scope of this paper. It is important to note, however, that the vast majority of previous research refers to econometric models whose domain of application is to advanced economies or specific sub-sectors. As noted by Danninger (2005) among others, both the quality and practice of economic forecasts in developing countries has received comparatively little attention.[1] The few studies that are of direct empirical interest fall into two categories. The first group are evaluations of multi-country, international economic outlook forecasts such as the IMF's World Economic Outlook series (e.g. Timmermann, 2006). The second are evaluations of forecasts associated with IMF-supported programmes (e.g. Phillips and Musso, 2002). In part driven by data restrictions as well as the objective of being of broad interest, the predominant focus in each of these categories is at the regional or international level. These studies tend to limit themselves to major fiscal indicators rather than the detail of revenue projections which is of interest here. As a consequence, and excluding research for advanced countries (for a brief survey see Golosov and King, 2002), there appear to be no available examples of evaluations of public sector forecasts at the country-level.

[1] The focus of this study must be distinguished from research into the impact of IMF and World Bank programmes in developing countries. While this research often has an interest in the plausibility of macroeconomic forecasts (expectations) associated with adjustment programmes, they are not motivated by an explicit interest in the practice of forecasting *per se*.

Despite the above, it is useful to note that multi-country studies by no means give a favourable assessment of forecasts for low income countries. Verbeek's (1999) review of the World Bank's Unified Survey projections for 1991-1997, for example, finds certain projections are 'quite disastrous' according to standard forecast criteria. In particular, he notes that forecasts for sub-Saharan African countries are the least accurate overall. Other studies find evidence for an optimistic bias in developing country fiscal forecasts, although this is by no means consistent to all forecast variables or countries. Golosov and King (2002), for example, find a significant bias of over one percentage point in forecasts of total tax revenue as a percentage of GDP. Phillips and Musso (2002) find that forecasts of inflation and international reserves are typically optimistic. In general terms, a consensus seems to be that forecasts in developing country situations are prone to error.

The existence of persistent forecast errors brings into focus the problem of what forecast are used for. As noted by Auerbach (1999), forecasts are not only technical point estimates used to provide information. They also act as (contested) inputs into complex policy processes. Thus, while it is typical to start from the proposition that errors from optimal forecasts *should be* randomly distributed around zero and incorporate all available information, it may well be that political actors, for whom the forecasts are produced, hold an alternative view of what constitutes an optimal forecast. This accords with Danninger (2005), who argues that government forecasts may not be pure unconditional expectations of future events but may have a normative content, especially when used as a means to influence the unobserved activity of agents such as the revenue collection agencies. This perspective is particularly apposite to developing countries under IMF programmes where projections of core fiscal variables can represent formal or informal programme targets. Not only are these targets typically conditional on the implementation of certain policies, but also they derive from a process of negotiation which may in itself be a source of bias (Phillips and Musso, 2002). In sum, the use of forecasts in policy processes may conflict with their ideal-technical characteristics. The extent to which this is the case will depend on the independence of the forecasting agency as well as the practical distinction between rules, targets and forecasts in fiscal policy. This issue is developed further in Section 5.

3 Data and Methods

3.1 Dataset

The variables in focus are real GDP growth, inflation as measured by a consumer price index, and sources of government revenues disaggregated by principal revenue categories. The latter includes eight internal revenue lines and two external financing items as well as their respective totals.[2] For each of the variables annual observations covering an 11 year period (1995-2005) of both outcomes (a_t) and their respective forecasts (f_t) made in the previous period are held. The dataset is constructed from a diverse array of documents including the annual government budget, official execution figures and published government planning documents. With respect to revenue items the forecast values represent the projections in the government budget while outcomes are those reported in the official government accounts (*Conta Geral do Estado*). Prior

[2]The internal revenue lines are: income taxes (personal and corporate combined), VAT (internal and external), import duties, excise duties, taxes on luxury goods, fuel tax, other taxes and non-tax revenues; the external financing items are split between grants and credits.

to 1998, however, although government budget documents were produced, no official public record was made of actual collected amounts (or expenditures). Thus, for such information one must resort to unpublished data produced by the collection agencies. For the macroeconomic variables, actual values can be gleaned from the national statistics agency (*Instituto Nacional de Estatística*) while forecasts are taken from an array of official planning documents.

With respect to the measurement quality of the variables, the outcome numbers are the final published official figures. While the quality of macroeconomic data is frequently questioned in developing country settings, for simplicity one must assume that these figures are accurate. The choice of which forecast for a given variable to include, however, is not straight-forward. This does not refer to the existence of competing forecasts made by different agencies, but rather to the issue that forecasts typically undergo rounds of revision within the same forecasting agency. Formalized or mature forecasting processes typically regularize and document these revisions, creating a richer data set on which forecast evaluation can take place. For Mozambique this is not the case and forecast revisions often are made on an *ad hoc* basis and are not consistently documented. As a result, one must rely on forecasts that have been either formally published by the government and/or form part of flagship internal documents. Where possible, forecasts which have been generated by standardized processes according to a regular cycle are preferred, the government budget being a case in point.[3]

[3]A remark should be made about budget revisions. In certain years stated budget targets can become misleading and even unhelpful due to unexpected changes and/or errors in the original budget. In extreme cases this necessitates a formal revision which, for the period under analysis, occurred in Mozambique in 2001. Given the magnitude of the revisions to the revenue forecasts

3.2 Forecast error definition

The evaluation of forecasts naturally focuses on the forecast error. For a given fiscal variable n_t, define the forecast error at time t as:

$$e_t = n_t^f - n_t^a \quad (1)$$

where superscript f denotes the forecast made at $t-1$ for the outcome at t, and a indicates the actual outcome value. As noted above, this is a simplified notation which ignores forecasts made over different horizons for the same period as these are not in focus here. With respect to the comparison of forecast errors both over time and across items, one must be sensitive to the choice of n. In the case of the macroeconomic variables, (real GDP growth and inflation), the range of actual and forecast values in the data is relatively small and scale concerns are not an issue. Thus, the relevant forecast errors refer to the difference in these rates expressed in percentage point terms. Revenue items are more problematic as in their nominal, local currency form they are non-stationary and not easily comparable across items. Moreover, nominal revenue forecasts often incorporate forecasts for other variables such as GDP and inflation.

Various responses to the above problem are encountered in the literature, such as the transformation of nominal revenue to a ratio of GDP (see for example, Golosov and King, 2002). For the purposes here, where different revenue lines are to be compared, none of these methods are unproblematic. As a result, it is useful to decompose the revenue variable into its component parts. Thus, define the observed nominal revenue

made at this time, the dataset uses the revised as opposed to original figures. For all other years original budget figures are used.

at time t as:

$$n_t^a = n_{t-1}^a \cdot z_t^a \cdot p_t^a \cdot g_t^a \tag{2}$$

where p reflects the rate of inflation, g the real growth of output and z the overall elasticity of n to changes in nominal GDP, described here as the tax effort. The forecast for the same period is defined in exactly the same way, substituting the a superscripts for their forecast values for all variables other than the common base given by n_{t-1}^a. Second, taking natural logarithms and subtracting the forecast from the actual equations one gets:

$$e_t = (z_t^f - z_t^a) + (p_t^f - p_t^a) + (g_t^f - g_t^a) \tag{3}$$

where the parentheses on the right-hand side respectively represent forecast errors in the tax effort, inflation and growth.

