ISSN 0265-8003

FISCAL SOLVENCY AND FISCAL FORECASTING IN EUROPE

Michael Artis and Massimiliano Marcellino

Discussion Paper No. 1836
March 1998

Centre for Economic Policy Research
90–98 Goswell Rd
London EC1V 7DB
Tel: (44 171) 878 2900
Fax: (44 171) 878 2999
Email: cepr@cepr.org

CEPR Discussion Paper No. 1836

March 1998

ABSTRACT

Fiscal Solvency and Fiscal Forecasting in Europe*

This paper analyses two features of concern to policy-makers in the countries of the prospective European Monetary Union: the solvency of their government's finances; and the accuracy of fiscal forecasts. Extending the existing methodology of solvency tests, the paper finds that, with few exceptions, EU governments are insolvent, albeit debt/GDP ratios show signs of stabilizing. The accuracy of official short-term fiscal forecasts (those of the OECD) is analysed, using conventional techniques, and found to be reassuring.

JEL Classification: C22, D99, E62, H63
Keywords: solvency, deficits, fiscal policy, Maasticht Treaty, Stability and Growth Pact, forecast accuracy

Michael Artis
Department of Economics
European University Institute
Badia Fiesolana
Via dei Roccettini 9
50016 San Domenico di Fiesole
Florence
ITALY
Tel: (39 55) 468 5777
Fax: (39 55) 468 5202
Email: artis@datacomm.iue.it

Massimiliano Marcellino
IGIER
Università Bocconi
Via Salasco 5
20136 Milan
ITALY
Tel: (39 2) 5836 3327
Fax: (39 2) 5836 3302
Email: massimiliano.marcellino@uni-bocconi.it

*The authors are indebted to Norbert Wuthe for invaluable research assistance and to Ramon Marimon and Andy Rose for helpful comments. Financial support from the Research Council of the European University Institute under the programme, 'Asset Markets, Structural Convergence and European Monetary Policy', is gratefully acknowledged. The remaining errors are entirely the authors' responsibility.

Submitted 16 January 1998

NON-TECHNICAL SUMMARY

Discussion of the economic policy issues associated with the operation of European Monetary Union (EMU) has clearly identified fiscal policy as among the most important. In this paper, we examine two features that are of particular significance in the debate. First, we analyse the *solvency* of the public finances of the EU member states; and second, we analyse the accuracy of fiscal *forecasts*.

Solvency is simply the capacity of governments to repay their debts. In the normal way, the capacity of governments in the last resort to create money is a guarantee – along with their extensive taxing powers – that their debt will be redeemed, at least in domestic currency. This assurance allows governments to maintain and roll over debt over long periods of time. On joining EMU individual governments will be obliged to cede their money-creating powers to the European Central Bank, however, the statutes of which expressly forbid the financing of national governments. At the same time, European governments' taxing powers are being continually eroded as the process of integration advances and tax bases become more mobile, a development which is likely to be accelerated by the advent of monetary union. The result is that debt/GDP ratios, which were tolerable in previous conditions, may begin to look rather risky. While the Maastricht Treaty advanced a debt/GDP ratio of 60% and a deficit/GDP ratio of 3% as 'reference values' in the convergence targets for potential participants in EMU, these figures have no analytical justification as either safe or optimal ratios. It is true that, on the assumption of 5% growth in nominal GDP, the two ratios are consistent, but only in the sense that a continual 3% deficit ratio will support a stable 60% debt ratio. In the new circumstances in which the member governments of EMU will find themselves, caution suggests that lower ratios should be targeted. It is perhaps for this reason that the Stability and Growth Pact indicates that the target for the medium-run average fiscal deficit should in fact be zero. A target of this type (which is for the deficit inclusive of interest payments) is consistent in the long run with a debt/GDP ratio of zero; that is, it would fulfil the solvency condition.

In this paper, we examine data on debt/GDP ratios for the EU countries over the last three decades, to see whether the behaviour of these ratios is consistent with the solvency condition. The paper reviews the econometric methods used in previous investigations of solvency and applies a number of these, together with a new approach to different versions of the debt/GDP ratio (both net and gross debt data are available). The principal finding is that, in our sample period, EU government finances were not solvent; while debt/GDP ratios were not generally exploding, their behaviour is consistent with stabilization at positive non-zero values. A policy change – such as that implied in following the mandate of the Stability and Growth Pact – is required to bring about solvency.

The second issue dealt with in the paper is the accuracy of fiscal forecasts (specifically deficit ratio forecasts). Forecasts of the deficit ratio could prove critical, both in the decision-making to take place in the Spring of 1998 on the initial composition of the Monetary Union and, subsequently, in the implementation of the Stability and Growth Pact. In both cases, consideration of the durability of current fiscal performance and the credibility of fiscal policy require accurate forecasts. The paper therefore subjects the OECD's fiscal deficit forecasts for the EU countries to a battery of tests for their accuracy. Perhaps surprisingly, in view of the sometimes politically charged nature of fiscal forecasts, the OECD's short-term predictions emerge from these tests as reasonably accurate, on the whole unbiased and at least weakly 'efficient' in a technical sense – though there are exceptions and there is room for improvement.

1. Introduction

Fiscal issues have come to command first-order importance in the discussion of economic policy in the prospective European Monetary Union. At the official level concern for these issues has been seen in the provisions of the 1992 Treaty of European Union (Maastricht Treaty), which set out as convergence criteria *inter alia*, a reference value for the budget deficit in ratio to GDP of 3 per cent and for the ratio of government debt to GDP a value of 60 per cent. The Stability and Growth Pact of 1997 carries the Maastricht provisions through to the operation of the Monetary Union itself, reinforcing the "excessive deficit procedure" set out in the Treaty and *inter alia* calling on Member States to "commit themselves to respect the medium-term budgetary position of close to balance or in surplus set out in their stability or convergence programmes" (European Council, 1997).

