Do fiscal variables affect fiscal expectations? Experiments with real world and lab data

Michele Bernasconi[†] Oliver Kirchkamp[*] Paolo Paruolo[†]

August 12, 2006

Abstract

We generate observable expectations about fiscal variables through laboratory experiments using real world data from several European countries as stimuli. We compare a VAR model of expectations for data which is presented in a fiscal frame with one for neutrally presented data. We find that participants understand the meaning of the fiscal variables, but also that their ability to perceive the correct characteristics of fiscal policy is limited. We tie the VAR analysis to specific models of forming expectations. We find that agents' expectations are neither consistent with rational nor with purely adaptive expectations but, instead, follow an augmented-adaptive scheme.

Keywords: Experiments, fiscal policy, expectations, causality, cointegration, panel data.

JEL classification: C91, D89, E62, H31.

---

[*]corresponding autor: Oliver Kirchkamp, University of St Andrews, School of Economics and Finance, St Salvator's College, St Andrews, Fife KY16 9AL, Scotland, UK. Phone:+44-1334-462440, Fax:+44-1334-462444, oliver@kirchkamp.de

We are grateful to two referees for their detailed comments. Financial support from Deutsche Forschungsgemeinschaft through Sonderforschungsbereich 504 is gratefully acknowledged.

[†]Department of Economics, University of Insubria, Via Ravasi 2, I-21100 Varese, Italy. michele.bernasconi@uninsubria.it, paolo.paruolo@uninsubria.it. Partial financial support from Italian MIUR grants ex 60% is gratefully acknowledged by both authors.

1 Introduction

Expectations on fiscal variables are crucial to understand the effect of fiscal policy on the private sector. Little, however, is known on the actual way people form expectations on fiscal variables. While many models are based on the hypothesis of rational expectations, expectations themselves are not easily observable. Empirical tests of the rational expectations hypothesis are, therefore, often indirect, involving tests of predicted relations between observables, like relationships between fiscal variables and components of output. From there the effect which the unobservable expectations might have played is inferred. Examples of this approach range from classical tests of the Ricardian equivalence (see Seater, 1993, for a review), to more recent analyses of the so called 'anti-Keynesian' (i.e. expansionary) effects of fiscal adjustments (Giavazzi, Jappelli, and Pagano, 2000).

A problem in this approach is that the identification of the effects of expectations is model-dependent, and model comparison is very hard. Factors which are not part of the model may affect expectations and choices in real economies. This makes identificating the effects of expectations difficult, because "economists cannot observe all the data that economic agents do" (Seater, 1993, p. 164).

The latter limitation is also relevant for expectation measures derived from opinion surveys[1]. Moreover surveys suffer from lack of economic incentives to reveal true opinions, so that for various reasons respondents "may express judgements that are different from the ones they choose to act upon" (Pesaran, 1987, p. 209).

A third approach is to measure expectations in a controlled experiment. Expectations experiments have been conducted since as early as the sixties in a wide range of economic contexts. One has looked at expectations for prices (Fisher, 1962, and Schmalensee, 1977), expectations about artificially constructed time series (Hey,

[1]For example, surveys conducted by Grun (1991) and Allers, de Haan, and de Kam (1998) found widespread evidence of misinformation on the conduct of government fiscal policy.

1994), expectations in normal-form games (Costa-Gomes and Crawford, 2001, and Nyarko and Schotter, 2002), expectations regarding the provision of public-goods (Offerman, Sonnemans, and Schram, 1996, and Croson, 2000), expectations in the cobweb economy (Hommes, Sonnemans, Tuinistra, and van de Velden, 2003), on monetary policies (e.g. Marimon and Sunder, 1993, and Bernasconi and Kirchkamp, 2000) or even fiscal policy (Swenson, 1997).

While the experimental approach allows to implement the *ceteris paribus* condition as none of the other two (and correct for the lack of incentives of the survey approach), it easily suffers from the critique that the stimuli are given based on abstract models in situations that are far from the real economic world, thus questioning the validity of the connection between the lab and the real economy.

In this paper we study the process of forming expectations on fiscal policy. We combine field and lab data to be able to control information that subjects observe, while still giving them concrete data taken from the real world[2]. Here we limit our control to the effect of fiscal variables.

The approach innovates on the existing literature in various respects. First of all, stimuli are represented by annual time series of fiscal variables from 15 OECD countries. Unlike in previous experiments, they refer to bivariate, not univariate, times series, namely taxes and public expenditures. This allows us to include a richer set of models in our comparison and to extend the analysis to augmented adaptive models.

We check the internal validity of the experimental set-up by way of a control group of participants who form expectations without knowing the fiscal origin of the times series and we discuss the extent of the external significance of the experiment.

We develop a modelling strategy which is consistent both with current practice in macro-econometrics and micro-econometrics. We assume that the joint data generating process (DGP) of stimuli and expectations data is a

[2]The experiments of Fisher (1962) and Schmalensee (1977) used real price data, though the focus there was on purely statistical, rather than economic meaning of data.

VAR. Various econometric results in the area are collected and we show how to analyse first the field data and subsequently the expectation data.

Many macroeconomic investigations use cointegration techniques and Granger causality tests to investigate the sustainability of fiscal policy and the type of causality between taxes and public expenditure (see, e.g., Trehan and Walsh, 1991 and Ahmed and Rogers, 1995, as classical references; Garcia and Henin, 1999, as a more recent example); the present two stage approach contains this analysis as the first step.

The approach encompasses many micro-models for the formation of expectations in the second step. We find that a major component of the process depends on past forecast errors. We find that subjects substantially fail to perceive the fiscal properties which we estimate in the field.

The evidence neither supports the 'rational expectation hypothesis', nor a purely adaptive scheme; rather, expectations fall within a class of so called 'augmented-adaptive models', introduced in the early eighties by various authors (see Pesaran, 1987). These models then become the starting point for a growing literature of 'bounded rationality' (Sargent, 1993) and 'adaptive learning' in self referential economies (see Evans and Honkapohja, 2001, for a comprehensive survey).

The VAR approach we take to analyse the data excludes non-linear behaviour in the DGP. This may be disputable, since discretionary interventions and exogenous shifts may introduce non linearities in fiscal policy (as for example documented for the US by Bohn, 1998, and Sarno, 2001). The latter case is of interest since it may also generate specific anti-Keynesian effects of fiscal policy (see e.g. the models surveyed in Giavazzi, Jappelli, and Pagano, 2000, and the empirical investigation conducted therein). We, therefore, test for nonlinearity and find that the VAR specification is robust against it.

The paper is organised as follows. Section 2 presents the setup of the experiment and discusses issues of internal and external validity. Section 3 develops the econometric approach to analyse both the stimulus and the expectations data. Inference results are presented in Section 4. The last section summarises and discusses the more general implications of the findings.

2 Experimental setup

The experiment has a time-structure, $t = 1, \ldots, n$; the setup nests a simple (two-periods) representative agent small economy. Participants are exposed to graphical representations of time series of fiscal variables, taken from various European countries. The stimuli refer to gross total taxes T_t, total public expenditure (inclusive of interest payments) G_t, public debt B_t, and change in the debt level $\Delta B_t = B_t - B_{t-1}$ at time t, all expressed as yearly percentage of GDP. Here and in the following Δ is the time difference. In this paper we focus on the relationship and the direction of perceived causality between taxes and expenditure which are denoted by the vector $x_t := (T_t, G_t)'$.

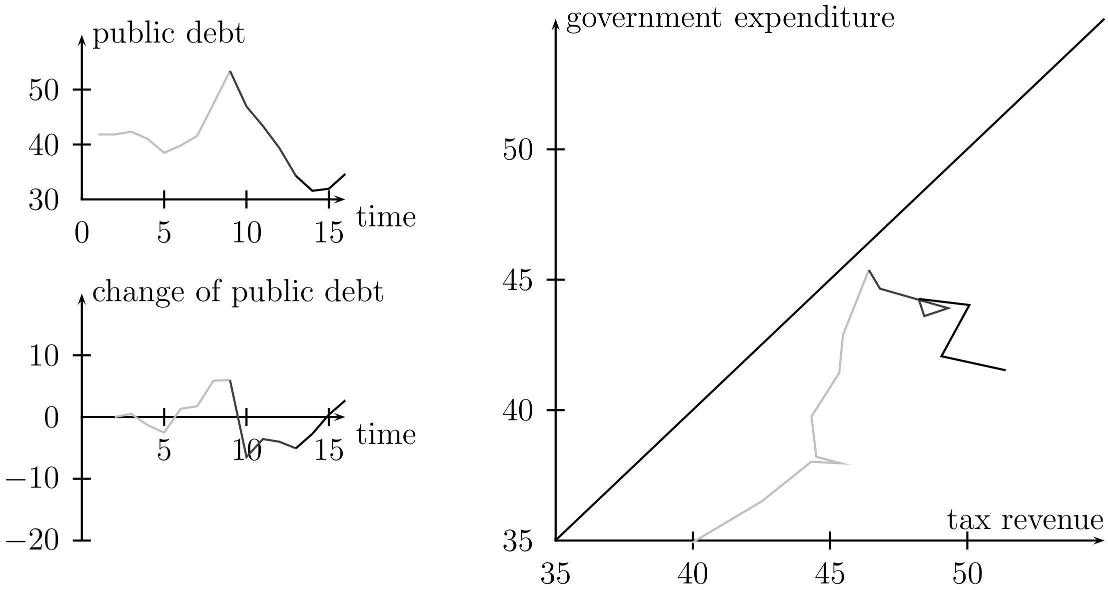
Agents do not know which country and which period the series refer to. Utility in the experiment is derived from consumption over two subsequent periods:

$$u_t = \prod_{i=t-1}^{t} \frac{3}{4}C_i + \frac{1}{4}G_i \tag{1}$$

subject to the budget constraint

$$\sum_{i=t-1}^{t} \underbrace{(1 - C_i - T_i)}_{\text{savings}} \cdot (1 + r)^i = 0 \qquad \text{with } r = 0.1 \tag{2}$$

(where 1 stands for the normalised total GDP). Agents receive initial information on the first seven values of stimuli which was for most countries the period from 1970 to 1976. Let $t - 1$ be the last available year and $X_{t-1} := (x_1, \ldots, x_{t-1})'$ the available information, then forecasts are made for each subsequent year t. Three experimental treatments were performed: in one of the treatments participants forecast both T_t and G_t (see figure 1), in a second treatment participants forecast T_t only (see figure 2). The third treatment is a control treatment that is designed similarly to the first (T_t and G_t) treatment. The only difference is that participants are not informed about the economic significance of the variables. Instead, variables are simply called A, B, and C. Everything else is as in the original T_t and G_t treatment. We call this treatment the 'neutral' treatment. Agents express forecasts by

