CENTER FOR FINANCIAL STUDIES

CFS Working Paper Series

No. 614

Christos Koulovatianos and Dimitris Mavridis

Increasing Taxes After a Financial Crisis: Not a Bad Idea After All...

Center for Financial Studies | Goethe University | Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de

The CFS Working Paper Series

presents ongoing research on selected topics in the fields of money, banking and finance. The papers are circulated to encourage discussion and comment. Any opinions expressed in CFS Working Papers are those of the author(s) and not of the CFS.

The Center for Financial Studies, located in Goethe University Frankfurt's House of Finance, conducts independent and internationally oriented research in important areas of Finance. It serves as a forum for dialogue between academia, policy-making institutions and the financial industry. It offers a platform for top-level fundamental research as well as applied research relevant for the financial sector in Europe. CFS is funded by the non-profit-organization Gesellschaft für Kapitalmarktforschung e.V. (GfK). Established in 1967 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia. GfK members comprise major players in Germany's financial industry. The funding institutions do not give prior review to CFS publications, nor do they necessarily share the views expressed therein.

Center for Financial Studies | Goethe University | Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de

Increasing Taxes After a Financial Crisis: Not a Bad Idea After All...

Christos Koulovatianos[a,b,*] and Dimitris Mavridis[c]

University of Luxembourg

October 28, 2018

[a] Department of Economics, University of Luxembourg

[b] Center for Financial Studies (CFS), Goethe University Frankfurt

[c] European Commission, Joint Research Centre

* Corresponding author: Department of Economics, University of Luxembourg, 162A avenue de la Faïencerie, Campus Limpertsberg, BRA 3.05, L-1511, Luxembourg, Email: christos.koulovatianos@uni.lu, Tel.: +352-46-66-44-6356, Fax: +352-46-66-44-6341. We have been helped by discussions with Günter Beck, Aleksander Berentsen, Gautam Tripathi, Volker Wieland, and participants in seminars and conferences at various places.

Increasing Taxes After a Financial Crisis: Not a Bad Idea After All...

Abstract

Based on OECD evidence, equity/housing-price busts and credit crunches are followed by substantial increases in public consumption. These increases in unproductive public spending lead to increases in distortionary marginal taxes, a policy in sharp contrast with presumably optimal Keynesian fiscal stimulus after a crisis. Here we claim that this seemingly adverse policy selection is optimal under rational learning about the frequency of rare capital-value busts. Bayesian updating after a bust implies massive belief jumps toward pessimism, with investors and policymakers believing that busts will be arriving more frequently in the future. Lowering taxes would be as if trying to kick a sick horse in order to stand up and run, since pessimistic markets would be unwilling to invest enough under any temporarily generous tax regime.

Keywords: Bayesian learning, controlled diffusions and jump processes, learning about jumps, Gamma distribution, rational learning

JEL classification: H30, D83, C11, D90, E21, D81, C61

1. Introduction

The 2008 credit crunch has reminded that optimal fiscal policy after a crisis is a poorly understood subject. Looking at empirical facts, Table 1 shows a post-crisis fiscal-policy regularity in OECD countries that can be considered as unconventional: government consumption increases after credit crunches and asset-price/house-price busts.[1] The massive loss in capital value brought by such severe incidents affects the productive capacity of the private sector. So, common intuition suggests that governments should avoid spending in unproductive public consumption which further implies higher distortionary marginal tax rates.[2] An increase in marginal tax rates seems to be in sharp contrast with the presumably optimal Keynesian policy of tax cuts in order to stimulate demand during a crisis.[3]

Δ(Government Consumption): ΔG (%)		
Event	**Ordinary**	**Severe (Bust/Crunch)**
Recession	1.79	2.16
Credit Contraction	2.83	5.98***
House-price decline	3.39	8.75***
Equity-price declines	3.59	7.48***

Table 1 - Data (Claessens, Kose, Terrones, EP 2009): 21 OECD countries from 1960-2007

[1] Table 1 contains selected values from Claessens et al. (2009, Table 9, p. 685), reporting medians of peak-to-trough changes in government consumption corresponding to each category of crunch/bust/recession event. Symbol ΔG denotes changes in government consumption, while symbol "***" indicates significance at the 1% level. Table A.1 in Appendix A provides some more estimates based on Claessens et al. (2009).

[2] The UK experience after the 2008 global financial crisis is one of the most characteristic examples of the implied tax-burden increases. As reported in Institute for Fiscal Studies (2013, Figure 5.8, p. 132), although drastic tax cuts had been attempted in fiscal years 2008 and 2009, since 2011 fiscal consolidation and tax increases have been far more drastic and dramatic.

[3] In 2009, the International Monetary Fund (IMF) announced that most countries in the Group of Twenty (G-20) had attempted discretionary Keynesian-type fiscal-stimulus policies, including drastic tax cuts, see International Monetary Fund (2009, p. 26) and several other references in that report.

We pose two questions. First, is fiscal stimulus through tax cuts an optimal policy after a capital-bust crisis? Second, is the message of Table 1 toward the direction of an optimal fiscal policy, or do we need to resort to any complicated political-economy mechanics in order to explain it? We argue that limited information, not limited rationality, about the frequency of capital-value busts can corroborate that increasing taxes after a bust incident is part of an optimal policy. Specifically, we demonstrate that after a capital bust, a pessimistic spell pervades markets. For a long period, investors think that capital busts will be arriving more frequently in the future. Investments become weak, due to the belief that part of each dollar invested will end up destroyed more frequently in the future. After a bust, benevolent policymakers shift to providing public goods because fiscal stimulus is ineffective in such a pessimistic weak-investment environment. We argue that this intuition is dominant using a model of endogenous growth with especially strong permanent negative effects of taxation on future investments and with especially clear analytics concerning optimal-fiscal policy (see Xie, 1997). Our main message is that pessimism and weak investment after a capital bust may dominate any fiscal-stimulus effects, and that fiscal-stimulus should be reconsidered during such periods.

We build an endogenous-growth model with busts (disasters) in capital's productive-capacity, driven by a Poisson process. We assume that the disaster-frequency parameter of the Poisson process is not known with perfect confidence by any agents in the model. So, both the private sector and welfare-maximizing policy makers must learn about this disaster-frequency parameter. A virtual econometrician in the model collects and processes data which are records of the dates and the cumulative count of past disasters. We obtain a result that simplifies such a Bayesian-learning analysis: we prove that this data combined with knowledge that the data-generating process is Poisson, implies that prior beliefs about the disaster-frequency parameter are Gamma-distributed. Using Bayes' rule, the virtual econometrician finds that posterior beliefs about the disaster-frequency parameter are also Gamma-distributed. So, such a Bayesian-learning setup can become a handy and tractable modeling ingredient.[4]

The subjective expectation of the disaster-frequency parameter plays a key role in both individual decision-making and in optimal-policy setting. This subjective expectation appears directly in all Hamilton-Jacobi-Bellman equations of decision making, and possesses intuitive dynamics during the process of learning. In particular, after a disaster occurs, Bayes' rule implies that average disaster-frequency beliefs jump to pessimistic levels. This new level of pessimism persists for several years, and optimism is rebuilt gradually during periods that no other disaster occurs, until the arrival of the next disaster.[5] These endogenous spells of pessimism that follow each disaster episode are what makes the study of optimal fiscal policy after capital-value busts special: in a pessimistic environment investment decisions may be lukewarm under most policy regimes, as tax cuts may not be as effective as they would be during times of non-extreme beliefs. To understand the interplay between elevated rare-disaster pessimism and policy effectiveness is the main goal of this paper.

We select an appropriate vehicle in order to perform a thought experiment that pushes the concept of distortionary taxation towards its limits. We choose the "AK" endogenous-growth model which emphasizes the permanent effect of any period's tax distortion on capital

[4] A study which assumes Gamma-distributed jump-frequency-parameter prior beliefs in the context of an asset-pricing model is Comon (2001). While Comon (2001) *assumes* such a prior distribution (perhaps due to the facilitating fact that these prior and posterior distributions are in the same family, i.e., conjugate), we *prove* that jump-frequency-parameter prior beliefs *are necessarily Gamma-distributed* under our sampling assumptions.

[5] Survey evidence using subjective statements regarding the likelihood of an imminent stock-market crash after the year-2000 dot-com bust and the year-2008 Lehman-Brothers default, corroborates that pessimistic disaster expectations followed these crash episodes. This survey database is the Crash-Confidence Index which is constructed using the survey method in Shiller et al. (1996) – for example, see the plot in Koulovatianos and Wieland (2011, Figure 2).

accumulation.[6] We introduce a public-consumption externality in the utility function in an additively-separable way, in order to disjoin any direct impact of public goods on the marginal utility of private consumption at any time. Within this framework we study private-sector economic decisions under rational learning. Subject to these private-sector decisions and competitive market clearing, we study the optimal provision of public consumption by a benevolent rational-learning planner who levies marginal income taxes.

Beyond mere-technicality issues, we assume logarithmic momentary utility functions. In the deterministic version of our setup, Xie (1997, pp. 416-9) has demonstrated that the optimal fiscal policy (open-loop policy with commitment) is time-consistent. Time consistency of optimal policy survives in our stochastic model with rational learning, offering analytical tractability. In a numerically solved deterministic setup which is similar to ours (yet, without endogenous growth), Klein et al. (2008) report quantitative differences in optimal time-inconsistent (open-loop) vs. time-consistent (closed-loop) fiscal policies.[7] Instead, the functional forms we use in our model make these two types of fiscal policies to coincide. So, our analysis clarifies the *qualitative responses* of these fiscal-policy types under the optimism/pessimism swinging in a rational-learning environment with disaster risk.[8] Finally, as in Klein et al. (2008), in the interest of simplicity and tractability, we do not allow any government to issue fiscal debt, given that our approach is one of setting the optimal size of government and not the way of financing exogenous government spending.[9]

[6] Rebelo (1991) demonstrates that the absence of productivity externalities in aggregate production makes capital taxation to have an excessively high negative imact on welfare through capital-accumulation distortions in endogenous-growth environments with "AK" type of production technology.

[7] In addition, Klein and Rios-Rull (2003) report quantitative differences in the cyclical responses of time-inconsistent (open-loop) vs. time-consistent (closed-loop) fiscal policies in a setup that studies the optimal financing of useless government spending.

[8] Examining the quantitative differences between time-consistent and time-inconsistent fiscal policies under learning in more general setups is an interesting, yet demanding extention, which is not pursued here.

