

University of
St Andrews

School of Economics & Finance Discussion Papers

Anticipation, Learning and Welfare: the Case of Distortionary Taxation

Emanuel Gasteiger and Shoujian Zhang

School of Economics and Finance Discussion Paper No. 1309

26 Aug 2013

JEL Classification: E32, E62, D84

Keywords: Fiscal Policy, Adaptive Learning, Oscillations

Anticipation, Learning and Welfare: the Case of Distortionary Taxation*

Emanuel Gasteiger[†]
Instituto Universitario de Lisboa

Shoujian Zhang[‡]
University of St Andrews

AUGUST 27 2013

Forthcoming in *Journal of Economic Dynamics and Control*

Abstract

We study the impact of anticipated fiscal policy changes in a Ramsey economy where agents form long-horizon expectations using adaptive learning. We extend the existing framework by introducing distortionary taxes as well as elastic labour supply, which makes agents' decisions non-predetermined but more realistic. We detect that the dynamic responses to anticipated tax changes under learning have oscillatory behaviour that can be interpreted as self-fulfilling waves of optimism and pessimism emerging from systematic forecast errors. Moreover, we demonstrate that these waves can have important implications for the welfare consequences of fiscal reforms. **JEL codes:** E32, E62, D84

*We are indebted to Alejandro Cunat, Robert Kunst and Gerhard Sorger for sound supervision and patience. We are also grateful to the editors and two anonymous referees, Szabolcs Deak, Seppo Honkapohja, Torsten Lisson, Albert Marcet, Paul Pichler, and Monika Turyna as well as the participants at the 4th RGS Doctoral Conference in Economics, the 2011 Annual Conference of the Scottish Economic Society and the Quantitative Economics Doctorate (QED) Jamboree 2011 for valuable comments. All remaining errors are the responsibility of the authors.

[†]Corresponding author: Instituto Universitario de Lisboa (ISCTE - IUL), Av. das Forças Armadas, 1649-026 Lisboa, Portugal, +351-926-785229, emanuel.gasteiger@iscte.pt

[‡]School of Economics and Finance, Castlecliffe, The Scores, University of St Andrews, Fife KY16 9AR. shoujian.zhang@st-andrews.ac.uk.

1. Motivation

Nowadays, fiscal policy is usually accompanied by legislation and implementation lags. These lags create a non-negligible span of time between the announcement and effective date of fiscal policy changes. This gives economic agents the opportunity to anticipate tax changes. The economic literature denotes this aspect of fiscal policy anticipated fiscal policy.¹

When agents anticipate, their resulting actions may to some extent depend on the way they form expectations about the future. The standard assumption of expectations in economics is perfect-foresight / rational expectations (RE). This assumption might be questioned on the grounds of its unrealistically strong restrictions. One prominent deviation of RE that imposes weaker requirements on the agent's information set when making his decisions, is the learning literature (see Evans and Honkapohja (2001) for the foundations of this approach). The main idea is that agents form expectations about the future evolution of contemporaneously unobservable variables by engaging in a kind of statistical inference, when making their economic choices.

Although the learning approach has gained significant popularity in some areas of macroeconomics, (anticipated) fiscal policy has, until recently, been neglected. A pioneering contribution to the study of anticipated fiscal policy under learning has been made by Evans et al. (2009). They demonstrate the adaptive constant gain learning approach in several deterministic economic environments including the popular Ramsey model. In this set-up, it is assumed that agents understand the structure of government financing but have to forecast

¹Recently Leeper (2009, p.11ff.) has listed empirical evidence for anticipated fiscal policy/fiscal foresight and reemphasized the relevance of expectations for sound fiscal policy. Furthermore, Leeper et al. (2009) is another good example of empirical evidence of fiscal foresight. Therein they also demonstrate the challenges for econometricians that aim to quantify the impact of fiscal policy actions and at the same time account adequately for fiscal foresight.

factor-prices on decentralized markets. Their key result is that for an anticipated balanced-budget permanent tax change the impact effects on key variables under learning are similar to the perfect foresight case, but the transition paths are remarkably different from the latter. Their result, at least with regard to the volatility of key variables' time paths may not come as a surprise. It is well known that constant gain learning causes excess volatility (see Evans and Honkapohja, 2001, p.49).²

Building on the contribution of Evans et al. (2009), we aim to generalize their analysis of anticipated fiscal policy under learning by studying an economy featuring distortionary taxes and elastic labour supply. Thus, our theoretical key contribution is to derive the dynamic paths of key variables for anticipated permanent changes in distortionary taxes in the prominent Ramsey model.³

Note that there are fundamental differences between lump-sum taxation and distortionary taxes such as labour income, capital income, or consumption tax.⁴ Furthermore, the assumption of elastic labour supply implies that endogenous variables such as factor prices as well as employment and consumption are not predetermined as in Evans et al. (2009, p.943ff.), but determined simultaneously. Thus, one can expect that changes in distortionary taxes may yield dynamics fundamentally different from the ones in the lump-sum tax case.

One of our main results supports this hypothesis. When we assume that agents use adaptive learning rules to forecast factor prices, our model predicts oscillatory dynamic responses to pre-announced permanent tax changes. More-

²In subsequent work, Evans et al. (2010) focus on Ricardian equivalence in the basic Ramsey model with anticipated fiscal policy under learning. Most important, Evans et al. (2010, p.8ff.) formally proof that the assumption of RE is not necessary for the Ricardian equivalence result.

³Another extension is to consider stochastic set-ups. Recently, Evans et al. (2011b) have pioneered the study of anticipated lump-sum tax changes under learning in the RBC model.

⁴Find the details in Ljungqvist and Sargent (2004, p.323ff.).

over, the case of distortionary taxation is particularly different with regard to the effects on impact and the volatility throughout the transition period. The source of the oscillations are persistent systematic forecast errors, that lead agents to incorrectly anticipate the effects of pre-announced tax changes. Thus, the oscillations can be interpreted as self-fulfilling waves of optimism and pessimism.

Farmer (1999, p.141ff.) and others have argued that these waves are a feature of US data and can be replicated by modified RBC models under RE. Evans et al. (2011b) replicate the waves as a consequence of fiscal policy changes under the assumption of adaptive learning in the standard RBC model. We demonstrate that these waves exist even in the deterministic version of this model.

Confronted with this result, we then ask, to what extent the striking differences in the dynamics between perfect foresight and learning, affect the welfare consequences of a pre-announced tax reform at the presence of several tax instruments. For this purpose, we make use of the welfare measure proposed by Lucas (1990) and also applied by Cooley and Hansen (1992) (for discrete time), which takes into account the whole transition path between the initial and new steady-states associated with initial and changed tax rates.

This welfare analysis is our second key contribution. It links the learning literature to the part of the public finance literature that is concerned with the welfare consequences of tax reforms.⁵ However, the existing literature evaluates and ranks various distortionary tax reforms according to their welfare consequences under perfect foresight, but do not consider the case of learning. We fill this gap and illustrate that tax reforms designed to improve welfare, do so

⁵See Chamley (1981) for an example of a comparative statics analysis, Judd (1987) for differences in unanticipated and anticipated changes in factor taxes, or Cooley and Hansen (1992) for a study in a stochastic set-up. Moreover, Garcia-Milà et al. (2010) have recently conducted research on welfare consequences of fiscal policy reforms in a heterogeneous agents model.

to a much lower extent under learning compared to perfect foresight. Thus, the learning perspective on tax reforms provides fundamental different insights for benevolent policy makers.

The remainder of the paper is organized as follows. In Section 2 we outline the economic model. We use Section 3 to detail our approach of learning and to derive the equations that govern the dynamics under learning. Section 4 provides a simple intuitive example of lump-sum tax changes. This section also contains a sensitivity analysis for some structural parameters. In Section 5 we focus on distortionary taxation and present a numerical welfare analysis of an exemplary tax reform. Section 6 concludes and points out directions for further research.

2. The Model

Our economy is a version of the Ramsey economy (see Ljungqvist and Sargent, 2004, p.323ff.). k_t evolves according to the economy-wide resource constraint

$$k_{t+1} = F(k_t, n_t) - c_t - g_t + (1 - \delta)k_t, \quad (1)$$

where $F(k_t, n_t)$ is the economy's production function showing that the firm sector uses the stock of capital k_t and labour n_t as inputs to produce the single good of the economy (see Section 2.2 for the details). Output can be purchased by households (c_t) or the government (g_t) or added to k_t , which is assumed to depreciate at a constant rate δ .