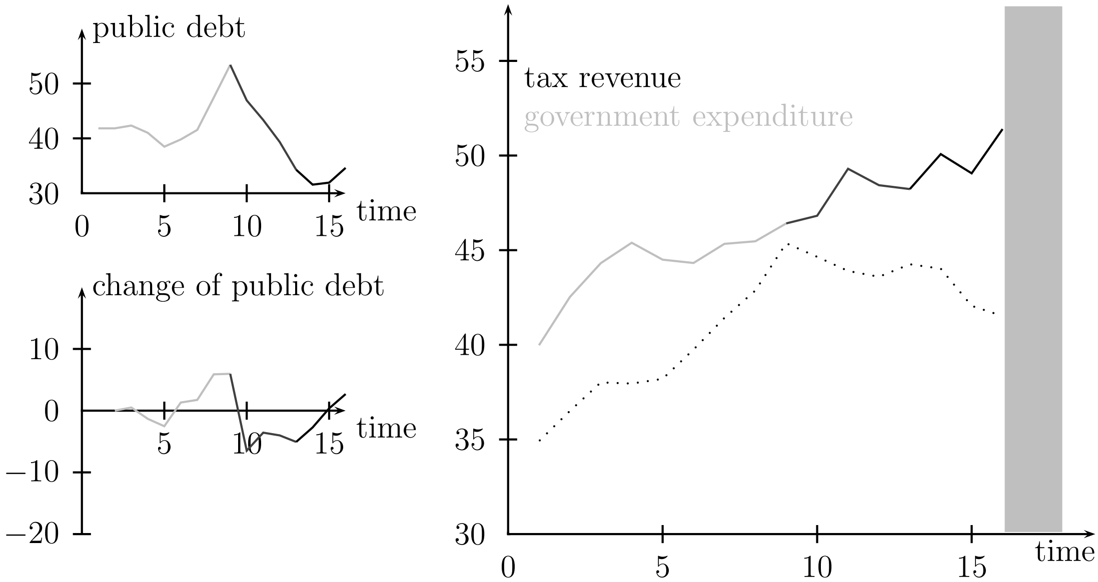


Values are given as percentage of GDP. Gray lines indicate past, black lines indicate recent years.

FIGURE 1: T_t and G_t treatment

clicking with the mouse directly into the diagrams. Instructions can be found in appendix A. Forecasts are called $T_t^{E_i}$ and $G_t^{E_i}$, where i indicates agent i and E stands for expectation. Let $y_{i,t}$ indicate all the forecasts of agent i that refer to time t; in the T_t and G_t treatment $y_{i,t} := (T_t^{E_i}, G_t^{E_i})'$ while $y_{i,t} := T_t^{E_i}$ in the T_t only treatment.

The time series of the stimuli are updated recursively in each period after forecasts are made, so that subjects learn about realization of the stimuli as the economy moves on. More specifically, given subjects' forecasts $y_{i,t}$ for year t, the computer determines an optimal consumption level C_{t-1} given eq. (1) and (2). In period t, $x_t := (T_t, G_t)'$ become available and are communicated to the participant. The computer uses equation (2) to determine C_t and then equation (1) to calculate the participant's utility. The participant's per



Values are given as percentage of GDP. Gray lines indicate past, black lines indicate recent years.

FIGURE 2: T_t only treatment

minute wage is

$$w = 0.66 \cdot (u_t/u_t^*)^\eta \text{ where } \eta = \begin{cases} 12000 \text{ in the } T_t \text{ and } G_t \text{ treatment} \\ 15000 \text{ in the } T_t \text{ only treatment} \end{cases} \quad (3)$$

where u_t^* is the utility the participant would obtain with forecasting the true values. This transformation from utilities into wages is monotonic and, hence, does not affect the maximisation problem of the individual. The transformation, however, creates steeper incentives to make good forecasts[3]. Participants got feedback not only about the development of the fiscal variables but also about their payoff and quickly understood the payment scheme in the experiment.

Participants are payed this wage up to two minutes for each forecast. If a

[3]Parameters of the payment scheme have been developed in pilot studies previous to the experiment.

Country	Sample period	number of participants T_t and G_t	 T_t only	 neutral
Austria	1970-1998	34	14	28
Belgium	1970-1998	52	20	48
Denmark	1971-1995	43	14	43
Finland	1970-1998	33	12	24
France	1977-1998	29	10	31
Germany	1970-1998	14	13	5
Greece	1975-1998	26	13	22
Ireland	1970-1995	45	21	44
Italy	1970-1998	35	16	29
Netherlands	1970-1998	32	9	23
Norway	1970-1998	32	14	26
Portugal	1970-1998	25	13	23
Spain	1970-1998	30	11	24
Sweden	1970-1998	53	16	41
United Kingdom	1970-1995	27	15	19
non-economists (54)		23	5	26
economists (116)		53	20	43
Total (170)		76	25	69

Table 1: Summary of the experimental treatments
Since participants made forecasts for more than one country the total number is not the sum of the number of participants in all countries.

participant needs more time to complete a forecast only the first two minutes are paid[4].

Different agents participated in the three treatments. 76 of our participants took part in the T_t and G_t treatment, 25 in the T_t only treatment, and 69 in the neutral treatment. Of these participants 116 were students in economics or business (we will call them economists), 54 were from a different field (natural sciences, languages, law, and social sciences; we will

[4]We have introduced this payment scheme to simultaneously encourage participants to think about their forecasts, but also to remain active.

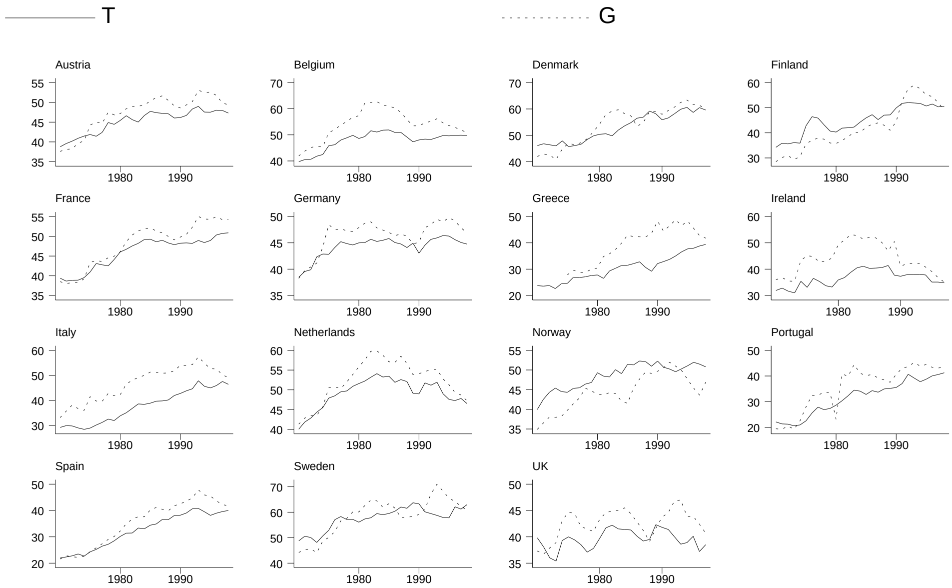


FIGURE 3: Stimulus data (as G and T over t)

call them non-economists). Each agent made predictions for several countries within each treatment. Stimulus data were from 15 European countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden and UK[5]. Table 1 summarises the parameters of the treatments with the number of participants in the various conditions. For the majority of countries the sample period of stimulus data was 1970-98; few exceptions (Denmark, France, Greece, Ireland and UK) are due to limits in the availability of the fiscal time series. For all countries, expectations recording started from the seventh year of the stimulus (which was then 1977 for most countries). A graph of the stimulus data for the different countries is shown in figure 3.

The experiment were carried out in the experimental laboratory of the SFB 504 in Mannheim. Experimental sessions were conducted individually.

[5]All stimulus data used in the experiment were taken from the OECD (2000) database "Fiscal Positions and Business Cycle".

All 170 participants spent about 2 hours in the laboratory. They made, on average, 136 forecasts (between 28 and 309), and completed on average one forecast every 39 seconds.

2.1 Internal validity

Laboratory experiments are often inspired by the real world, and often use in their setup parameters which are based on field data. Recent examples from macroeconomic experiments include tax compliance experiments (see, e.g., Alm and McKee, 2004) and decisions over monetary policy (Lombardelli, Talbot, and Proudman, 2002; Blinder and Morgan, 2005). Using not only parameters, but datasets from the field in the laboratory is novel and we believe innovative, but poses various questions. A first one is about the internal validity of the experimental setup.

In particular, can we trust that participants understand the difference between the economic and non-economic context? Or are they simply behaving alike in the two experimental frames?

Notice that we don't necessarily expect that subjects in the fiscal frame can make very accurate expectations or can perceive the correct properties of the fiscal variables. More simply, we only want to verify whether or not subjects in the economic frame understand the meaning of the fiscal variables and feel confident that they can make forecasts about them.

To answer this question we compare the T_t and G_t treatment with the neutral treatment. If participants understand and make use of the economic context we might expect smaller forecast errors in the fiscal frame than in the neutral frame. Furthermore, we can also check whether only participants from the economists group are confident with the fiscal variables.

To test this formally we explain the mean squared errors as a function of a dummy d_{e} that is one for economists and zero otherwise, and a dummy d_{n} that is one in the neutral treatment and zero otherwise.

$$(T - T^E)^2 = \beta_{\mathrm{e}}^T d_{\mathrm{e}} + \beta_{\mathrm{n}}^T d_{\mathrm{n}} + c + \epsilon \tag{4}$$

$$(G - G^E)^2 = \beta_e^G d_e + \beta_n^G d_n + c + \epsilon \quad (5)$$

Results of the robust regression are shown in table 2. In both equations we

estimation of equation 4 (145 independent observations)						
$(T - T^E)^2$	β	σ	t	$P > \|t\|$	95% Conf. Interval	
β_n	.5586159	.2216946	2.52	0.013	.12042	.9968119
β_e	−.0226633	.2255656	−0.10	0.920	−.4685105	.423184
c	2.791348	.1996356	13.98	0.000	2.396754	3.185943
estimation of equation 5 (145 independent observations)						
$(G - G^E)^2$	β	σ	t	$P > \|t\|$	95% Conf. Interval	
β_n	1.071765	.4808198	2.23	0.027	.1213885	2.022141
β_e	.0039117	.5156949	0.01	0.994	−1.015398	1.023221
c	6.366691	.5207256	12.23	0.000	5.337437	7.395944

TABLE 2: Test for internal validity
All standard deviations are calculated using a robust regression method taking into account correlations of observations for each participant.

have β_n positive and significantly different from zero. β_e is not significantly different from zero[6].