[9] Dealing with both the choice of government size and a mixture of financing government spending can lead to a government maximization problem that is not well defined. In order to avoid such technical problems and policy indeterminacy, see the relevant discussion in Klein et al. (2008, p. 804), and especially the two papers of Stockman (2001, 2004). Two recent papers that study the problem of how to optimally finance

The key message revealed by our setup is sharp. After a disaster has occurred, pessimistic beliefs, that rare disasters are more frequent than thought before, prevail. During such periods of elevated pessimism optimal-policy setting under rational learning implies higher public-consumption provision and higher distortionary taxes. This happens because private-capital accumulation is expected to be less profitable than before the disaster, and investment rates are low. So, the policy maker appreciates that the responsiveness of private-sector's investment to marginal-tax changes is weakened by this type of disaster-frequency pessimism. Hence, at times of disaster-expectations pessimism there is a weakening of the distortionary effect of marginal taxation, which motivates increasing marginal tax rates at exactly these times in our model.

Our analysis corroborates that reducing tax rates at times after a disaster has occurred is not necessarily effective for stimulating investment. In particular, given the severely distortionary nature of marginal taxes in "AK"-type endogenous growth models, most models with capital accumulation would support what our optimal-policy experiment suggests: tax cuts in periods of elevated disaster-risk pessimism may not be the optimal policy to follow.

2. Bayesian Learning about the Jump Frequency of Poisson-Driven Jumps

Before we proceed with the formal description of the model, we state and prove a key result regarding Bayesian learning about jumps which are generated by Poisson processes. Specifically, consider that in a continuous-time economic model the Poisson process, $q(t)$,

exogenously given government spending in overlapping-generations models that consider fiscal debt as well are Conesa et al. (2009) and Fehr and Kindermann (2015).

characterized by,

$$dq(t) = \begin{cases} \nu & \text{with Probability } \lambda^* dt \\ 0 & \text{with Probability } 1 - \lambda^* dt \end{cases}, \tag{1}$$

with $\lambda^* > 0$, drives the occurrence of jumps of size $\nu \neq 0$, which apply to some of the model's variables. The Poisson process $q(t)$ is independent from other exogenous random variables in the model. The jump size ν can be an independent random variable, too. So, at any particular time $t \geq 0$, $q(t) = \sum_i \mathbf{1}_{\{\hat{t}_i \leq t\}} \nu$ in which $\hat{t}_i = \{t \geq 0 \mid dq(t) \neq 0\}$, and $\mathbf{1}_{\{\hat{t}_i \leq t\}} = \{1, \text{ if } s = \hat{t}_i \text{ ; } 0, \text{ if } s \neq \hat{t}_i \text{ ; for all } s \leq t\}$.

We assume that all decision makers know all parameters of the model except λ^*. In order to deal with this parameter uncertainty let's assume that the economy is equipped with an invisible virtual econometrician who collects the following jump data: dates at which jumps occur. This data-collection assumption seems minimal, yet there is no more raw information one can collect.[10] If the virtual econometrician assumes that the data-generating process is Poisson, then the jump-frequency parameter is indeed the one and only parameter to learn about.[11]

In this section we demonstrate an explicit and tractable result. If the virtual econometrician indeed makes the Poisson assumption about the data-generating process and interprets her track record of jump data through the Poisson lens, then there is a necessary implication: the virtual econometrician's beliefs about the jump-frequency parameter can only be Gamma-distributed.

[10]See, for example, Claessens et al. (2009), and Barro and Ursua (2009) who collect cross-country data on rare disasters. That database consists of disaster dates and also jump magnitudes. Here we assume that the random process driving the magnitude of disasters is both known and independent from the Poisson process which governs the random occurrence dates of disasters.

[11]Ross (2003, p. 275) demostrantes that the concept of memorylessness coincides with the concept of exponentially-distributed arrival times of discrete events with a single constant parameter driving the associated density function. So, the virtual econometrician's Poisson assumption is essentially the assumption that the data-generating process is memoryless, i.e., this process self regenerates the same forward probabilities of future arrival times as actual time passes by.

The two parameters of the Gamma distribution of the disaster-frequency-parameter beliefs are explicitly specified, and have a natural interpretation. The one Gamma-distribution parameter is the total time elapsed since sample collection started.[12] The other Gamma-distribution parameter is the total cumulative count of past jumps since sampling started. Certainly, what makes the jump-frequency-parameter perceptions a random variable is the randomness of any collected data sample. Here, sharply, both Gamma-distribution parameters perfectly describe the track-record data sampled.

A well-known result is that the Bayesian posterior of a Gamma prior is also Gamma. This statistical conjugacy leads to recursive tractability in models of decision making. Both this tractability and the intuitiveness of the underlying assumptions that imply Gamma-distributed disaster-frequency-parameter beliefs, make this section's learning setup a natural and promising building block for studying Bayesian learning about rare disasters in a wide range of decision-making problems.

2.1 Data collection specification and implied beliefs

Without loss of generality, consider that the time instant at which the virtual econometrician starts keeping a track record is $t = 0$. Let any integer $n \geq 1$, and denote by T_n the arrival time of the n-th jump. Let also ΔT_i denote the time elapsed between the $(i-1)$-th and the i-th jump for all $i \in \{2, ..., n\}$. Assume that T_1 is known and fixed, let $\Delta T_1 \equiv T_1 - 0 = T_1$ by convention, and notice that,

$$T_n = \sum_{i=1}^{n} \Delta T_i \ . \tag{2}$$

Assumption 1 *For all $t \geq 0$, the virtual econometrician keeps a track record of all inter-arrival times $\{\Delta T_i\}_{i=1}^{n}$ if $n \geq 1$.*

[12] While it is not necessary that sampling started with the first actual jump observation, it is necessary to have at least one jump observation in the sample in order that the Gamma distribution be well-defined.

Assumption 1 formally states that the virtual econometrician keeps track of dates at which disasters occur. In addition, the virtual econometrician assumes that the data-generating process is Poisson. This Poisson perception is equivalent to stating that the virtual econometrician thinks that inter-arrival times, $\{\Delta T_i\}_{i=1}^n$, are exponentially distributed.[13] Let's state this distributional perception formally and define what exactly the econometrician learns about.

Assumption 2 *For all $t \geq 0$, the virtual econometrician thinks that all inter-arrival times $\{\Delta T_i\}_{i=1}^n$ if $n \geq 1$, are exponentially distributed but does not know this exponential distribution's true associated parameter λ. The non-informative priors of the virtual econometrician are Haldane's agnostic priors. On the contrary, the virtual econometrician knows the probability distribution of disaster magnitudes, ζ, and knows that ζ is independent from the exponential process that generates all disaster inter-arrival times.*

Let the virtual econometrician focus on some certain level of the exponential distribution's parameter, say $\tilde{\lambda} > 0$.[14] Naturally, the virtual econometrician wants to examine the likelihood that $\tilde{\lambda}$ reflects the inter-arrival-time data, $\{\Delta T_i\}_{i=1}^n$, accurately. The rest of this section builds on the analysis of Ross (2003, pp. 293-4) and focuses on this specific concern: what is the probability that $\tilde{\lambda}$ perfectly explains all available data $\{\Delta T_i\}_{i=1}^n$?[15]

So, fix any subjective parameter level $\tilde{\lambda} > 0$ and under the condition that nature's true parameter $\lambda = \tilde{\lambda}$, define the point process which counts any past and future jumps (counting

[13]For a formal proof of this statement see, for example, Ross (2003, Proposition 5.1 p. 293).
[14]Throughout, a tilde denotes subjective perception of a certain variable.
[15]Ross (2003, pp. 293-4) derives the probability distribution of waiting time until the n-th jump, which is different from the question we pose here. Yet, we borrow some solution arguments from Ross (2003, pp. 293-4) in order to build our proofs.

process),

$$N\left(t\right) \equiv \#\left\{ \hat{t} \in [0,t] \;\middle|\; d\tilde{q}\left(\hat{t}\right)|_{\lambda=\tilde{\lambda}} \neq 0 \right\} \;\; , \quad \text{for all } t \geq 0 \; . \tag{3}$$

in which

$$d\tilde{q}\left(t\right)|_{\lambda=\tilde{\lambda}} = \begin{cases} -\zeta & \text{with Probability } \tilde{\lambda}dt \\ 0 & \text{with Probability } 1-\tilde{\lambda}dt \end{cases} \; .$$

Under the convention that $\lambda = \tilde{\lambda}$, and since the Poisson process has independent increments, all time intervals ΔT_i are independent from each other, and each interval ΔT_i is exponentially distributed with parameter $\tilde{\lambda} > 0$. Given this independent-increments property, within a time interval $[0,t]$ for any $t \geq 0$, the average frequency of disasters is $\tilde{\Lambda} = \tilde{\lambda} \cdot t$. So, by the definition of the Poisson distribution and (3),

$$\Pr\left\{N\left(t\right) \geq n\right\} = \sum_{i=n}^{\infty} e^{-\tilde{\Lambda}} \frac{\tilde{\Lambda}^i}{i!} = \sum_{i=n}^{\infty} e^{-\tilde{\lambda}t} \frac{\left(\tilde{\lambda}t\right)^i}{i!} \;\; , \;\; n = 1, 2, ..., \text{ for all } t \geq 0 \; . \tag{4}$$

Notice that definition (2) implies

$$\Pr\left\{N\left(t\right) \geq n\right\} = \Pr\left\{T_n \leq t\right\} \;\; , \;\; n = 1, 2, ..., \text{ for all } t \geq 0 \; . \tag{5}$$

Yet,

$$T_n \leq t \Leftrightarrow \tilde{\lambda} \cdot T_n \leq \tilde{\lambda} \cdot t \Leftrightarrow \frac{\tilde{\lambda} \cdot T_n}{t} \leq \tilde{\lambda} \;\; , \;\; n = 1, 2, ..., \text{ for all } t > 0 \; . \tag{6}$$

The term $\left(\tilde{\lambda} \cdot T_n\right)/t$ is the point estimate of the disaster frequency that is induced by the cumulative history of collected data, T_n, at any point in time $t > 0$, under the assumption that the data generating process is driven by parameter $\tilde{\lambda}$. Notice that (5) and (6) imply,

$$\Pr\left\{N\left(t\right) \geq n\right\} = \Pr\left\{\frac{\tilde{\lambda} \cdot T_n}{t} \leq \tilde{\lambda}\right\} \equiv F_{T_n}\left(\tilde{\lambda}\right) \;\; , \tag{7}$$

in which $F_{T_n}\left(\tilde{\lambda}\right)$ is the data-induced cumulative distribution function (c.d.f.) of $\tilde{\lambda}$. With these observations at hand, we can prove the main result of this section which is stated by Theorem 1.