The economic rationale for the inclusion of the fiscal criteria in the Treaty among the convergence requirements has been much discussed (see, for example, Buiter and Kletzer (1992) and Buiter, Corsetti and Roubini (1992)), whilst the provisions of the Stability and Growth Pact have also been hotly debated (Artis and Winkler (1997) attempt a review). Certain points are clear and relatively uncontroversial, however. In particular, as far as debt/GDP ratios are concerned, the move to monetary union raises question marks about the sustainability of debt/GDP ratios at previous, or even at much-reduced levels: as Mongelli (1996) has pointed out, within the framework of European Monetary Union (EMU) member states lose the power of money creation with which to guarantee the repayment of their debts, whilst at the same time these same member states face increasing restrictions on their taxing powers due to the rising mobility of tax bases within the Union. Further, participation in EMU itself, with the added pressure this will bring on remaining obstacles to the free flow of finance and financial services within the euro-zone and the greater transparency imparted to transactions, is likely to liberate investors from captive home markets. In these circumstances it is not difficult to appreciate the point that McKinnon is making when he draws attention (McKinnon, 1996, chapter 19) to the disparity between national debt/GDP ratios in Europe (even at Maastricht-levels of 60 per cent) and the comparable debt/State product ratios in the United States (which are nearer to sixteen per cent). There is no particular reason to think that any set ratio of debt to

GDP is sustainable; the transition from the Maastricht 60 per cent debt ratio with a 3 per cent interest-inclusive deficit limit to the Stability and Growth Pacts emphasis on a target of a zero or surplus (interest-inclusive) budget balance in the medium run is perhaps indicative of this concern. A fresh examination of the solvency of EU government finances is equally justified and it is such an examination which occupies the first part of this paper.

Since solvency is literally the condition that future primary surpluses can be foreseen which are sufficient to repay all existent (and any future) debt, this might seem at first sight a rather stringent criterion to implement. However, it has to be conceded that solvency, as judged by the behaviour of fiscal variables within the sample period, is by no means a sufficient condition for pronouncing the government finances healthy (or otherwise), and can only be regarded as an indication in this respect. In particular, solvency is in essence a forward-looking concept. The future path of fiscal policy may not, due to a regime change, resemble that of the sample period; a regime change that has taken place only near the end of the sample may not have sufficient weight, when pooled with earlier observations, to produce the "correct" verdict on the solvency of the State[1]. In these respects solvency analysis is very much a first step, and one that needs to be accompanied by other forms of analysis. In particular, as Perotti et al. (1997) stress, controllability is a key issue. They place their emphasis on searching for the causes of fiscal errors and on reform of the fiscal process to check these, though in doing so they start from Maastricht deficit and debt ratio criteria. These latter have no particular analytical basis, other than being consistent with one another on reasonable assumptions about growth and inflation. A bottom-line justification for the solvency criterion, by contrast, is that its meaning is unambiguous, at least, in principle.

Current arrangements provide for the identification of potential EMU members to take place in April 1998 on the basis of out-turn data for 1997; but it is clear that fiscal forecasts for further years will play a role in determining whether those out-turns reflect durable and credible achievements, as required by both the Maastricht Treaty

[1] Hansen et al. (1991) conclude that the intertemporal budget balance condition yields no useful restrictions. But they reach this conclusion on the basis of additional assumptions which are inappropriate to our data set.

and the resolution of the German Parliament[2] which governs Germanys acceptance of partner countries in the European Monetary Union. Thus fiscal forecasting accuracy is at a premium. In the latter part of this paper we address directly the question of the accuracy of official fiscal forecasts - in this case, those of the OECD.

The data base used in this study and some of its concerns resemble those in Uctum and Wickens (1997). As do these authors, we also use data derived from the OECD with some, mostly minor, differences. We also share a concern for testing intertemporal budget constraints - in our case, exclusively for solvency (infinite horizon) whilst in their case the emphasis is on testing for sustainability (finite horizon achievement of specified debt/GDP positions). Whilst Uctum and Wickens incorporate forecast values in the sample they analyse for sustainability, we test for the accuracy of these (or related) forecasts.

2. The evolution of debt

A test for solvency is simply a check on whether debt can be repaid. So the solvency condition for government debt requires that there be a prospect for future budget surpluses sufficient to pay off current debt. To clarify these points, some algebra will be useful.

The accounting identity describing the evolution of government debt at constant prices is

$$B_t \equiv (1 + r_t)B_{t-1} - S_t, \tag{2.1}$$

where B_t and S_t indicate the debt and primary surplus inclusive of seigniorage, while r_t is the real interest rate. Assuming that $r_t \geq 0$ in all time periods, (2.1) is an unstable non-homogenous difference equation which can be solved forward to yield

$$B_t = \lim_{n\to\infty} E_t\left(\prod_{s=1}^{n}(\frac{B_{t+n}}{1+r_{t+s}})\right) + \sum_{s=1}^{\infty} E_t\left(\prod_{j=1}^{s}(\frac{1}{1+r_{t+j}})S_{t+s}\right), \tag{2.2}$$

where E_t is the expectation operator conditional on information available at time t. When

$$\lim_{n\to\infty} E_t\left(\prod_{s=1}^{n}(\frac{B_{t+n}}{1+r_{t+s}})\right) = 0, \tag{2.3}$$

[2] Deutscher Bundestag: Bundestagsdruck Sache 12/3906, 2 Dezember 1992.

the debt at time t equals the sum of discounted future surpluses, the intertemporal government budget constraint is satisfied, and the *solvency* condition (or no Ponzi game condition) is met. The government will not then be indulging in perpetual debt refinancing.