It is good news that β_n is significant and positive in both equations. This allows us to conclude that we have obtained internal validity in our experiment. It is also good news that β_e is not significantly different form zero. We can conclude that students from non-economic fields understood the context of the experiment as well as those from economics.

[6]The estimation of an extended model that controls also for country specific effects yields similar results. Since the country specific effects are not significant we have omitted the results of this estimation here.

2.2 External significance

An other important question about the experimental setup concerns its external significance. Consumers in the field form expectations on fiscal policy not only on the basis of spending and taxes. Expectations might be affected by the composition of the government, the date of the next election, economic development, foreign policy, natural disasters, to name only a few.

We are neglecting all these factors in the experiment to be able to focus on a class of simple models. We want to make sure that subjects' expectations in our experiments only depend on the dynamics of taxation and public expenditure and are not tainted by other variables. Indeed, it is a strength, not a weakness, of laboratory experiments that we can concentrate on a small number of factors neglecting others. In the field we would not know whether a change in expectations is due to a change in the fiscal policy or due to some other factor. In the lab changes in our participants' expectations can only be the result of the information we provide them.

The limited information provided to subjects in our experiment represents the typical information used by a vast empirical literature to analyse the fiscal conduct of countries and to derive judgements regarding the sustainability of their fiscal policies. We in particular refer to an important stream of literature which since at least the mid-1980s have used cointegration techniques and the econometrics of Granger causality to study the dynamics of the public budget intertemporal constraint[7]. In the remaining body of the paper, we develop an empirical model of stimuli and of expectation data, which coherently ties the present experiment to such econometric practice. Through it we discuss the meaning and implication of general expectation schemes (rational, adaptive, augmented adaptive), we consider and estimate properties of field data and we check how these features are perceived by subjects in the experiment.

[7] See section 3.1 below for various references to the literature.

3 A VAR approach for the field-lab data

This section discusses our modelling strategy for the stimuli and for expectation data. The approach is based on the assumption that the data generation processes for the field and the lab data can be taken to be two Vector Autoregressive Processes (VAR). The VAR assumption is particularly advisable in the present context for various reasons. First of all, it permits a specific analysis of the stimuli which in the present context is just as important as the one of the responses. Furthermore, regarding the latter, it will allows us to model explictly the links between the information set available to subjects from the field and their expectations in the lab[8].

3.1 The VAR for the field data (stimuli)

The field model concerns the stimuli data on taxes and public expenditure, namely the vector $x_t := (T_t, G_t)'$, treated for each country separately. A natural and standard VAR representation of the field systems is given by:

$$\Delta x_t = \mu_x + \alpha_x \beta'_{1x} x_{t-1} + \sum_{\ell=1}^{k-1} \Gamma_{\ell,xx} \Delta x_{t-\ell} + \epsilon_{xt}, \tag{6}$$

The expression takes in particular the stimuli data $x_t := (T_t, G_t)'$ as integrated of order 1, I(1) (this is also confirmed by visual inspection of the series in the differences); μ_x is a vector of constants; $\beta'_{1x} x_{t-1}$ is an error correction term (cointegrating relation CI, see below) expressing the long run disequilibrium between taxes and expenditure in the field; $\alpha_x = (\alpha_T : \alpha_G)$ is the associated vector of the adjustment coefficients; $\Gamma_{\ell,xx}$ are the matrices of the coefficients of partial adjustments, and ϵ_{xt} are error terms i.i.d. $N(0, \Omega_{xx})$.

Versions of this basic model have been analysed quite extensively in the

[8] In fact, we emphasise that if even we present here the VARs for the field and for the lab as two separate models, they can be shown to be consistently derived as a marginal model (the field model) and as a conditional model (the lab model) of a joint general VAR process. Such formal derivation, omitted here for sake of simplicity, is available from our working paper (Bernasconi, Kirchkamp, and Paruolo, 2004).

empirical time series analysis on fiscal policy. Two issues have been considered with particular attention: the sustainability of fiscal policy and the direction of causality between fiscal variables.

In greater detail, the issue of sustainability of fiscal policy, which is of an ever ending concern in most countries, has been tested in the literature in various ways. One of the most popular rests on the cointegration analysis between taxes and public expenditure inclusive of interest payments: in short, cointegration tests of sustainability are based on the idea that solvency requires the budget deficit to be stationary (see, e.g. Trehan and Walsh, 1991, Hakio and Rush, 1991, Ahmed and Rogers, 1995)[9]. Even if T_t and G_t are non-stationary in most cases, there may be linear combinations of the form $T_t + \gamma G_t + \rho$ (where γ and ρ are constants) which are stationary. In this case, T_t and G_t are said to be cointegrated of order 1, with cointegrating vector (CI) $(1, \gamma, \rho)$. Stationarity of the budget then further requires $\gamma = -1$ (homogeneity condition).

We follow a standard procedure to test for cointegration in system (6) based on Johansen (1995). We start determining the lags order k of the field VAR: for each country, we begin with $k = 5$ and then restrict the order eliminating lags which are not statistical significant and checking for absence of correlation in the residuals. The Johansen (1995) procedure is used to test for the rank r of the system (that is, the number of CI vectors). We use the LR trace test for $H_0 : r = 0$ versus $H_1 : r = 1$, and exclude the case of a stationary system $r = 2$, where both taxes and expenditure are stationary in levels. This assumption is justified by previous studies and

[9] A slight modification of the same idea works on tests of unit root in the debt/GDP ratio, with failure to reject the hypothesis of unit root implying that fiscal policy doesn't satisfy an intertemporal budget constraints (see e.g. Hamilton and Flavin, 1986 as a classical reference). We also note that, empirically, it has been proven difficult to accept sustainability in most countries (see references in Bohn, 1998, for the US; and see, e.g., Manasse, 1996, and the results section below for international evidence). Recently, Bohn (1998), and Sarno (2001) have adopted an econometric approach which allow for non-linearity in the adjustment process of fiscal policy; and have reach results supporting sustainability. We consider the issue of nonlinearity in some details in Section 4.3.

informal inspection of the graphs of the series (see figure 3 with stimulus data in Section 2). We rely on standard n-asymptotic tables, despite the limited time span of the data set. This reflects the unavailability of finite sample size quantiles, and it is also consistent with the inference agents could possibly perform in the experiment.

For the countries for which $r = 1$, we estimate the CI vector in the form (β'_{1x}, ρ) with $\beta'_{1x} x_t = T_t + \gamma G_t$ and impose the restriction $\mu_x = \alpha_x \cdot \rho$ to exclude linear trends in the data (see Johansen, 1995)[10]. We then test whether the parameter $\gamma = -1$.

The second empirical issue analysed within equation (6) concerns the direction of causality between taxes and public expenditure. This is also a classical theme in public finance. We recall four basic hypotheses. The first is that taxes cause spending; this has for example been advocated by the Leviathan State writers (Buchanan, 1977; Buchanan and Wagner, 1978), and as an implication of improvements in the technological capacity of raising revenues (Friedman, 1978). The opposite view that expenditure proceeds taxes is rooted in the theory of fiscal illusion, dating back to the nineteenth century 'Italian School of Public Finance' (Buchanan, 1960); it is also implied by Barro's models of exogenous public spending (Barro, 1974, 1979). Bidirectional causality may follow when taxes and expenditure are simultaneously determined according to the standard economic calculus of weighting the marginal costs and the marginal benefits of public services (Musgrave, 1966; Meltzer and Richard, 1981). Lack of causality may finally arise when taxes and public expenditure are decided upon by distinct institutional authorities (Hoover and Sheffrin, 1992).

Tests of the presence and direction of causality can be based on the field model (6). In particular, when the series in $x_t := (T_t,\, G_t)'$ are cointegrated, we know that at least one between T_t and G_t adjusts to disequilibrium with respect to the long run relation. The four cases above correspond to vector $\alpha_x = (\alpha_T : \alpha_G)$ of the forms $(0, *)'$, $(*, 0)'$, $(*, *)'$ and $(0, 0)'$ (with $*$ indicat-

[10]Because x_t contains the ratio of taxes and expenditure to total GDP, we expect the system $x_t := (T_t, G_t)'$ not to contain a linear trend.

ing a non-zero coefficient), and provide tests of Granger long-run causality. Similarly, short run-causality can be simply checked looking for significant off-diagonal coefficients in the matrices $\Gamma_{\ell,xx}$. Mixed results have been obtained regarding causality; see, e.g. recently, Garcia and Henin (1999) and Payne (1998).

Again we emphasise that in the present paper we are *not* interested in studying the *true* causality (or the true sustainability of fiscal policy) but rather the *perceived* properties of fiscal policy in the experiments.

3.2 The VAR for the lab data (expectations)

For the lab model about subjects' expectations, let m indicate the number of individuals, $i = 1, \ldots, m$, in a treatment. Recall that $y_{i,t} = (T_t^{E_i}, G_t^{E_i})'$ in the T_t and G_t treatment and $y_{i,t} = T_t^{E_i}$ in the T_t only treatment. We now illustrate the model and the inference procedure for $y_{i,t} := (T_t^{E_i}, G_t^{E_i})'$, which we then also apply with the obvious modifications to the T_t only treatment.