Theorem 1 *Under Assumptions 1 and 2, for all $t > 0$, the probability that some $\tilde{\lambda}^o > 0$ perfectly explains all available data is given by the density of a Gamma distribution with parameters t (the elapsed time from the beginning of sampling), and $N(t)$. Consequently, disaster-frequency-parameter beliefs of a virtual econometrician follow this Gamma distribution.*

Proof

From (7) and (4) it is,

$$F_{T_n}\left(\tilde{\lambda}\right) = \sum_{i=n}^{\infty} e^{-\tilde{\lambda}t} \frac{\left(\tilde{\lambda}t\right)^i}{i!} .$$

Yet, since $F_{T_n}\left(\tilde{\lambda}\right)$ is a c.d.f.,

$$\Pr\left\{\tilde{\lambda} = \tilde{\lambda}^o\right\} = F'_{T_n}\left(\tilde{\lambda}\right)|_{\tilde{\lambda}=\tilde{\lambda}^o} = f_{T_n}\left(\tilde{\lambda}^o\right) ,$$

in which $f_{T_n}\left(\tilde{\lambda}^o\right)$ is the data-induced density function conditional upon n observed past disasters. In particular, for $t \geq T_1$,

$$\begin{aligned}
f_{T_n}\left(\tilde{\lambda}^o\right) &= F'_{T_n}\left(\tilde{\lambda}\right)|_{\tilde{\lambda}=\tilde{\lambda}^o} \\
&= -t\sum_{i=n}^{\infty} e^{-\tilde{\lambda}^o t} \frac{\left(\tilde{\lambda}^o t\right)^i}{i!} + t\sum_{i=n}^{\infty} e^{-\tilde{\lambda}^o t} \frac{\left(\tilde{\lambda}^o t\right)^{i-1}}{(i-1)!} \\
&= -t\sum_{i=n}^{\infty} e^{-\tilde{\lambda}^o t} \frac{\left(\tilde{\lambda}^o t\right)^i}{i!} + t\sum_{i=n+1}^{\infty} e^{-\tilde{\lambda}^o t} \frac{\left(\tilde{\lambda}^o t\right)^{i-1}}{(i-1)!} + t \cdot e^{-\tilde{\lambda}^o t} \frac{\left(\tilde{\lambda}^o t\right)^{n-1}}{(n-1)!} \\
&= t \cdot e^{-\tilde{\lambda}^o t} \frac{\left(\tilde{\lambda}^o t\right)^{n-1}}{(n-1)!}
\end{aligned}$$

and since the econometrician exploits up to the last observation until time $t > 0$, we can set $n = N(t)$, i.e.

$$f_{T_n}\left(\tilde{\lambda}^o\right) = t \cdot e^{-\tilde{\lambda}^o t} \frac{\left(\tilde{\lambda}^o t\right)^{N(t)-1}}{[N(t)-1]!} , \qquad \text{if } t \geq T_1 \text{, and 0 otherwise,} \tag{8}$$

which is the density function of a Gamma-distributed variable. Since the virtual econometrician has Haldane's agnostic priors, these can be described by a Gamma distribution with hyperparameters equal to 0.[16] Combining (8) with these hyperparameters which are equal to 0, the hyperparameters of (8) remain as in equation (8), i.e., as described by the theorem, completing the proof. □

Theorem 1 gives sufficient structure for the modeling of rare disasters. Specifically, Theorem 1 demonstrates that $N(t)$ and the elapsed sampling time are sufficient statistics for identifying the whole probability density function of beliefs about the disaster frequency. Next section provides the key implications of Theorem 1 which provide facility to Bayesian-learning modeling.

2.2 Posterior beliefs, posterior-belief moments, and long-run learning

From this section and on we view the elapsed sampling time as a state variable, for convenience in dynamic-programming applications. We denote the length of the elapsed sampling time by $T(t)$ for all $t \geq 0$. In addition, we assume that $T(0) > 0$ and $N(0) \geq 1$. Theorem 1 implies that the posterior distribution of the virtual econometrician's beliefs at any time $t \geq 0$, is given by,

$$\Pr(\lambda \mid \mathcal{F}_t) = f_{\left(N(t)\ ,\ [T(0)+t]^{-1}\right)}(\lambda) = \begin{cases} [T(0)+t]\, e^{-[T(0)+t]\lambda} \frac{\{[T(0)+t]\lambda\}^{N(t)-1}}{\Gamma(N(t))} & ,\ \text{if } \lambda \geq 0 \\ 0 & ,\ \text{if } \lambda < 0 \end{cases}, \tag{9}$$

in which $\Gamma(a) \equiv \int_0^\infty e^{-v} v^{a-1} dv$ is the Gamma function, and $\mathcal{F}_t$ is the filtration at time $t \geq 0$.[17] Equation (9) demonstrates the well-known result that the posterior distribution of

[16] See, e.g. Kerman (2011, pp. 1455-6).
[17] For a formal definition of filtration see, for example, Stokey (2008, pp. 17-18).

a Gamma prior is also Gamma.[18] Based on standard results about the moments of Gamma-distributed variables, the mean and variance of the posterior distribution for all $t \geq 0$ are,[19]

$$E\left[\tilde{\lambda}(t) \mid N(t) = N(0) + n\right] = \frac{N(0) + n}{T(0) + t}, \tag{10}$$

and

$$Var\left[\tilde{\lambda}(t) \mid N(t) = N(0) + n\right] = \frac{N(0) + n}{[T(0) + t]^2}, \tag{11}$$

for all $n \in \{0, 1, ...\}$.

In the context of optimization through HJB equations, the key result to use is the mean jump-frequency belief. In the related formula given by (10), the denominator is a continuously and linearly growing variable, while the numerator is a discrete point process. The point process in the numerator means that once a jump occurs, average jump frequency beliefs jump upwards to a pessimistic level. After a period without further busts, average jump frequency beliefs decay, implying that the learning agents become more optimistic. In brief, the trajectory of average jump-frequency beliefs will exhibit spikes which coincide with the occurrence of busts.

In Appendix B, we show that, in the long run, beliefs converge to rational expectations. Specifically, as time passes and more jumps are revealed, a learning agent gradually acquires more confidence in $E\left[\tilde{\lambda}(t) \mid N(t) = N(0) + n\right]$. The spiky trajectory of average beliefs implied by equation (10) is characterized by dampened spikes over time, which eventually disappear as average beliefs converge to the true parameter λ^* with infinite confidence asymptotically.

18See, for example, Gelman et al. (2004, p. 53).

19See, for example, Papoulis and Pillai (2002, p. 154) for the moment-generating function of the Gamma distribution.

3. Model and Optimal Policy under Rational Learning

The model extends Xie's (1997, pp. 416-9) parametric example in which optimal fiscal policy is time-consistent. Our extensions are two: (a) the inclusion of shocks to capital value and the study of the role of rational-expectations uncertainty on optimal fiscal policy, studied in this section, and (b) the introduction of model-parameter uncertainty (uncertainty about jump frequencies), and the study of optimal fiscal policy, studied in the next section.

The key reason for building on Xie's (1997, pp. 416-9) parametric example is that it delivers optimal fiscal policy which, technically, does not suffer from the time-inconsistency problem.[20] So, this setup allows us to study optimal fiscal policy using optimal-control techniques. This facility allows us to retain an intuitive state-space representation of the optimal control problem, to obtain analytical solutions, and to identify how shocks and model parameters affect optimal policies. In turn, identifying such analytical interconnections between parameters and optimal policies allows us to identify how subjective perceptions of parameter uncertainty affect optimal policies in the context of rational learning. We can go this far, because the state-space representation we obtain in the rational-expectations setup allows for the introduction of rational learning in the open-loop differential Stackelberg game of optimal-fiscal policy setting.

3.1 Description of the benchmark model

3.1.1 Production

Time is continuous. A final consumable good is produced under perfect competition, by a large number of identical firms of total mass equal to one. Production uses only (a composite form of) capital. The production function is linear, and the aggregate production function

[20]Xie (1997) fully explains why the time-inconsistency problem can be avoided in some cases, using insights from the theory of differential Stackelberg games.

is,

$$Y(t) = A \cdot K(t) \ , \tag{12}$$

for all $t \geq 0$, with $A > 0$. Capital is rented by firms in perfectly competitive markets. So, the rental rate of capital equals capital's marginal product, A. With δ being the capital depreciation rate, the interest rate is constant over time, i.e.,

$$r(t) = R \equiv A - \delta \ , \quad \text{for all } t \geq 0 \ . \tag{13}$$

3.1.2 Government

The government continuously provides a single public-consumption good, G, levying marginal-income taxes and using a balanced fiscal budget. Capital depreciation is tax-exempt, so for all $t \geq 0$,

$$G(t) = \tau(t)(A - \delta) K(t) = \tau(t) R K(t) \ , \tag{14}$$

in which $\tau(t)$ is the marginal tax rate on income. The government is a benevolent social planner who maximizes social welfare (the social-welfare function is provided below), subject to levying marginal-income taxes and to the balanced-fiscal budget rule given by (14).

