Life-cycle Horizon Learning, Social Security Reform, and Policy Uncertainty

Erin Cottle Hunt*
University of Oregon

This Version: January 12, 2018
Current Version Online

Abstract

The US social security system faces funding pressure due to the aging of the US population. Future social security benefits may be reduced or future taxes may be increased. This paper examines the welfare cost of social security reform and social security policy uncertainty under rational expectations and under learning. I develop two theories of bounded rationality called Life-cycle Horizon Learning (LCH) and Finite Horizon Life-cycle Learning (FHL). In both models, agents use adaptive expectations to forecast future aggregates, such as wages and interest rates. The expectations are updated each period as agents acquire additional information. This adaptive learning feature introduces cyclical dynamics along a transition path, which magnifies the welfare cost of changes in policy and policy uncertainty. I model policy uncertainty as stochastic process in which reform takes place in one of two periods as either a benefit cut or a tax increase. I find the welfare cost of this policy uncertainty is equivalent to less than 0.3% of period consumption for the cohort of agents most harmed in a standard, rational expectations framework. The welfare cost of policy uncertainty is larger in the learning models; the worst-off cohort in the Life-cycle Horizon Learning model would be willing to give up 1.98% of period consumption to avoid policy uncertainty, and would be willing to give up 0.98% of period consumption to avoid the cyclical dynamics introduced via learning.

Key words: Learning, Bounded Rationality, Policy Uncertainty, Social Security Reform, Retirement Savings, Overlapping Generations Model, Fiscal Sustainability, Bond-Financed Deficits

*ecottle@uoregon.edu, https://blogs.uoregon.edu/erinhunt/

1 Introduction

"Uncertainty about Social Security's future magnifies the anxiety that many Americans experience as they plan and prepare for retirement."

- Commission on Retirement Security and Personal Savings
Bipartisan Policy Center, June 2016

The aging of developed nations will strain unfunded pension systems like the U.S. social security system. The Old Age Survivor and Disability Insurance program (OASDI), commonly referred to as social security, is one of the largest transfer programs in the world, with total expenditures of nearly $922 billion in 2016. Social security provides benefits to 61 million people, the majority of whom receive retirement benefits. Social security's cost has exceeded its tax income in each year since 2010 and the Social Security Administration (SSA) projects this relationship to extend through the long-term. The SSA Board of Trustees project that an immediate and permanent payroll tax rate increase of 2.76 percentage points (12.4 percent to 15.16 percent) or an immediate and permanent benefit cut of 17 percent applied to all current and future beneficiaries would be needed to insure sufficient tax revenue to cover promised benefits over the long-term (SSA, 2017).

Given the projected shortfall between social security tax revenues and promised benefits, reform seems likely. However, it is not clear if or when the government will act. The uncertainty regarding when and how social security will be reformed is interesting and economically important. A survey conducted by the American Life Panel in 2011 found that 56% of respondents are "not too confident" or "not confident at all" that the social security system will be able to provide them the level of benefits currently promised. Wariness in the government's ability to pay benefits is decreasing in age, with 75% of respondents under the age of 40 reporting they are not confident they will receive benefits (47% "not too confident" and 28% "not confident at all").[1] Social security benefits are a major source of retirement income for many Americans, and uncertainty regarding future benefits could impact savings decisions. The economy wide uncertainty regarding social security deficits and debt accumulation is also interesting, as rapid debt growth could strain the economy (see, for example, Davig et al. (2010)).

I explore the macroeconomic consequences of social security policy uncertainty in a general equilibrium lifecycle model. Agents' expectations about future social security policy drive agent-level decision making. These decisions, combined with

[1] The RAND American Life Panel is a nationally representative, probability-based panel of over 6000 members ages 18 and older who are regularly interviewed over the internet for research purposes. This question comes from the 2011 survey *Well-being ms179*. More information about the ALP is available at alpdata.rand.org.

realized government policy, determine macroeconomic output and prices. The way in which agents form expectations has a large impact on both the short-run responses to changes in policy (or possible changes in policy) and the long-run level of capital and output in the economy. I consider rational expectations and adaptive expectations. My paper contributes to a growing literature that examines the response to anticipated fiscal policy in models with adaptive expectations (see Evans et al. (2009), Mitra et al. (2013), Gasteiger and Zhang (2014), and Caprioli (2015)).