A suitable VAR specification for the expectations data can in particular be taken to be composed by m individual sub-systems of the following form:

$$\begin{pmatrix} \Delta T_t^{E_i} \\ \Delta G_t^{E_i} \end{pmatrix} = \mu_i + \alpha_i \begin{pmatrix} T_{t-1} + \widehat{\gamma} G_{t-1} \\ T_{t-1}^{E_i} - T_{t-1} \\ G_{t-1}^{E_i} - G_{t-1} \end{pmatrix} + \sum_{\ell=1}^{k-1} (\Gamma_{\ell,ii} : \Gamma_{\ell,ix}) \begin{pmatrix} \Delta y_{i,t-\ell} \\ \Delta x_{t-\ell} \end{pmatrix} + \widehat{\varepsilon}_{yt}, \tag{7}$$

Equation (7) includes: a vector of constant terms μ_i; a matrix (2×3) of adjustment coefficients α_i, decomposed as $\alpha_i := (\alpha_{i1} : \alpha_{i2})$, where $\alpha_{i1} = (\alpha_{iT^E} : \alpha_{iG^E})$ shows how the expectation variables may react to the disequilibrium vector estimated in the field, namely $T_{t-1} + \widehat{\gamma} G_{t-1}$ (also labeled ${}_{\text{ECM}}TG$); while $\alpha_{i2} = ((\alpha_{iT^E(T^ET)}, \alpha_{iT^E(G^EG)}) : (\alpha_{iG^E(T^ET)}, \alpha_{iG^E(G^EG)}))$ includes the responses to two further possible CI relations given by the errors in expectations $(T_{t-1}^{E_i} - T_{t-1}, G_{t-1}^{E_i} - G_{t-1})'$[11]; matrices $(\Gamma_{\ell,ii} : \Gamma_{\ell,ix})$ of partial adjustments coefficients to, respectively, past fiscal expectations $\Delta y_{i,t-\ell}$ and past fiscal re-

[11]System (7) can in particular be obtained as a restricted specification of the more

alizations $\Delta x_{t-\ell}$; a vector of error terms $\widehat{\varepsilon}_{yt} := \varepsilon_{yt} - \alpha_{i2}(\widehat{\gamma} - \gamma)G_{t-1}$, with ϵ_{yt} i.i.d. $N(0, \Omega_{yy})$[12].

The analysis of eq. (7) permits to discriminate among different expectation formation processes. In particular, if subject i has *rational expectations*, one would expect the equations for $\Delta y_{i,t}$ to collapse to the specification of the marginal field system Δx_t. This can be checked very simply by testing if the coefficients of the variables that are present in (7) and not in (6) are equal to zero.

Of specific interest will be to investigate whether and how the long-run adjustment in the lab $\alpha_{i1} = (\alpha_{iT^E} : \alpha_{iG^E})$ relate to the one in the field model $\alpha_x = (\alpha_T : \alpha_G)$, since this coefficient indicates if agents perceive and adjust to the actual cointegration characteristics in the data.

One alternative class of expectation formation processes is the adaptive

---

general system:

$$\Delta y_{i,t} = \mu_i + \alpha_{i1}(\beta_{1x}' x_{t-1}) + \alpha_{i2}\beta_{2yx_i}' \begin{pmatrix} y_{i,t-1} \\ x_{t-1} \end{pmatrix} + \sum_{\ell=1}^{k-1} (\Gamma_{\ell,ii} : \Gamma_{\ell,ix}) \begin{pmatrix} \Delta y_{i,t-\ell} \\ \Delta x_{t-\ell} \end{pmatrix} + \epsilon_{yt}, \quad (8)$$

including the CI relation $\beta_{1x}' x_{t-1}$ from the field and up to 2 further free CI relations through term $\alpha_{i2}\beta_{2yx_i}'$. Inference on the number of additional CI vectors can in principle be performed for fixed values of $\beta_{1x}' x_t = T_t + \gamma G_t$, using the ML estimate $\widehat{\gamma}$ of γ taken from the field system. (Note that the generated regressor bias has no effect on the n-asymptotics for the lab system, because $\widehat{\gamma}$ is superconsistent, $\widehat{\gamma} - \gamma = O_p(n^{-1})$ compared to the $n^{1/2}$ consistency of the parameters of stationary variables). The hypothesis of adaptive behaviour in the formation of expectations has suggested that we could take in this paper a simpler approach and calculate the expectation errors $(T_t^{E_i} - T_t, G_t^{E_i} - G_t)'$ as a possible choice of extra CI relations $\beta_{2yx_i}'(y_{i,t-1}', x_{t-1}')'$. Because these extra relations do not contain any parameter to be estimated, we have inspected directly the time-series of the forecast errors in order to infer if they are stationary or I(1). We have done this visually and through univariate unit root tests. Both tests (not reported for brevity) confirm that the forecast errors are stationary, which has led us to conclude for the system (7) estimated in the text.

[12]Notice also that given individual forecasts are performed independently of each other, both the matrices $\Gamma_{\ell,ii}$ and Ω_{yy} are block-diagonal, namely $\Gamma_{\ell,ii} = \text{diag}(\Gamma_{\ell,11}, \ldots, \Gamma_{\ell,mm})$ for all $\ell = 1, .., k-1$, and $\Omega_{yy} = \text{diag}(\Omega_{11}, \ldots, \Omega_{mm})$.

scheme, as originated in the 1950s by the works of different scholars (see Pesaran, 1987, for references). We emphasise the bivariate nature which can take here the adaptive schemes.

In particular, under *univariate* or *purely adaptive* expectations, one would expect that only past values of the forecasted variable x_t^1 and of its forecast $y_{i,t}^1$ enter in the expectation process for $y_{i,t}^1$, where a superscript 1 indicates one stimulus variable and the corresponding forecast. In the present experiment, if other variables enter in the estimated equation for $\Delta y_{i,t}^1$, this is evidence against a purely adaptive scheme and in favour of a more general class of models known as *augmented-adaptive* (see again Pesaran, 1987, for the many variations nested within this definition).

In most of the earlier experiments studying expectation formation, subjects had to forecast a univariate time series (as for example in Schmalensee, 1977 or Hey, 1994). They generally found support for adaptive expectations (though not necessarily of the first order, e.g Hey, 1994). Also in more recent experiments we don't know of explicit tests of augmented-adaptive schemes, though there is a growing experimental literature testing reduced-univariate schemes arising from augmented-adaptive forms. That is for example the case of investigations of price expectations in the classical cobweb economy (see, e.g. Hommes, Sonnemans, Tuinistra, and van de Velden, 2003, as recent example), or in experiments on inflation expectations in overlapping generation economies (as in Marimon and Sunder, 1993 or Bernasconi and Kirchkamp, 2000). The question of specific interest in the latter wave of studies is whether reduced form augmented-adaptive schemes, sometimes also referred to as 'boundedly rational' learning models (see Sargent, 1993 and Evans and Honkapohja, 2001), can bring convergence to rational expectations equilibria. In the partial equilibrium approach of the present experiment, the question of convergence is of a lesser relevance, though it is interesting to verify whether augmented-adaptive schemes are robust outside the reduced-univariate specifications.

In augmented-adaptive schemes one can also address questions of perceived causality between taxes and public expenditure. In particular, by inspection of the parameters in (7) one should note the following: the α_{i1}

coefficients and the off-diagonal elements in the $\Gamma_{\ell,ix}$ matrices determine the direction of causality from the field to the lab, while the α_{i2} coefficients and the off-diagonal elements in the $\Gamma_{\ell,ii}$ matrices regulate the ones from the past expectations on present expectations.

The individual lab sub-systems (7) may be estimated one at the time or jointly. Joint estimation under some homogeneity restrictions allows to exploit the panel dimension of the data to increase efficiency. In the empirical analysis we will first assume all individual-specific parameters to be equal across agents $(\mu_i, \alpha_i, \Gamma_{\ell,ii}, \Gamma_{\ell,ix}, \Omega_{ii}) = (\mu_*, \alpha_*, \Gamma_{\ell,*}, \Gamma_{\ell,*x}, \Omega_{**})$, obtaining the maximal reduction in the number of parameters. We estimate for each country a model of a representative-agent's expectations, which we compare with the corresponding country field model. This comparison will primarily focus on the expectations obtained from the experiment conducted under the fiscal frame. Next we will compare results from different treatments, including experiments conducted under the neutral frame. There we will also reconsider the homogeneity condition imposed through the representative agent assumption and will perform alternative estimations at more disaggregated levels.

4 Empirical evidence

4.1 Analysis on the level of countries

Cointegration in the field: The result of the cointegration analysis of the stimulus data is given in table 3. We find that two lags (that is, models with one lagged difference, the VAR order is $k = 2$) are enough to characterise the dynamic structure of the series for most of the countries considered in the experiments. For 9 countries (Austria, Finland, Germany, Italy, Netherlands, Norway, Portugal, Sweden and UK), we find that taxes and public expenditures are cointegrated; for 6 (Belgium, Denmark, France, Greece, Ireland and Spain) we find that they are not. Among the former and consistently with the general evidence reported in the literature (see, e.g., Manasse, 1996), we found that the condition for stationarity of the budget $\widehat{\gamma} = -1$ is rejected for

Table 3: Results of cointegration analysis on stimulus data

Country (sample period)	VAR order	$H_0(r=0)$ *versus* $H_1(r=1)$	Rank	Cointegrating Vector (β'_{1x}, ρ)	Test of the homogeneity condition $\widehat{\gamma}=-1$
Austria	$k=3$	42.5^{***}	$r=1$	$(1;-0.739;-10.385)$	-13.23^{***}
Belgium	$k=2$	9.05	$r=0$		
Denmark	$k=2$	15.92	$r=0$		
Finland	$k=5$	27.99^{***}	$r=1$	$(1;-0.565;-21.436)$	-28.29^{***}
France	$k=2$	15.92	$r=0$		
Germany	$k=2$	19.40^{*}	$r=1$	$(1;-0.572;-17.863)$	-4.32^{**}
Greece	$k=2$	14.37	$r=0$		
Ireland	$k=2$	16.67	$r=0$		
Italy	$k=4$	18.19^{*}	$r=1$	$(1;-0.892;0)$	-1.72
Netherlands	$k=2$	20.47^{**}	$r=1$	$(1;-0.606;-17.630)$	-13.51^{***}
Norway	$k=2$	18.42^{*}	$r=1$	$(1;-1.051;0)$	-3.52^{**}
Portugal	$k=3$	49.12^{***}	$r=1$	$(1;-1.177;13.667)$	-2.09^{*}
Spain	$k=2$	16.10	$r=0$		
Sweden	$k=2$	18.18^{*}	$r=1$	$(1:-0.950;0)$	-2.13^{*}
UK	$k=2$	23.42^{**}	$r=1$	$(1;-0.922;0)$	-9.18^{***}

Legend: *, **, *** denote rejection at, in the order, 10%, 5%, 1% significance level.

most countries. In fact, it is accepted only for Italy, somehow surprisingly given that Italy is notoriously considered a country with very easy public spending[13].

Causality in the field and in the lab: Table 4 summarises the main findings of the inference on both the field and the lab systems in the two experimental treatments under the fiscal frame. The complete parameter estimates, together with standard errors and level of statistical significance are available in Bernasconi, Kirchkamp, and Paruolo (2004), other evidence on the estimates will be provided below.