A number of aspects of this decomposition are of interest. First, from equation (2) it is trivial to calculate the individual tax effort measures (elasticities) for the forecast and actuals. This is useful as it permits a deeper analysis of collection behaviour; an effort (elasticity) equal to one (or 100%) indicates nominal revenues adjust in line with nominal GDP, equivalent to an unchanged tax effort or a stable weight in GDP. Second, from (3) the tax effort error can be calculated, assuming it is the only unknown error component in the equation. This is the approach adopted for the analysis of revenue forecasts due to the desirable properties of the resulting errors – they are comparable across revenue lines, invariant to scale changes over time, and cleansed of other observable errors. Third, the method emphasises the extrapolative properties of revenue forecasts. Although this may not be appropriate in all countries, it is relevant for Mozambique as revenue forecasts employ a model of this form. Fourth, an assumption of the decomposition is that both the forecasts and actuals are extrapolated from the same base. In practice this does not always holds as forecasts are elaborated before the final value of n^a_{t-1} is known. As a result, the measure of the tax effort forecast error will incorporate any measurement noise. However, this unknown element can be treated as an integral part of the forecasters problem and does not seriously undermine the validity of the approach. Finally, it should be stressed that z may be determined by a large number of intermediate variables including tax policy instruments, the applicable tax base and administrative factors. Thus, any interpretation of movements in this measure must be undertaken with care.

3.3 Forecast evaluation

It is typical to review three distinct characteristics of forecast errors, namely their unbiasedness, efficiency and accuracy.[4] *Unbiasedness* holds when the expected value (mean) of the forecast error is equal to zero. Holden and Peel (1990) show that the preferred test for bias is a simple regression of the forecast error on a constant. Under the null of unbiasedness one expects to find the coefficient on the constant is equal to zero. For this test to be valid, however, the forecast errors should be approximately Gaussian. As this is by no means guaranteed, basic distributional tests on the forecast errors must be conducted in advance.

Efficiency refers to the full use of information available to the forecaster at the time of forecasting. This concept is discussed at length in Nordhaus (1987) who

[4] It is beyond the scope of this paper to discuss evaluation methodologies. For an overview see Schuh (2001) and the contributions in Hendry and Ericsson (2001).

distinguishes between strong and weak forms of efficiency. While the former tends to focus on the full information set pertinent to the forecast, it is more typical to test for the latter via a simple regression of the current forecast error on its past values. For weak efficiency one expects to find no relationship between the dependent variable and the regressors:

$$e_t = \alpha + \beta_1 e_{t-1} \tag{4}$$

the point being that if a significant joint relationship is found then, in principle, past forecasts could have been improved by an adjustment proportionate to β_1. Related to this test, two further issues can be highlighted. First, it is evident that equation (4) nests a test for unbiasedness, a main difference with a simple test being the reduction in statistical degrees of freedom due to an additional regressor. Secondly, the time series econometric literature distinguishes between different types of trend effects. Among these is a trend-stationary process, defined as a random walk around an underlying (linear) trend. This is relevant as there is no *a priori* reason to assume that any bias in the errors will be constant. Slow changes to the determinants of forecast quality, including political variables, may translate into forecast error trends. Thus, while equation 4 remains a sufficient test for weak efficiency, it does not distinguish between trended and non-trended inefficiency.

As a result, a general test for bias, trend effects and seriality in the forecast errors derives from a single regression of the form:

$$e_t = \alpha + \beta_1 e_{t-1} + \beta_2 t + \gamma Z \tag{5}$$

where Z is a vector of additional information available to the forecaster at time $t-1$ and therefore represents an expansion of the efficiency concept to information not contained in the lagged forecast errors. With regards to significance testing, simple parameter tests on the coefficients are appropriate to identify the source of any regularities in the forecast errors. Tests applied to the constant term, $\alpha = 0$, do not represent strict tests for bias (i.e., that $E(e_t) = 0$), but rather refer only to the bias remaining after adjusting for the effects of other variables.[5] Given unbiasedness is a logically necessary condition for efficiency, an overall test for weak efficiency is given by the joint significance of the entire equation. Finally, although it is common in the literature to run tests for bias and weak efficiency via two separate regressions, a general specification is preferred for both analytical and econometric reasons. With respect to the latter, a 'complete' specification helps to reduce both omitted variable bias and the risk of spurious results arising from running a large number of regressions. Also, a general specification permits the analyst to distinguish between the separate (partial) influences of bias, serial correlation or trend effects that would not be evident from separate regressions.

Accuracy refers to the comparison of errors from two forecast models, namely the employed model and a naive alternative which only uses information from past observations. This evaluation is useful because the existence of bias and inefficiency does not necessarily imply that a more simple forecast model is to be preferred. A number of statistical measures have been proposed for this purpose and their properties have been reviewed extensively (for example, see Hyndman and Koehler, 2005). The approach employed here is based on Diebold and Mariano (1995), hereafter DM. Their method is employed widely in the literature (e.g., Artis and Marcellino, 2001), pre-

[5] For further discussion see Barrionuevo (1992). Note that bias here is also tested after the effect of any time-trend; thus in order to ensure the average bias is captured, the trend variable is set to zero for the mid-point in the time series.

ferred for its focus on the statistical properties of the goodness-of-fit series generated by subtracting each naive forecast error from its non-naive (employed) counterpart according to a specified loss function. Thus, defining this loss function as $g(\cdot)$, the series of interest is calculated from $d_t = g(e_t^a) - g(e_t^b)$, where a and b respectively refer to the employed and naive forecast models. The formal DM measure is akin to a t-test applied to the series $\{d_t\}$, the main modification referring to the calculation of variance which adjusts for serial correlation in the series. However, as multiple forecast horizons are not in focus here, simple parametric t-tests can be used in place of the formal DM test, as long as $\{d_t\}$ is approximately Gaussian and free of other unwanted features. Even where this is not the case, other distribution-free approaches may be used, as discussed in Diebold and Mariano (1995).

For this study the above framework is applied for two loss functions: $g(e) = |e|$ and $g(e) = \ln(|e|)$. These are chosen as being most pertinent to the Mozambican context where our choice variables may exhibit substantial volatility arising from policy shifts, macroeconomic shocks or reliance on a poorly diversified tax base. As such, and in contrast to the more common use of a quadratic loss function, this does not give undue weight to large errors. The use of absolute error values indicates that it is only the size rather than the direction of the forecast errors that is of interest; while this may not be true for policy makers it represents our best prior in a Bayesian sense.

3.4 Time series cross-section considerations

Before moving to the presentation of results, it is necessary to reflect on the properties of the data from an econometric perspective. As the dataset contains both time-series and cross-section (TSCS) elements one cannot assume all observations are identically and independently distributed either over time or contemporaneously. Rather, and for the revenue data in particular, it is reasonable to expect substantial contemporaneous correlation across the forecast errors given the common influence of macroeconomic variables as well as political pressures, notwithstanding the use of a common forecasting model. Our intended focus on the overall tax elasticity forecast error represents only a partial attempt to adjust for the contemporaneous impact of errors in the macroeconomic forecasts. Thus, the meaningful application of any regression-based tests, must not disregard TSCS econometric issues.[6]

Given the structure of the data it would not be appropriate or efficient to pool all observations, assuming common coefficients, or to run separate regressions for each of the choice variables. In contrast, in order to estimate equation (5) a full information approach, in the spirit of Zellner's (1962) Seemingly Unrelated Regressions (SUR) technique, is necessary to provide unbiased and consistent estimates. For the macroeconomic variables and the revenue aggregates a standard SUR analysis is applied, incorporating different dependent and independent variables. For the disaggregated revenue items the same model applies to each line in cross-section; thus a general-to-specific approach is pursued. This involves, first, estimating a single general specification in which the slope and constant coefficients are permitted to vary for each line. On the basis of the results, groups of similarly behaved revenue lines can be identified and the model re-estimated employing these groups as pools. At each stage, Hansen tests for parameter stability (Hansen, 1992) are used to verify the results. Finally, while there is some controversy regarding the best estimator for this kind of model, a Prais-Winsten panel-corrected estimator (PCSE) is used, which

[6]For a discussion of these see Beck (2001).

incorporates adjustments for the time series and cross-sectional properties of the data (e.g., see Blackwell, 2005). This is preferred due to concerns that (alternative) feasible GLS estimators may underestimate the standard errors of the explanatory variables and thus over-play the significance of the results.[7]

4 Forecast error analysis

4.1 Description

To contextualise the analysis, Figure 1 shows trends in aggregate internal and external revenues as a percentage of GDP for the entire period. The rough equivalence of the two sources of income illustrates the weight of external financing in the government budget and, consequently, the importance of including these items in the analysis. Figures 2 and 3 plot the behaviour over time of the forecast errors for the aggregate variables, including both internal and external revenues. Three initial observations can be made. First, while external revenues as a percentage of GDP appear to have followed a relatively stable medium-term trend, changes in internal revenues can be divided into two periods — a period of consistent growth from 1996-2000, followed by an absence of any growth trend for 2000-05. Second, in both levels and forecast error terms, the dissimilar behaviour of internal and external revenues is noticeable, warning against a simplistic aggregation of these two revenue sources. Internal revenue forecast errors appear to follow a systematic trend; however, the errors for external revenues would seem to be a random walk. Finally, with respect to the two macroeconomic variables the large error for 2001 can be explained (in part) by the unforeseen rapid economic recovery following catastrophic floods in 2000, the

[7] All estimations are implemented in STATA via either the *sureg* or *xtpcse* commands.

actual effect of which also is indicated in the GDP error for that year.