3.1.3 Households

A continuum of identical infinitely-lived households with total mass equal to one maximizes lifetime expected utility derived by the stochastic flows of private and public consumption, namely

$$E_0 \left\{ \int_0^\infty e^{-\rho t} \left\{ \ln[c(t)] + \theta \ln[G(t)] \right\} dt \right\} \ , \tag{15}$$

in which $\rho > 0$ is the rate of time preference and $\theta > 0$. We denote all variables referring to individual households by lowercase symbols, while we use uppercase symbols for aggregate variables. Each household has an initial endowment $k(0) > 0$ of private capital holdings that it rents to firms which produce. Capital is of the same type across all individual households. Capital value (and its production efficiency) is subject to shocks which are revealed right after the consumption/investment decision has been made. To make the timing of events clearer, consider the discrete-time version of the household's budget constraint, which is,

$$\hat{k}_t = [1 + (1 - \tau_t) R] \underbrace{(1 + x_t) \hat{k}_{t-1}}_{\substack{\shortparallel \\ k_t}} - c_t \ , \tag{16}$$

in which x_t is some shock to the value of capital, which is realized after the consumption decision, c_t, has been made. The symbol $\hat{k}_{t-1}$ denotes the units of capital inherited from period $t-1$, before the shock x_t has been realized and embodied in the productive capacity of capital. After the embodiment of x_t takes place, the value of capital units in terms of consumer baskets in period t is $k_t = (1 + x_t) \hat{k}_{t-1}$. It is this new, ex-post capital value, k_t, that determines the after-tax income of the household, $(1 - \tau_t) k_t$, and also the capital units of next period, $\hat{k}_t$, measured in embodied efficiency according to the latest shock x_t. In addition, the value of the provided public good, G_t, is determined according to the aggregate ex-post value of capital, $K_t = (1 + x_t) \hat{K}_{t-1}$. So, in discrete time, c_t is chosen in the beginning of a period, while G_t is determined in the end of the same period. Re-arranging terms in (16) gives,

$$\Delta \hat{k}_t = \hat{k}_t - \hat{k}_{t-1} = (1 - \tau_t) R \hat{k}_{t-1} - c_t + [1 + (1 - \tau_t) R] \hat{k}_{t-1} x_t \ . \tag{17}$$

Equation (17) reveals the form of the continuous-time version of the household's budget constraint which corresponds to the timing of events explained above. We assume that the value of capital is hit by two shocks, one diffusion shock and one jump shock, namely,

$$dk(t) = \{[1 - \tau(t)] R k(t) - c(t)\} dt + \{1 + [1 - \tau(t)] R\} k(t) [\sigma dz(t) + dq(t)] \ , \tag{18}$$

in which $z(t)$ is a standard Brownian motion, i.e., $dz(t) = \varepsilon(t) \sqrt{dt}$, with $\varepsilon(t) \sim N(0, 1)$, for all $t \geq 0$ (so $z(t) = \int_0^t dz(s)$ with $\int$ being the stochastic Itô integral), $\sigma > 0$, while variable

$q(t)$ is a Poisson process driving random downward jumps in income of size $\zeta \in (0,1)$. The random variable ζ has a given time-invariant distribution having compact support, $\mathcal{Z} \subset (0,1)$. In particular, the Poisson process $q(t)$ is characterized by,

$$dq(t) = \begin{cases} -\zeta & \text{with Probability } \lambda^* dt \\ 0 & \text{with Probability } 1-\lambda^* dt \end{cases}, \tag{19}$$

in which λ^* is the jump-frequency parameter. So, $q(t) = -\sum_i \mathbf{1}_{\{\hat{t}_i \leq t\}} \zeta$ in which $\hat{t}_i = \{t \geq 0 \mid dq(t) \neq 0\}$, and $\mathbf{1}_{\{\hat{t}_i \leq t\}} = \{1, \text{ if } s = \hat{t}_i \text{ ; } 0, \text{ if } s \neq \hat{t}_i \text{ ; for all } s \leq t\}$. The Brownian motion, the Poisson process and ζ are all independent from each other.[21] Finally, the analogy between the discrete-time representation of the shock x_t in equations (16) and (17) and the presence of the stochastic processes $z(t)$ and $q(t)$ in (18) is given by $1+x_t = e^{\sigma\varepsilon_t+\nu_t}$, in which $\varepsilon_t \sim N(0,1)$ for all $t \in \{0,1,...\}$, and $\nu_t = \{\ln(1-\zeta) \text{ , with Prob. } \lambda^* \text{ ; } 0 \text{ , with Prob. } 1-\lambda^*\}$.

Switching from the discrete-time representation in (17) to the continuous-time representation in (18), the distinction between discrete-time variables k_t and $\hat{k}_{t-1}$ is suppressed through the presence of the infinitesimal influence of the stochastic processes $z(t)$ and $q(t)$ on the capital stock. Yet, the timing of events is crucial for specifying the continuous-time budget constraint appearing in (18). If consumption decisions were made after the realization of shock x_t in equation (16), then the last term of the budget constraint in (18) would have been $\{\{1+[1-\tau(t)]R\}k(t)-c(t)\}[\sigma dz(t)+dq(t)]$ instead of the simpler term which now appears in (18), i.e., $\{1+[1-\tau(t)]R\}k(t)[\sigma dz(t)+dq(t)]$.[22] The timing-of-events choice

[21] The quadratic covariation of $z(t)$ and $q(t)$ is zero by construction, since $q(t)$ is a pure jump process.

[22] If consumption decisions were made after the realization of the capital efficiency shock, x_t, then the household's uncertainty concerns would focus on next period's shock, x_{t+1}. To address concerns about x_{t+1}, equation (16) could be written as $(1+x_{t+1})^{-1}k_{t+1} = [1+(1-\tau_t)R]k_t - c_t$, in which we have used $(1+x_{t+1})^{-1}k_{t+1} = \hat{k}_t$. This form of equation (16) becomes $(1+x_{t+1})^{-1}[k_{t+1}-(1+x_{t+1})k_t] = (1-\tau_t)Rk_t - c_t$, which leads to, $k_{t+1}-k_t = (1-\tau_t)Rk_t - c_t + \{[1+(1-\tau_t)R]k_t - c_t\}x_{t+1}$. This last expression corresponds to $dk(t) = \{[1-\tau(t)]Rk(t)-c(t)\}dt + \{\{1+[1-\tau(t)]R\}k(t)-c(t)\}[\sigma dz(t)+dq(t)]$ in continuous time, which would be the expression to use if timing was different, i.e. if consumption decisions were made after the instantaneous revelation of the shocks (notice also that $1+x_{t+1} = e^{\sigma\varepsilon_{t+1}+\nu_{t+1}}$ in the discrete-time version of our model).

we make here slightly simplifies paper-and pencil analysis but does not affect any of the model's qualitative results.

3.2 Policy setting under rational expectations

Given the presence of marginal taxes, the first welfare theorem fails, and we need to solve for the decentralized equilibrium. Individual households maximize lifetime utility given by equation (15), subject to the budget constraint (18), given a sequence of variables they do not control: taxes, τ, the level of publicly provided goods, G, as well as aggregate capital, K. But in order to solve an individual household's problem, we also need to keep track of the motion of aggregate capital, K, since K affects the level of the publicly provided good G.[23]

Since the fiscal budget is continuously balanced, the level of $K(t)$ and the marginal-tax level, $\tau(t)$, jointly determine the motion of $G(t)$ at all times, $t \geq 0$, through equation (14). So, henceforth, let $(G(s))_{s\geq t}$ comply with $(\tau(s), K(s))_{s\geq t}$ and the balanced fiscal rule (14). In addition, the path $(K(s))_{s\geq t}$ must comply with optimal investment decisions of individual households, subject to a policy stream $(\tau(s), G(s))_{s\geq t}$ for all $t \geq 0$, and the aggregate version of the resource constraint given by (18). These features allow us to use a simple recursive representation of the individual household's problem which discards the sequence $(G(s))_{s\geq t}$, and focuses on the impact of the sequence $(\tau(s))_{s\geq t}$ (subject to (14)) on investment decisions. In turn, optimal-policy setting focuses on the determination of the marginal-tax sequence $(\tau(s))_{s\geq t}$. The following section provides the problem's recursive formulation.

[23]If utility was not linear, then aggregate capital, K, would affect the formation of the return to capital, too. With linear production, (13) shows that the return to capital is idependent from the level of K.

3.2.1 Optimization problem of an individual household and its solution

Let the decision rule $c^{RE}(t) = \mathbb{C}^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right)$, denote the solution to the individual household's problem, in which the superscript "RE" denotes a "rational-expectations" equilibrium. Market clearing implies that equilibrium aggregate consumption is $C^{RE}(t) = \mathbb{C}^{RE}\left(K, K, t \mid (\tau(s))_{s\geq t}\right)$. Substituting the rule $\mathbb{C}^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right)$ into the objective function (15), with variables k and K evolving according to the resource constraint given by (18) and its aggregate version, the value function of an individual household, denoted by $J^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right)$, for all $t \geq 0$, satisfies the Hamilton-Jacobi-Bellman (HJB) equation,

$$\rho J^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right) = \max_{c(t)\geq 0} \Bigg\{ \ln(c) + \theta \ln[\tau(t) RK] +$$

$$+ \{[1-\tau(t)] Rk - c\} \cdot J_k^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right) +$$

$$+ \frac{(\sigma k)^2}{2} B(\tau(t))^2 \cdot J_{kk}^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right) +$$

$$+ \left\{[1-\tau(t)] RK - \mathbb{C}^{RE}\left(K, K, t \mid (\tau(s))_{s\geq t}\right)\right\} \cdot J_K^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right) +$$

$$+ \frac{(\sigma K)^2}{2} B(\tau(t))^2 \cdot J_{KK}^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right) +$$

$$+ \sigma^2 kK B(\tau(t))^2 \cdot J_{kK}^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right) +$$

$$+ J_t^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right) +$$

$$+ \lambda^* \Bigg\{ E_\zeta \left[J^{RE}\left(k \cdot [1 - \zeta B(\tau(t))]\ , K \cdot \{1 - \zeta B(\tau(t))\}\ , t \mid (\tau(s))_{s\geq t}\right)\right] -$$

$$- J^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right) \Bigg\} \Bigg\}\ , \quad (20)$$

in which $B(\tau(t)) \equiv 1 + [1 - \tau(t)] R$ and with E_ζ being the expectations operator focusing on the uncertainty of variable ζ alone. In order that the HJB equation (20) be well-defined we need to ensure that $1 - \zeta B(\tau) > 0$ for all $\zeta \in \mathcal{Z}$, and all $\tau \geq 0$.[24] So, throughout the rest of the paper we assume that $\mathcal{Z}$ and R are such that,

$$\sup(\mathcal{Z}) < \frac{1}{1+R} \ . \tag{21}$$

For ensuring that $J^{RE}\left(k, K, t \mid (\tau(s))_{s \geq t}\right)$ be well-defined we also assume that

$$0 \leq \tau(t) \leq 1 \ , \quad \text{for all } t \geq 0 \ . \tag{22}$$

In Appendix B we show that, under the parametric constraint given by (21) and a tax profile $(\tau(t))_{t\geq 0}$ restricted by (22), the solution to the maximization problem expressed by the HJB equation (20) is,

$$c^*(t) = \mathbb{C}^{RE}\left(k, K, t \mid (\tau(s))_{s \geq t}\right) = \rho k \quad \text{for all } t \geq 0 \ , \tag{23}$$

and the value function is given by,

$$J^{RE}\left(k, K, t \mid (\tau(s))_{s \geq t}\right) = \frac{1}{\rho}\left[\ln(\rho) + \frac{1+\theta}{\rho}(R - \rho) + \theta \ln(R) + \ln(k) + \theta \ln(K)\right] + \\ + \int_t^\infty e^{-\rho(s-t)} M(\tau(s))\, ds \ , \tag{24}$$

in which

$$M(\tau) = \theta \ln(\tau) - \frac{1+\theta}{\rho}\left\{R\tau + \frac{\sigma^2}{2} B(\tau)^2 - \lambda^* E_\zeta \{\ln[1 - \zeta B(\tau)]\}\right\} \ . \tag{25}$$

3.2.2 Maximizing social welfare under rational expectations

The social-welfare function at time 0 is derived through setting $k = K$ and $t = 0$ in function J^{RE}. In particular, equation (24) implies that the social-welfare function is,

$$J^{RE}\left(K, K, 0 \mid (\tau(t))_{t \geq 0}\right) = \frac{1}{\rho}\left[\ln(\rho) + \frac{1+\theta}{\rho}(R - \rho) + \theta \ln(R) + (1+\theta)\ln(K)\right] +$$

[24] Notice that τ cannot be negative, as it only finances a public good which satisfies $G(t) \geq 0$ for all $t \geq 0$.