The rational expectations hypothesis is standard in macroeconomics and I develop a rational expectations model as the baseline for my analysis. I also relax the assumption of rational expectations and explore the consequences of agents using adaptive expectations in a model with an aging society and social security reform. One criticism of the rational expectations hypothesis is that it requires agents to possess more knowledge about the structure of the economy than any econometrician possesses about the actual economy. Models of adaptive learning relax that requirement, and allow for boundedly rational agents (Sargent (1993), Evans and Honkapohja (2001)).

Adaptive learning has several benefits. Models of adaptive learning produce more realistic dynamics that better fit observed data. Eusepi and Preston (2011) demonstrate that infinite-horizon learning amplifies technology shocks in a real business cycle model and creates more realistic dynamics. Branch and Mcgough (2011) show that a calibrated model with heterogenous beliefs (some agents are fully rational, others use N-step ahead adaptive learning) provides a closer fit to business cycle data than the rational expectations baseline. Models of adaptive learning are also better able to match survey data on expectations (Branch (2004)), asset price volatility (Bullard and Duffy (2001), Adam et al. (2016)), and experimental data on expectation formation (Adam (2007) and Pfajfar and Santoro (2010)). Additionally, adaptive learning can be viewed as a robustness check to examine the stability properties of a rational expectations equilibrium (see Evans and Honkapohja (2001) and (2009)).

In this paper, I introduce a two new frameworks for modeling bounded rationality that I call life-cycle horizon learning, and finite horizon life-cycle learning. I simulate the underlying demographic changes and the two policy changes that the SSA suggests would be sufficient to fund future benefits. I find the welfare cost of implementing a reform is much larger for some cohorts in the learning models, particularly in the life-cycle horizon learning model, compared to a model fully rational agents. Cohorts of agents alive before and after reforms can be negatively impacted by the cyclical dynamics adaptive expectations introduce into the model.

I also explore the consequences of policy uncertainty by modeling reform that can take place in one of two different dates and can take the form of a tax increase or a benefit cut. I find that the welfare cost of this type of policy uncertainty is small in the rational expectations model. The consumption equivalent variation that makes an agent indifferent between an announced policy reform or policy uncertainty (between four different options, tax increases or benefit cuts possible in two different periods) is at most 0.3% of period consumption. I find that the welfare cost to be much larger for agents in a life-cycle horizon learning model. The most harmed agents who experience an announced policy change in a rational expectations framework would have to give up 1.9% of their period consumption to be indifferent between announced policy in a life-cycle horizon model and policy uncertainty in a life-cycle horizon learning model.

I present the model in section 2 and introduce Life-cycle Horizon Learning and Finite Horizon Learning in sections 3 and 4. The model is calibrated in section 5. Policy applications, including social security reform uncertainty are explored in section 6. Welfare analysis is in sections 7 and 8. Section 9 includes alternative modeling assumptions, and section 10 concludes.

2 Model

2.1 Households

Households live for J periods, choose asset allocation (a^j for $j = 1, \cdots J-1$) in the first $J-1$ periods of life, and consumption (c^j for $j = 1, \cdots, J$) in all J stages of life, to maximize utility, taking price, and government social security policy (tax rates τ, and benefit z) as given. Agents receive wage w_t for labor provided in period t, and retire at date $T \leq J$. The gross real return on savings in period t is given by R_{t+1}.