Table 1 provides a summary of the forecast errors, including the disaggregated external and internal revenue lines. The majority of errors are symmetric and approximately Gaussian. For those errors where the null hypothesis of normality can be rejected at the 5% level or below this is often associated with a relatively large skewness measure, indicative of the presence of extreme values. Even so, there are substantial dissimilarities between the revenue lines. With respect to internal revenues, the mean error spans a wide range from a low of -4.71% (income taxes) to a maximum of 24.53% (other taxes). For external revenues lines the mean errors have opposite signs and that their standard errors are comparatively large, being greater than the mean in the case of credits. Thus, while it appears valid to apply the econometric methodology outlined above, one needs to be sensitive to the influence of extreme values and due care must be taken with regard to the cross-sectional nature of the data.

4.2 Bias and efficiency

Based on the techniques described in Sections 3.3 and 3.4 it is possible to formally analyse the characteristics of the forecast errors. Starting with the aggregate variables, an SUR regression of the form given by equation (5) is employed, adding a dummy variable to account for the effects of the shock recovery in 2001. The results, summarized in Table 2, confirm a number of the findings suggested above. First, after adjusting for the effects of the other regressors, the two macroeconomic variables are found to be biased in an optimistic direction – the average downward bias of the inflation forecasts being around 1.25 percentage points, while real growth forecasts are biased upwards by approximately 1 percentage point. In addition, the inflation forecast errors are serially correlated, the negative sign here reflecting the direction of bias. The overall explanatory power of the regressions for these variables is good, ($R^2 = 87.3\%$ for inflation; and $R^2 = 63.7\%$ for real growth), testifying to the systematic nature of these errors despite the small sample size.

The results of the SUR for the aggregate revenue variables are also insightful, but moderately less strong. For total external financing, there is some evidence of an underlying negative bias but this cannot be relied upon as the overall specification is not significant. For internal revenues on the other hand, and despite no evidence of a fixed level of bias throughout the period, the positive time-trend on the errors indicates they have been increasing in size by an average of one unit (0.01) per year. These forecasts also are serially correlated and do not pass the overall test for weak efficiency; i.e., at the 5% level the null hypothesis that all parameters are zero can be rejected. It follows that improved forecasts could have been made by better incorporating available information.

Turning to the component revenue lines, a full general specification, as described in Section 3.4, supports the contention that there are important dissimilarities across revenue lines such that pooling all internal and external revenues, either separately or jointly, is not advisable. Note the specification includes a dummy for 2001 (as above) as well as lagged forecast errors of the macroeconomic variables to capture other information available to the forecaster. Due to the large number of regressors, however, these latter variables do not enter the respective SUR equation for each revenue line; rather they enter as separate regressions in the SUR system as a whole.

Even so, the coefficients on these terms are allowed to vary according to the two-period division described above. Table 3 summarizes the results, showing for each variable only the values of those coefficients for the bias, trend and serial correlation terms of equation (5) which are significant at the 10% level. The table also states the probability that these coefficients are jointly equal to zero, giving the overall test for weak efficiency.

Excluding the effects of the macro variables, the two external financing lines are absent from the table as no consistent pattern is found, meaning that the forecasts for these lines cannot be deemed to be inefficient. In contrast, all the internal revenue line forecast errors show some form of systematic, significant tendency and in no case can one accept the null hypothesis of weak forecast efficiency. A strong result is the significant positive bias across all revenue lines excluding income taxes (this is after adjusting for any trend and serial correlation effects; see footnote 5). Bias ranges from approximately 10% (luxury items) to over 20% for excise taxes and the group of small unclassified taxes. Evidence for trend or serial correlation is not consistent across internal revenue lines; however, there is clear evidence of an upward linear time-trend for three core domestic taxes which, together, represent well over one half of total internal revenues (VAT, income and luxury goods taxes). In contrast, excise taxes show a declining trend, perhaps evidence of a forecast-response to the large forecast errors of the early part of the period (depicted in Figure 4). There is no evidence for consistent serial correlation across the revenue lines, this being restricted only to VAT and luxury goods.

Table 3 also shows significant coefficients on the lagged macroeconomic variables. While this finding is explored further in Section 5, at this point it is sufficient to note that the partial correlation of the tax elasticity with these lagged errors is particularly marked in the second period, running in the opposite direction to the mean error of the variables themselves. In other words, while the real growth forecast errors are positive on average, the coefficient in the SUR specification is negative – meaning that a one unit increase (improvement) in the growth error in the previous period is associated with an average 12.2 percentage point reduction in the predicted elasticity forecast error in the current period. Of course, one might question the assumption of a common coefficient across all revenue lines for these macroeconomic regressors. Hansen tests for parameter stability, however, indicate one cannot reject the null hypothesis that the results are stable across different sub-samples of the data. This finding also extends to all other variables as well the specification as a whole.[8]

Tests for parameter equality across the revenue lines suggest a more parsimonious, reduced-form specification can be used as compared to the general SUR described above. This is useful not only from the point of view of reducing the number of regressors, thus increasing the overall degrees of freedom, but also to identify common trends across groups of revenue lines. Tests identify four broad groups of similarly-behaved revenues – (i) core domestic taxes (VAT, income and luxury goods taxes); (ii) taxes at the border; (iii) other taxes and non tax income; and (iv) external

[8]According to Hansen (1992), one can reject at the 5% level the null hypothesis of parameter stability for those individual coefficients for which the given (L) statistic exceeds 0.47. For the macro variables in the SUR general specification the largest L statistic is 0.133, relating to the second period real growth error; for the specification as a whole, the joint L statistic is 1.618 which with 36 degrees of freedom cannot be taken as indicative of instability at any reasonable confidence level.

financing items. Due to the weak findings as regards the serial correlation term, the reduced specification restricts the slope on this term to be equal across all revenue groups.[9] The results of this parsimonious specification are shown in Table 4; as expected, they do not present anything substantially new but rather summarize and affirm our previous results. Of note is the significant negative slope coefficient on the border-related group of taxes, suggesting that as a group these errors have followed a declining trend during the period. The overall strength of the regression ($R^2 = 47\%$) underlines the broad finding that forecast errors for the internal revenue lines are not efficient but rather contain persistent levels of bias, show time-trend effects, and do not fully incorporate information available at the time of forecasting. Finally, the validity of this specification is supported by Hansen tests – for none of the coefficients can the null of stability be rejected, either individually or jointly.

4.3 Decomposition

The error decomposition technique described in Section 4.3 can be used to further explore the forecast errors. Figure 4 compares the mean actual and forecast tax elasticities (efforts) for each internal revenue line across two sub-periods (1995-99 and 2000-05). Aside from the evident difference between the (mean) tax elasticity forecast versus the actual value, an important result is the reduction in variation across the lines between the two periods for both forecast and actual values. The plot for the second period has a flatter topography, suggesting that both the tax system and the forecast errors have come to behave in a more stable or regular fashion over time.