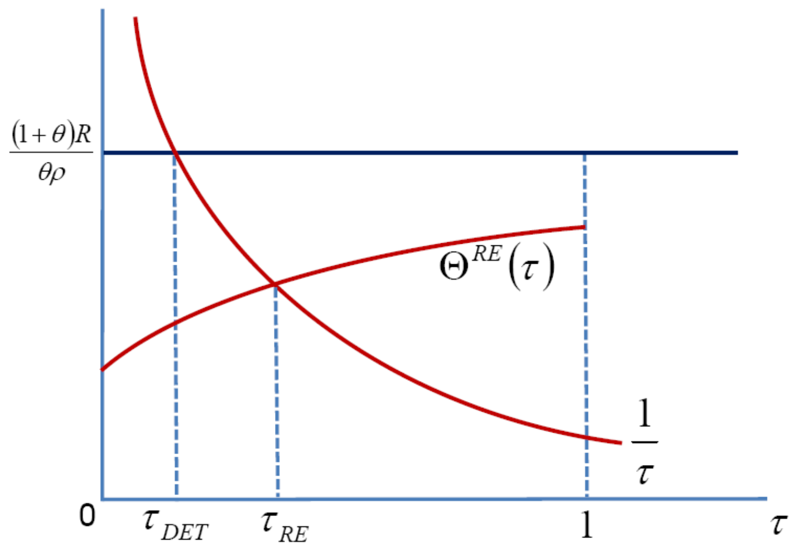

Figure 1

$$+\int_0^\infty e^{-\rho t} M(\tau(t))\, dt \ . \quad (26)$$

Setting the optimal-policy function of time $(\tau^{RE}(t))_{t\geq 0}$ is a task that can be performed analytically since the social-welfare function given by (26) is time-separable. In particular, optimal policy is determined by setting $M'(\tau) = 0$, which implies that

$$\tau^{RE}(t) = \tau_{RE} \ , \quad \text{for all } t \geq 0 \ ,$$

with τ_{RE} solving the following equation,

$$\frac{1}{\tau} = \Theta^{RE}(\tau) \ , \quad (27)$$

in which

$$\Theta^{RE}(\tau) \equiv \frac{1+\theta}{\theta\rho} R \left\{ 1 - \sigma^2(1+R) + \sigma^2 R\tau - \lambda^* E_\zeta \left[\frac{\zeta}{1-\zeta(1+R)+\zeta R\tau} \right] \right\} \ . \quad (28)$$

Equations (27) and (28) show that the optimal rational-expectations policy $(\tau^{RE}(t))_{t\geq 0}$ is both time-consistent and time-invariant. Notice also that, under the parametric constraint given by (21), the policy τ_{RE} solving (27) is also the unique global optimum, since $M''(\tau) < 0$ for all $\tau \geq 0$.[25] In the deterministic version of the model, which corresponds to the case of $\sigma = \lambda^* = 0$, the optimal tax rate, denoted by $\tau^{DET}(t)$ in which "DET" stands for "deterministic", is

$$\tau^{DET}(t) = \tau_{DET} = \frac{\theta\rho}{(1+\theta)R} , \quad \text{for all } t \geq 0 ,$$

which corresponds to Xie's (1997, eq. 23, p. 418) solution.[26] The comparison between the deterministic solution and the stochastic version under rational expectations $(\sigma, \lambda^* > 0)$ is given by Figure 1.[27] In the presence of uncertainty, as the values of σ and λ^* increase, function $\Theta^{RE}(\tau)$ in Figure 1 shifts downwards, demonstrating that the optimal marginal tax rate always increases (this is also why $\tau_{RE} > \tau_{DET}$ if we define τ_{RE} as corresponding to the stochastic version of the model with $\sigma, \lambda^* > 0$).[28]

One of the reasons contributing to higher taxation as the value of σ increases is the insurance effect of marginal taxation. Higher marginal taxes reduce the effective income fluctuations caused by shocks, thus generating an insurance effect. The insurance effect of marginal taxation has been theoretically suggested by Mirrlees (1974) and Varian (1980), and has also been empirically demonstrated by Grant et al. (2008, 2010). Another reason why marginal taxes increase as the value of λ^* increases is the fact that the average expected

[25] In particular,

$$M''(\tau) = -\frac{1}{\tau^2} - \frac{1+\theta}{\theta\rho} R^2 \left\{ \sigma^2 + \lambda^* E_\zeta \left\{ \left[\frac{\zeta}{1-\zeta B(\tau)} \right]^2 \right\} \right\} < 0 .$$

[26] Notice that in Xie's (1997, pp. 416-418) example, $\theta = 1$.

[27] In Figure 1 we assumed that all parameters are such that $\Theta^{RE}(1) > 1$, namely

$$\frac{1+\theta}{\theta\rho} R \left[1 - \sigma^2 - \lambda^* E_\zeta \left(\frac{\zeta}{1-\zeta} \right) \right] > 1 .$$

[28] The function $\Theta^{RE}(\tau)$ shifts downwards as σ increases if and only if $R\tau < 1 + R$, i.e., with $\tau < 1 + 1/R$. As we focus on parameterizations that imply $\tau < 100\%$ (as the case in Figure 1), equation (28) together with Figure 1 imply that $\partial\tau_{RE}/\partial\sigma > 0$ for all τ belonging to the interval $[0, 1]$.

return of capital decreases, making capital investment less attractive and tax distortions on capital less severe. That taxes are perceived as being less distortionary when capital is expected to have lower returns, is the cornerstone of the belief mechanism affecting optimal policy that we suggest in the rest of this paper.

4. Optimal Fiscal Policy under Rational Bayesian Learning

Here information is incomplete, and the policy game is a Stackelberg game with instantaneous precommitment, as in Cohen and Michel (1988). In the incomplete-information setup the policymaker needs to re-evaluate priors before re-setting policy instantaneously. So there is an additional set of state variables entering the rational-learning problem in order to describe the state of beliefs and model uncertainty related to this informational limitation. As we have shown throughout Section 2, the pair $(N(t), T(t))$, with $T(t)$ denoting the elapsed sampling time up to time instant $t \geq 0$, and with $N(t)$ denoting the cumulative count of jumps up to $T(t)$, is a sufficient-statistics set. Having $(N(t), T(t))$ as a sufficient-statistics pair facilitates the incorporating of learning in a HJB equation through the introduction of N_t as the only additional state variable to the set of state variables used in the rational-expectations problem. Specifically, in the rational-expectations problem the set of state variables for an individual household is $\left(k, K, t \mid (\tau(s))_{s \geq t}\right)$, while in the rational-learning problem the corresponding set of state variables is $\left(k, K, N, T \mid (\tau(s))_{s \geq T-T_0}\right)$, because the time state variable is incorporated through the linear transformation $t = T(t) - T_0$.

Let the superscript "RL" denote a "rational-learning" equilibrium. Fix an initial condition T_0 and let the decision rule $c^{RL}(t) = \mathbb{C}^{RL}\left(k, K, N, T \mid (\tau(s))_{s \geq T-T_0}\right)$ denote the solution to the individual household's problem. Accordingly, equilibrium aggregate consumption is given by $C^{RL}(t) = \mathbb{C}^{RL}\left(K, K, N, T \mid (\tau(s))_{s \geq T-T_0}\right)$. The rational-learning HJB equation is,

$$
\begin{aligned}
&\rho J^{RL}\left(k,K,N,T \mid (\tau(s))_{s\geq T-T_0}\right) = \max_{c(t)\geq 0}\Bigg\{\ln(c)+\theta\ln\left[\tau(t)RK\right]+ \\
&+\left\{\left[1-\tau(t)\right]Rk-c\right\}\cdot J_k^{RL}\left(k,K,N,T \mid (\tau(s))_{s\geq T-T_0}\right)+ \\
&+\frac{(\sigma k)^2}{2}B(\tau(t))^2\cdot J_{kk}^{RL}\left(k,K,N,T \mid (\tau(s))_{s\geq T-T_0}\right)+ \\
&+\left\{\left[1-\tau(t)\right]RK-\mathbb{C}^{RL}\left(K,K,N,T \mid (\tau(s))_{s\geq T-T_0}\right)\right\}\cdot J_K^{RL}\left(k,K,N,T \mid (\tau(s))_{s\geq T-T_0}\right)+ \\
&+\frac{(\sigma K)^2}{2}B(\tau(t))^2\cdot J_{KK}^{RL}\left(k,K,N,T \mid (\tau(s))_{s\geq T-T_0}\right)+ \\
&+\sigma^2 kKB(\tau(t))^2\cdot J_{kK}^{RL}\left(k,K,N,T \mid (\tau(s))_{s\geq T-T_0}\right)+ \\
&+J_T^{RL}\left(k,K,N,T \mid (\tau(s))_{s\geq T-T_0}\right)+ \\
&+\frac{N}{T}\Bigg\{E_\zeta\left[J^{RL}\left(k\cdot\left[1-\zeta B(\tau(t))\right],K\cdot\left\{1-\zeta B(\tau(t))\right\},N+1,T \mid (\tau(s))_{s\geq T-T_0}\right)\right]- \\
&\qquad\qquad -J^{RL}\left(k,K,N,T \mid (\tau(s))_{s\geq T-T_0}\right)\Bigg\}\Bigg\}. \qquad (29)
\end{aligned}
$$

In equation (29), the belief change over time is captured by two elements. First element is the partial derivative J_T^{RL}, which takes into account the impact of a change in elapsed sampling time on beliefs, since T is one of the two sufficient statistics for deriving posterior beliefs. The second element is having "$N+1$" replacing "N" in the last term in brackets on the right-hand side of equation (29), expressing that sufficient statistics will add up one more jump in case a jump occurs within dt time from the present instant. These two elements incorporate Bayes' rule in the HJB equation, making learning rational instead of adaptive Bayesian learning.[29]

[29]For a distinction between rational learning and adaptive learning see, for example, Koulovatianos, Mirman, and Santugini (2009), and Koulovatianos and Wieland (2011).