Several proposals have been put forward in the literature to test whether government debt histories meet this condition. They are briefly reviewed in the first subsection, while in the second one we present the econometric methodology that will be followed in the paper.

2.1. Literature review

Hamilton and Flavin (1986) assume a constant real interest rate, r, and maintain an assumption that the deviation of the debt from the sum of discounted future surpluses grows at the rate r. In this case (2.2) would become

$$B_t = c(1+r)^t + \sum_{s=1}^{\infty}(1+r)^{-s}E_t(S_{t+s}), \tag{2.4}$$

and the intertemporal budget constraint would only be satisfied if $c = 0$. They suggest three procedures for testing whether $c = 0$. The first one relies on the observation that if both the debt and the sum of discounted surpluses are stationary, then indeed $c = 0$.[3] The other ones are tests for the significance of $(1+r)^t$ in the regression equation which is obtained after substituting the expected values in (2.4) with alternative extrapolative approximations. When applied to annual US data for the period 1960-1984, the hypothesis that $c = 0$ cannot be rejected by any methods, providing support for the validity of the intertemporal budget constraint. A similar conclusion is achieved by Haug (1990), with quarterly data for the period 1960-1987.

Wilcox (1989) relaxes the assumption of a constant interest rate. He discounts the variables back to period zero, so that we rewrite equation (2.1) as

$$q_t B_t = q_{t-1}B_{t-1} - q_t S_t, \tag{2.5}$$

[3] c can be equal to zero even if both the debt and the sum of discounted surpluses are I(1) variables, but they are cointegrated with a cointegration vector equal to (1,-1). Smith and Zin (1991), using monthly data for Canada for the period 1946:1-1984:12 and assuming that the surplus follows an AR(1) process, test for cointegration and reject it.

where

$$q_t = \prod_{j=0}^{t} \frac{1}{1+r_j}; \quad q_0 = 1. \tag{2.6}$$

Equation (2.2) then becomes

$$q_t B_t = \lim_{n\to\infty} E_t(q_{t+n}B_{t+n}) + \sum_{s=1}^{\infty} E_t(q_j S_j), \tag{2.7}$$

and the relevant issue for solvency is whether the infinite horizon forecast of the discounted debt, the first term on the right hand side of (2.7), is equal to zero or not. As we will see in more detail in the next section, a necessary condition for the limit to exist is that the discounted debt is not integrated of order one, $I(1)$, while the expectation is equal to zero if the variable is stationary and its unconditional mean is equal to zero. Both these hypotheses are rejected by Wilcox, using unit root tests with the Hamilton and Flavin (1986) dataset.

Ahmed and Rogers (1995) show that under mild conditions the first term on the right hand side of (2.7) is equal to zero if and only if the deficit inclusive of interest payments is a zero mean stationary process. If receipts (T), expenditures (G), and interest payments are I(1) variables, the latter condition is satisfied if and only if

$$T_t - G_t - r_t B_{t-1}$$

is a cointegration relationship. Hence, they test for cointegration using a very long sample (1792-1992 for the US and 1692-1992 for the UK), and accept this hypothesis. Similar results were obtained by Trehan and Walsh (1988, 1991), while Hakkio and Rush (1991) rejected cointegration over the period 1975-1988 with quarterly data.

Other authors have focused on the behaviour of the debt to gdp ratio. This seems natural in a growth economy. In this case (2.5) can be rewritten as

$$d_t b_t = d_{t-1} b_{t-1} - d_t s_t, \tag{2.8}$$

where b and s are the debt and surplus to gdp ratios, g is the rate of growth, while d, the discount factor, is now

$$d_t = \prod_{j=0}^{t} \frac{1}{1+r_j-g_j}; \quad d_0 = 1. \tag{2.9}$$

Equation (2.2) becomes

$$d_t b_t = \lim_{n\to\infty} E_t(d_{t+n} b_{t+n}) + \sum_{s=1}^{\infty} E_t(d_j s_j), \tag{2.10}$$

and the transversality condition

$$\lim_{n\to\infty} E_t(d_{t+n} b_{t+n}) = 0 \tag{2.11}$$

is satisfied if $d_{t+n} b_{t+n}$ is a stationary zero mean process. Uctum and Wickens (1997) test for the validity of (2.11) using unit root tests with annual data for the period 1965-1994 and get mixed results for EU countries, while its validity is rejected for the US. The latter result contrasts with the finding of a bounded debt to gnp ratio by Kremers (1989).

Notice that convergence to zero of the discounted debt ratio is in general not sufficient for convergence of the undiscounted ratio. Actually, if the debt ratio is positive and we consider $r - g$ as a random variable with positive support whose lower bound is $\underline{r-g}$, it is

$$d_{t+n} \le \underline{r-g}^{-(t+n)}$$

and

$$E_t(d_{t+n} b_{t+n}) \le \underline{r-g}^{-(t+n)} E_t(b_{t+n}).$$

The term $\underline{r-g}^{-(t+n)}$ converges to zero exponentially, so that the discounted debt ratio can converge to zero even if the undiscounted ratio diverges at a lower than exponential rate. This suggests that both quantities should be analysed and not only the discounted one.

Equations (2.1), (2.5), and (2.8) can be also used to track the behaviour of debt, possibly discounted or as a ratio of gdp, over a finite horizon. This is particularly relevant when there is a medium term target in terms of a certain level of debt, and it is of interest to evaluate whether the current economic policy will allow the target be achieved or not. For example, from (2.8), the expected value of the debt ratio in period $t + m$ is

$$E_t(d_{t+m} b_{t+m}) = d_t b_t - \sum_{s=1}^{m} E_t(d_j s_j).$$

From this formula, given a desired value for $d_{t+m} b_{t+m}$ and a future path for expenditures, growth and the real interest rate, it is possible to determine a path of receipts which is expected to satisfy the target. If the current expected path of receipts already satisfies the target, the policy is usually said to be *sustainable*.