[9] From the general specification, a Wald test for equality of the serial correlation slope coefficients across revenue lines gives a χ^2 statistic of 23.6 which is relatively small, meaning one cannot reject the null hypothesis of equality at the 0.5% significance level.

Further evidence for a stabilization effect is given by changes in the pooled distributions of the forecast and actual tax elasticities. As shown in Figure 5, both distributions show a general tightening (sharpening) between sub-periods with a much more pronounced change in the distribution of forecast elasticities. In light of the persistence of forecast errors throughout the two periods (see above), it follows that the more recent forecast errors derive from a more systematic error generating process than those of the first sub-period. It is also of interest that the actual tax elasticity measures are distributed around a mean (and median) of approximately one for both periods. For both aggregate and individual revenue lines, statistical tests reveal that one cannot reject the hypothesis that, on average, the tax elasticity is equal to one. This is indicative of long-term stability (inertia) in the actual tax system and confirms the point that that measured forecast errors derive from the forecast component as compared to erratic actual behaviour.[10]

4.4 Accuracy

The final analytical step refers to the accuracy of the forecasts, answering the question whether a naive forecast model could have out-performed the employed forecasts. As described in Section 4.4, the first step is to calculate and review the goodness-of-fit series based on a specified loss function. Note that while a large number of naive models might be used, for simplicity the naive forecast is defined as the value observed in the last period; for the revenue lines, therefore, this refers to the tax effort.

[10]This result derives from a simple regression of the actual tax elasticity on a constant plus dummies for each revenue line (internal only).

The goodness-of-fit calculated using the absolute loss function is analytically messy and unsuitable for parametric analysis. This derives from the existence of a small but not insignificant number of larger forecast errors. As a result, the non-parametric Wilcoxon (matched-pairs) signed-ranks test (see Siegel, 1956) can be used to investigate whether the median difference between the employed and naive forecast errors is equal to zero. With respect to each of the variables in the dataset, the only significant result from this exercise is for real growth. This shows a significant negative mean difference, meaning that the employed forecast consistently outperforms the naive model; in other words, the employed growth forecasts have been more accurate than a martingale strategy. For all other variables, the performance of the employed and naive models are indistinguishable at the 5% significance level (two-tailed).

The logarithmic loss function generates a goodness-of-fit series which is approximately normal both across revenue lines and on a pooled basis. As a result, a standard t-test can be applied meaningfully, as discussed in Section 3.3. Figure 6 plots the pooled distribution of this series for all revenue items, distinguishing between the first and second sub-periods. The vertical lines represent the interval within which one can expect to find at least 75% of all observations regardless of the distribution, based on Chebyshev's inequality (Abramowitz and Stegun, 1972). This interval is also approximate to the 95% confidence interval associated with a student's t-distribution. Whatever the benchmark adopted, the vast majority of observations fall within the confidence interval, suggesting there is no robust statistical difference between the employed and naive tax effort forecasts either between the two periods or across lines. In fact, the sharper distribution for the second period signifies that despite reduced variation of the mean actual tax elasticities across lines in the second period, (see Figure 4), there has been no corresponding improvement in forecast accuracy vis-à-vis the naive model; rather, the tighter distribution indicates the very opposite is the case. This confirms the argument that despite reduced variation in tax effort (both forecast and actual), a corresponding improvement in forecast quality has not materialised.

Continuing with the same loss function, Figure 7 plots the relevant t-statistic, (based on the logarithmic loss goodness-of-fit series), for each revenue line over the entire period. This provides no substantially new information compared to the tests for the absolute loss function, but at least confirms the consistency of the results. It can be pointed out, however, that the majority of the t-statistics for internal revenues are greater than zero, indicative of a marginally but not significantly better performance of the naive model in most cases. The significance of the t-statistics for the macroeconomic variables confirm the previous result for real growth; however, the statistic for inflation is now also found to be negative and significantly different from zero at the 10% confidence level (two-tailed). The overall message, therefore, is that only forecasts for the macroeconomic variables appear to broadly outperform a simple martingale strategy. Of course this does not mean that our forecast model outperforms the hypothetical set of all naive models. For example, when real growth is volatile around a stable central tendency, an arguably better prior is not the last period's value, as tested here, but rather a long-term average. Indeed, it is simple to verify that an alternative naive forecast rule, which posits that real growth will always be 7%, significantly outperforms the employed growth forecasts (based on the absolute loss function tested via the Wilcoxon signed-ranks test).

5 Explanations

The previous sections have analysed the characteristics of the forecast errors over time. While a comprehensive discussion of the underlying causes of the identified regularities is beyond the scope of this paper, some important lines of argument can be sketched. Following the discussion in Section 2, the distinction between technical and normative influences on the forecasting process is important. From a technical perspective it is clear that instability in the economic environment, as well as structural changes or shocks, can undermine the quality of forecasts. Improvements in environmental factors of this kind certainly help to explain the reduced variation in tax elasticities over the period, as well the observed trends in the tax effort errors for certain revenue lines. As described elsewhere (Arndt et al., 2007), the mid-1990s were characterised by substantial and strategic public sector reforms, including privatizations and the stream-lining and eventual independence of the Central Bank. Since 1997 a programme of public financial management reforms has been implemented, focusing on the quality, coverage and transparency of the government budget. The enhanced predictability of the macroeconomic environment, rationalisation of government intervention in the economy and the more prominent role given to fiscal planning during this period are likely to be valid explanations of the relative stabilization, including the reduction of extreme errors in revenue forecasts as shown in Figure 4. The enhanced predictability of the exchange rate, as well as an ongoing programme of customs reforms including trade liberalization, also represent plausible explanations for the trend reduction in forecast errors of border-related taxes compared to other domestic revenues.

Working in the opposite direction, similar arguments can be employed to explain the trend increase in forecast errors for major internal taxes. During the period under analysis these tax instruments underwent substantial reform – namely, the replacement of a cascading sales tax by value-added taxation (VAT) in 1999 and the implementation of a radically revised income taxation code from 2002 (see Chapter 3).[11] Not only do major tax policy reforms often take time to 'bed in', but also forecasters may have difficulty in the absence of historical data on collections or knowledge of the effective (new) tax base. With respect to income taxes, for example, the reforms appear to have been broadly positive as estimates of the actual collection effort have risen to over one (see Figure 4), meaning that collections are now growing faster than real output. Even so, the forecast errors for these taxes have switched from being trend negative in the first period, to being trend positive in the second, perhaps evidence that forecasters do not have a detailed understanding of these new taxes.

Although insightful, the latter arguments do not give a convincing explanation for the optimistic bias of the forecasts. Rather, they point towards the effect of expectations or normative influences. At least superficially, such normative forces would appear to be at work given the nature of the forecasting process. Projections incorporated in the government's annual budget cannot be considered *pure* technical estimates as they are the outcome of both internal political processes and discussions

[11] In the previous analysis no distinction is made between the sales and VAT tax as the latter simply replaced the former. Thus data for the sales tax are incorporated within the VAT line

with the IMF. As such, forecasts are consistent with fulfilment of the applicable IMF programme, which repeatedly have stressed the need for increases in domestic revenue as a percentage of GDP (e.g., IMF, 2005). Other government documents also manifest this conflation of normative targets and technical forecasts. For example, the macroeconomic position taken in the government's first poverty reduction strategy paper (República de Moçambique, 2001) emphasises the objective of reducing dependence on external revenues via growth in the tax ratio from 12.4% in 2000 to 15.4% by the end of 2005, or an average annual increment of 0.6 percentage points. As shown in Figure 1, since 2000 the domestic revenue ratio has oscillated around a mean of 12.3% of GDP with no robust growth trend. However, this has not engendered any major revision of expectations – the second strategy paper (República de Moçambique, 2006) projects growth in domestic revenues of around 0.5 percentage points per year for the period 2005-09.