Parameters reported in table 4 are significant at a 5% level at least[14]. In considering the results, recall that in the T_t only treatment, in which agents forecast taxes, expectations on $G_t^{E_i}$ are not available. Thus, inference results on the lab system for the T_t only treatment is limited to the equation for $\Delta y_t^1 = \Delta T_{t+1}^{Ei}$.

The leftmost part of the table reports, for the field and for the lab, the coefficients of responses to the field error correction term ${}_{\text{ECM}}TG = T_t + \widehat{\gamma} G_t$. Consider first the field evidence (the first column in the table). Among the nine countries for which cointegration between taxes and expenditure was found, expenditure is long-run adjusting to taxes in four cases (Austria, Norway, Sweden and UK); taxes are adjusting to expenditure in four others (Finland, Germany, Italy and Netherlands); and in one country (Portugal) there is bidirectional adjustment.

Results from the experiments show that subjects fail completely to perceive the CI characteristics of the field data. In both experimental treatments the lab responses α_{i1} to ${}_{\text{ECM}}TG$ are often not significantly different from zero and even when they are, α_{i1} are in any case quite small[15].

---

[13]Notice, however, that the hypothesis of cointegration for Italy was accepted only marginally with a p-value of 0.09.

[14]All models presented below have been selected performing a computerised strategy, which started to drop coefficients with higher p-values. We have checked that the selected models are robust to alternative procedures.

[15]In greater details, in the T_t and G_t treatment, in which subjects forecasted both taxes

	Vectors of responses to $_{\text{ECM}}TG$: $T_t+\widehat{\gamma}G_t$			Vectors of responses to $(_{\text{ECM}}T^ET,\ _{\text{ECM}}G^EG)'$: $(T_t^{E_i}-T_t, G_t^{E_i}-G_t)'$		Direction of short run causality inferred from $\Gamma_{l,xx}$ and $\Gamma_{\ell,ix}$		
	Field α_x	T_t and G_t treatment α_{i1}	T_t only treatment α_{i1}	T_t and G_t treatment α_{i2}	T_t only treatment α_{i2}	Field	T_t and G_t treatment	T_t only treatment
Austria	(0,1.308)	(0,0)	(0,*)	((-0.921,0), (0.084,-0.695))	((-0.863,*) ,*)	$T \leftrightarrow G$		$T \leftarrow G$
Belgium				((-0.838,0), (0,-0.656))	((-0.810,*) ,*)	$T \leftarrow G$	$T \leftrightarrow G$	$T \leftarrow G$
Denmark				((-0.885,0.138), (0.096,-0.622))	((-0.872,*) ,*)		$T \leftrightarrow G$	$T \leftarrow G$
Finland	(-0.953,0)	(-0.145,0)	(-0.164,*)	((-0.923,0.171), (0,-0.512))	((-0.825,*) ,*)	$T \leftarrow G$	$T \leftrightarrow G$	$T \leftarrow G$
France				((-0.709,-0.124), (0,-0.849))	((-0.655,*) ,*)	$T \leftarrow G$	$T \leftrightarrow G$	$T \leftarrow G$
Germany	(-1.007,0)	(0,0)	(0.165,*)	((-0.842,0), (0,-0.695))	((-0.961,*) ,*)		$T \rightarrow G$	$T \leftarrow G$
Greece				((-0.915,0), (0,-0.717))	((-1.058,*) ,*)		$T \leftarrow G$	$T \leftarrow G$
Ireland				((-0.739,0), (-0.141,-0.799))	((-0.676,*) ,*)		$T \rightarrow G$	$T \leftarrow G$
Italy	(-0.155,0)	(-0.072,-0.084)	(-0.034,*)	((-0.771,0.074), (0,-0.744))	((-0.864,*) ,*)		$T \rightarrow G$	
Netherlands	(-0.623,0)	(0,0)	(0,*)	((-0.736,0), (0,-0.800))	((-0.809,*) ,*)		$T \leftarrow G$	$T \leftarrow G$
Norway	(0,0.308)	(0,0.083)	(0,*)	((-0.732,0), (0.166,-0.751))	((-0.731,*) ,*)	$T \rightarrow G$	$T \rightarrow G$	$T \leftarrow G$
Portugal	(0.290,0.670)	(0,0.175)	(0,*)	((-0.744,0), (0.240,-0.633))	((-0.758,*) ,*)	$T \leftrightarrow G$	$T \leftrightarrow G$	$T \leftarrow G$
Spain				((-0.732,0), (0,-0.657))	((-0.548,*) ,*)		$T \rightarrow G$	$T \leftarrow G$
Sweden	(0,0.280)	(-0.034,0.086)	(0,*)	((-0.692,0), (0.111,-0.743))	((-0.670,*) ,*)	$T \rightarrow G$	$T \leftarrow G$	$T \leftarrow G$
United-Kingdom	(0,0.435)	(0,0)	(0,*)	((-0.883,0), (0.175,-0.780))	((-0.675,*) ,*)		$T \rightarrow G$	$T \leftarrow G$

TABLE 4: Summary of inference results

In the rightmost part of table 4 we show the results of tests on Granger causality in the field as it can be inferred from inspection of the off-diagonal coefficients of the matrix $\Gamma_{l,xx}$ from equation (6), and compare them with tests of perceived causality in the lab as inferred from the off-diagonal coefficients of matrix $\Gamma_{\ell,ix}$.

Here also results speak strongly against subjects correctly perceiving the properties of field stimuli data. Field evidence in particular shows that in eight countries, causality runs in neither direction[16], in three countries causality runs from expenditure to taxes; in two countries it runs from taxes to expenditure[17]; and in other two causality is bidirectional[18].

Lab results are not consistent with the directions of causality in the field. In the T_t and G_t treatment only in one case (Portugal) the causality matches the causality from the field data[19]. In the T_t only treatment the causality from the field is consistent with the one in the lab in six cases[20] and not consistent in the remaining nine.

---

and expenditure, we find that in 4 countries (Austria, Germany, Netherlands, UK) subjects fail to perceive any long run adjustment; in one country (Finland), they perceive some low adjustment of taxes to expenditure, in two (Norway and Portugal) of expenditure to taxes, and in other two (Italy and Sweden) a bi-directional long run adjustment. In the T_t only treatment, we don't observe how participants perceive the dynamics of the series of public expenditure; on the other hand, we see that out of the six countries in which adjustment in the field was from expenditure to taxes, only in the case of Finland and Italy participants seem to perceive some small adjustment in the right direction, while for Germany the coefficient of the long run adjustment has the wrong sign.

[16]The countries were causality in the field runs in neither direction are Denmark, Germany, Greece, Ireland, Italy, Netherlands, Spain, UK.

[17]These countries are Norway and Sweden.

[18]Bidirectional causality is found in Austria and Portugal in the field data.

[19]For one country (Austria), subjects don't perceive any causality in the T_t and G_t treatment while we estimate a bidirectional causality in the field. For the remaining 14 countries we estimate a causality with the lab data which is matched by the field causality only in the case of Portugal.

[20]These countries are Austria, Belgium, Finland, France, Italy, Portugal.

Expectation schemes: The middle part of table 4 shows the estimates of the vector α_{i2} describing how individuals react in the long run to errors in the process of expectations forming $(T_t^{E_i} - T_{t-1}, G_t^{E_i} - G_{t-1})'$ (also denoted $({}_{\text{ECM}}T^ET, {}_{\text{ECM}}G^EG)'$). The results support the importance of adaptive expectations. In the T_t and G_t treatment we find the coefficients for ${}_{\text{ECM}}T^ET$ in the equation for $\Delta T_t^{E_i}$, and the coefficients on ${}_{\text{ECM}}G^EG$ in the equation for $\Delta G_t^{E_i}$ to be positive and close to 1. The same holds in the T_t only treatment for the coefficients for ${}_{\text{ECM}}T^ET$ in the equation for $\Delta T_t^{E_i}$[21]. We also see that in both experimental treatments the estimated parameters are typically less than 1. (This is also confirmed by formal t-tests conducted on the parameters). Conversely, we should note that in the T_t and G_t treatment, the off-diagonal diagonal of α_{i2} are in most cases equal to or very close to zero. In other words, subjects do not adjust expectations for taxes to errors in the expectations on expenditure, and vice versa.

Taken as a whole, the aggregate evidence from table 4 thus indicates that subjects are far from a model of rational expectations; their behaviour seems rather to follow adaptive expectations, but possibly not quite a purely univariate adaptive model. While the results suggest that participants assume some causality, this causality changes over countries. We take this to be some evidence for model of augmented-adaptive expectations.

4.2 Analysis on the level of individuals

The results discussed in section 4.1 have been estimated assuming all individual-specific parameters to be equal across agents within each country. This restriction may be natural when all individual behaviour satisfies rational expectations. In the discussion of the estimates on the level of countries we have seen that subjects in aggregate violate rational expectations. Still, there may be some subjects behaving differently or more rational than others. We have therefore also estimated equation (7) separately for each individual in each country.

[21] In the T_t only treatment subjects do not form expectations on G, hence, we can not measure the other coefficients.

To obtain further evidence on the issue, in figure 4 we compare the distributions of the coefficients for the two fiscal treatments summarised in table 4 with similar aggregate estimates obtained from the experiments conducted under the neutral frame. The left column of diagrams in figure 4 shows estimated coefficients for $\Delta T_t^{E_i}$, the right column of diagrams shows the coefficients for $\Delta G_t^{E_i}$. We see some difference among treatments, but also a good deal of similarity.

Coefficients of responses to the field error correction term ${}_{\mathrm{ECM}}TG$ are zero or close to zero for all treatments and for both equations for $\Delta T_t^{E_i}$ and $\Delta G_t^{E_i}$.

The distributions of the adjustment coefficients to the error terms $({}_{\mathrm{ECM}}T^ET, {}_{\mathrm{ECM}}G^EG)'$ confirm that in all treatments subjects follow in the aggregate a model of adaptive expectation, with on-diagonal coefficients close to but generally lower than one, while off-diagonal coefficients are basically zero. That is, subjects don't adjust adaptively across variables.