It is also possible to construct indicators of fiscal sustainability based on the divergence between current and required fiscal paths, see for example Blanchard *et al.* (1990) or Mongelli (1996). In this case the crucial element is the formulation of expectations on the future path of relevant variables such as growth, inflation and interest rates. One possibility is to construct time series models for these variables and use them for forecasting future values, see e.g. Chouraqui et al. (1986). As an alternative, official forecasts can be used, as e.g. in Wickens and Uctum (1997). These authors show that, even in those EU countries where the solvency condition is satisfied, current fiscal policy may prove unsustainable, in the sense of being inconsistent with the achievement of a particular debt ceiling by a given short-medium term target date; the required fiscal contraction can be rather substantial.

2.2. Econometric methodology

The main statistical result that can be derived from the previous subsection is that, once the proper debt measure is chosen, the validity of the transversality (solvency) condition requires the debt measure to be a stationary zero mean random variable. The alternative procedures which have been suggested in the literature and some new ones that we propose below, all aim at testing this hypothesis. It is useful to distinguish five sets of tools for the analysis of the debt ratio.

1. *Descriptive analysis.* A graph of the evolution of the debt measure over time can provide a first indication about its stationarity around a zero mean, besides conveying useful information on the existence of structural breaks in the series. The shape of the correlogram is also a useful indicator. For the variable to be stationary, it should decay rapidly, starting from a relatively low value. Moreover, if the debt measure is stationary, differencing will induce a unit moving average root, so that the spectrum of the first difference of the variable, i.e. the deficit, should be equal to zero at frequency zero.

2. *Unit root tests.* These are the most common tool for testing for stationarity of the debt measure. Apart from the usual warnings concerning the sensitivity of

the results to the choice of the deterministic component and of the lag length (or of the band-width for spectral based tests) (e.g. Schwert (1989)), the low power in finite samples (e.g. Podivinski (1997)), and the bias in the results if there are structural breaks (e.g. Perron (1989)), a subtle but important statistical issue should be recalled because it is usually overlooked when interpreting the results. Since we are interested in testing for stationarity of the debt measure, it is this that should be the null hypothesis of the test, while the null hypothesis of the usually adopted unit root tests, e.g. the ADF test or the Phillips test, is that the variable is non-stationary, I(1) in particular. Moreover, strictly speaking, rejection of the I(1)-ness hypothesis should not be interpreted as acceptance of stationarity, because the test is not symmetric. Hence, either it should be clearly specified that it is believed that the debt measure is I(1) and the usual tests run with the aim of providing support for this hypothesis, or different tests have to be used, whose null hypothesis is stationarity, e.g. tests for a unit MA root in the generating mechanism of the first differenced variable. We will also present a simple spectral based test for the latter hypothesis.

3. *Cointegration tests.* Focusing for example on the discounted debt ratio, we can think of at least two alternative possibilities for applying cointegration tests in order to check whether the solvency condition is satisfied. Actually, from (2.8), $d_t b_t$ is (non-) stationary if and only if either $log(d_t)$ and $log(b_t)$, or $d_{t-1}b_{t-1}$ and $-d_t s_t$ are (non-) cointegrated, in both cases with cointegration vectors equal to $(1,1)$. There is again the warning that the null hypothesis of the usual tests is lack of cointegration, see e.g. Engle and Granger (1987), Johansen (1988).

4. *Policy simulations.* The techniques that we have described so far are extrapolative, in the sense that they rely on historical data and assume that the generating mechanism of these data will remain unchanged in the future. Yet, economic policy in most EU countries has been changing, mainly in order to satisfy the Maastricht criteria, which have now been complemented by the more ambitious injunctions of the Stability and Growth Pact. Hence, we will also try to determine and comment on the likely consequences of these criteria for the validity of the solvency condition.

5. *Forecasting.* Apart from the infinite horizon forecast which is neccessary to assess the validity of the solvency condition, it is often of interest to evaluate the short-medium term behaviour of the debt or deficit measures. One possibility is to produce ARMA based forecasts for the variables of interest, but in this way it is not possible to take into account expected changes in policy. To do this requires the use of large macroeconometric models or the forecasts produced by official agencies, such as the OECD, IMF or EC. Uctum and Wickens (1997) include OECD final forecast data in their sustainability analysis. In this paper we assess the reliability of these (deficit) forecasts.

We conclude with two warnings. First, the derivation of the solvency condition relies on the assumption that the real interest rate, possibly after subtracting the growth rate, is positive. That this is the case is usually taken for granted, likely because otherwise the economy would be dynamically inefficient according to standard economic theory, see e.g. Diamond (1965).[4] Yet, Ball *et al.* (1995) find that the average of $r_t - g_t$ is slightly negative for the US, even if there are differences across periods, and a similar result was achieved by Mishkin (1984) for other countries. This implies that the government, in the presence of balanced primary budgets, could repay its debt without tightening fiscal policy. Second, expectations play a key role in the ex ante analysis of debt behaviour, but realizations are what really matter. Hence, the ex ante results need not be valid ex post, even if they can provide useful insights on what could happen.

3. Debt ratios in EU countries

In this section we analyse the behaviour of (government) debt to gdp ratios for EU countries, except Greece and Luxembourg for which the OECD, our data source, does not provide debt figures. The exact sample ranges are indicated in the figures and tables; they often start in the early '70s and end in 1994. More detailed information on the variables involved in the analysis are contained in a data appendix which is available upon request. The first subsection presents a descriptive analysis of the

[4] Abel et al. (1989) show that the latter claim is incorrect in a stochastic environment.

relevant variables. In the second subsection the issue of whether the debt ratios are stationary or not is addressed.