In light of the above, the motivation for persisting with optimistic forecasts needs to be uncovered. Taking the principal-agent framework proposed by Danninger (2005) (see also Danninger et al., 2005), one hypothesis is that (high) targets are used to influence collection effort when the true effort of the collection agency cannot be fully controlled by its principal. On the assumption that revenue shortfalls are somewhat costly to the agent, optimistic or overstated targets will be employed as long as the benefits of higher collection surpass the costs to the principal of overstated budgets. Recalling that revenue forecasts in Mozambique are the outcome of negotiations with the IMF, it can be assumed for simplicity that the agent is the government while the (final) principal is the IMF.[12] From this perspective, there are various costs to the government of failing to meet the revenue target, including jeopardising the country's relationship with the IMF.

In order to examine these issues, the relationship between the actual and forecast tax effort measures, as well as other proximate variables which may affect collection performance, come into focus. The null hypothesis, incompatible with a principal-agent dynamic, is that the actual level of collection effort is uncorrelated with both its forecast and other strategic variables. This rests on the previous findings of substantial stability in the actual tax effort (elasticity) and persistently optimistic forecast errors. To test this hypothesis the actual effort is regressed on its forecast, controlling for the previous period actual effort as well as current macroeconomic conditions; formally:

$$z_t^a = \alpha_t + z_t^f + z_{t-1}^a + p_t + g_t + t \tag{6}$$

where z^f and z^a are the forecast and actual levels of tax effort, p the (change) in prices for the period, g real growth, t time and α a constant. Note that inclusion of the macroeconomic variables, as well as the time trend, reflects earlier results concerning the influence of these factors on revenue forecast errors. The model is estimated for internal revenues only, allowing the slope coefficients on the forecast effort regressor to vary for each of the revenue groups identified in Section 4.2. The panel-based estimator discussed in Section 3.4 is applied once again, (although the feasible GLS estimator is also used for comparison).

[12] Note that this is an alternative to the principal-agent model outlined in Danninger (2005) where the principal is defined as the central government and the agent a semi-autonomous revenue collection authority.

The results, summarized in Table 5, indicate that approximately 80% of variation in the actual tax elasticity (effort) is explained by the model. Hansen tests for parameter stability confirm the validity of the results despite the partial pooling applied. Importantly, the positive and significant coefficients on the forecast effort suggests there is a positive association between forecast and actual tax effort across revenue lines and over time. Although a simple interpretation of this result would be that the forecasts are accurate, this can be rejected on the basis of previous analysis. Indeed, it is precisely the combination of these results alongside persistent forecast optimism that is highly suggestive of the employment of forecasts-as-targets. Integral to this interpretation, which runs against the null hypothesis, is that the coefficients on the forecast effort regressors are positive but less than one (≈ 0.3), showing only a partial (imperfect) effect of the forecasts on collection effort.

Results for the other regressors in the model confirm that the agent retains a considerable degree of strategic flexibility in his choice of behaviour. First, the absence of a significant relationship between the current tax effort and its lagged value indicates the absence of longer-term (unexplained) determinants of actual tax effort. The coefficients for the macroeconomic variables demonstrate that when the macroeconomic environment is conducive to a higher nominal tax take, via high inflation or robust output growth, the tax effort falls. This reading is particularly apposite as the Mozambican tax authorities employ nominal revenue targets (stated in the budget) to monitor their progress. Thus, high levels of real growth and/or inflation enable these targets to be met passively, without extra effort. In summary, the results provide convincing support for the proposition that revenue targets are employed to influence taxation effort, but that the collection agency retains considerable flexibility in his choice of final effort. Of course, this is not the only possible explanation of the relationships in the data. However this thesis represents a highly intuitive explanation for the characteristics of the forecast errors.

It remains pertinent to ask whether this is the best possible strategy to stimulate collection effort. From a theoretical position (Danninger, 2005), there is no doubt that mechanisms of this kind are second best (at most) given they do not fully internalize the principal-agent problem. This is proven in this instance by the remaining strategic behaviour of the agent. The very absence of sustained growth in the domestic revenue ratio and the negative coefficient on the time trend regressor also supports this view. However, it has been shown that the use of *nominal* targets enables the agent to choose his effort depending on macroeconomic fluctuations. This suggests that simple modifications to revenue targets could at least reduce some elements of strategic behaviour. Even so, it may be the case that the use of forecasts-as-targets suffers from a rational expectations critique – knowing that the forecasts are deliberately optimistic, the agent has no incentive to fully meet the targets and they lose effectiveness. This also may be behind the negative coefficient on the time trend regressor.

According to the above framework, optimistic targets will be applied as long as the benefits outweigh the costs from the principal's point of view. This says nothing of the aggregate social costs of optimistic targets, which may be significant. These include undermining the efficiency and quality of government services, as well as the credibility of the overall budget process. In addition, where normative factors dominate the forecasting process, technical contributions may be undervalued which, in turn, weaken their influence on fiscal decisions. While these points cannot be developed here, the assumption that the principal is an external actor, as opposed to the general public, would be compatible with a failure to take these wider costs into account.

A final issue refers to *how* improved forecasting might be achieved without prejudicing taxation effort. On the one hand, mechanisms which separate technical forecasts from normative targets may be useful. These might include formal forecasting rules which limit the scope of normative interferences. On the other hand, a valuable contribution may come from efforts to develop a consensus view among all key actors, including the private sector, as to realistic medium-term macroeconomic and fiscal projections. According to the analysis in Chapter 6, Mozambique's domestic taxation performance is neither an outlier nor is it far below its expected level in comparative terms. If this is broadly correct then harmonization of medium-term taxation expectations around such a view could reduce short-term pressures on revenue forecasts.

6 Conclusion

This study has undertaken a careful review of the quality of government fiscal forecasts in Mozambique, focusing on revenue forecasts in particular. A methodological toolkit has been presented and its considerable practical value demonstrated. Persistent forecast bias in an optimistic direction has been found to affect all major variables, excluding external revenues. There is also robust evidence for an upward time trend in the forecast errors for core domestic revenue lines. Despite enhanced macroeconomic and fiscal stability, neither revenue nor macroeconomic forecasts have shown any marked improvement overall. In turn this suggests that forecast errors derive from a more systematic process than in the past. Differences in forecast performance across revenue lines and over time, however, suggest the need to develop a more sophisticated understanding of the dynamics of individual tax instruments. This is especially important for larger revenue lines, such as the VAT and income taxes, which have experienced policy reforms during the period.

It has been argued that government forecasts are not pure unconditional expectations but can be influenced by normative considerations. As such, persistent optimistic forecast bias may be explained by political pressure to realise sustained growth in the tax ratio and/or the desire to sustain an image of effective macroeconomic and fiscal management. With respect to revenue behaviour, the analysis provides strong evidence of a principal-agent problem in which optimistic nominal targets are employed to induce a higher collection effort. However, this does not appear to be wholly effective in light of the continued flexibility of the collection agency to choose his final effort. Moreover, this approach may undervalue the social costs of inferior budget credibility and execution efficiency arising from overstated forecasts. In response, alternative mechanisms are required to both stimulate strong collection performance and reinforce budget quality. Further consideration should be given to actions which separate technical and normative forecasts in government planning, possibly by adopting formal forecast rules.

References

Abramowitz, M. and Stegun, I. A., editors (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* New York: Dover, 9th printing edition.

Arndt, C., Jones, S., and Tarp, F. (2007). *Theory and Practice of Foreign Aid*, volume 1 of *Frontiers of Economics and Globalization*, chapter Aid and development: the Mozambican case, pages 235–288. Elsevier B.V.

Artis, M. and Marcellino, M. (2001). Fiscal forecasting: The track record of the IMF, OECD and EC. *Econometrics Journal*, 4(1):S20–S36.

Auerbach, A. (1999). On the performance and use of government revenue forecasts. Berkeley Olin Program in Law & Economics, Working Paper Series 1000, Berkeley Olin Program in Law & Economics.

Barrionuevo, J. M. (1992). How accurate are the World Economic Outlook projections? In *Staff Studies for the World Economic Outlook*, chapter II. International Monetary Fund.

Beck, N. (2001). Time-series cross-section data: what have we learned in the past few years? *Annual Review of Political Science*, 4:271–93.