The next graphs of the figure report the distribution of the coefficients of matrices $\Gamma_{\ell,ix}$ summed over lags. In both equations, the diagonal coefficients (hence the own-effects of $\Delta T_{t-\ell}$ on $\Delta T_t^{E_i}$ and of $\Delta G_{t-\ell}$ on $\Delta G_t^{E_i}$) are larger than the off-diagonal terms, which corroborates the view that perception of Granger causality in the T_t and G_t treatment is neither systematic nor strong. A noteworthy exception is perceived causality from expenditure to taxes in T_t only treatment which is often significant and positive, which we view as support to include short run causalities for this treatment in table 4.

The last two diagrams show the coefficients from matrix $\Gamma_{\ell,ii}$, summed-up over lags, which refer to the short run effects of past expectations on present expectations. The results show a moderate tendency of subjects smoothing out peaks in past expectations, as these (diagonal) effects are sometimes negative, but generally close to zero.

Let us look more closely that the regression coefficients for ${}_{\mathrm{ECM}}G^EG$, ${}_{\mathrm{ECM}}T^ET$ and $\Gamma_{\ell,ix}$ which are not zero or near zero. To visualise heterogeneity among individuals we divide individuals into three groups with the help of a kmeans cluster analysis based on these regression coefficients[22]. Figure

[22]The cluster analysis is here just a convenient way to identify similar groups of indi-

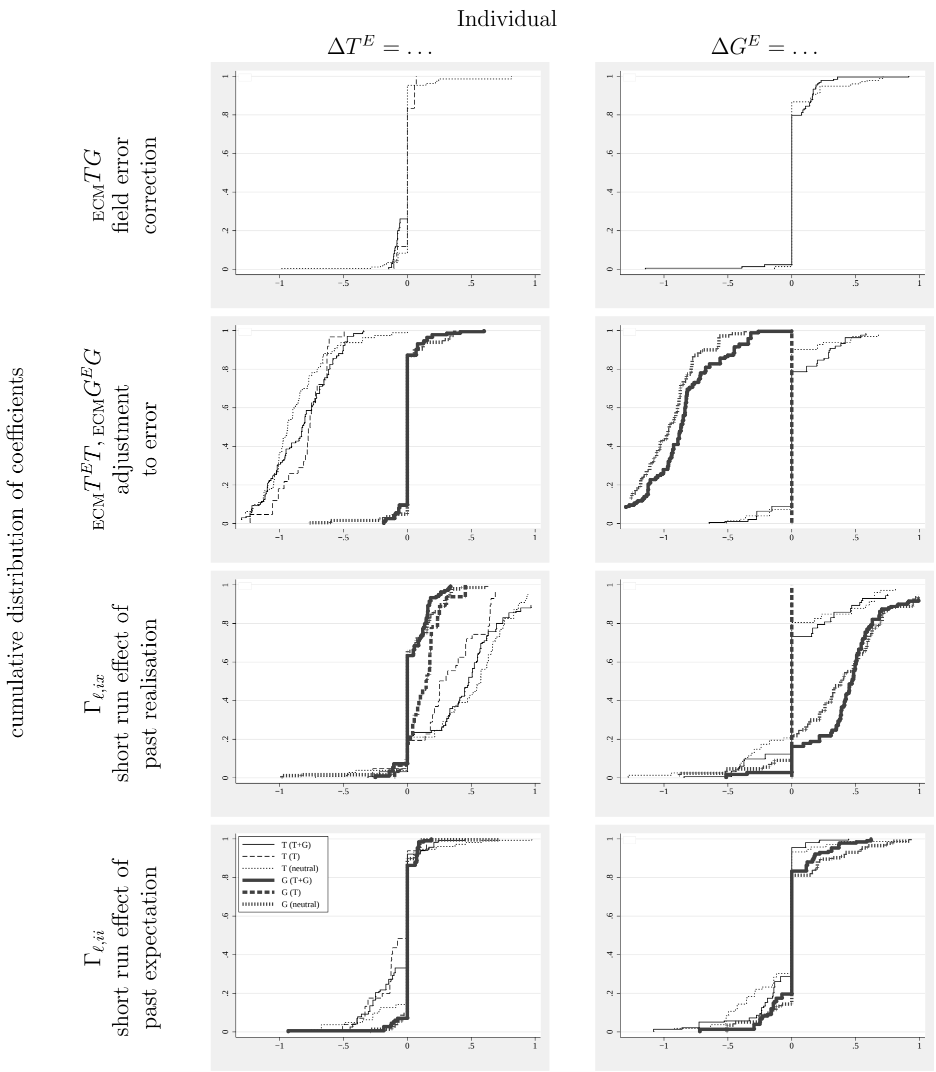


FIGURE 4: Results of estimating equation (7) for each individual

Graphs show the distribution of the estimated coefficients of equation (7). We impose the restriction that each individual uses the same model for all countries. The distribution in the $T + G$ treatment is always shown as a solid line, the distribution in the T treatment as a dashed line, and the distribution in the neutral treatment is shown as a dotted line. Distributions of coefficients for ${}_{\text{ECM}}T$, ΔT, and ΔT^E are shown as thin lines, distributions of coefficients for ${}_{\text{ECM}}G$, ΔG, and ΔG^E are shown as thick lines.

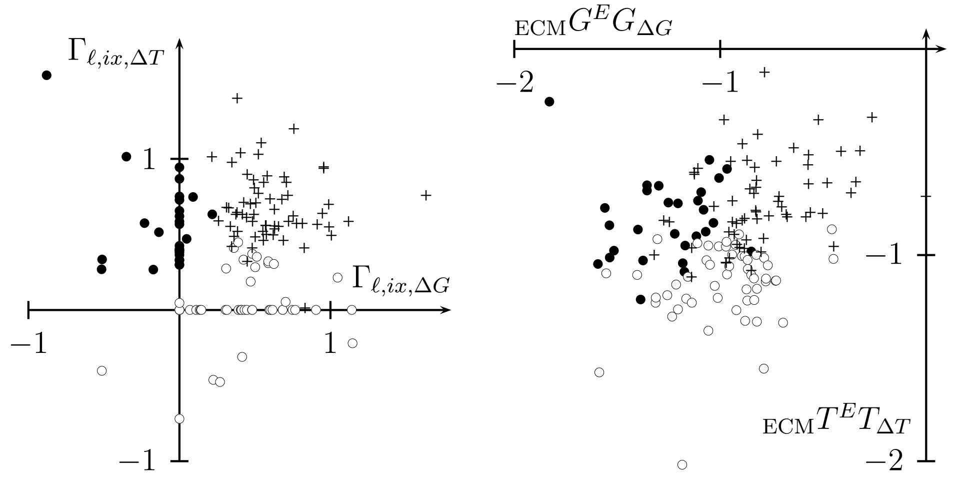


FIGURE 5: Scattergram of $\Gamma_{\ell,ix}$ and $_{\text{ECM}}G^EG$ coefficients from equation (7) for each individual

In this diagram we impose the restriction that each individual uses the same model for all countries. Individuals from the T_t and G_t treatment and the neutral treatment are pooled. Similar symbols denote individuals from similar clusters from the cluster analysis in both diagrams.

5 shows a scattergram of the coefficients where the data from T_t and G_t treatment and the neutral treatment is pooled for simplicity. We see that individuals, indeed, specialise. In the graph on the left we find several participants with a zero (or close to zero) coefficient for $\Gamma_{\ell,ix}$ either in the ΔT or in the ΔG equation. These participants are identified with • and ◦. In the graph on the right we see that these participants choose a comparatively larger error correction term for T and G respectively. Participants with positive $\Gamma_{\ell,iG}$ and $\Gamma_{\ell,iT}$ (which are marked with a + in both graphs) respond less to error correction.

4.3 Misspecification tests for nonlinearity and anti-Keynesian effects

Overall, the results of the previous sections are rather disappointing in regard to the view that subjects can correctly perceive the properties of the fiscal time series from the field. In fact, even if some differences across the various

viduals in the different diagrams.

countries' fiscal policy can be spotted visually (for example, going back to the stimulus data in figure 3, it is easy to see how countries like Italy and Greece have persistent budget deficits relative to the records of countries such as UK, Germany or Sweden), we have seen no systematic difference in the way in which subjects have treated the various countries in the experiment[23].

This may due to various reasons. In particular the VAR approach pursued in the previous sections works under the maintained hypothesis of a linear process for the time series considered. As intuition and some current literature has documented (see Bohn, 1998, and Sarno, 2001), this hypothesis may conflict with the fact that fiscal policy is subject to various possible structural shifts and discretionary interventions, which may introduce nonlinearity in fiscal policy.

We have controlled for misspecification biases due to nonlinearity in the field systems by way of standard RESET tests. We haven't found evidence of misspecification for the spans of data given as stimulus to subjects in the experiment.

When the field systems are correctly specified, still the conduct of fiscal policy might change at some point in time. As emphasised by some recent literature, a large intervention on the fiscal variables, perhaps addressed to correct disequilibrium in the public budget, may be perceived by the public as to imply lower taxation in the future; and, hence, also generate an expansion in economic activity, rather than a contraction as predicted by a standard Keynesian perspective.

Proposed originally by the Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung 1981, this view has been made popular by Giavazzi and Pagano (1990), who brought to the attention of the profession the astonishing expansionary fiscal consolidations occurred in the mid 1980s in Denmark and Ireland. Various subsequent literature has been developed on the circumstances and conditions under which expectations may react nonlinearly to change in fiscal policy (e.g. Bertola and Drazen, 1993, and Sutherland, 1997).

[23]This is also confirmed by the size of the parameters estimated for the various countries.

Taking a pragmatic approach, Giavazzi, Jappelli, and Pagano (2000) have defined periods of 'large and persistent' fiscal contractions and fiscal expansions as situations in which, for at least two consecutive years, the budget balance as a percentage of GDP increase or decrease, respectively, by at least 1.5 point per year.

We control for the possibility of nonlinear effects in the present experiment introducing dummies in the equation (7) for expected taxes. Two dummies are introduced on the intercept: one (d_{CONTR}) activated in periods of "sizeable and persistent" fiscal contractions, and the other (d_{EXP}) activated in periods of "sizeable and persistent" fiscal expansions. According to the view, we should expect the coefficients on d_{CONTR} to be negative and that on d_{EXP} to be positive[24].

Table 5 reports the periods in which the two dummies were activated, with the related evidence. We counted 11 episodes of fiscal expansions (in 10 countries, with Sweden counting for two), and 10 episodes of fiscal contractions (in 7 countries, with Ireland counting for two and Sweden for three). In the majority of cases the dummies are not significant, in some the sign of the dummies are opposite to that predicted by theory. The evidence rejects the hypothesis of nonlinear responses of subjects' expectations in the experiment[25].