3.1. Descriptive analysis

The starting point is the choice of the proper debt measure to use. Most of the previous studies focus on net debt, which is the relevant measure from an economic point of view because it takes into account the financial assets held by the Government. Yet, policy makers are often more interested in gross figures, e.g. the Maastricht criteria are in terms of gross debt. Moreover, for Denmark and Portugal only gross debt data are available.. Hence, when possible, we will study the behaviour of both net and gross debt to gdp ratios, b and gb respectively.

From figure 1, the two debt ratios present a similar evolution, which implies that the difference between the two debt measures has been rather stable over time, and of relevant size, around 20% of gdp on average. Belgium, Italy and Ireland have the highest ratios, but while all three ratios steeply increased up to the late '80s, the Irish one substantially decreased afterward, the Belgian one decreased its rate of growth, while the Italian one continued rising. Sweden and Finland are instead characterized by a negative net debt for most of the sample period, and by a rapid increase in the ratios in the final part of the sample, which also takes place in Denmark, Austria, France and Germany, and is partly due to the consequences of the recession of the early '90s on government deficits.

The rate of growth in the debt ratios is in general higher in the '80s than in the '70s, with the exception of Finland and, in particular, of the U.K. As we will see later on, the main determinant of such a pattern is the different behaviour of the interest and growth rates in the two periods. But the most important feature of the graphs of the debt ratios is that they provide very little support for a convergence of the ratios to zero in the long run.

So far we have used measures of debt which are available at face value, but it may be more appropriate to use its market value. The proper discounting requires premultiplying the debt figures by $1/(1+i)$, where i is the nominal interest rate on government debt. A proxy for i is the ratio of net interest payments to net debt lagged one period (see e.g. Uctum and Wickens (1997)). Yet, this measure mixes the interest rates paid by the Government on its liabilities with those received on its assets. Hence, it may be more appropriate to use the ratio of gross interest payments to gross debt lagged one period, which is called gi. We will discount b with i and gb with gi, and refer to the resulting debt ratios as mb and mgb. Their behaviour is very similar, respectively, to that of b and gb. [5]

A comparison of i and gi provides useful information on the "financial efficiency" of the Government. Actually, it can be easily shown that when i is higher than gi the government is paying a higher average interest rate on its liabilities than it receives on its assets. From the graphs in figure 1, this seems to be the case for Austria, the Netherlands, and UK in the final part of the sample. The reverse relationship is more reasonable and reflects the lower risk premium that the government has to pay (and, the use by the government of zero coup on financial instruments). Actually, both i and gi are lower on average than market rates.

Several authors have also suggested discounting the debt measure back to the beginning of the sample period. In the case of the debt to gdp ratio, the proper discount factor is d in equation 2.9. To construct the required difference between the real rate of interest and the real growth rate $(r - g)$, we can subtract the nominal rate of growth, ng, from either i or gi. We use $i - ng$ to discount mb, and $gi - ng$ to discount mgb. The resulting measures are labelled dmb and $dmgb$. The alternative definitions of the debt ratio that we have introduced so far are summarised in table 1.

[5] There are problems in the calculation of i for those countries which experience a negative net debt in some periods. In fact, the net interest payments are sometimes positive in the same periods, reflecting an inefficient financial managament and/or measurement errors. In these cases we have preferred to rely on gi.

Table 1: Alternative definitions of the debt ratio

$b_t = \frac{\text{Net Debt}_t}{\text{Gdp}_t}$	$gb_t = \frac{\text{Gross Debt}_t}{\text{Gdp}_t}$
$mb_t = \frac{1}{1+i_t} b_t$	$mgb_t = \frac{1}{1+gi_t} gb_t$
$dmb_t = \prod_{j=0}^{t} \frac{1}{1+i_j-ng_j} mb_t$	$dmgb_t = \prod_{j=0}^{t} \frac{1}{1+gi_j-ng_j} mgb_t$

Notice that d is a proper discount factor only if $r-g$ is positive, otherwise d is larger than one and increasing in time. This is also the condition that ensures that the forward solution of the equation which governs the evolution of the debt ratio (equation 2.8) is not explosive. As we mentioned in the previous section, this is often an untested assumption in empirical analyses on debt sustainability. Yet, the graphs in figure 1 show that the real rate of interest was higher than growth in most countries during the '70s, sometimes also in the early '80s, and for Spain, Ireland and Finland for most of the sample period..

In this case, and in the presence of a balanced primary budget, the debt ratio can decrease without any need for restrictive fiscal policy, as can be immediately derived from equation 2.8, and eventually converges to zero. In fact, the ratio started decreasing in several countries in the early '70s, e.g. Austria, Belgium and the Netherlands. But then it remained rather stable or even increased. This is consistent with the fact that governments ran budget deficits, mainly in order to offset the negative effects of the two oil crises and the related recessions, which more than cancelled out the beneficial effects of low real interest rates.

The situation changed in the '80s, when $r-g$ become positive in several countries, and this helps to explain the aforementioned higher growth of the debt ratio in this subperiod. It seems reasonable to regard the '70s as a rather particular period and thus to assume that $r-g$ will remain positive in the future. Thus, the forward solution of equation 2.8 is stable, and whether solvency holds remains an issue. The graphical analysis in this subsection casts serious doubts that this is the case, even if the discounted debt ratios start decreasing in the final part of the sample in some countries.