We make the same assumptions as in the case of rational expectations, that the parametric constraint given by (21) holds, and that any tax profile $(\tau(t))_{t\geq 0}$ is restricted by (22). Using similar techniques as in the case of rational expectations, but taking special care of the belief evolution over time, in Appendix B we show that, under the solution to the maximization problem expressed by the HJB equation (20) is,

$$c^*(t) = \mathbb{C}^{RL}\left(k, K, N, T \mid (\tau(s))_{s\geq T-T_0}\right) = \rho k \quad \text{for all } t \geq 0 \ , \tag{30}$$

and the value function is given by,

$$J^{RL}\left(k, K, N, T \mid (\tau(s))_{s\geq T-T_0}\right) = \frac{1}{\rho}\left[\ln(\rho) + \frac{1+\theta}{\rho}(R-\rho) + \theta\ln(R) + \ln(k) + \theta\ln(K)\right] + \\ + E_T\left[\int_{T-T_0}^{\infty} e^{-\rho t}\hat{m}(\tau(t))\,dt\right] + \frac{N}{T}\frac{1+\theta}{\rho}E_\zeta\left\{\ln\left[1-\zeta B(\tau(T-T_0))\right]\right\} \ , \tag{31}$$

in which

$$\hat{m}(\tau(t)) = \theta\ln(\tau) - \frac{1+\theta}{\rho}\left[R\tau + \frac{\sigma^2}{2}B(\tau)^2\right]\left\{R\tau + \frac{\sigma^2}{2}B(\tau)^2 - \frac{N}{T}E_\zeta\left\{\ln\left[1-\zeta B(\tau)\right]\right\}\right\} . \tag{32}$$

4.1 Maximizing social welfare under limited information

Since a benevolent planner does not have a different information set compared to all other agents in the economy, the social-welfare function at time 0 is derived through setting $k = K$ and $T = T_0$ in function J^{RL}. So,

$$J^{RL}\left(K, K, N, T \mid (\tau(t))_{t\geq T_0}\right) = \frac{1}{\rho}\left[\ln(\rho) + \frac{1+\theta}{\rho}(R-\rho) + \theta\ln(R) + (1+\theta)\ln(K)\right] + \\ + E_0\left[\int_{T_0}^{\infty} e^{-\rho t}\hat{m}(\tau(t))\,dt\right] + \frac{N}{T}E_\zeta\left\{\ln\left[1-\zeta B(\tau(T_0))\right]\right\} \ . \tag{33}$$

In Appendix B we show that optimal policy is determined by setting $M'(\tau) = 0$, which implies that

$$\tau^{RL}(t) = \tau_{RL} \ , \quad \text{for all } t \geq 0 \ ,$$

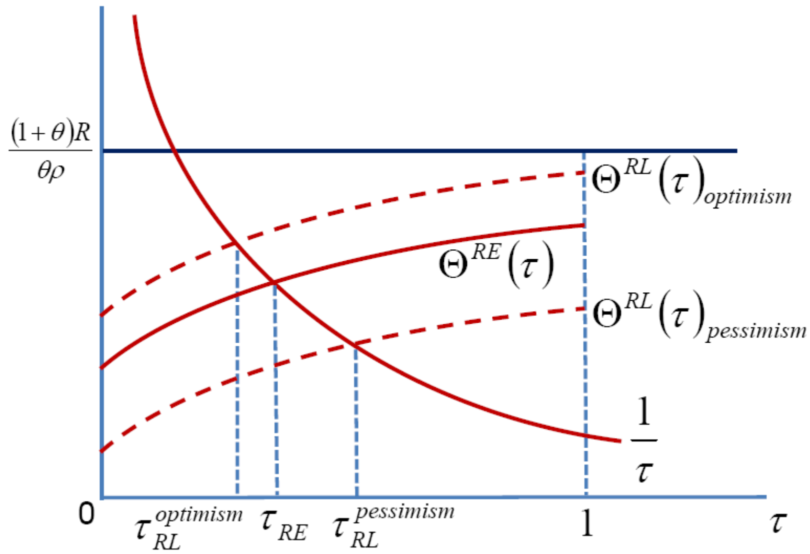

Figure 2

with τ_{RL} solving the following equation,

$$\frac{1}{\tau} = \Theta^{RL}(\tau) = \hat{\Theta}^{RL}(\tau \mid N, T) \ , \tag{34}$$

in which

$$\hat{\Theta}^{RL}(\tau \mid N, T) \equiv \frac{1+\theta}{\theta\rho} R \left\{ 1 - \sigma^2(1+R) + \sigma^2 R\tau - \frac{N}{T} E_\zeta \left[\frac{\zeta}{1 - \zeta(1+R) + \zeta R\tau} \right] \right\} \ . \tag{35}$$

Equations (34) and (35) show that, under limited information, $\left(\tau^{RL}(t)\right)_{t\geq 0}$ is not time-invariant any more. Instead, the level $\tau^{RL}(t)$ varies with the sufficient statistics $(N(t), T(t))$, depending on the history of past realizations. So, despite that we preserve rationality, under limited information there is path dependence, driven by the evolution of the information set. Based on equations (28) and (35) Figure 2 shows how optimal tax rates relate to rational-expectations taxes. Whenever there is optimism $(N(t)/T(t) < \lambda^*)$, optimal tax rates are lower than τ_{RE}. Since agents in the model think that capital accumulation is not disrupted frequently by capital busts they wish to have a higher investment rate. This vivid propensity to invest comes from the tendency of agents in this economy to smooth consumption irrespective of tax rates and pessimistic/optimistic beliefs, by keeping a pre-tax consumption rate to $c(t)/y(t) = (\rho/A)\,k(t)$ for all $t \geq 0$. The social planner, who has the same level of optimism, understands that under optimism the economy is more responsive to investment stimulus and encourages this high propensity to invest by taxing less. On the contrary, if there is pessimism, $(N(t)/T(t) > \lambda^*)$, then optimal tax rates are higher than τ_{RE}. In this case, the social planner understands that the economy is not responsive to tax incentives for investment because everyone's belief is that capital busts occur too often. So, at times of pessimism the planner finds it optimal to provide more public goods by immediately financing those goods with higher taxes.

So, beliefs about the frequency of capital busts act as a taste shifter for the social welfare function. Even if the fiscal balance was not balanced, every time that a government provides a higher ratio of public goods marginal taxes have to increase soon thereafter. Apart from the standard argument that, after a massive capital-value bust a government provides more public goods in order to alleviate poverty, we believe we have captured another aspect of post-crisis policy setting. In a model which has abstracted from income heterogeneity and poverty-line considerations, we have shown that the responsiveness of the economy to tax incentives for investment matters as well. This is a closed-economy model, not allowing for capital flight. Certainly extensions in open-economy models may shed more light on why public consumption rises so much after capital-value busts.

5. Conclusion

In a deterministic "AK" endogenous growth model a tax cut has permanent positive effects, as it can boost future growth opportunities and welfare. The reason is, taxes discourage investment, and investment builds capital, which is the productive backbone of the economy.

Importantly, in the "AK" endogenous growth model, past distortions are not wiped out with time, they leave a permanent negative mark. We have shown that all above responses of the "AK" endogenous growth model to marginal-tax increases are retained after introducing uncertainty, even if this uncertainty involves rare capital-value busts. Yet, we have shown that everything changes once we introduce a natural feature: due to the inherently infrequent occurrence of rare capital-value busts, markets, market analysts, and policy makers may not be confident about the key parameter, the disaster-frequency parameter. So, all form subjective beliefs about the disaster-frequency parameter.

It turns out that the average subjective belief about the disaster-frequency parameter plays a critical role for all investment decisions. Most importantly, it plays a key role in determining the responsiveness of investment to marginal taxes. Before a massive capital loss, markets and policymakers have optimistic beliefs about investment-benefit prospects: they think of an economy as a race horse. Policymakers do not tax as much, because chopping off a leg from a race horse is an enormous waste. On the contrary, after a massive capital loss, rational/informationally-efficient revision of beliefs by both markets and policymakers dictates that the economy is not a race horse, but a sick horse, prone to such pathological events in the future. Here is what brings the policy-regime switch: you can beat a sick horse in order to stand up, but you will not do much; so, no major harm is done by taxing more after a major capital bust, because nobody wants to invest in an economy where capital frequently gets busts. Yet, after a long period without disasters, rational/informationally-efficient revision of beliefs about the frequency of capital busts by both markets and policymakers dictates that optimism is reasonable to return: everyone thinks that capital busts are more rare events than before, and so optimal fiscal policy returns to a lower marginal taxation on capital.

We believe our study contributes to the understanding of optimal or actual policymaking in periods shortly after rare capital busts. We have emphasized the role of Bayesian updating of beliefs about the frequency of jumps, and we have demonstrated that any future studies that can incorporate our analysis will, (a) retain rationality and a rationalization of common priors despite the limited information, (b) retain a solid welfare criterion for understanding first-best and second-best policy, and (c) use the same dynamic-programming tools in order to model markets in such forward-looking dynamic environments with or without endogenous fiscal policy.