5 Concluding discussion

Laboratory experiments are often inspired by the real world, and often use in their setup parameters which are based on field data. Recent examples from macroeconomic experiments include tax compliance experi-

[24]This would in particular signal a negative and positive change in the drift of agents' expectations following periods of "large and persistent" fiscal contraction and fiscal expansion, respectively.

[25]We conducted tests even with different measures of changes in the budget-GDP ratio to define 'large and persistent' fiscal contractions, or using other notions of intervention inducing nonlinear effects (as based on Bertola and Drazen, 1993, or on Sutherland, 1997). In all cases we found similarly negative evidence.

	Vectors of responses to $_{\text{ECM}}TG$: $T_t+\widehat{\gamma}G_t$		Vectors of responses to $(_{\text{ECM}}T^ET,\ _{\text{ECM}}G^EG)'$: $\beta'_{w1}\cdot(y_{t-1},\ x_{t-1})'$		Direction of short run causality inferred from $\Gamma_{l,xx}$ and $\Gamma_{\ell,ix}$		Fiscal Expansions coefficients of d_{EXP}			Fiscal Contractions coefficients of d_{CONTR}		
	T_t and G_t treatment α_{i2}	T_t only treatment α_{i2}	T_t and G_t treatment α_{i1}	T_t only treatment α_{i1}	T_t and G_t treatment	T_t only treatment	Episodes	T_t and G_t treatment	T_t only treatment	Episodes	T_t and G_t treatment	T_t only treatment
Austria	(0,0)	(0,*)	((-0.921,0), (0.084,-0.695))	((-0.863,*) ,*)		$T \leftarrow G$						
Belgium			((-0.838,0), (0,-0.656))	((-0.780,*) ,*)	$T \leftrightarrow G$	$T \leftarrow G$	80, 81	0	-0.303			
Denmark			((-0.885,0.138), (0.096,-0.622))	((-0.872,*) ,*)	$T \leftrightarrow G$	$T \leftarrow G$	74, 75, 80, 81, 82	0	0	83, 84, 85, 86	0	0
Finland	(-0.238,0)	(-0.164,*)	((-0.932,0.210), (0,-0.512))	((-0.825,*) ,*)	$T \leftrightarrow G$	$T \leftarrow G$	91, 92, 93	0.623	0	75, 76, 88, 89	0	0
France			((-0.712,0), (0,-0.849))	((-0.655,*) ,*)	$T \leftrightarrow G$	$T \leftarrow G$	92, 93	0.466	0			
Germany	(0,0)	(0.165,*)	((-0.842,0), (0,-0.695))	((-0.961,*) ,*)	$T \rightarrow G$	$T \leftarrow G$	74, 75	0	0			
Greece			((-0.915,-0.067), (0,-0.717))	((-1.058,*) ,*)		$T \leftarrow G$	88, 89, 90	0	0	96, 97	0	0
Ireland			((-0.739,0), (-0.141,-0.799))	((-0.676,*) ,*)	$T \rightarrow G$	$T \leftarrow G$	74, 75, 78, 79, 80	0	0	83, 84, 87, 88	0	0
Italy	(-0.072,-0.084)	(-0.034,*)	((-0.771,0.074), (0,-0.744))	((-0.864,*) ,*)	$T \rightarrow G$		71, 72	0	0			
Netherlands	(0,0)	(0,*)	((-0.736,0), (0,-0.800))	((-0.809,*) ,*)	$T \leftarrow G$	$T \leftarrow G$						
Norway	(0,0.083)	(0,*)	((-0.736,0), (0.166,-0.751))	((-0.731,*) ,*)	$T \rightarrow G$	$T \leftarrow G$	91, 92	0.324	0	94, 95, 96	0	0
Portugal	(0,0.175)	(0,*)	((-0.744,0), (0.240,-0.633))	((-0.758,*) ,*)	$T \leftrightarrow G$	$T \leftarrow G$	74, 75, 76	0	0			
Spain			((-0.712,0), (0,-0.657))	((-0.548,*) ,*)	$T \leftrightarrow G$	$T \leftarrow G$	81, 82	0	0	96, 97	0.312	0
Sweden	(-0.034,0.086)	(0,*)	((-0.715,0), (0.111,-0.743))	((-0.673,*) ,*)		$T \leftarrow G$	78, 79, 91, 92, 93	0.575	0	83, 84, 86, 87, 94, 95, 96	0	0.238
United-Kingdom	(0,0)	(0,*)	((-0.883,0), (0.175,-0.780))	((-0.897,*) ,*)	$T \rightarrow G$		72, 73, 92, 93	0	0.791			

Table 5: Evidence on expectations in episodes of "large and persistent" fiscal adjustments

ments (see, e.g., Alm and McKee, 2004) and decisions over monetary policy (Lombardelli, Talbot, and Proudman, 2002; Blinder and Morgan, 2005). In this paper we are going a step further. We are not only following parameters from the real world, from a methodological perspective, the main novelty of the present approach is the idea of using entire datasets from field and laboratory data complementarily. Experimental economics has grown substantially over the last two or three decades, as it is now a well-acknowledged method through which decision theorists, game theorists and microeconomists have tested and refined theoretical models in their respective fields of interest.

Relatively few experiments have been conducted in the field of macroeconomics. The reason, probably, is that macroeconomists deal with real world questions to a much greater degree than other economists, in the conviction that laboratory experiments cannot really answer such type of questions. Using parameters from the real world in the experiment is one elegant way to answer this question. The approach pursued in this paper is different. We follow the idea that using real world data as stimulus for subjects in the experiments helps us understanding which variables and forces are essential when individuals take actions of macroeconomic relevance.

The paper takes a first step in studying the reliability of the hypothesis of rational expectations and the process people use to form expectations on fiscal variables more generally. We have conducted an experiment in which participants receive as stimuli only real world data on fiscal variables and where they form expectations on the basis of that information alone. This simple structure replicates the tradition of models on formation of expectations. We have followed an econometric approach for the process of formation of expectations, which coherently arises from the data generating process of the field stimuli; we have distinguished between long-run and short-run effects, both of the stimuli and of past expectations.

We found that subjects behave adaptively, though they do not adjust perfectly to past expectation errors, not even in the long run. We found that subjects follow an augmented-adaptive model, which hasn't however revealed a general pattern of behaviour regarding causal relationships between fiscal variables. If at all, we can identify a weak causality from expenditure to taxes.

A possible next step would be to include further factors, like political events, announcements, etc., in a richer experimental setup, which gives information on political events or announcements from historical records, as additional stimulus data.

References

Ahmed, S., and J. H. Rogers, 1995, Government Budget Deficits and Trade Deficits: Are Present Value Constraints Satisfied in Long-Term Data?, *Journal of Monetary Economics*, 36, 351–74.

Allers, M., J. de Haan, and F. de Kam, 1998, Using Survey Data to Test for Ricardian Equivalence, *Public Finance Review*, 26, 565–82.

Alm, J., and M. McKee, 2004, Tax Compliance as a Coordination Game, *Journal of Economic Behavior and Organization*, 54(3), 297–312.

Barro, R. J., 1974, Are Government Bonds Net Wealth?, *Journal of Political Economy*, 82, 1095–1117.

——— , 1979, On the Determination of the Public Debt, *Journal of Political Economy*, 87, 940–71.

Bernasconi, M., and O. Kirchkamp, 2000, Why do monetary policies matter? An experimental study of saving and inflation in an overlapping generations model, *Journal of Monetary Economics*, 46, 315–343.

Bernasconi, M., O. Kirchkamp, and P. Paruolo, 2004, Do fiscal variables affect fiscal expectations? Experiments with real world and lab data, Discussion Paper 04-27, University of Mannheim, SFB 504.

Bertola, G., and A. Drazen, 1993, Trigger Points and Budget Cuts: Explaining the Effects of Fiscal Austerity, *The Americal Economic Review*, 83(1), 11–26.

Blinder, A. S., and J. Morgan, 2005, Are two heads better than one? Monetary policy by committee, *Journal of Money, Credit and Banking*, 37(5), 789–811.

Bohn, H., 1998, The behavior of US public debt and deficits, *Quarterly Journal of Economics*, 113, 949–963.

Buchanan, J. M., 1960, *Fiscal theory and political economy*. The University of North Carolina Press, Chaple Hill.

——— , 1977, *Democracy in deficit*. Academic Press, New York.

Buchanan, J. M., and R. Wagner, 1978, Dialogues concerning fiscal religion, *Journal of Monetary Economics*, pp. 626–627.

Costa-Gomes, M., and V. P. Crawford, 2001, Cognition and Behavior in Normal-Form Games: An Experimental Study, *Econometrica,*, 69(5), 1193–1235.

Croson, R., 2000, Thinking like a game theorist: factors affecting the frequency of equilibrium play, *Journal of Economic Behavior and Organization*, 41(3), 299–314.

Evans, G. W., and S. Honkapohja, 2001, *Learning and expectations in macroeconomics*. Frontiers of Economic Research. Princeton and Oxford: Princeton University Press.

Fisher, F. M., 1962, *A Priori Information and Time Series Analysis*. North-Holldan, Amsterdam.

Friedman, M., 1978, The limitations of tax limitations, *Policy Review*, pp. 7–14.

Garcia, S., and P.-Y. Henin, 1999, Balancing Budget through Tax Increases or Expenditure Cuts: Is It Neutral?, *Economic Modelling*, 16, 591–612.

Giavazzi, F., T. Jappelli, and M. Pagano, 2000, Searching for non-linear effects of fiscal policy: evidence from industrial and developing countries, *European Economic Review*, 44, 1259–1289.

Giavazzi, F., and M. Pagano, 1990, Can severe fiscal contraction be expansionary? Tales of two small European countries, *NBER Macroeconomics Annual*, pp. 75–111.

Grun, D., 1991, What people know and what economists think they know: surveys on Ricardian equivalence, *Australian Economic Papers*, 30, 1–9.

Hakio, C., and M. Rush, 1991, Is the budget deficit too large?, *Economic Enquiry*, 29, 449–445.

Hamilton, J., and M. Flavin, 1986, On the Limitations of Government Borrowing: a Framework for Empirical Testing, *American Economic Review*, 76, 808–819.

Hey, J. D., 1994, Expectations formations: rational or adaptive or ..., *Journal of Economic Behavior and Organization*, 25, 329–349.

Hommes, C., J. Sonnemans, J. Tuinistra, and H. van de Velden, 2003, Learning in Cobweb Experiments, Discussion Paper 020/1, Timbergen Institute.