2.1. Households

We assume a continuum of households, where we normalize the size of the economy to unity and each household faces the problem

$$\max_{c_t, n_t} E_t^* \left\{ \sum_{t=0}^{\infty} \beta^t [\log(c_t) + \eta \log(\bar{L} - n_t) + \phi \log(g_t)] \right\} \text{ s.t. (2)}$$

$$k_{t+1} + \frac{b_{t+1}}{R_t} + (1 + \tau_t^c)c_t = (1 - \tau_t^l)w_t n_t + (1 - \tau_t^k)r_t k_t + (1 - \delta)k_t + b_t - \tau_t + \pi_t, (3)$$

where all variables are in per capita terms. Thus, b_{t+1} is the level of government debt holdings chosen in period t . Furthermore, r_t is the rental rate of capital and R_t is the gross real interest rate in period t . Next, τ_t^\bullet denotes a distortionary tax either on consumption, labour income or capital income⁶. The real wage in period t is given by w_t and $(\bar{L} - n_t)$ denotes leisure. In consequence, n_t is labour supply of the household. τ_t is a per capita lump-sum tax, and $\pi_t = 0$ is the profit under perfect competition among firms. Furthermore, the parameter $\eta \geq 0$ measures the elasticity of labour supply, ϕ measures the degree of substitution between private consumption and government spending, see Ambler and Paquet (1996), and β is the common discount rate. $E_t^*\{\bullet\}$ denotes subjective period t expectations for future values of variables.⁷

Given this set-up, the household Euler condition, the usual no-arbitrage condition for capital and bonds, and the consumption leisure trade-off are respectively

⁶We use the symbol \bullet as a placeholder throughout our analysis.

⁷Households apply this operator, if they do not have perfect foresight. This assumption is commonly used in the learning literature. Furthermore, note that we abstract from aggregate uncertainty, i.e. we conduct our analysis in a deterministic economy. Thus, if households do not have perfect foresight, their expectations are so-called point expectations, i.e. agents base their economic choices on the mean of their expectations, see Evans and Honkapohja (2001, p.61). In Section 3.1 below we outline our concept of learning. An important aspect of this concept is that forecasts of single variables are independent of each other. In consequence, we can assume that for any two variables X and Y it is true that $E_t^*\{XY\} = E_t^*\{X\}E_t^*\{Y\}$ holds.

given by

$$\bar{c}_t^{-1} = \beta R_t E_t^* \left\{ c_{t+1}^{-1} \frac{(1 + \tau_t^c)}{(1 + \tau_{t+1}^c)} \right\}, \quad (4)$$

$$R_t = [(1 - \delta) + (1 - E_t^* \{ \tau_{t+1}^k \}) E_t^* \{ r_{t+1} \}], \quad (5)$$

$$n_t = \bar{L} - \frac{\eta(1 + \tau_t^c)c_t}{(1 - \tau_t^l)w_t}. \quad (6)$$

2.2. Firms

We assume a unit continuum of firms who compete perfectly. Each firm in each period t rents capital at given price r_t and labour at given price w_t and produces the numeraire good with constant returns to scale production function

$$y_t = F(k_t, n_t) = A k_t^\alpha n_t^{(1-\alpha)}, \quad (7)$$

where $\alpha \in (0, 1)$. The optimal firm behaviour requires that

$$r_t \stackrel{!}{=} \frac{\partial y_t}{\partial k_t} = A \alpha k_t^{\alpha-1} n_t^{1-\alpha}, \quad (8)$$

$$w_t \stackrel{!}{=} \frac{\partial y_t}{\partial n_t} = A(1 - \alpha) k_t^\alpha n_t^{-\alpha}, \quad (9)$$

i.e. each production factor earns its marginal product. Finally, we have the per capita national income identity $\pi_t = y_t - r_t k_t - w_t n_t = 0$.

2.3. Government

The government finances its expenses on goods and debt repayment by tax revenues and the issuance of new bonds in each period t ,

$$g_t + b_t = \tau_t^c c_t + \tau_t^l w_t n_t + \tau_t^k r_t k_t + \tau_t + \frac{b_{t+1}}{R_t}.$$

For the remainder, we will assume that the government operates a balanced-budget rule in each period t , thus tax revenues will fully cover expenses such that bonds are in zero net supply as a direct consequence. Thus, in each period t the government sets $g_t, \tau_t^c, \tau_t^l, \tau_t^k$ and τ_t constrained by

$$g_t = \tau_t^c c_t + \tau_t^l w_t n_t + \tau_t^k r_t k_t + \tau_t. \quad (10)$$

3. Learning and Learning Dynamics

3.1. Learning

The concept of learning applied herein was elaborated first in Evans et al. (2009, p.943ff.). For completeness we restate the crucial assumptions. Under learning, households are supposed to know the entire history of endogenous variables. They observe the current period value of exogenous variables and they know the state variables. Furthermore, they know the structure of the economy with regard to the fiscal policy sector. Agents understand the implications of any pre-announced policy change for the government budget constraint. They are also convinced that the intertemporal government budget constraint will always hold (see Evans et al., 2009, p.944). We assume decentralized markets for labour and capital, where agents are not in possession of perfect foresight. Actual factor prices are not observable. Thus, agents forecast factor prices $r_{t+j}^e(t)$ and $w_{t+j}^e(t)$ for $j \geq 1$, by a constant-gain steady-state adaptive learning rule⁸

$$z_{t+j}^e(t) = z^e(t) = r^e(t-1) + \gamma(z_{t-1} - z^e(t-1)), \quad (11)$$

⁸Here we apply the same short-hand notation as Evans et al. (2009). Thus for any variable say z , its period t expected future value in period $t+j$ derived by a learning rule may either be denoted $E_t^* \{z_{t+j}\}$ or equivalently $z_{t+j}^e(t)$. An additional notation we introduce is $z_{t+j}^p(t)$, which denotes the agent's planned choice of the variable z in period $t+j$ based on expected values formed via the learning rule in period t .

where $0 < \gamma \leq 1$ is the gain parameter.⁹ Our choice of this specific learning rule is motivated by two well known arguments in the learning literature. First, as Evans and Honkapohja (2001, p.332) outline, choosing a constant gain learning rule is the appropriate choice for agents, when they are aware of structural change, as in such a learning rule agents discount past data exponentially. Note that rule (11) is equivalent to $z^e(t) = \gamma \sum_{i=0}^{\infty} (1-\gamma)^i z_{t-i-1}$.¹⁰ Second, the timing of the learning rule, i.e. that agents' update in period t uses data up to period $t-1$, is chosen in order to avoid simultaneity between $r^e(t)$ and r_t as well as $w^e(t)$ and w_t (see for example Evans and Honkapohja, 2001, p.51). Simultaneity in this context means that agents' expectations affect current values of aggregate endogenous variables and vice versa, which could introduce strategic behaviour.

Such a learning rule yields a sequence of so-called *temporary equilibria*, which consist of sequences of (planned) time paths for all endogenous variables. These sequences satisfy the learning rule above, the expectation history, household and firm optimality conditions, the government budget constraint, and the economy-wide resource constraint given the exogenous variables as well as the current stock of capital in each period. These plans are revisited and potentially altered in each period after expectations have been updated.

3.2. Learning Dynamics With Distortionary Taxation

Now, we derive the dynamic paths under learning in presence of multiple types of taxes, i.e. $\tau_t^c, \tau_t^l, \tau_t^k \in [0, 1]$ and $\tau_t \neq 0$ for all t . Equations (4)-(5) yield

$$c_t^{-1} = \beta(c_{t+1}^p(t))^{-1} \left[\frac{(1 + \tau_t^c)}{(1 + \tau_{t+1}^{c,e}(t))} \right] [(1 - \delta) + (1 - \tau_{t+1}^{k,e}(t))r_{t+1}^e(t)]$$

⁹The gain parameter measures the responsiveness of the forecast to new observations, see Evans and Honkapohja (2001, p.18). Be aware that in our model the gain parameter is exogenous. See Branch and Evans (2007) for an example where agents choose the gain parameter.