6. Appendix A - More detailed evidence on ΔG and busts

21 OECD countries from 1960-2007	
Events	**% Change in Government Consumption**
Recessions	1.79
Severe recessions	2.16
Credit contractions	2.83
Credit crunches	5.98***
House-price declines	3.39
House-price busts	8.75***
Equity-price declines	3.59
Equity-price busts	7.48***
Recessions without credit crunches	1.57
Recessions with credit crunches	3.23***
Recessions with severe credit crunches	4.57***
Recessions without house-price busts	1.73
Recessions with house-price busts	1.82
Recessions with severe house-price busts	2.12
Recessions without equity-price busts	1.62
Recessions with house-price busts	2.14
Recessions with severe house-price busts	2.16

Table A.1 – Selected numbers from Claessen et al. (2009, Table 9, p. 685). The reported numbers are medians of peak-to-trough changes in government consumption corresponding to each event. "Severe" events are those in the top quartile of the Claessen et al. (2009) sample of crunches/busts/shocks. Symbol "***" indicates significance at the 1% level.

7. Appendix B - Proofs

7.1 Proof of the long-run convergence to rational expectations

From a modeler's perspective, the expected value of $N(t)$ in (3) is $\lambda^* \cdot t$ (denote the expected realization from a modeler's perspective by $E_m[N(t)] = \lambda^* \cdot t$). After applying the law of iterated expectations on equations (10) and (11) it is,

$$E_m\left[\tilde{\lambda}(t)\right] = \frac{N(0) + \lambda^* \cdot t}{T(0) + t} \ , \tag{36}$$

and

$$Var_m\left[\tilde{\lambda}(t)\right] = \frac{N(0) + \lambda^* \cdot t}{[T(0) + t]^2} \ . \tag{37}$$

The asymptotic distribution is directly characterized from (36) and (37) which imply,

$$\lim_{t \to \infty} E_m\left[\tilde{\lambda}(t)\right] = \lambda^* \ , \tag{38}$$

and

$$\lim_{t \to \infty} Var_m\left[\tilde{\lambda}(t)\right] = 0 \ . \tag{39}$$

Equation (39) implies infinite confidence asymptotically, and together with the unbiasedness implied by (36) the result is proved. $\square$

7.2 Finding the solution to the RE problem given by the HJB equation (20)

We start with the guess that $J^{RE}\left(k, K, t \mid (\tau(s))_{s \geq t}\right)$ takes the form,

$$J^{RE}\left(k, K, t \mid (\tau(s))_{s \geq t}\right) = \kappa + \int_t^{\infty} e^{-\rho(s-t)} m(\tau(s))\, ds + a \ln(k) + b \ln(K) \ , \tag{40}$$

for all $t \geq 0$, in which κ, a, and b are undetermined coefficients, while $m(\tau(s))$ is an unknown function. The guess given by (40) declares that $J^{RE}\left(k, K, t \mid (\tau(s))_{s \geq t}\right)$ depends on time t and also on the current and future policies $(\tau(s))_{s\geq t}$. The first-order conditions of the problem given by (20) are $c^{-1} = J_k^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right)$, while (40) implies $J_k^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right) = ak^{-1}$, so,

$$\mathbb{C}^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right) = a^{-1}k \ . \tag{41}$$

Based on (41) it is straightforward to linearly aggregate individual consumption rules and obtain,

$$\mathbb{C}^{RE}\left(K, K, t \mid (\tau(s))_{s\geq t}\right) = a^{-1}K \ . \tag{42}$$

Substituting the guess (40) and its derivatives, together with (41) and (42) into (20) we obtain,

$$\rho\kappa + \rho h\left(t \mid (\tau(s))_{s\geq t}\right) + \rho a \ln(k) + \rho b \ln(K) = \ln(k) + \theta \ln(K) + \ln\left(a^{-1}\right) + \theta \ln(R) +$$
$$+\theta \ln[\tau(t)] + (a+b)\left\{[1-\tau(t)]R - a^{-1} - \frac{\sigma^2}{2}B(\tau(t))^2 + \lambda^* E_\zeta\{\ln[1-\zeta B(\tau(t))]\}\right\} +$$
$$+h_t\left(t \mid (\tau(s))_{s\geq t}\right) \ , \tag{43}$$

in which

$$h\left(t \mid (\tau(s))_{s\geq t}\right) \equiv \int_t^\infty e^{-\rho(s-t)} m(\tau(s))\, ds \ . \tag{44}$$

In order that the terms $\ln(k)$ and $\ln(K)$ in equation (43) be eliminated, it is necessary and sufficient to set,

$$a = \rho^{-1} \quad \text{and} \quad b = \theta\rho^{-1} \ . \tag{45}$$

Substituting the coefficients for a and b implied by (45) into (43), and after setting

$$\kappa = \frac{1}{\rho}\left[\ln(\rho) + \frac{1+\theta}{\rho}(R-\rho) + \theta \ln(R)\right] \ , \tag{46}$$

in order to eliminate the constants from (43), equation (43) reduces to,

$$\rho h\left(t \mid (\tau(s))_{s\geq t}\right) - h_t\left(t \mid (\tau(s))_{s\geq t}\right) = M(\tau(t)) \ . \tag{47}$$

According to the definition of $h\left(t \mid (\tau(s))_{s\geq t}\right)$, which is given by (44), equation (47) implies,

$$m(\tau(t)) - \int_t^{\infty} e^{-\rho(s-t)} m'(\tau(s))\,\dot{\tau}(s)\,ds = M(\tau(t)) \ , \tag{48}$$

in which $\dot{\tau}(s)$ is the derivative of function $\tau(s)$ with respect to time. According to Xie (1997) we can have a solution with $\dot{\tau}(t) = 0$ for all $t \geq 0$, i.e., a constant tax over time. This possibility is corroborated by equation (48). We can investigate whether setting $\dot{\tau}(t) = 0$ for all $t \geq 0$ leads to a contradiction. Specifically, fix, for the moment, an optimal plan such that $\dot{\tau}(t) = 0$ for all $t \geq 0$. Under the $\dot{\tau}(t) = 0$ assumption, equation (48) implies that

$$m(\tau(t)) = M(\tau(t)) \ , \tag{49}$$

with function $M(\tau(t))$ given by equation (25). After substituting (49) into (44), the tax plan $(\tau(s))_{s\geq t}$ enters the value function $J^{RE}\left(k, K, t \mid (\tau(s))_{s\geq t}\right)$ additively, through the time-separable term $\int_t^{\infty} e^{-\rho(s-t)} M(\tau(s))\,ds$. This time separability of the tax-rate function $M(\tau(s))$ in $\int_t^{\infty} e^{-\rho(s-t)} M(\tau(s))\,ds$ implies that optimal-tax setting is achieved through solving $M'(\tau(t)) = 0$ for all $t \geq 0$ (it is verifiable that $M''(\tau(t)) < 0$). Since function $M(\tau(t))$ is time-invariant, the original claim that $\dot{\tau}(t) = 0$ for all $t \geq 0$ is reconfirmed.[30] After substituting (45), (46), (49), and (44) into (40) we obtain the value function given by (24). Finally, substituting (45) into (41) we obtain the decision rule given by (23), proving the result. □

[30] Xie (1997) provides an alternative argument proving that the optimal tax rate for this model is time-consistent and constant. Xie (1997) derives a representative household's Lagrange multipliers that pertain to its optimization. Then Xie (1997) inserts these household Lagrange multipliers into the optimization problem of the government and proves that the boundary conditions on the government Lagrange multipliers determine whether a model has a time-consistent solution or not Xie (1997, Proposition 1, p. 416). Xie's (1997) solution approach is accommodated by the deterministic environment he analyzes, which allows for using Hamiltonians. Extending Xie's (1997) argument to a stochastic environment is possible, but it is more cumbersome: it requires proving boundary-condition requirements for composite functions involving partial derivatives of value functions characterized by HJB equations. The proof we use here is simpler and straighforward for the purposes of the stochastic model of this paper.

7.3 Finding the solution to the RL problem given by the HJB equation (29)

Our guess about the functional form of $J^{RL}\left(k,K,N,T \mid (\tau(s))_{s\geq T-T_0}\right)$ is,

$$J^{RL}\left(k,K,N,T \mid (\tau(s))_{s\geq T-T_0}\right)=\hat{\kappa}+\int_t^{\infty} e^{-\rho(s-t)}E_t\left[\hat{m}\left(\tau(s)\right)\right]ds$$

$$+\hat{a}\ln(k)+\hat{b}\ln(K)+\frac{N}{T}\xi\hat{B}\left(\tau(t)\right)\ , \tag{50}$$

in which $t=T-T_0$,

$$\hat{B}(\tau)\equiv E_\zeta\left\{\ln\left[1-\zeta B(\tau)\right]\right\}\ , \tag{51}$$

while $\hat{\kappa}$, $\hat{a}$, and $\hat{b}$ are undetermined coefficients, and $\hat{m}(\tau,N,T)$ is an unknown function. The first-order conditions of the problem given by (29) imply,

$$c=\hat{\alpha}^{-1}k\ \text{ and }\ C^{RL}\left(K,K,N,T \mid (\tau(s))_{s\geq T-T_0}\right)=\hat{\alpha}^{-1}K\ . \tag{52}$$

Using the guess given by (50) and (52), the HJB equation (29) becomes,

$$\rho\hat{\kappa}+\rho\int_t^{\infty} e^{-\rho(s-t)}E_t\left[\hat{m}\left(\tau(s)\right)\right]ds+\rho\hat{a}\ln(k)$$

$$+\rho\hat{b}\ln(K)+\rho\frac{N}{T}\xi\hat{B}\left(\tau(t)\right)=$$

$$=\ln\left(\hat{a}^{-1}\right)+\ln(k)+\theta\ln(R)+\theta\ln\left[\tau(t)\right]+\theta\ln(K)$$

$$+\left\{\left[1-\tau(t)\right]R-a^{-1}\right\}k\hat{a}k^{-1}$$

$$+\frac{(\sigma k)^2}{2}B\left(\tau(t)\right)^2(-\hat{a})k^{-2}$$

$$+\left\{\left[1-\tau(t)\right]R-a^{-1}\right\}K\hat{b}K^{-1}$$

$$+\frac{(\sigma K)^2}{2}B\left(\tau(t)\right)^2\left(-\hat{b}\right)K^{-2}$$

$$+\sigma^2 kKB\left(\tau(t)\right)^2\cdot 0$$

$$
\begin{aligned}
&+\frac{d\int_t^{\infty} e^{-\rho(s-t)} E_t\left[\hat{m}\left(\tau\left(s\right)\right)\right] ds}{dt} \\
&-\frac{N}{T^2}\xi \hat{B}\left(\tau\left(t\right)\right) \\
&+\frac{N}{T}\left[\hat{a}+\hat{b}+\left(\frac{N+1}{T}-\frac{N}{T}\right)\xi\right]\hat{B}\left(\tau\left(t\right)\right) \ .
\end{aligned} \tag{53}
$$