3.2. Is there a unit root?

In Section 2 we provided formulae which describe the evolution of (possibly discounted) debt ratios conditional on the behaviour of interest and growth rates, and of primary deficits. The solvency condition requires convergence to zero of the expected value of the debt ratio. Such an expected value can be also obtained from a univariate representation of the debt ratio, namely, one where the evolution of the ratio only depends on its own lags. Let us consider for simplicity the model

$$b_t = c + \phi b_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim i.i.d.(0, \sigma_\varepsilon). \tag{3.1}$$

If $\phi = 1$, it is

$$E(b_{t+n}) = b_t + cn,$$

so that the expected value diverges linearly if $c \neq 0$, or is equal to the current debt ratio if $c = 0$.[6] Thus, $|\phi| < 1$ is a necessary condition for solvency. It is not sufficient, because $c = 0$ must also hold for the expected value to converge to zero. It can be easily demonstrated that these conditions are also valid for more general univariate generating mechanisms, and this explains the interest in the literature in testing whether there is an autoregressive unit root in the generating mechanism of the debt ratio.

In order to test for such an hypothesis, we start by applying the Augmented Dickey Fuller (ADF) tests, whose null hypothesis is that $\phi = 1$, i.e., that the solvency condition is *not* satisfied. We include a constant in the regression and up to four lags, when the coefficient of the highest lag must be significant according to a t-test. Usually only one or two lags are necessary, which is coherent with the annual frequency of the data, but the results appear to be robust to the choice of the lag

[6]When $\phi > 1$, divergence is even explosive but this case is usually ruled out a priori in economic applications. In our case it would happen, for example, if the difference between the real interest and growth rate were constant and positive, and the deficits deviated randomly from a constant, as can be derived from equation 2.8. From an empirical point of view, we never found autoregressive roots larger than one, even if there were a few spurious cases where they were due to a different behaviour of the variables at the beginning or end of the sample period.

length and to whether the constant is present or not. We also check that the resulting residuals are uncorrelated, homoskedastic and normally distributed, which are the required conditions for the statistic to be the maximum likelihood test for a unit root. These hypotheses are usually accepted, and when they are not we have modified the models by including additional lags or dummy variables in order to evaluate whether the result of the unit root test changed, but in most cases it did not.

In Table 2 we report the ADF tests for b, gb, dmb, $dmgb$, run on the basic autoregressive models, together with the chosen lag length and the sample period.[7] For none of the countries and debt ratio measures can the hypothesis of a unit root be rejected. The estimated values of the root are often higher than 0.9, which is coherent with the slowly decaying autocorrelation functions whose starting values are also often above 0.9. Hence, the alternative to a unit root should be a root very close to one, but the power of the test in discriminating between these two possibilities is very low.

A rather subtle issue is whether the parameters of the models are stable in time. Usual statistical techniques for testing for such an hypothesis are hardly applicable in our context because of the small sample size which, e.g., makes recursive estimation often infeasible or unreliable. As an alternative, structural breaks could be imposed a priori. In particular, from the graphs in figure 1 to 4, the hypothesis of a segmented trend in the generating mechanism of the debt ratios could be a plausible alternative to that of a unit root for several series, e.g., it could capture the change in the growth rate of the debt ratios due to the reversal of the relationship between real interest and growth rates. Yet, the implementation of tests to distinguish between the two hypotheses (e.g. Perron (1989)) is not particularly interesting in our context, because both of them imply that the debt ratios grow linearly, so that solvency could not be satisfied.

We now exploit cointegration theory to provide additional evidence on whether the debt ratios are stationary or not. The debt ratio can be decomposed into

$$b_t = (1 + r_t - g_t)b_{t-1} - s_t. \tag{3.2}$$

[7]The results for mb and mgb are very close to those for b and gb and are not reported to save space. The values for net ratios are close to those in table 1 of Uctum and Wickens (1997) who analyse a similar dataset, minor differences being due to the choice of the lag length and the sample period.

Thus, cointegration between $u_t = (1+r_t-g_t)b_{t-1}$ and $v_t = -s_t$ is a necessary condition for the debt ratio to be stationary, while a cointegration vector equal to $(1,1)$ is also sufficient. Moreover, from the definition of the discounted debt ratio and invariance of stationarity to the logarithmic transformation, the discounted ratio is stationary if and only if $y_t = \log(d_t/(1+i_t))$ and $w_t = \log b_t$ are cointegrated with cointegration vector equal to $(1,1)$.

The maximum likelihood statistics suggested by Johansen (1988, 1991, 1995) allow us to test these hypotheses. Notice that the null hypotheses of his trace and $\lambda-\max$ statistics are no cointegration, i.e., that solvency does *not* hold. The starting point for the construction of the tests is the specification of a VAR model for the variables, whose residuals are uncorrelated, homoskedastic and normally distributed. In our case, VARs with one or two lags and an unrestricted constant usually satisfy these requirements and therefore provide a proper framework for cointegration testing. The results are reported in table 3, where gu, gv, gy and gw are defined as u, v, y and w but using gross variables.[8]

With respect to unit root tests, the results from cointegration tests for the undiscounted debt ratios are the same; for all countries the null hypothesis that they are integrated cannot be rejected, possibly with the exception of Finland. Actually, the hypothesis that the cointegration vector is $(1,1)$ is rejected, even if cointegration is often accepted. This is mainly due to stationarity of the primary surplus, i.e. the cointegration vector is $(0,1)$. Instead, for the discounted measures there are some differences. Discounted net debt ratios appear to be stationary for Belgium, Spain and, marginally, Italy, and gross ratios for Austria, Belgium, U.K., and the Netherlands.

The different outcomes may be due to the fact that cointegration tests are applied to the logarithms of the ratios, while unit root tests are referred to their levels.