¹⁰The learning rule is similar to an exponential smoothing method.

and forward substitution of this expression yields

$$c_{t+j}^p(t) = \beta^j D_{t,t+j}^{k,e}(t) \left[\frac{(1 + \tau_t^c)}{(1 + \tau_{t+j}^{c,e}(t))} \right] c_t, \quad (12)$$

where we define $D_{t,t+j}^{k,e}(t) \equiv \prod_{i=1}^j [(1 - \delta) + (1 - \tau_{t+i}^{k,e}(t))r_{t+i}^e(t)]$. One can think of this term as “expectations of the interest rate factor $D_{t,t+j}$ at time t ” (see Evans et al., 2009, p.933). Furthermore, notice that the consumption leisure trade-off is now given by (6). If agents plan to satisfy an adequate transversality condition

$$\lim_{T \rightarrow \infty} \left(D_{t,t+T}^{k,e}(t) \right)^{-1} k_{t+T+1}^p(t) = 0, \quad (13)$$

the inter-temporal budget constraint of the consumer is

$$(1 + \tau_t^c)c_t + \sum_{j=1}^{\infty} \frac{1}{D_{t,t+j}^{k,e}(t)} (1 + \tau_{t+j}^{c,e}(t))c_{t+j}^p(t) = [(1 - \delta) + (1 - \tau_t^k)r_t]k_t \\ + (1 - \tau_t^l)w_t n_t - \tau_t + \sum_{j=1}^{\infty} \frac{1}{D_{t,t+j}^{k,e}(t)} [(1 - \tau_{t+j}^{l,e}(t))w_{t+j}^e(t)n_{t+j}^p(t) - \tau_{t+j}^e(t)],$$

which by the virtue of (12) as well as (6) yields the consumption function

$$c_t = \frac{(1 - \beta)}{(1 + \eta)(1 + \tau_t^c)} [(1 - \delta) + (1 - \tau_t^k)r_t]k_t + (1 - \tau_t^l)w_t \bar{L} - \tau_t \\ + SW_2 - ST_2 - ST_3, \text{ where} \quad (14)$$

$$SW_2 = \sum_{j=1}^{\infty} \frac{1}{D_{t,t+j}^{k,e}(t)} w_{t+j}^e(t) \bar{L}, \quad (15)$$

$$ST_2 = \sum_{j=1}^{\infty} \frac{1}{D_{t,t+j}^{k,e}(t)} \tau_{t+j}^{l,e}(t) w_{t+j}^e(t) \bar{L}, \text{ and} \quad (16)$$

$$ST_3 = \sum_{j=1}^{\infty} \frac{1}{D_{t,t+j}^{k,e}(t)} \tau_{t+j}^e(t) \quad (17)$$

are the expected present value of labour income, the labour income taxes and lump-sum taxes in turn.

We now need to think about the policy experiment we will study. We are looking at a scenario of a credible permanent (simultaneous) change in (some) taxes announced at the outset of period $t = 1$ and effective from period $t = T_p$ onwards. The dynamics under perfect foresight are standard.¹¹ Under learning, given that households know the future path of taxes, they can explicitly calculate SW_2 , ST_2 , and ST_3 .¹² Given a calibration of the structural parameters, we can then compute the dynamics responses for c_t and the other endogenous variables.

4. A Simple Intuitive Example: Lump-Sum Tax Increase

We would like to illustrate the applied methodology for the simple case of lump-sum taxation ($\tau_t^c = \tau_t^l = \tau_t^k = 0$) for three reasons: first, the rather complex consumption function (14)-(17) under learning simplifies in this case and facilitates to develop some intuition; second, we want to illustrate the consequences of the introduction of elastic labour supply compared to the case of inelastic labour supply as assumed in Evans et al. (2009, p.943ff.) and its effect on the dynamic paths of the key variables such as c_t and k_t , given their calibration (see Table 1 below); third, below in Subsection 4.2, we aim to present a sensitivity analysis for the very basic version of the model to deepen the intuition for the impact of the learning parameters on the dynamics.

¹¹Ljungqvist and Sargent (2004, p.323ff.) illustrate the analytical derivations and numerical simulation alternatives for the perfect foresight case. We will simply make use of the DYNARE toolbox throughout all calculations to compute dynamics under perfect foresight. Note that this toolbox employs linearization methods.

¹²See Appendices A.1, A.2, and A.3 for details.

For an anticipated lump-sum tax change under learning (14) simplifies to

$$c_t = \frac{(1 - \beta)}{(1 + \eta)} \{[(1 - \delta) + r_t]k_t + w_t \bar{L} - \tau_0 + SW_1 - ST_1\}, \quad (18)$$

where the expected present value of labour income is given by

$$SW_1 = \sum_{j=1}^{\infty} \frac{1}{D_{t,t+j}^{k,e}(t)} w_{t+j}^e(t) \bar{L} = \frac{w^e(t) \bar{L}}{r^e(t) - \delta}, \quad (19)$$

and the expected present value of lump-sum taxes is given by $ST_1 =$

$$\sum_{j=1}^{\infty} \frac{1}{D_{t,t+j}^{k,e}(t)} \tau_{t+j}^e(t) = \begin{cases} \frac{\tau_0}{r^e(t) - \delta} + (\tau_1 - \tau_0) \frac{[(1-\delta) + r^e(t)]^{t-T_p}}{1 - [(1-\delta) + r^e(t)]^{-1}} & \text{for } 1 \leq t < T_p \\ \frac{\tau_1}{r^e(t) - \delta} & \text{for } t \geq T_p. \end{cases} \quad (20)$$

4.1. Inelastic Labour Supply vs. Elastic Labour Supply

We believe that it is of importance to use a model that features elastic labour supply in order to study the implications of fiscal policy reforms adequately. Completely inelastic labour supply is a quite unrealistic assumption itself and at least some moderately elastic labour supply should be considered.¹³ Moreover, inelastic labour supply implies that agents' choices of current period endogenous variables are in fact predetermined as is pointed out in Evans et al. (2009, p.944). In order to illustrate differences in the dynamics of endogenous variables based on the assumption of inelastic and elastic labour supply, we return to the simulation exercise of Evans et al. (2009, p.943ff.).¹⁴ The calibration for this subsection is given in Table 1 below.¹⁵ The policy experiment considered in Evans et al. (2009,

¹³The empirical evidence discussed in Chetty et al. (2012) suggests $\eta \approx 4$.

¹⁴Note that $\tau_t^c = \tau_t^l = \tau_t^k = \delta = 0$ and $\eta = 0$ imply that $n_t = \bar{L}$ (i.e. inelastic labour supply) for all t . Therefore, we are exactly in the same scenario as in Evans et al. (2009, p.943ff.).

¹⁵Note that we do not fully agree with the calibration of Evans et al. (2009), but we will stick to their calibration in this subsection to keep our results comparable. The basic reason

Parameter	Value	Parameter	Value
A	1.00	δ	0.00
α	0.33	T_p	20
β	0.95	γ	0.10
ϕ	0.00		

Table 1: Parameters similar as in Evans et al. (2009, p.945).

p.943ff.) is a permanent increase in government purchases from $g_0 = \tau_0 = 0.9$ to $g_1 = \tau_1 = 1.1$ that is announced credibly in period $t = 1$ and will be effective from period $T_p = 20$ onwards. It is assumed that the economy is in steady-state in period $t = 0$. Simulations in Evans et al. (2009, p.943ff.) for consumption and capital are recalculated (with $\eta = 0, \bar{L} = 0.5182$) and displayed in Figures 1(a) and 1(b) below. Next, Figures 1(c) and 1(d) exhibit the dynamics for elastic labour supply (with $\eta = 2.00, \bar{L} = 1.00$) such that $n_0 = 0.5182$ and $g_0 = 0.9$.

Two distinct features emerge from Figure 1. First, when we compare the dynamic paths of c_t , as well as k_t , under perfect foresight and learning, they are different from each other no matter with or without elastic labour supply. Therefore, it may be quite important to consider learning when evaluating fiscal policies as learning is a more realistic assumption of human behaviour from our point of view.¹⁶ Second, obviously the learning paths in Figures 1(a) and 1(b) for inelastic labour supply are strikingly different to the ones under elastic labour supply in Figures 1(c) and 1(d). In particular, elastic labour supply yields more volatility in the time paths of c_t and k_t (as well as other variables in the model) compared to the inelastic labour supply case. In fact, the variables oscillate

for this disagreement is the combination of parameters $\beta = 0.95$ and $T_p = 20$. These parameter choices imply that a government, which in reality is usually in charge of a legislation period of four to six years, may announce a tax policy change that will be effective in 20 years' time. From our perception of political execution and our confidence in fiscal policy makers' ability to commit, this appears to be unrealistic in most cases.