After rearranging terms and simplifying (53) we obtain,

$$
\begin{aligned}
&\rho\int_t^{\infty} e^{-\rho(s-t)} E_t\left[\hat{m}\left(\tau\left(s\right)\right)\right] ds - \frac{d\int_t^{\infty} e^{-\rho(s-t)} E_t\left[\hat{m}\left(\tau\left(s\right)\right)\right] ds}{dt} \\
&+\left(\rho\hat{a}-1\right)\ln\left(k\right)+\left(\rho\hat{b}-\theta\right)\ln\left(K\right)+\left[\rho\xi-\left(\hat{a}+\hat{b}\right)\right]\frac{N}{T}\hat{B}\left(\tau\left(t\right)\right) \\
&+\rho\hat{\kappa}-\ln\left(\hat{a}^{-1}\right)-\theta\ln\left(R\right)-\left(R-a^{-1}\right)\left(\hat{a}+\hat{b}\right) \\
&=\theta\ln\left[\tau\left(t\right)\right]-\left(\hat{a}+\hat{b}\right)\left[R\tau\left(t\right)+\frac{\sigma^2}{2}B\left(\tau\left(t\right)\right)^2\right] \ .
\end{aligned} \tag{54}
$$

In order to get rid of variables K, k, in equation (54) we can set,

$$
\hat{a}=\frac{1}{\rho} \ \text{ and } \hat{b}=\frac{\theta}{\rho} \ . \tag{55}
$$

Combining (55) with the need to eliminate dependence on the ratio N/T in (54) implies,

$$
\xi=\frac{1+\theta}{\rho^2} \ . \tag{56}
$$

In order to eliminate constant terms in (54), equations (55) and (56) further imply that

$$
\hat{\kappa}=\frac{1}{\rho}\left[\ln\left(\rho\right)+\theta\ln\left(R\right)+\frac{1+\theta}{\rho}\left(R-\rho\right)\right] \ . \tag{57}
$$

After imposing (55), (56), (57) into (54) the remaining terms are,

$$
\begin{aligned}
&\rho\int_t^{\infty} e^{-\rho(s-t)} E_t\left[\hat{m}\left(\tau\left(s\right)\right)\right] ds - \frac{d\int_t^{\infty} e^{-\rho(s-t)} E_t\left[\hat{m}\left(\tau\left(s\right)\right)\right] ds}{dt} \\
&=\theta\ln\left[\tau\left(t\right)\right]-\frac{1+\theta}{\rho}\left[R\tau\left(t\right)+\frac{\sigma^2}{2}B\left(\tau\left(t\right)\right)^2\right] \ ,
\end{aligned}
$$

which implies,

$$\hat{m}(\tau(t)) = \theta \ln[\tau(t)] - \frac{1+\theta}{\rho}\left[R\tau(t) + \frac{\sigma^2}{2}B(\tau(t))^2\right] + \int_t^{\infty} e^{-\rho(s-t)} E_t[\hat{m}(\tau(s))\hat{m}'(\tau(s))\dot{\tau}(s)]\,ds\ . \tag{58}$$

It remains to show that $E_t[\hat{m}(\tau(s))\hat{m}'(\tau(s))\dot{\tau}(s)] = 0$ for all $s \geq t$ in equilibrium. Specifically, we show that, at any time instant $t \geq 0$, future taxes are not expected to change from the current level $\tau(t)$, not even instantaneously, i.e., $E_t[\dot{\tau}(s)] = 0$ for all $s > t$.

Consider that $E_t[\dot{\tau}(s)] = 0$ for all $s > t$. Since all shocks in the model are independent over time, $E_t[\dot{\tau}(s)] = 0$ for all $s > t$ implies that $E_t[\hat{m}(\tau(s))\hat{m}'(\tau(s))\dot{\tau}(s)] = 0$ for all $s \geq t$ as well. So, the integral on the right-hand side of (58) vanishes, and (58) implies,

$$\hat{m}(\tau(t)) = \theta \ln[\tau(t)] - \frac{1+\theta}{\rho}\left[R\tau + \frac{\sigma^2}{2}B(\tau(t))^2\right]\ . \tag{59}$$

REFERENCES

Barro, Robert. J., and José F. Ursua (2009), "Stock-Market Crashes and Depressions", NBER Working Paper No. w14760.

Claessens, S., M. A. Kose and M. E. Terrones (2009): What happens during recessions, crunches and busts? Economic Policy, 24, 653-700.

Cohen, D., and P. Michel (1988): How Should Control Theory be Used by a Time-Consistent Government? Review of Economic Studies, 55, 263-274.

Comon, E. (2001): Extreme Events and the Role of Learning in Financial Markets. Chapter 1 of PhD thesis titled "Essays on investment and consumption choice", Harvard University.

Conesa, J. C., S. Kitao, and D. Krueger (2009): Taxing Capital? Not a Bad Idea After All!, American Economic Review, 99, 25–48

Cox., D. R., and H. D. Miller (1965): The Theory of Stochastic Processes, London: Methuen & Co. Ltd.

Dixit, A. K., and R. S. Pindyck (1994), Investment under Uncertainty, New Jersey: Princeton University Press.

Dreyfous, S. E. (1965): Dynamic Programming and the Calculus of Variations. *R-441-PR, A Report Prepared for United States Airforce Project Rand*, The Rand Corporation, August 1965.

Fehr, H., and F. Kindermann (2015): Taxing capital along the transition-Not a bad idea after all?, Journal of Economic Dynamics and Control, 51, 64-77.

Gelman, A., J. B. Carlin, H. S. Stern and D. B. Rubin (2004): Bayesian Data Analysis. Second Edition, Chapman and Hall/CRC.

Grant, C., C. Koulovatianos, A. Michaelides, and M. Padula (2008): Evidence on the Insurance Effect of Marginal Income Taxes. Centre for Economic Policy Research Discussion Paper No. 6710.

Grant, C., C. Koulovatianos, A. Michaelides, and M. Padula (2010): Evidence on the Insurance Effect of Redistributive Taxation. Review of Economics and Statistics, 92, 965-973.

Institute for Fiscal Studies (2013), The IFS Green Budget: February 2013, Edited by Carl Emmerson, Paul Johnson and Helen Miller, DOI: 10.1920/re.ifs.2013.0074, downloadable from, http://www.ifs.org.uk/budgets/gb2013/gb2013.pdf

International Monetary Fund (2009), Global Economic Policies and Prospects, Group of Twenty, Meeting of the Ministers and Central Bank Governors, March 13–14, 2009, London, U.K., Note by the Staff of the International Monetary Fund, downloadable from http://www.imf.org/external/np/g20/pdf/031909a.pdf

Karlin, S. and Taylor, H.M. (1981): A Second Course in Stochastic Processes. New York: Academic Press.

Kerman, J. (2011): Neutral noninformative and informative conjugate beta and gamma prior distributions, Electronic Journal of Statistics, 5, 1450-1470.

Koulovatianos, C., L. J. Mirman, and M. Santugini (2009): Optimal growth and uncertainty: Learning, Journal of Economic Theory, 144, 280-295.

Koulovatianos, C. and V. Wieland (2011): Asset Pricing under Rational Learning about Rare Disasters. Centre for Economic Policy Research (CEPR) Discussion Paper No. 8514.

Kushner, H. J. (1967): Stochastic Stability and Control. New York: Academic Press.

Liptser, R. S. and Shiryayev, A. N. (1977): Statistics of Random Processes I. New York: Springer-Verlag.

Merton, R.C. (1971): Optimum Consumption and Portfolio Rules in a Continuous-Time Model. Journal of Economic Theory, 3, 373-413.

Mirrlees, J. (1974): Notes on Welfare Economics, Information and Uncertainty, in M. Balch, D. McFadden and S.Wu (Eds.), Essays on Economic Behaviour under Uncertainty, Amsterdam: North-Holland.

Papoulis, A. (1991), Probability, Random Variables, and Stochastic Processes. 3^{rd} Edition, New York: McGaw-Hill, Inc.

Papoulis, A. and S. U. Pillai (2002), Probability, Random Variables, and Stochastic Processes, 4^{th} Edition, New York: McGaw-Hill, Inc.

Ross, S. M. (2003): Introduction to Probability Models, 8^{th} Edition, New York: Academic Press.

Stockman, D. R. (2001): Balanced-Budget Rules: Welfare Loss and Optimal Policies, Review of Economic Dynamics, 4, 438-459.

Stockman, D. R. (2004): Default, Reputation and Balanced-Budget Rules, Review of Economic Dynamics, 7, 382-405.

Stokey, N. L. (2008): The Economics of Inaction, Princeton, NJ: Princeton University Press.

Varian, H. R. (1980): Redistributive Taxation as Social Insurance, Journal of Public Economics, 14, 49-68.

Xie, D. (1997): On time inconsistency: a technical issue in Stackelberg differential games. Journal of Economic Theory, 76, 412-430.

CENTER FOR FINANCIAL STUDIES

CFS Working Paper Series

Recent Issues

All CFS Working Papers are available at *www.ifk-cfs.de*.

No.	Authors	Title
613	John Donaldson, Christos Koulovatianos, Jian Li and Rajnish Mehra	*Demographics and FDI: Lessons from China's One-Child Policy*
612	Hans Gersbach	*Contingent Contracts in Banking: Insurance or Risk Magnification?*
611	Christian Leuz	*Evidence-Based Policymaking: Promise, Challenges and Opportunities for Accounting and Financial Markets Research*
610	Christian Leuz and João Granja	*The Death of a Regulator: Strict Supervision, Bank Lending and Business Activity*
609	Christian Leuz, Steffen Meyer, Maximilian Muhn, Eugene Soltes, and Andreas Hackethal	*Who Falls Prey to the Wolf of Wall Street? Investor Participation in Market Manipulation*
608	Brandon Gipper, Luzi Hail, and Christian Leuz	*On the Economics of Audit Partner Tenure and Rotation: Evidence from PCAOB Data*
607	Vanya Horneff, Raimond Maurer and Olivia S. Mitchell	*Putting the Pension Back in 401(k) Retirement Plans: Optimal versus Default Longevity Income Annuities*
606	Lutz Kilian and Xiaoqing Zhou	*The Propagation of Regional Shocks in Housing Markets: Evidence from Oil Price Shocks in Canada*