[8]Notice that the equivalent of equation (3.2) for discounted variables is:

$$d_t b_t = d_t(1+r_t-g_t)b_{t-1} - b_t s_t = d_{t-1}b_{t-1} - d_t s_t.$$

Thus, we cannot test for cointegration between $d_{t-1}b_{t-1}$ and $d_t s_t$ with the Johansen procedure, because the equation for $d_{t-1}b_{t-1}$ in a VAR would be the identity

$$d_{t-1}b_{t-1} = d_{t-2}b_{t-2} - d_{t-1}s_{t-1}.$$

Hence, in the case of discounted variables, we focus on testing for cointegration between y_t and w_t.

Whether logs or levels are used is irrelevant asymptotically, but it can be important in small samples. In order to evaluate whether this is the case, we ran ADF tests for the logs of the discounted debt ratios, with results that agree with those from the cointegration tests. Hence, the logarithmic transformation matters and, given that we are interested in the levels of the ratios, where conclusions differ it seems safer to rely on the ADF tests.

The null hypothesis of the test statistics that we have used so far is that solvency does not hold. Because of the available small sample size and the persistence of the variables, the power of the tests is rather low, so that the null hypothesis is likely to be accepted even if it is not true. This suggests that in order to have a fair evaluation of whether the debt ratios are stationary or not, we should also apply tests that maintain stationarity as the null hypothesis. For example, from the ADF regressions, the null hypothesis that the highest autoregressive root is 0.9 can always be accepted. But the choice of the stationary value of the root to be tested for is arbitrary. As an alternative, we recall that differencing a stationary series will induce a unit moving average root in the generating mechanism of the first differenced variable. Hence, we can test for stationarity of the debt ratio by testing for a unit moving average root in the generating mechanism of its first difference.

Unfortunately, the distribution of the likelihood ratio test for this hypothesis has not been derived so far, while Lagrange multiplier tests (e.g. Tanaka (1990), Saikkonen and Lukkonen (1993)) are rather complex and their small sample performance still has to be thoroughly evaluated. Thus, we adopt a simpler procedure which is based on the observation that the spectrum at frequency zero of a variable has to be equal to zero if there is a unit moving average root in its generating mechanism. In table 4 we report estimates of the spectrum at frequency zero with standard errors, using the Bartlett kernel and setting the bandwith at double the square root of the number of observations. Similar results are obtained with the Tukey and Parzen kernels, and with different values of the bandwith. Even if the distribution of the estimator is not exactly normal, for all the variables the value zero always falls well within the 95% confidence interval based on the normal distributution (estimated value $\pm$ 1.96 * standard error). We think that this provides reliable evidence that if stationarity is the maintained hypothesis, it cannot be rejected.

In summary, there is substantial uncertainty on whether the debt ratios are stationary or not. Fortunately, an exact answer to this question is not necessary for our aim, and the reason for this statement is the role of the constant term. The t-statistics for its significance in the ADF regressions, which are quite often based on congruent univariate representations of the variables, are also reported in Table 2. The critical values are different under the hypotheses of integration and stationarity, higher in the former case (see Dickey and Fuller (1981)). With few exceptions, the constant is not significant if it is accepted that the debt ratio is non stationary, while it is significant if the debt ratio is stationary. In both cases the implication is that the ratio will converge to a constant value, but not to zero, and therefore solvency can be expected not to hold.

4. Solvency and the Maastricht Criteria

So far we have used historical data to make inferences on whether the debt ratio will converge to zero or not. This provided useful information, but it is also important to take into account expectations of the future behaviour of the variables, which mainly reflect announced changes in fiscal policy. Uctum and Wickens (1997) report OECD forecasts for net debt ratios up to the year 2000, and these in general show a decline in the ratios, in particular for those countries whose ratio was over 60% in 1994, or at least a non marked increase[9].

This is probably the result of the fiscal requirements in the Maastricht Treaty, and in particular of the deficit and debt ratios ceilings. Actually, we can rewrite the equation for the evolution of the gross debt ratio as

$$gb_t - gb_{t-1} = (gi_t gb_{t-1} - gs_t) - ng_t gb_{t-1}, \tag{4.1}$$

where the term in parentheses is the gross deficit ratio, $dr_t = gi_t gb_{t-1} - gs_t$, and ng is the nominal rate of growth. If dr is set equal to 0.03 (the 3% of gdp), real growth to 0.03, and inflation to 0.02 (so that $ng = 0.05$), the equilibrium value of the gross

[9] Actually, there is a suggestion in the data that the 60% debt ratio has become an 'attractor', both for high ratio and low ratio countries. A model rationalizing this behaviour can be found in Giovannetti et al. (1997).

debt ratio is 0.6 (i.e. 60% of gdp), which coincides with the Maastricht requirement.[10] Countries that start with a higher debt ratio will experience a gradual reduction toward this value, while there can be an increase in low ratio countries, unless their deficit ratio is lower than 0.03, e.g. because of the lower interest payments burden.

If the ratio of government assets to gdp remains constant, a similar pattern will emerge for net debt, which is in fact coherent with the OECD forecasts. Instead, if the real interest rate is even only marginally higher than real growth, the discount factor will converge to zero, and this will also drive the discounted debt ratios toward this value.

Of course these results are sensitive to changes in the forecasts of inflation, growth, real interest rates, and the primary deficit ratio. Inflation plays a minor role because, apart from minor receipts from seigniorage, higher nominal interest rates are compensated by higher nominal growth. Higher real interest rates do not also affect the evolution of the debt ratio if the deficit ratio remains the same, but this requires lower primary deficits or higher surpluses. Otherwise they lead to an increase in the equilibrium ratio. Instead, lower growth always exerts a negative effect through the term $ng_t gb_{t-1}$, and it can also lead to a temporary relaxation of the deficit criterion, according to the rules set up in the Treaty. While higher interest rates and lower growth have a negative effect on the debt ratio, they speed up the convergence to zero of the discount factor, and of the discounted ratio, notwithstanding the increase in the raw figures. Finally, higher primary deficit ratios are possible in the presence of lower interest payments; otherwise they will lead to an increase in the debt ratio.