¹⁶This is the core message of Evans et al. (2009).

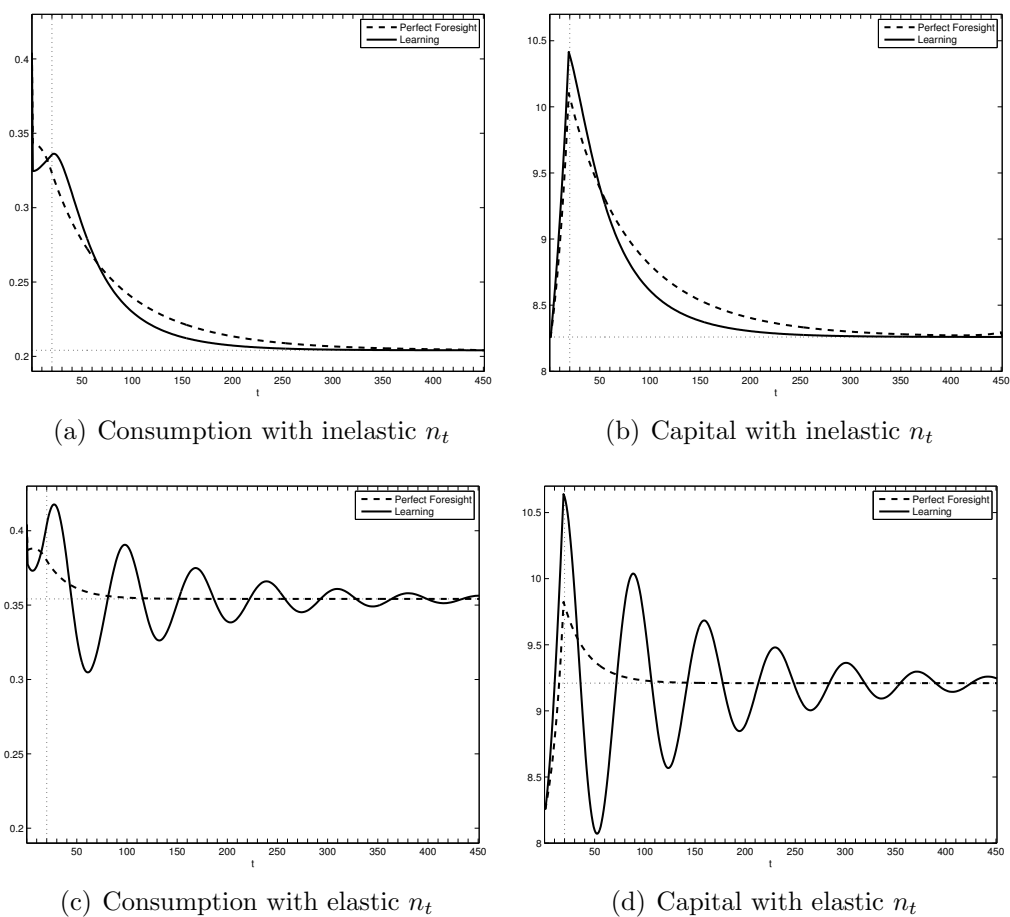


Figure 1: Consumption (a) and capital (b) dynamics under learning (solid curve) and perfect foresight (dashed curve) with inelastic labour supply as in Evans et al. (2009, p.943ff.) as well as consumption (c) and capital (d) dynamics under learning (solid curve) and perfect foresight (dashed curve) with elastic labour supply. The dotted horizontal line indicates the (new) steady state, the dotted vertical line indicates period T_p .

around their steady-state until they converge to it.

In order to develop some intuition for the transition dynamics illustrated in Figure 1, we start with the familiar case of *perfect foresight*.¹⁷ Given *inelastic labour supply*, agents respond to the credible pre-announced tax increase with a decrease in c_t on impact. The policy change means a negative wealth effect to the agents, as expected future taxes increase. Due to their perfect foresight, they precisely quantify this negative wealth effect. In the subsequent period, this increases k_t beyond its steady-state equilibrium value. Out of the steady-state equilibrium, the transitional dynamics lead to a further decrease in c_t and a temporary investment boom that increases k_t throughout the pre-implementation period. The rise in k_t affects marginal products and increases w_t , decreases r_t , and increases y_t . Once the policy change becomes effective in T_p , the levels of c_t and k_t are consistent with the new saddle-path and thereafter decrease monotonically to the new steady-state equilibrium. The evolution of w_t , r_t , and y_t is reversed and they converge monotonically to their respective steady-state values.

Also under *learning*, agents anticipate a negative wealth effect, and decrease c_t on impact. However, they do so to a lower extent, as they are not able to precisely quantify the negative wealth effect. The reason is that their expectations about factor prices are predetermined point estimates. In fact, on impact they project the initial steady state values into the entire future. Thereby they are too pessimistic about $w^e(t)$ and too optimistic about $r^e(t)$. The latter induces them to underestimate the present value of future taxes, as one can see from (20). However, the negative impact effect on c_t causes an investment boom that yields an increase in k_t during the pre-implementation period. A higher k_t also yields a higher level of y_t . Moreover, via marginal products w_t increases and r_t decreases.

¹⁷Our intuition is developed along the lines of Evans et al. (2011b) and the references therein.

Now, these changes in factor prices, in turn affect c_t via two distinct channels, as (18) makes clear. First, there is a direct effect on c_t . A higher w_t can increase c_t and a lower r_t can decrease c_t . However, the net effect in our case is positive. Second, there is a self-fulfilling channel. The changes in actual factor prices trigger a learning process due to systematic forecast errors. Rule (11) leads to upward revisions of $w^e(t)$ and downward revisions of $r^e(t)$. The consumption function (18) makes clear that this leads to an increase in the sum of expected future labour income and taxes, where the increase in the former outweighs the increase in the latter. Thus, c_t also increases via this channel in the pre-implementation period.

When the increase in lump-sum taxes becomes effective in period T_p , the consumption function (18) indicates that it is almost fully anticipated by that date. Only the self-fulfilling channel, due to its persistence, is weakly driving up c_t . Thus, the tax increase mostly materializes in a drop in k_t via (1).

Throughout the post-implementation period, y_t decreases due to the shrinking k_t . In addition, w_t decreases and r_t increases. Over time, this yields downward revisions of $w^e(t)$ and upward revisions of $r^e(t)$. This reverses the evolution of c_t and decreases it. In the long run, the self-fulfilling channel gains importance. If agents manage to learn the new steady-state factor prices by continuous upward and downward revisions of factor prices, the economy eventually converges to the new steady-state.

In principal, the convergence is characterized by a sequence of waves of optimism and pessimism about factor prices and other variables in consequence. The only reason, why those waves are not very apparent in the learning paths in Panels 1(a)-1(b) is the particular calibration, as we will argue in Subsection 4.2 below.

In the case of *elastic labour supply* under *perfect foresight*, the dynamics are qualitatively similar to the ones with inelastic labour supply, which is well known from the standard Ramsey model. However, under *learning*, elastic labour supply has two major consequences. First, the impact effect on c_t is smaller, as $\eta \geq 0$ lowers the initial drop via (18). Second, actual factor prices and expectations about factor prices affect both c_t and n_t . The latter two are substitutes, and therefore the evolution of n_t is now subject to two opposing influences. Everything else equal, first, n_t tends to the opposite direction of c_t . Thus, given the drop in c_t on impact, households increase n_t , as can be seen from (6). Second, n_t in turn, now affects actual factor prices. On impact, the increase in n_t not only increases r_t and y_t , but decreases w_t , which in turn lowers n_t .

Lower c_t and larger y_t again create an investment boom before T_p . The intuition throughout the pre-implementation, implementation and post-implementation period is similar to the case of inelastic labour supply. However, now actual factor prices are exposed to two potentially opposing influences via k_t and n_t .

Nevertheless, the self-fulfilling channel is quantitatively more important right away and oscillations are amplified. The reason is that actual factor prices now react on impact, and expectations therefore already react in the period afterwards. In particular, $w^e(t)$ are first lowered, and, once the investment boom materializes, upward revised even in the pre-implementation period. Likewise, $r^e(t)$ are first increased and subsequently downward revised, even before T_p . These movements in expectations also explain, why c_t decreases further after T_p , before it recovers.