Hoover, K. D., and S. M. Sheffrin, 1992, Causation, Spending, and Taxes: Sand in the Sandbox or Tax Collector for the Welfare State?, *American Economic Review*, 82, 225–48.

Johansen, S., 1995, *Likelihood-based inference in cointegrated vector autoregressive models.* Advanced Texts in Econometrics. Oxford and New York: Oxford University Press.

Lombardelli, C., J. Talbot, and J. Proudman, 2002, Committees versus Individuals: An Experimental Analysis of Monetary Policy Decision-Making, *Bank of England Quarterly Bulletin*, 42(3), 262–273.

Manasse, P., 1996, Are Taxes Too Low?, *Journal of Economic Dynamics and Control*, 20, 1263–88.

Marimon, R., and S. Sunder, 1993, Indeterminacy of Equilibria in a Hyperinflationary World: Experimental Evidence, *Econometrica*, 61, 1073–1107.

Meltzer, A. H., and S. F. Richard, 1981, A Rational Theory of the Size of Government, *Journal of Political Economy*, 89, 914–27.

Musgrave, R. A., 1966, Principles of budget determination, in *Public finance: selected readings*, ed. by H. Camson, and W. Henderson. Random House, New York.

Nyarko, Y., and A. Schotter, 2002, An Experimental Study of Belief Learning Using Elicited Beliefs, *Econometrica*, 70(3), 971–1005.

Offerman, T., J. Sonnemans, and A. J. H. C. J. Schram, 1996, Value Orientations, Expectations and Voluntary Contributions in Public Goods, *Economic Journal*, 106(437), 817–845.

Payne, J. E., 1998, The Tax-Spend Debate: Time Series Evidence from State Budgets, *Public Choice*, 95, 307–20.

Pesaran, M. H., 1987, *The limits to rational expectations*. Oxford and New York: Blackwell.

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung, 1981, *Vor Kurskorrekturen — Zur finanzpolitischen und währungspolitischen Situation im Sommer 1981*Special Report of July 4, 1981.

Sargent, T. J., 1993, *Bounded Rationality in Macroeconomics*. Clarendon-Oxford University Press, Oxford.

Sarno, L., 2001, The behavior of US public debt: a nonlinear perspective, *Economics Letters*, 74, 119–125.

Schmalensee, R., 1977, An experimental study of expectation formation, *Econometrica*, 44, 17–41.

Seater, J. J., 1993, Ricardian Equivalence, *Journal of Economic Literature*, 31, 142–90.

Sutherland, A., 1997, Fiscal crises and aggregate demand: can high public debt reverse the effects of fiscal policy, *Journal of Public Economics*, 65, 147–162.

Swenson, C. W., 1997, Rational expectation and tax policy: experimental market evidence, *Journal of Economic Behavior and Organization*, 32, 433–455.

Trehan, B., and C. E. Walsh, 1991, Testing Intertemporal Budget Constraints: Theory and Applications to U.S. Federal Budget and Current Account Deficits, *Journal of Money, Credit, and Banking*, 23, 206–23.

A Instructions to the experiment

The experiment was conducted in German. In this section you find a translation of the instructions:

1. Please read the instructions carefully. Only if you have understood them well you can successfully participate in the experiment and gain money.

2. **Thereafter** fill in the questionnaire at the screen.

Welcome to the strategy experiment

Welcome to the strategy experiment

This strategy experiment is financed by the University of Mannheim and the German research council.

The instructions are simple, and if you carefully pay attention to it and decide deliberately, you will win a considerable amount of money, which is disbursed to you at the end of the game.

The payment is dependent on your success. In the experiment you forecast the development of public expenditures and taxation in several European countries. For that purpose there are past data about budget debt, annual change of budget debt, government expenditure and taxes made available for you. Dependent on the quality of your forecast you receive a payment for each period.

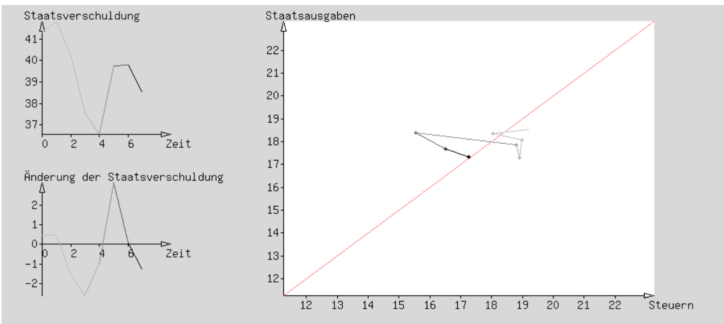
Please note that we do not have any interest in paying less money than you are entitled to. We must return all the money, which we do not disburse to you, to the German research council.

Please note that we will not deceive you in this experiment. Everything you read in these instructions is correct. You may take this for granted, but actually there are occasionally experiments in psychology, where experiment participants are deceived about parts of the experiment. This is not the case in economic experiments like this. In the beginning we explain exactly the rules to you, and we will also adhere to them.

Rules

You will play several rounds in turns. In each round it is your task to forecast the development of two variables. These variables refer to the development of government expenditure and taxes in several European states between 1950 and 2000. Which states you play in each case will be specified randomly and is not made known to you. These data are shown graphically.

Top of the screen

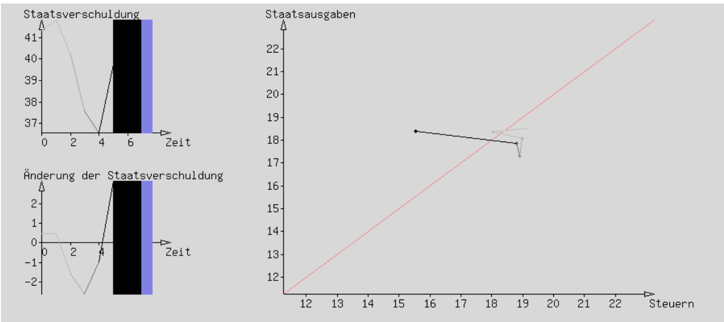


On the left you can see the development of the budget debt and annual change of budget debt, each in per cent of the gross national product. The horizontal axis shows time in years. You may use this data to obtain a reference point how government expenditure and taxes will change in the future. Current periods are shown in black, past periods are shown in gray.

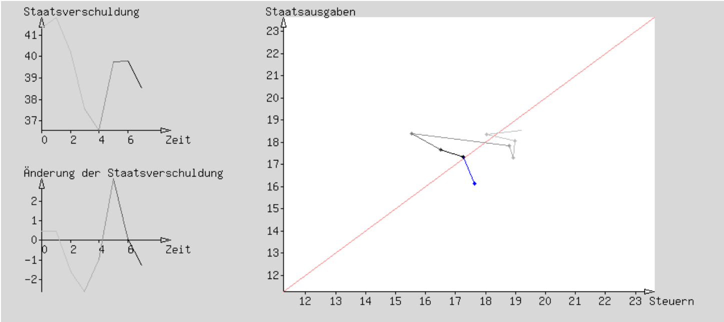
On the right you can see the government expenditure and taxes, again as percentage share of the gross national product. The vertical axis shows government expenditure, the horizontal axis shows taxes. Government expenditure is higher than taxes above the red diagonal; below, government expenditure is lower than taxes. Past periods are shown in a lighter shade of gray than current periods.

Partial representation of the past development You can present yourself also only one part of the past periods to get a better overview.

In order to do so click on the diagrams budget debt and annual change of budget debt. In these diagrams the area to the right of your click is covered black. Also in the diagram of government expenditure and taxes the covered periods are not shown. Each click onto the black area of the diagrams budget debt and annual change of budget debt uncovers one period after another. A click on the blue area uncovers all periods.



Forecasts In order to make a forecast about the development of government expenditure and taxes, click onto the white area. Your forecast is shown in blue.



If you are content with your forecast, please confirm it by clicking on confirm forecast. If you want to correct your forecast, please click on delete forecast.

Payment Given your forecast the computer determines a consumption decision, which would be optimal for a person who lives in the period. From your consumption-decisions you derive a certain utility. This utility is compared with the utility you would have obtained if you had forecasted the true future development of taxes and government expenditure.

You receive a wage of 0.45€ per minute for a correct forecast. Worse forecasts result in smaller wages.

It is worth to spend some time to make a good forecast. Example: You need 2 minutes in order to make a very good forecast and therefore receive wages of 0.45€ per minute. Your income in the 2 minutes is thus 0.90€.

Another person, who makes forecasts for e.g. 4 periods in these 2 minutes, which are not so good, may only receive a wage of 0.10€ per minute for each forecast. The income of this person in the 2 minutes is thus only 0.20€.

You should settle your forecast within 2 minutes. If you need more time for a forecast, you are paid only for the first 2 minutes.

A warning on the left side will remind you, as soon as you need more than 2 minutes.

Furthermore you get a list about the income of your past forecasts on the left side.

Duration of the experiment The experiment takes 90 minutes, regardless whether you made many or few forecasts in this time. That requires, however, that you take yourself at least 20 seconds time for each forecast on the average. If you take yourself less time, you are finished with the experiment sooner, but earn fewer money, accordingly.

Should you have any questions, you now have the opportunity to ask them. In addition, you can ask questions at any time during the experiment.

Appendix to the instructions

To determine your payoff we use the following model. **It is not necessary to understand this model to participate successfully in the experiment.** The model is shown only in case you want to control us.

In two subsequent periods you consume c_0 and c_1 and pay taxes t_0, t_1. You save the remaining part:

$$s_i = 1 - c_i - t_i \tag{9}$$

Your total income in each period is $Y = 1$ (note that all values are relative to the gross domestic product Y).

We call government expenditure g_i. Then your utility in two subsequent periods is

$$u = \prod_{i=0}^{1} \gamma c_i + (1-\gamma) g_i \tag{10}$$

In your case $\gamma = 0.75$.

Your budget restriction is

$$\sum_{i=0}^{1} s_i \cdot (1+r)^i = 0 \tag{11}$$

with an interest rate $r = 0.1$.

Based on your forecast for t_1 and g_1 we determine your optimal consumption c_0.

In the next period t_1 and g_1 are realised. Your actual consumption c_1, and, hence, your utility u, follows from the budget restriction. This utility is compared with the utility u^* that you could have obtained with the correct forecast for t_1 and g_1. Your wage is $(u/u^*)^\eta$. In your case $\eta = 12000$. This normalisation does not change your utility maximisation problem.