In summary, the Maastricht criteria are compatible with a constant debt ratio and, under rather plausible assumptions about the average future behaviour of the variables, the equilibrium value coincides with the debt ratio criterion. Thus, solvency is not implied by the Maastricht criteria if the debt ratio is undiscounted, while if it is discounted solvency can be expected to hold. The zero medium run deficit requirement of the Stability and Growth Pact clearly implies more. Since the deficit is interest-inclusive, maintenance of the target will imply, on average, primary deficit surpluses and hence a reduction towards zero in the debt ratio, discounted or not.

[10] This consistency has been pointed out elsewhere (e.g. in Buiter, Corsetti and Roubini (1992)) though it is not clear that these figures were chosen for their consistency (see Bini-Smaghi and Padoa Schioppa (1994)).

5. Deficit forecasts

The previous sections mainly dealt with long/medium horizon forecasts of the debt ratio, but short term forecasts are also of interest. Their reliability depends on that of the determinants of the evolution of the debt ratio, dr and ng from equation (4.1). Artis (1989, 1997) analyses official forecasts from IMF and OECD for growth and inflation, and concludes that they are in general quite accurate for EU countries, which implies that ng in (4.1) can also be accurately forecast[11]. In this section we evaluate the reliability of gross deficit ratio forecasts, drf, which is also interesting per se, given the considerable relevance that the deficit ratio has for economic policy.

We focus on OECD forecasts, which is coherent with our choice of analysing OECD data on debt ratios, and consider year-ahead and current year forecasts, $drfya$ and $drfc$ respectively. Year-ahead forecasts are identified with those published in the OECD's Economic Outlook for December of year t for t+1, whilst current-year forecasts are those made in the June issue of the Economic Outlook in year t for year t. They are compared with first released actual data on gross deficit ratios. Such a comparison is the most interesting from a policy perspective, and further revisions of actual data usually do not affect the results, see e.g. Artis (1989). We analyse countries for which at least ten forecasts are available, namely, all EU countries except Ireland and Portugal.

The year ahead and current year forecasts errors, $eya = drfya - drf$ and $ec = drfc - drf$ respectively, are graphed in figure 2. The mean error (ME), mean absolute error (MAE) and root mean squared error (RMSE) are reported in Tables 5 and 6. On average the MAE is about 1%; the worst performance is for Finland, Norway, Sweden and Greece with around 2% MAE. The values for the MEs are smaller and negative, which implies that on average the actual values for these countries were under-predicted, but also that there were periods of substantial over-prediction. Actually, from the graphs of the forecast errors, it appears that under-prediction in the '80s was followed by over-prediction in the '90s. Such a pattern, in particular over-prediction in the '90s, seems to be also present for other countries, e.g. Austria, Belgium, Denmark, France and Spain. Such an asymmetric pattern could be

[11] In fact, there is some evidence that as errors in real growth and inflation forecasts are negatively correlated, forecasts of nominal output growth are the most accurate.

due to the introduction of the Maastricht Criteria, which may have increased the loss associated with underprediction of the gross deficit ratio.

A comparison of the RMSEs for current year and year ahead forecasts indicates that the former are smaller than the latter for all countries, suggesting that the additional six months' information is indeed useful in forecasting the deficit ratios. MAEs are also smaller for current year forecasts, while the results on MEs are somewhat mixed. Again, the highest RMSEs are for Finland, Norway, Sweden and Greece.

We now formally analyse the unbiasedness and weak efficiency of the forecasts. It is conventional to claim that the forecasts are unbiased when $\alpha_0 = 0$, $\alpha_1 = 1$ in the regression

$$dr_h = \alpha_0 + \alpha_1 y_h + u_h, \tag{5.1}$$

where y is either $drfya$ or $drfc$, and u is an error term that under the null hypothesis of unbiasedness coincides with the forecast error (see, e.g., Clements and Hendry (1997a, Ch.3)). Yet, Holden and Peel (1990) showed that this condition is sufficient but not necessary for unbiasedness, and suggested to substitute for it the condition $\beta_0 = 0$ in the regression

$$x_h = \beta_0 + v_h, \tag{5.2}$$

where x is either eya or ec, and v is the demeaned forecast error. As is well known, weak efficiency instead requires the forecast error to be uncorrelated in time (see, e.g., Clements and Hendry (1997a, Ch.3)).

Table 7 reports, for year ahead and current year forecasts, the Wald test (W) for $\alpha_0 = 0$, $\alpha_1 = 1$, which is distributed as $\chi^2(2)$, the t-test (T) for $\beta_0 = 0$ and two Lagrange Multiplier tests ($L1$ and $L2$) for lack of first and second order autocorrelation in the forecast errors, which are distributed as $F(1, H-1)$ and $F(2, H-2)$, where H is the number of available forecasts. The main result is that for both year ahead and current year forecasts unbiasedness is rejected for most countries by the W test, with the exception of Belgium, France, UK, Spain and Sweden. Yet, when the T test is applied, unbiasedness is accepted for many more countries. Even if this is possible from a theoretical perspective, such a remarkable mismatch between the T and W tests is likely due to the small sample size. Another warning for the interpretation of the results is that the graph of the forecast errors suggested a possible change of regime in the '90s. There are too few observations for running a split-sample analysis,