In sum, allowing for elastic labour supply results in amplified oscillations starting at an earlier date. However, the oscillations remain a feature of the samples that agents utilize to learn the steady-state values of factor prices and their resulting persistent and systematic forecast errors.

4.2. Sensitivity Analysis

Compared to the previous literature, the learning approach herein introduces two additional structural parameters. One is γ , the gain parameter and a second one is the implementation date T_p . Therefore, we are interested in how these parameters affect the dynamics of the economy.

4.2.1. Sensitivity Analysis for the Gain Parameter

There exists no consensus for the choice of γ in the learning literature. We are only aware of a single empirical estimate provided by Milani (2007, p.2074) for quarterly frequency and this is $\gamma = 0.0183$.¹⁸ But a reason to be cautious to use the estimate of Milani (2007, p.2074) is that it is based on a data set containing output, inflation and the nominal interest rate, whereas in our setting agents forecast the rental rate of capital and the real wage. Next, Milani (2007, p.2074) mentions that a range of $\gamma \in [0.01, 0.03]$ is commonly used. Evans and Honkapohja (2009, p.154) note a range of $\gamma \in [0.01, 0.06]$ as known estimates.

Below we will present sensitivity of the dynamics under learning for $\gamma \in \{0.01, 0.02, 0.05, 0.08, 0.10\}$.¹⁹ We do so for the original numerical analysis of Evans et al. (2009, p.943ff.) ($\bar{L} = 1.00, \eta = 0.00$), as in this case, there is inelastic labour supply and we can focus solely on the possible fluctuations introduced by varying the gain parameter γ .²⁰

In Figure 2(a) we observe that the smaller γ , the smaller the increase in c_t until T_p (after the initial drop). Furthermore, as we recognize from Figure 2(b), the smaller γ , the larger the increase in k_t until T_p . However, in both Figure 2(a) and 2(b), we observe that, after T_p , with decreasing γ the dynamics fluctuate

¹⁸This number indicates that agents use approximately $1/\gamma \approx 55$ quarters of data.

¹⁹Values in the range $\gamma \in [0.002, 0.01)$ do not alter the conclusions in this subsection.

²⁰Note that the two thick lines in Figures 2(a) and 2(b) exactly replicate the Figures 8 and 9 in Evans et al. (2009, p.943ff.).

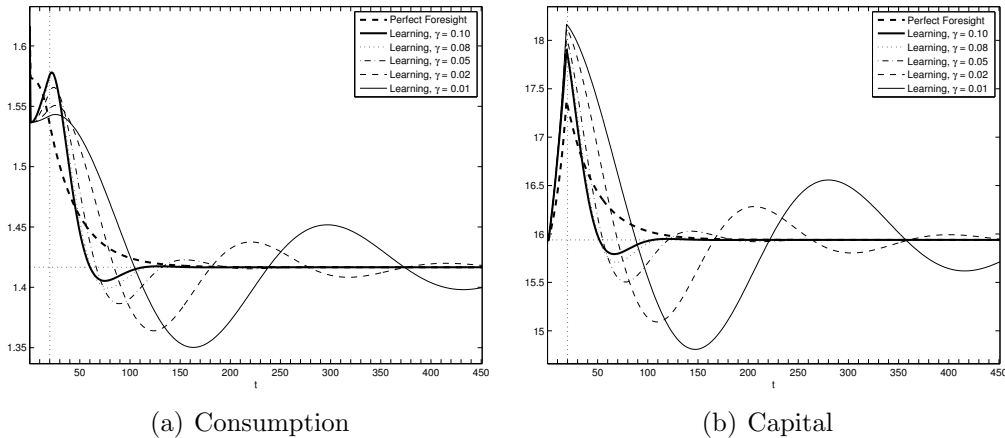


Figure 2: Consumption (a) and capital (b) dynamics under learning and perfect foresight with inelastic labour supply as in Evans et al. (2009, p.943ff.) for alternating values of γ . The dotted horizontal line indicates the (new) steady state, the dotted vertical line indicates period T_p .

around the steady-state with increasing amplitude and convergence slows down. These observations are partly at odds with what Evans and Honkapohja (2001, p.332) report for the statistical literature: “a larger gain is better at tracking changes but at the cost of a larger variance”. In our case it holds, that a smaller gain yields larger volatility.²¹

Inspection of the learning rule (11) explains this fact. The term $(1 - \gamma)^i$ indicates that a smaller γ , means stronger discounting of past observations. Thus, with a smaller γ , agents have more confidence in their initial expectations. Throughout the pre-implementation period, a smaller γ means less optimism about SW_1 and approximately similar pessimism about ST_1 . In consequence, the temporary increase of c_t is lower and the temporary investment boom peaks

²¹Evans et al. (2011a, p.23) confirm this observation with us in a similar model under educative learning. Both papers study a deterministic multivariate economy with capital accumulation, where agents form point expectations over an infinite-horizon. Exactly this combination of features distinguishes them from the existing literature, such as the references in Evans et al. (2011a, p.23), which reports the opposite result, and may be the starting point for providing a general explanation as the result of future research.

at a higher level of k_t . Moreover, a smaller γ yields slower convergence in the post-implementation period, as the expectational errors are more persistent.

4.2.2. Sensitivity Analysis for the Implementation Date

Next, we examine sensitivity of dynamics under learning for various implementation dates, in particular $T_p \in \{3, 10, 20\}$.

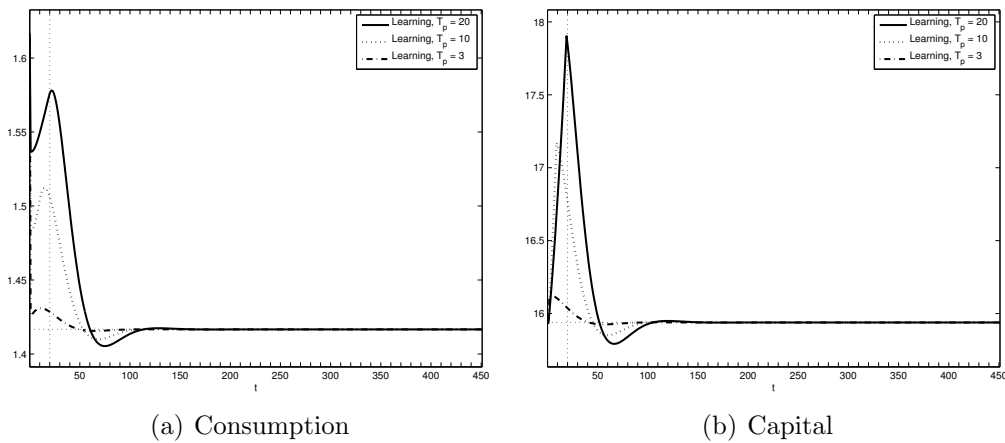


Figure 3: Consumption (a) and capital (b) dynamics under learning with inelastic labour supply as in Evans et al. (2009, p.943ff.) for alternating values of T_p and $\gamma = 0.10$. The dotted horizontal line indicates the (new) steady state, the dotted vertical line indicates period $T_p = 20$.

In Figure 3(a) we observe that the shorter the distance between the announcement date and T_p , the higher the initial drop in c_t and the lower the increase in c_t until T_p thereafter. Next, in Figure 3(b) we observe that with decreasing T_p , the peaks in k_t become smaller, but the learning dynamics remain qualitatively similar.

Inspection of SW_1 and ST_1 makes clear that T_p has only a direct impact on ST_1 . If T_p decreases, agents are less optimistic about ST_1 and reduce c_t more on impact. With lower T_p agents are also less optimistic about ST_1 throughout the pre-implementation, thus the temporary recovery in c_t is dampened, the in-

vestment boom ends earlier, and convergence speeds up. However, the nature of dynamics is not seriously affected.