Blackwell, J. L. (2005). Estimation and testing of fixed-effect panel-data systems. *The Stata Journal*, 5(2):202–207.

Danninger, S. (2005). Revenue forecasts as performance targets. IMF Working Papers 05/14, International Monetary Fund.

Danninger, S., Cangiano, M., and Kyobe, A. (2005). The political economy of revenue-forecasting experience from low-income countries. IMF Working Paper WP/05/2, IMF.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–63.

Golosov, M. and King, J. (2002). Tax revenue forecasts in IMF supported programmes. Working Paper WP/02/236, IMF, Fiscal Affairs Department.

Hansen, B. E. (1992). Testing for parameter instability in linear models. *Journal of Policy Modelling*, 12(4):517–533.

Hendry, D. F. and Ericsson, N. R. (2001). *Understanding Economic Forecasts*. MIT Press, Cambridge.

Holden, K. and Peel, D. (1990). On testing for unbiasedness and efficiency of forecasts. *Manchester School of Economic and Social Studies*, 58:120–127.

Hyndman, R. J. and Koehler, A. B. (2005). Another look at measures of forecast accuracy. Monash Econometrics and Business Statistics Working Papers 13/05, Monash University, Department of Econometrics and Business Statistics.

IMF (2005). Republic of Mozambique: Selected issues and statistical appendix. IMF Country Report 05/311, International Monetary Fund.

Nordhaus, W. D. (1987). Forecasting efficiency: Concepts and applications. *The Review of Economics and Statistics*, LXIX(4):667–674.

Phillips, S. and Musso, A. (2002). Comparing projections and outcomes of IMF-supported programs. IMF Working Papers 01/45, International Monetary Fund.

República de Moçambique (2001). Plano de Acção para a Redução da Pobreza Absoluta 2001-2005 (PARPA). Technical report, República de Moçambique.

República de Moçambique (2006). Plano de Acção para a Redução da Pobreza Absoluta 2006-2009 (PARPA II). Technical report, República de Moçambique.

Schuh, S. (2001). An evaluation of recent macroeconomic forecast errors. *New England Economic Review*, 2001:35–56.

Siegel, S. (1956). *Non-parametric statistics for the behavioral sciences*. New York: McGraw-Hill.

Timmermann, A. (2006). An evaluation of the World Economic Outlook forecasts. IMF Working Papers 06/59, International Monetary Fund.

Verbeek, J. (1999). The World Bank's Unified Survey projections : how accurate are they? an ex-post evaluation of US91-US97. Policy Research Working Paper Series 2071, The World Bank.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias. *Journal of the American Statistical Association*, 57:348–368.

Tables and Figures

Table 1: Summary forecast error statistics

	N	Weight	Mean	SE	Skew	Kurtosis	P(norm)
GDP growth	11	.	0.40	0.68	-0.74	1.32	0.57
Inflation rate	11	.	-1.69	0.99	-2.13	3.76	0.00
Internal revenues	11	1.00	6.93	2.34	0.11	-0.25	0.57
VAT	11	0.34	6.49	2.16	-0.35	-1.19	0.15
income taxes	11	0.17	-4.71	3.44	-0.41	-0.68	0.74
luxury items tax	11	0.06	3.43	5.62	-0.44	-0.83	0.66
trade taxes	11	0.17	8.84	5.30	-0.12	-0.49	0.90
fuel taxes	11	0.10	7.82	4.99	1.24	0.91	0.06
excise taxes	11	0.04	15.08	7.29	0.71	-1.05	0.02
other taxes	11	0.04	24.53	14.50	-0.23	0.50	0.42
non-tax income	11	0.08	12.88	8.82	-0.05	-0.68	0.48
Pooled (unweighted)	88	1.00	9.30	2.66	0.72	3.68	0.00
External finance	11	1.00	-9.99	8.67	-1.84	2.68	0.00
grants	11	0.66	-14.90	9.09	-0.98	-0.26	0.08
credits	11	0.33	6.66	14.00	-1.10	1.54	0.17
Pooled (unweighted)	22	1.00	-4.12	8.48	-0.70	0.98	0.15

Source: author's calculations

Note: N gives the number of observations; Weight is the average actual contribution of each revenue line to total revenues during the period; Mean is the forecast error (multiplied by 100); SE refers to the standard error of the mean; P(norm) gives the probability of falsely rejecting the null hypothesis that the variable is normally distributed based on the Shapiro-Wilks test.

Table 2: Forecast efficiency of aggregate variables, SUR results

	Inflation	**Growth**	**Internal**	**External**
Constant	-1.23	0.96	2.02	-17.75
	(0.40)***	(0.46)**	(1.69)	(8.59)**
Lagged error	-0.21	-0.03	0.37	-0.15
	(0.07)***	(0.21)	(0.11)***	(0.26)
Trend	-0.02	0.12	1.01	3.17
	(0.13)	(0.15)	(0.47)**	(2.73)
Year 2001	-9.60	-5.85	4.99	19.90
	(1.20)***	(1.70)***	(4.67)	(26.39)
F	26.71	5.96	4.67	0.81
R^2	0.87	0.64	0.47	0.20
N	10	10	10	10

* p<.1, ** p<.05, *** p<.01; standard errors in parentheses
Source: author's calculations
Note: columns refer to the results from the SUR in which the title variable is the regressand; model is as described in the text.

Table 3: Forecast efficiency of forecast errors for disaggregated revenue lines, summary of unrestricted SUR

	Bias	**Trend**	**Seriality**	**Joint**
VAT	19.27	4.43	-0.92	0.00
income taxes	.	5.30	.	0.00
luxury taxes	9.89	7.73	-0.48	0.00
trade taxes	13.88	.	.	0.00
fuel taxes	18.25	.	.	0.00
excise taxes	22.71	-4.17	.	0.00
other taxes	24.18	.	.	0.00
non-tax income	17.86	5.17	.	0.00
lag inflation error, '00-05	5.81	.	.	0.00
lag growth error, '95-99	26.38	.	.	0.04
lag growth error, '00-05	-12.24	.	.	0.00
Year 2001	40.9	.	.	0.00
R^2				0.55
N				94

Source: author's estimates

Note: forecast error is the regressand; columns 'bias', 'trend' and 'seriality' give the coefficients on these regressors for each revenue line in the model; 'joint' gives the joint probability that all coefficients are equal to zero; model is as described in the text; only significant results at 10% level shown.

Table 4: Forecast efficiency of revenues, restricted SUR model results

	PCSE	**FGLS**
Lag forecast error	0.13	0.23
	(0.07)*	(0.04)***
Bias g1	8.43	11.72
	(2.93)***	(1.79)***
Bias g2	17.52	16.75
	(3.78)***	(1.81)***
Bias g3	14.78	15.04
	(4.68)***	(3.02)***
Bias g4	0.15	-1.23
	(6.00)	(2.64)
Trend g1	4.12	3.86
	(0.75)***	(0.42)***
Trend g2	-2.41	-2.30
	(0.92)***	(0.52)***
Trend g3	2.44	3.15
	(1.17)**	(0.96)***
Trend g4	1.88	3.70
	(1.82)	(0.86)***
Lag growth error, '95-99	10.12	-5.41
	(9.86)	(4.01)
Lag growth error, '00-05	-10.90	-10.06
	(2.53)***	(1.10)***
Lag inflation error, '95-99	2.91	8.92
	(1.68)*	(0.78)***
Lag inflation error, '00-05	5.02	4.93
	(1.48)***	(0.63)***
Year 2001	39.60	38.07
	(8.88)***	(3.85)***
R^2	0.47	.
Wald χ^2	153.53	1244.56
N	94	100

* p<.1, ** p<.05, *** p<.01; standard errors in parentheses

Source: author's estimates

Notes: forecast error is the regressand; restricted SUR specification as described in the text; pools of revenue lines are denoted by: g1 (VAT, income and luxury taxes), g2 (border taxes), g3 (all other internal revenues), and g4 (external revenue items); PCSE and FGLS denote panel-corrected standard errors and feasible generalized least squares estimators respectively, both of which adjust for the time-series cross-sectional nature of the data; FGLS includes outliers.