In order to summarize, there are three important insights from the analysis above. First, there are at least qualitative differences between the case of inelastic labour supply ($\eta = 0$) and elastic labour supply ($\eta > 0$). Therefore, if one regards the latter assumption as more realistic, a model that allows for elastic labour supply is a more appropriate framework to study anticipated fiscal policy under learning. Second, our sensitivity analysis suggests that the choice of the gain parameter γ and the implementation date T_p does not affect the nature of transition paths so we consider ourselves free to choose any of the values considered in the sensitivity analysis.²² Finally and most notably, we observed at least a qualitative difference in the dynamics under learning compared to the dynamics under perfect foresight. The former appear to be much more volatile than the latter. This stylized fact motivates the quantification and comparison of welfare effects of anticipated fiscal policy reforms under learning and RE below.

5. The Case of Distortionary Taxation

Herein we are utilizing our derivations from Subsection 3.2 to conduct an exemplary tax reform. Our calibration now is given by Table 2. In particular, we consider a credible pre-announced cut in τ_t^k and adjust τ_t^l such that the new steady-state government revenue is the same as before the policy change. The change is effective in $T_p = 8$, which will correspond to a duration of 2 years.²³

²²In particular, in the subsequent analysis, we will choose $\gamma = 0.08$ and $T_p = 8$, which will correspond to 8 quarters.

²³This reform is comparable to the ones of Judd (1987), Lucas (1990), or Cooley and Hansen (1992). According to this literature, such a reform yields a welfare improvement. The learning perspective, despite its empirical support (see for example Milani (2011)), is unknown.

Parameter	Value	Parameter	Value
A	1.00	δ	0.025
α	0.33	T_p	8
β	0.985	γ	0.08
η	4.00	\bar{L}	0.15
ϕ	0.30		

Table 2: Model calibration for the tax reform.

Moreover, we assume that the budget is balanced in each period and that $g_t = \bar{g} = 0.25$. Thus, we require τ_t to adjust throughout the transition.

We evaluate the tax reform by welfare measures, which are computed following the approach of Cooley and Hansen (1992, p.301ff.).²⁴ Intuitively speaking, we compute the increase in consumption that an individual would require to be as well off as under the equilibrium allocation without taxes. We express that number in percentage of output. Non-zero tax rates as chosen according to Table 3 lead to distortions and reveals that the steady-state welfare loss due to our initially chosen tax rates amounts to $\mathcal{W}_0 = 4.75\%$.

The particular tax reform is a credible pre-announced cut in τ_0^k to $\tau_1^k = 0.25$ and an adequate increase in τ_1^l . Figure 4 indicates that the dynamics are familiar to ones observed for the simple example discussed above. Note the new steady-state c_t is higher and n_t is lower. This can be explained by the more severe distortions due to a larger τ_1^l . Next, Table 3 reveals that the reform yields a considerable welfare gain if agents have perfect foresight. The welfare loss decreases by 75.6% to $\mathcal{W}_1^P = 1.16\%$. However, the welfare loss under learning, $\mathcal{W}_1^L = 4.52\%$, indicates that the reform does not yield large welfare improvements when agents lack perfect foresight.

Some intuition for the transitory dynamics and the welfare loss can be devel-

²⁴We detail the computation in Appendix B.

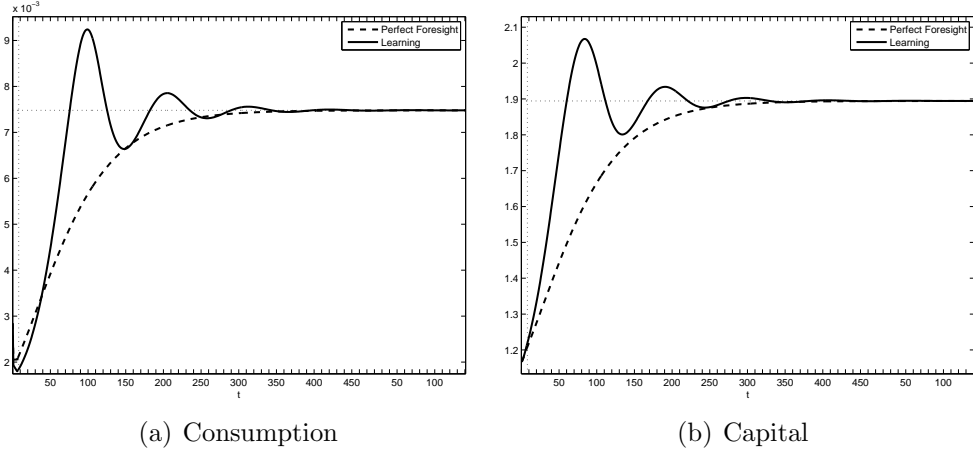


Figure 4: Consumption (a) and capital (b) dynamics under learning (solid curve) and perfect foresight (dashed curve). The dotted horizontal line indicates the (new) steady state, the dotted vertical line indicates T_p .

oped. Under *perfect foresight*, it is well known that, on impact, the pre-announced cut in τ_0^k yields to a drop in c_t and therefore to an increase in n_t . This increases r_t as well as y_t , and decreases w_t on impact. Agents anticipate that saving is more attractive in the future and that this will have a positive net effect on wealth. Therefore agents start accumulating k_t right away. The responses of non-predetermined variables to the simultaneously pre-announced increase in τ_t^l are in line with these impact effects. Subsequent accumulation of k_t yields increases in w_t and y_t and a decrease in r_t . The monotonic evolution of variables continues until the saddle-path is reached in T_p . Thereafter, c_t and k_t monotonically increase to the new steady-state. The increase in c_t yields a decrease in n_t , that in turn decreases r_t and increases w_t . The sustained growth in k_t causes y_t to grow as well until the new steady-state is reached. Finally, constant government expenditures imply that there is no kink in the evolution of k_t in T_p .

Under *learning*, agents again make mistakes, when anticipating the effects of the policy change. They underestimate the positive net effect on wealth.

However, as a feature of their sample and the learning rule (11), agents start forming either too optimistic or too pessimistic expectations about factor prices via the self-fulfilling channel, which has similar consequences throughout the pre- and post-implementation period as discussed above in Section 4. Thus, we observe waves of optimism and pessimism until the economy eventually converges to the new steady-state. Exactly those waves generated by systematic forecast errors cause excessive volatility of the economy and in turn are the source for the low welfare gain of the tax reform.

In sum, the tax reform indicates that the resulting welfare improvements of an anticipated tax reform might be much smaller in magnitude under learning compared to its improvements under perfect foresight.

Tax Reform				
τ_0	\rightarrow	τ_1	0.1596	
τ_0^k	\rightarrow	τ_1^k	0.5000	\searrow 0.2500
τ_0^l	\rightarrow	τ_1^l	0.2300	\nearrow 0.3180
τ_0^c	\rightarrow	τ_1^c	0.0500	
\mathcal{W}_0	\rightarrow	\mathcal{W}_1^P	0.0475	\searrow 0.0116
\mathcal{W}_0	\rightarrow	\mathcal{W}_1^L	0.0475	\searrow 0.0452

Table 3: Simulation results for the tax reform. The initial tax rates are the same as in Cooley and Hansen (1992, p.305) except for the consumption tax. The latter is chosen as in Giannitsarou (2007, p.1433).

6. Conclusion

We demonstrate within the Ramsey model that the responses to anticipated permanent tax changes when agents learn are remarkably different compared to their counterparts under perfect foresight. The learning dynamics appear to oscillate around the steady-state to which they converge slowly. Thus, there is more volatility under learning.

We argue that the observed oscillations are related to expectational errors. These systematic forecast errors are caused by the anticipated permanent tax change and lead agents to incorrectly quantify the effects of tax reforms on their life-time wealth. The persistence of the expectational errors in the learning rule of the agents is the fundamental reason for agents to alternately be either too optimistic or too pessimistic about the consequences of a pre-announced tax change.

These learning dynamics, despite the simplicity of the model, have the potential to capture important features of the empirical evidence on how the economy responds to policy changes. However, we believe that future research in this area needs to come up with more empirical evidence on whether or how agents learn about fiscal policy.

Apart from this, our sensitivity analyses show that a smaller gain parameter leads to higher volatility. This result is at odds with conventional wisdom about the link between the gain parameter and the dynamic responses in the literature and requires further investigation as well.

Finally, an exemplary tax reform indicates that the magnitude of welfare improvements appears to be substantially lower under the assumption of learning compared to the case of perfect foresight. This can be explained by the oscillations under learning. Thus, the learning perspective on the tax reform considered herein has a more general implication for benevolent policy makers. Tax reforms may not lead to a considerable aggregate welfare improvement compared to the status quo. In consequence, other criteria, like the distributional consequences of tax reforms in presence of heterogeneity, may become more important in the design of such tax reforms.

A. Model Derivations

A.1. Derivation of SW_2

We start from (15), i.e. $\sum_{j=1}^{\infty} \frac{1}{D_{t,t+j}^{k,e}(t)} w_{t+j}^e(t) \bar{L}$. Next, we recall the definition of $D_{t,t+j}^{k,e}(t)$. Given the learning rule (11) we get

$$D_{t,t+j}^{k,e}(t) = \begin{cases} \prod_{i=1}^j [(1-\delta) + (1-\tau_0^k)r^e(t)] = [(1-\delta) + (1-\tau_0^k)r^e(t)]^j & \text{for } \tau_{t+j}^{k,e}(t) = \tau_0^k \\ \prod_{i=1}^j [(1-\delta) + (1-\tau_1^k)r^e(t)] = [(1-\delta) + (1-\tau_1^k)r^e(t)]^j & \text{for } \tau_{t+j}^{k,e}(t) = \tau_1^k \end{cases} \quad (\text{A.1.1})$$

for $\tau_{t+j}^{k,e}(t) = \tau_1^k$. Thereafter, we split this infinite sum into

$$\begin{aligned} SW_2 &= \bar{L} \left[\sum_{j=1}^{T-1} \frac{1}{D_{t,t+j}^{k,e}(t)} w^e(t) + \sum_{j=T}^{\infty} \frac{1}{D_{t,t+j}^{k,e}(t)} w^e(t) \right] \\ &= \bar{L} \left[\sum_{j=1}^{T-1} ([(1-\delta) + (1-\tau_0^k)r^e(t)]^j)^{-1} w^e(t) + \sum_{j=T}^{\infty} ([(1-\delta) + (1-\tau_1^k)r^e(t)]^j)^{-1} w^e(t) \right], \end{aligned}$$

or

$$\begin{aligned} SW_2 &= \frac{w^e(t)\bar{L}}{[(1-\delta) + (1-\tau_0^k)r^e(t)]} \sum_{j=0}^{T-2} \left([(1-\delta) + (1-\tau_0^k)r^e(t)]^{-1} \right)^j \\ &\quad + \frac{w^e(t)\bar{L}}{[(1-\delta) + (1-\tau_1^k)r^e(t)]} \sum_{j=T-1}^{\infty} \left([(1-\delta) + (1-\tau_1^k)r^e(t)]^{-1} \right)^j. \end{aligned}$$

Now we exploit the properties of a geometric series, e.g. $\sum_{j=m}^n f^j = \frac{f^{n+1}-f^m}{f-1}$ for some constant f , and derive

$$SW_2 = \frac{w^e(t)\bar{L}}{[(1-\delta) + (1-\tau_0^k)r^e(t)]} \left(\frac{1 - [(1-\delta) + (1-\tau_0^k)r^e(t)]^{1-T}}{1 - [(1-\delta) + (1-\tau_0^k)r^e(t)]^{-1}} \right) \\ + \frac{w^e(t)\bar{L}}{[(1-\delta) + (1-\tau_1^k)r^e(t)]} \left(\frac{[(1-\delta) + (1-\tau_1^k)r^e(t)]^{1-T}}{1 - [(1-\delta) + (1-\tau_1^k)r^e(t)]^{-1}} \right).$$

When we respect the timing outlined in the experiment above, we get (15)

$$SW_2 = \begin{cases} \frac{w^e(t)\bar{L}}{[(1-\tau_0^k)r^e(t)-\delta]} + w^e(t)\bar{L} \left[\frac{[(1-\delta)+(1-\tau_1^k)r^e(t)]^{t-T_p}}{1-[(1-\delta)+(1-\tau_1^k)r^e(t)]^{-1}} - \frac{[(1-\delta)+(1-\tau_0^k)r^e(t)]^{t-T_p}}{1-[(1-\delta)+(1-\tau_0^k)r^e(t)]^{-1}} \right] & \text{for } 1 \leq t < T_p \\ \frac{w^e(t)\bar{L}}{[(1-\tau_1^k)r^e(t)-\delta]} & \text{for } t \geq T_p. \end{cases}$$

A.2. Derivation of ST_2

Starting from (16), i.e. $\sum_{j=1}^{\infty} \frac{1}{D_{t,t+j}^{k,e}} \tau_{t+j}^{l,e}(t) w_{t+j}^e(t) \bar{L}$, for (A.1.1) and $\tau_{t+j}^{l,e}(t)$ is either given by τ_0^l or τ_1^l , we may again split the infinite sum into

$$ST_2 = w^e(t)\bar{L} \times \\ \left[\sum_{j=1}^{T-1} \left([(1-\delta) + (1-\tau_0^k)r^e(t)]^j \right)^{-1} \tau_0^l + \sum_{j=T}^{\infty} \left([(1-\delta) + (1-\tau_1^k)r^e(t)]^j \right)^{-1} \tau_1^l \right], \\ = \frac{\tau_0^l w^e(t)\bar{L}}{[(1-\delta) + (1-\tau_0^k)r^e(t)]} \sum_{j=0}^{T-2} \left([(1-\delta) + (1-\tau_0^k)r^e(t)]^{-1} \right)^j \\ + \frac{\tau_1^l w^e(t)\bar{L}}{[(1-\delta) + (1-\tau_1^k)r^e(t)]} \sum_{j=T-1}^{\infty} \left([(1-\delta) + (1-\tau_1^k)r^e(t)]^{-1} \right)^j.$$

Now, the properties of the geometric series allow us to rewrite this as

$$ST_2 = \frac{\tau_0^l w^e(t) \bar{L}}{[(1-\delta) + (1-\tau_0^k)r^e(t)]} \left(\frac{[(1-\delta) + (1-\tau_0^k)r^e(t)]^{1-T} - 1}{[(1-\delta) + (1-\tau_0^k)r^e(t)]^{-1} - 1} \right) \\ + \frac{\tau_1^l w^e(t) \bar{L}}{[(1-\delta) + (1-\tau_1^k)r^e(t)]} \left(\frac{-[(1-\delta) + (1-\tau_1^k)r^e(t)]^{1-T}}{[(1-\delta) + (1-\tau_1^k)r^e(t)]^{-1} - 1} \right).$$

Given the timing outlined above, we get (16)

$$ST_2 = \begin{cases} \frac{\tau_0^l w^e(t) \bar{L}}{[(1-\tau_0^k)r^e(t) - \delta]} + w^e(t) \bar{L} \left[\frac{\tau_1^l [(1-\delta) + (1-\tau_1^k)r^e(t)]^{t-T_p}}{1 - [(1-\delta) + (1-\tau_1^k)r^e(t)]^{-1}} - \frac{\tau_0^l [(1-\delta) + (1-\tau_0^k)r^e(t)]^{t-T_p}}{1 - [(1-\delta) + (1-\tau_0^k)r^e(t)]^{-1}} \right] & \text{for } 1 \leq t < T_p \\ \frac{\tau_1^l w^e(t) \bar{L}}{[(1-\tau_1^k)r^e(t) - \delta]} & \text{for } t \geq T_p. \end{cases}$$

A.3. Derivation of ST_3

Starting from (17), i.e. $\sum_{j=1}^{\infty} \frac{1}{D_{t,t+j}^{k,e}} \tau_{t+j}^e(t)$, given (A.1.1) and $\tau_{t+j}^e(t)$ is either τ_0 or τ_1 , we again split the infinite sum into

$$ST_3 = \left[\sum_{j=1}^{T-1} \left([(1-\delta) + (1-\tau_0^k)r^e(t)]^j \right)^{-1} \tau_0 + \sum_{j=T}^{\infty} \left([(1-\delta) + (1-\tau_1^k)r^e(t)]^j \right)^{-1} \tau_1 \right], \\ = \left[(1-\delta) + (1-\tau_0^k)r^e(t) \right]^{-1} \left[\sum_{j=0}^{T-2} \left([(1-\delta) + (1-\tau_0^k)r^e(t)]^{-1} \right)^j \tau_0 \right] \\ + \left[(1-\delta) + (1-\tau_1^k)r^e(t) \right]^{-1} \left[\sum_{j=T-1}^{\infty} \left([(1-\delta) + (1-\tau_1^k)r^e(t)]^{-1} \right)^j \tau_1 \right].$$