Table 5: Determinants of the tax effort, internal revenue lines only

	PCSE	**FGLS**
Forecast effort, g1	0.33	0.31
	(0.07)***	(0.07)***
Forecast effort, g2	0.31	0.27
	(0.07)***	(0.07)***
Forecast effort, g3	0.31	0.29
	(0.07)***	(0.07)***
Lag actual tax effort	0.10	0.07
	(0.12)	(0.08)
Growth	-0.91	-1.00
	(0.33)***	(0.39)***
Inflation	-0.46	-0.46
	(0.09)***	(0.10)***
Trend	-0.01	-0.01
	(0.00)***	(0.00)***
Constant	2.07	2.23
	(0.42)***	(0.47)***
N	80	80
R^2	0.79	...
χ^2	42.72	46.65

* p<.1, ** p<.05, *** p<.01; standard errors in parentheses

Source: author's estimates

Notes: actual tax effort is the regressand; pools of revenue lines are denoted by: g1 (VAT, income and luxury taxes), g2 (border taxes), and g3 (all other internal revenues); PCSE and FGLS indicate panel-corrected standard errors and feasible generalized least squares estimators respectively, both of which adjust for the time-series cross-sectional nature of the data.

Figure 1: Total recorded revenues as % GDP by internal and external sources

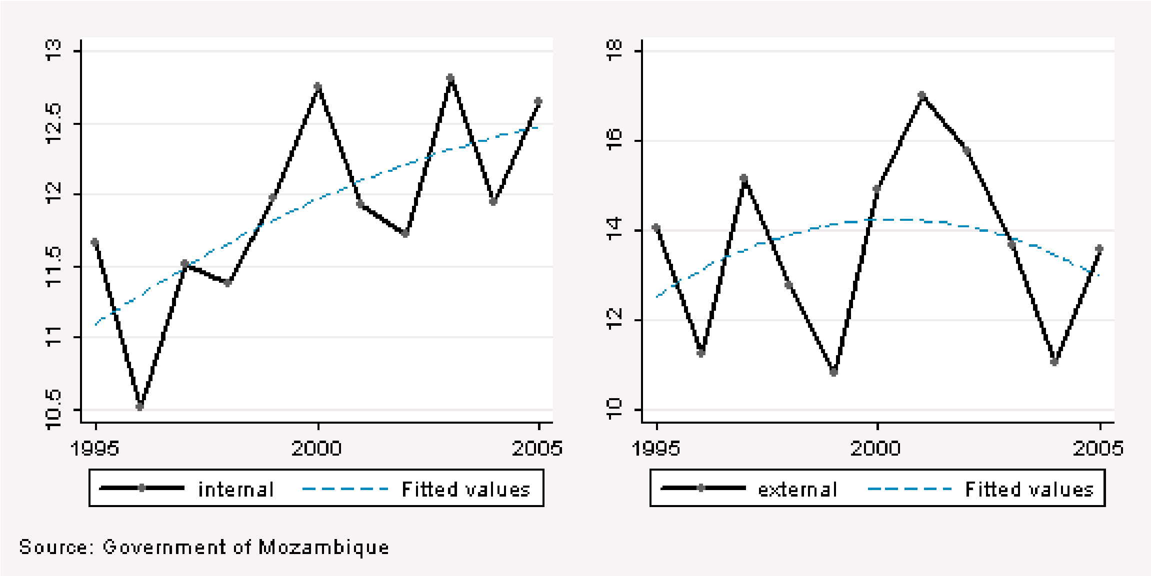

Figure 2: Scatter plots of real growth and inflation forecast errors (in percentage points)

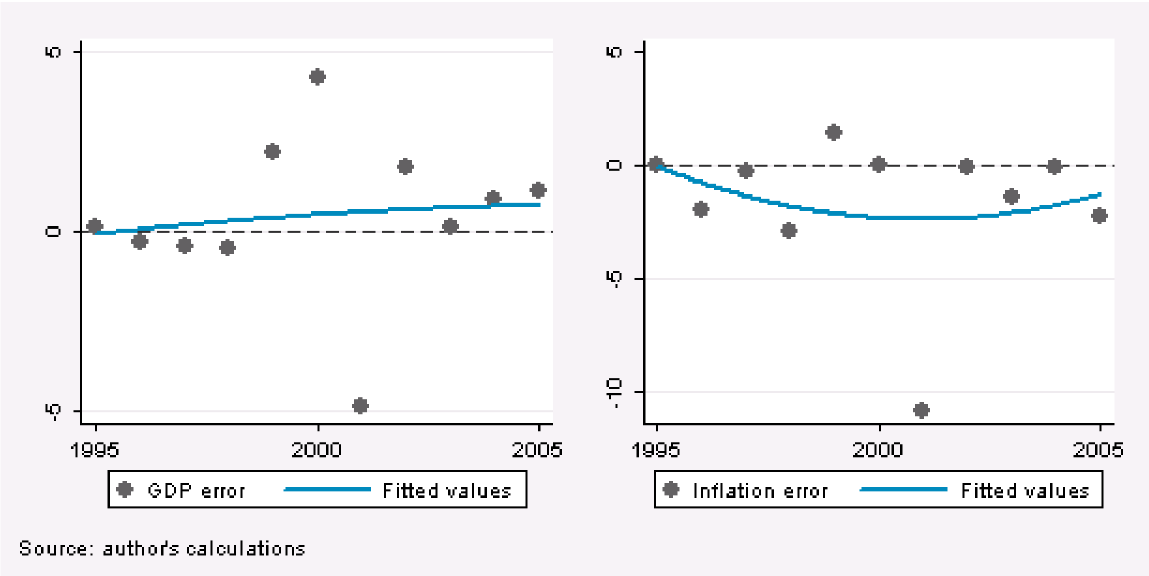

Figure 3: Tax ratio errors for total internal and external revenues

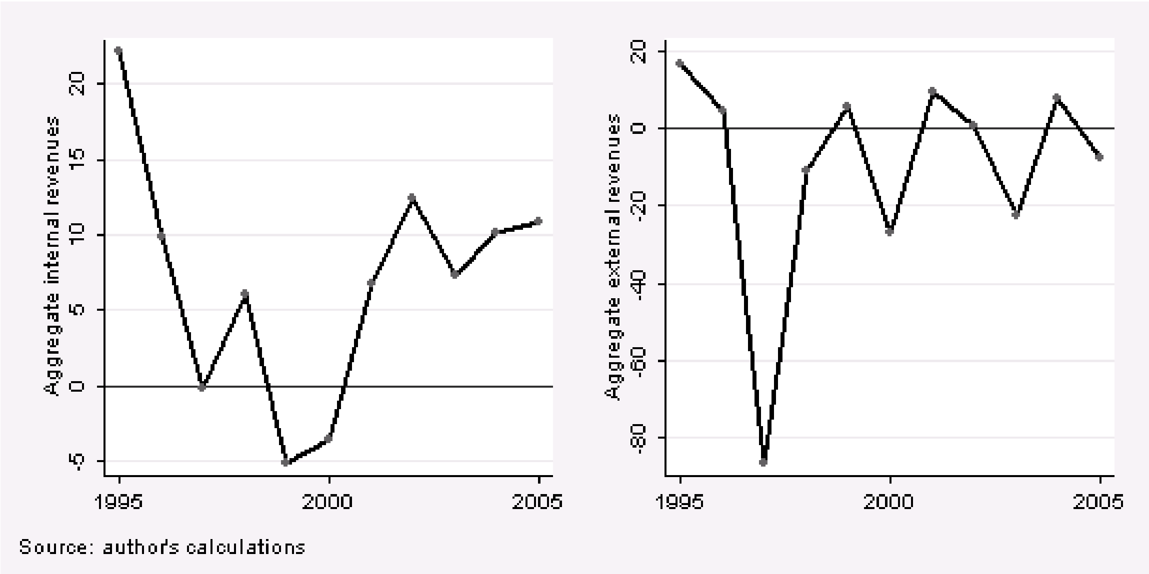

Figure 4: Actual and forecast tax elasticities by internal revenue line and period