Given the properties of geometric series we can rewrite the latter as

$$ST_3 = \left[(1-\delta) + (1-\tau_0^k)r^e(t) \right]^{-1} \left(\frac{[(1-\delta) + (1-\tau_0^k)r^e(t)]^{1-T} - 1}{[(1-\delta) + (1-\tau_0^k)r^e(t)]^{-1} - 1} \tau_0 \right) \\ + \left[(1-\delta) + (1-\tau_1^k)r^e(t) \right]^{-1} \left(\frac{-[(1-\delta) + (1-\tau_1^k)r^e(t)]^{1-T}}{[(1-\delta) + (1-\tau_1^k)r^e(t)]^{-1} - 1} \tau_1 \right).$$

Once more, given the timing outlined above, we get (17)

$$ST_3 = \begin{cases} \frac{\tau_0}{[(1-\tau_0^k)r^e(t)-\delta]} + \left[\frac{[(1-\delta)+(1-\tau_1^k)r^e(t)]^{t-T_p}}{1-[(1-\delta)+(1-\tau_1^k)r^e(t)]^{-1}} \tau_1 - \frac{[(1-\delta)+(1-\tau_0^k)r^e(t)]^{t-T_p}}{1-[(1-\delta)+(1-\tau_0^k)r^e(t)]^{-1}} \tau_0 \right] & \text{for } 1 \leq t < T_p \\ \frac{\tau_1}{[(1-\tau_1^k)r^e(t)-\delta]} & \text{for } t \geq T_p. \end{cases}$$

B. Computing Welfare Changes

B.1. Comparative Statics

We follow the approach of Cooley and Hansen (1992, p.301ff.). Their measure of welfare change for a given policy change is derived by solving

$$U_0 = \log[c_1(1+x^\bullet)] + \eta \log[1-n_1] + \phi \log[g_1] \quad (\text{B.1.1})$$

for x^\bullet .²⁵ U_0 is the utility a household obtains in the steady-state without any tax. c_1 and n_1 are the values of consumption and employment at the new steady-state after the tax change either under perfect foresight or learning. It follows that

$$x^\bullet = \frac{\exp(U_0)}{c_1(1-n_1)^\eta g_1^\phi} - 1. \quad (\text{B.1.2})$$

Thus, in general, we need to solve for x for the perfect foresight dynamics and another x^* for the dynamics under learning.²⁶ Given x^\bullet we can calculate

$$\bar{\mathcal{W}} = \frac{\Delta C}{y_1} = \frac{x^\bullet c_1}{y_1}, \quad (\text{B.1.3})$$

where ΔC is the restoration value of consumption, which in our case may be interpreted as the total change in consumption required to restore a household

²⁵ x^\bullet is either x under perfect foresight or x^* under learning.

²⁶This must yield the same $x = x^*$ both under perfect-foresight and under learning, but this number may be useful to compare different policy experiments.

to the level of utility obtained under the allocation associated with zero taxes. y_1 is the level of output at the new steady-state.

B.2. Transition Measure

Again we follow the approach of Cooley and Hansen (1992, p.301ff.) based on Lucas (1990). Their measure of welfare change accounting for transition given a policy change is derived by solving

$$\sum_{t=1}^{\infty} \beta^t \{ \log[c_t(1 + x^\bullet)] + \eta \log[1 - n_t] + \phi \log[g_t] - U_0 \} \stackrel{!}{=} 0 \quad (\text{B.2.1})$$

for x under perfect foresight and x^\bullet under learning. c_t , n_t , g_t , and y_t are period t consumption, employment, government spending and output respectively, either under perfect foresight or learning.

$$x^\bullet = \left[\frac{\exp\left(\frac{\beta}{(1-\beta)}[U_0 - \beta^T U_1]\right)}{\prod_{t=1}^T c_t^{\beta^t} \times \prod_{t=1}^T (1 - n_t)^{\eta\beta^t} \times \prod_{t=1}^T g_t^{\phi\beta^t}} \right]^{\frac{(1-\beta)}{\beta}} - 1. \quad (\text{B.2.2})$$

Hereby T is the terminal period of the simulation. For the $T < t \leq \infty$ it is assumed that agents' period utility is approximately given by U_1 , i.e. all variables involved are close to their new steady-state value. Thus, T needs to be sufficiently large such that x^\bullet does no longer change significantly. Given x^\bullet we can calculate

$$\mathcal{W}^\bullet = \frac{\sum_{t=1}^T \beta^t \{x^\bullet c_t\}}{\sum_{t=1}^T \beta^t \{y_t\}}, \quad (\text{B.2.3})$$

which will be reported as \mathcal{W} for the perfect foresight dynamics and as \mathcal{W}^* for the dynamics under learning.

References

- Ambler, S. and Paquet, A. (1996). Fiscal Spending Shocks, Endogenous Government Spending, and Real Business Cycles. *Journal of Economic Dynamics and Control*, 20(1-3):237–256.
- Branch, W. A. and Evans, G. W. (2007). Model Uncertainty and Endogenous Volatility. *Review of Economic Dynamics*, 10(2):207–237.
- Chamley, C. (1981). The Welfare Cost of Capital Income Taxation in a Growing Economy. *The Journal of Political Economy*, 89(3):468–496.
- Chetty, R., Guren, A., Manoli, D., and Weber, A. (2012). Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities. *NBER Macroeconomics Annual*, 27.
- Cooley, T. F. and Hansen, G. D. (1992). Tax Distortions in a Neoclassical Monetary Economy. *Journal of Economic Theory*, 58(2):290–316.
- Evans, G. W., Guesnerie, R., and McGough, B. (2011a). Eductive Stability in Real Business Cycle Models. *Working Paper*.
- Evans, G. W. and Honkapohja, S. (2001). *Learning and Expectations in Macroeconomics*. Frontiers of Economic Research. Princeton University Press, Princeton, NJ.
- Evans, G. W. and Honkapohja, S. (2009). Robust Learning Stability with Operational Monetary Policy Rules. In Schmidt-Hebbel, K. and Walsh, C. E., editors, *Monetary Policy Under Uncertainty and Learning*, volume XIII of *Series on Central Banking, Analysis, and Economic Policies*, pages 145–170, Santiago, Chile. Central Bank of Chile.
- Evans, G. W., Honkapohja, S., and Mitra, K. (2009). Anticipated Fiscal Policy and Adaptive Learning. *Journal of Monetary Economics*, 56(7):930–953.
- Evans, G. W., Honkapohja, S., and Mitra, K. (2010). Does Ricardian Equivalence Hold When Expectations are not Rational? *Centre For Dynamic Macroeconomic Analysis Working Paper Series*, 10/08.
- Evans, G. W., Honkapohja, S., and Mitra, K. (2011b). Policy Change and Learning in the RBC Model. *Bank of Finland Discussion Papers*, 22.
- Farmer, R. E. A. (1999). *Macroeconomics of Self-fulfilling Prophecies*. MIT Press, Second edition.
- Garcia-Milà, T., Marcet, A., and Ventura, E. (2010). Supply Side Interventions and Redistribution. *The Economic Journal*, 120(543):105–130.

- Giannitsarou, C. (2007). Balanced Budget Rules and Aggregate Instability: The Role of Consumption Taxes. *The Economic Journal*, 117(523):1423–1435.
- Judd, K. L. (1987). The Welfare Cost of Factor Taxation in a Perfect-Foresight Model. *The Journal of Political Economy*, 95(4):675–709.
- Leeper, E. M. (2009). Anchoring Fiscal Expectations. *NBER Working Paper*, 15269.
- Leeper, E. M., Walker, T. B., and Yang, S.-C. S. (2009). Fiscal Foresight and Information Flows. *NBER Working Paper*, 14630.
- Ljungqvist, L. and Sargent, T. J. (2004). *Recursive Macroeconomic Theory*. MIT Press, Cambridge, MA, Second edition.
- Lucas, Jr., R. E. (1990). Supply-Side Economics: An Analytical Review. *Oxford Economic Papers*, 42(2):293–316.
- Milani, F. (2007). Expectations, Learning and Macroeconomic Persistence. *Journal of Monetary Economics*, 54(7):2065–2082.
- Milani, F. (2011). Expectation Shocks and Learning as Drivers of the Business Cycle. *The Economic Journal*, 121(552):379–401.