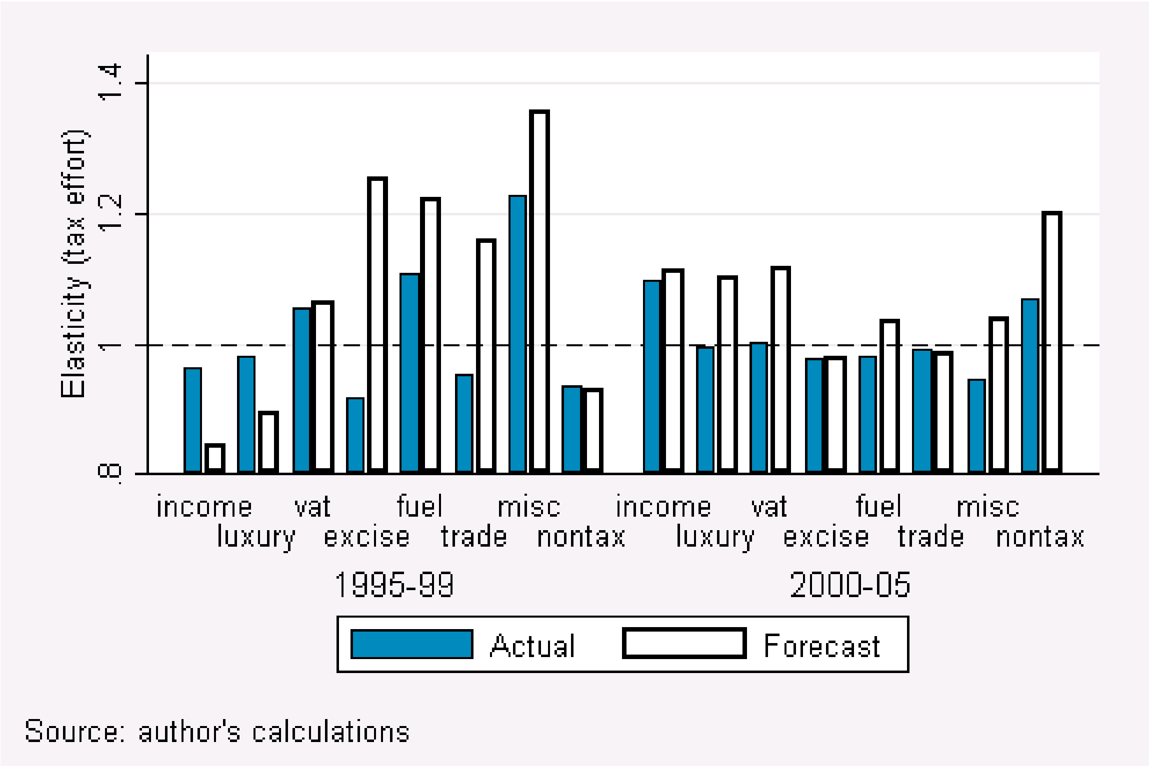

Figure 5: Kernel densities of pooled actual and forecast tax elasticities by period, internal revenues only

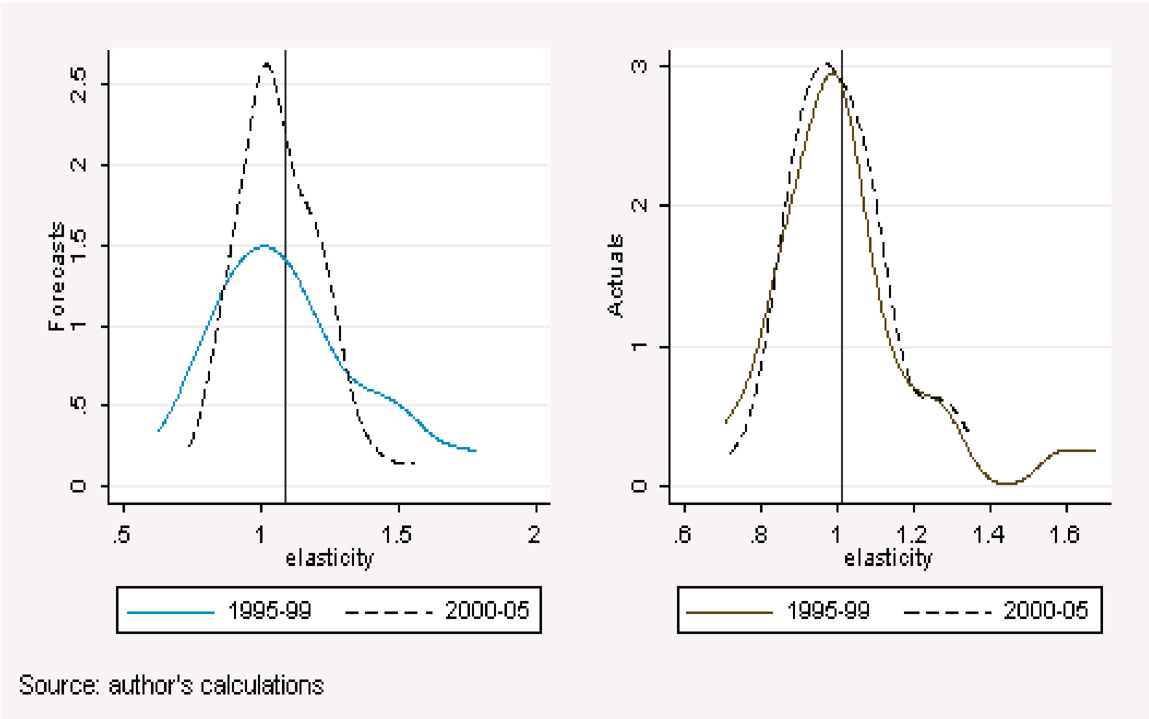

Figure 6: Distribution of logarithmic loss function goodness-of-fit series by period

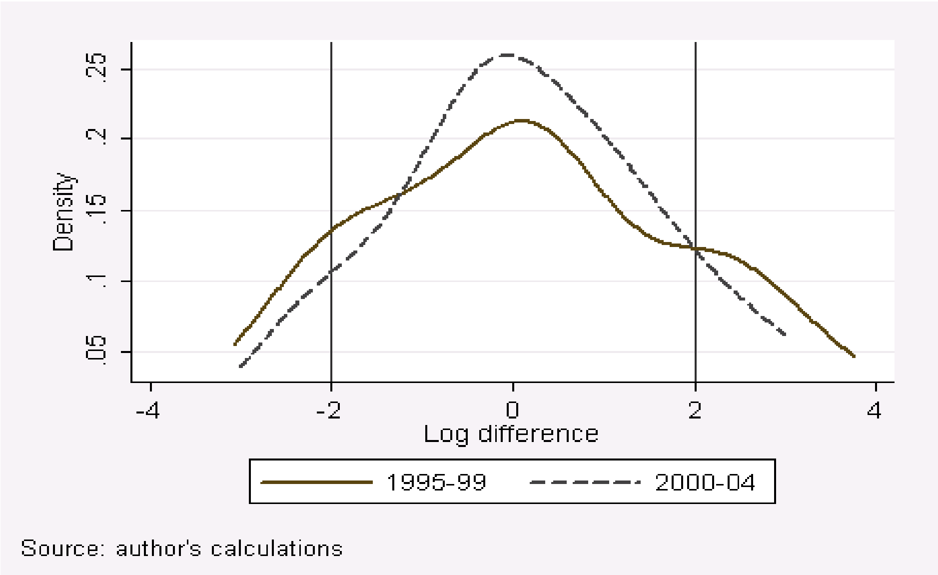

Figure 7: Individual t-statistics for logarithmic loss function by revenue line