Household's maximize discounted expected lifetime utility (1) subject to period budget constraints (2). Here $E_t^*(x)$ indicates the time t expectation of x. The star indicates that the expectations need not be rational.

$$\max_{a^j_{t+j-1}} E_t^* \sum_{j=1}^{J} \beta^{j-1} u(c^j_{t+j-1}) \tag{1}$$

$$c^j_{t+j-1} + a^j_{t+j-1} \leq R_{t+j-1} a^{j-1}_{t+j-2} + y^j_{t+j-1} \qquad \text{for } j = 1, \cdots, J \tag{2}$$

Here y^j indicates period labor income during working life and social security income after retirement:

$$y^j_{t+j-1} = (1 - \tau_{t+j-1}) w_{t+j-1} \quad \text{for } j < T$$
$$y^j_{t+j-1} = z^j_{t+j-1} \quad \text{for } j \geq T$$

The lifespan is certain and agents do not have a bequest motive, so they exhaust all of their resources in the final period of life.

$$a^J_{t+J-1} = 0 \tag{3}$$

The household first order conditions are:

$$u'(c_{t+j-1}^{j}) = \beta E_t^*[R_{t+j}u'(c_{t+j}^{j+1})] \quad \text{for } j = 1, \cdots, J-1 \tag{4}$$

The agent (trivially) chooses to consume all of her resources in the final period of life, that is $c_{t+J-1}^{J} = R_{t+J-1}a_{t+J-2}^{J-1} + z_{t+J-1}^{J}$.

Labor is supplied inelastically and preferences are given as the standard constant elasticity of substitution function:

$$\begin{aligned} u(c) =& \frac{c^{1-\sigma}}{1-\sigma} \quad \text{if } \sigma \neq 1 \\ u(c) =& \ln(c) \quad \text{if } \sigma = 1 \end{aligned}$$

2.2 Demographics

To start off the economy, I assume that in period zero, there are J cohorts who enter the economy with given asset holdings according to their age. I assume that the initial young enter the economy with zero assets, and all other cohorts enter with a_{-1}^{j} for $j = 1, \cdots, J-1$. Successive cohorts enter the model with zero assets when they are young.

N_t indicates generation t and is given by number of young (i.e., the generation born) at time t. The population grows at rate n_t such that

$$N_t = (1 + n_t)N_{t-1} \tag{5}$$

The population at any time t is the sum of all living cohorts. If the growth rate n is constant, the population can be expressed relative to the initial young or initial old:

$$\sum_{j=1}^{J} N_{t+1-j} = \sum_{j=1}^{J}(1+n)^{1-j}N_t = \sum_{j=1}^{J}(1+n)^{J-j}N_{t+1-J}.$$

2.3 Production

The consumption good in the economy (Y_t) is produced by single firm (or equivalently many small firms) using a constant elasticity of substitution technology that takes aggregate capital (K_t) and labor (H_t)[2] as inputs and produces the consumption good according to:

$$Y_t = F(K_t, H_t) = K_t^{\alpha}H_t^{1-\alpha}$$

[2]Labor is supplied inelastically by all working cohorts, thus $H_t = \sum_{j=1}^{T-1}(1+n)^{1-j}N_t$ which is less than the total population. Recall N_t is the number of young.

The parameter α measures the intensity of use of capital in production.[3] Factor markets are competitive and capital and labor (hours worked) are paid their marginal products. The gross real interest rate R_t is given by:

$$R_t = F_K(K_t, H_t) + 1 - \delta \tag{6}$$

where δ is the rate of depreciation. The wage rate w_t is given by

$$w_t = F_H(K_t, H_t) \tag{7}$$

2.4 Government

The government runs a modified pay-as-you-go social security system. The government pays retirement benefits to the retired generations by taxing the working generations and by (possibly) issuing debt.

The pay-roll tax rate is τ_t. The tax has two components, a baseline tax rate τ_t^0, and a Leeper tax τ_t^1 that responds to the level of government debt. Although actual social security taxes do not change based on government debt, a Leeper tax can be thought of as a way to capture legislative unease with increasing debt. I consider a Leeper tax of the following form:

$$\tau_t = \tau_t^0 + \tau_t^1(B_t/H_t). \tag{8}$$

where $\tau^0 \in [0, 1]$ is the baseline tax rate when government debt is zero, $\tau^1 \geq 0$ is the incremental tax, and B_t/H_t is government debt per labor hours. Notationally, a tax rate *without* a superscript (τ_t) will refer to the entire pay-roll tax $\tau_t = \tau_t^0 + \tau_t^1(B_t/H_t)$.[4]

Social security benefits z_t^j are paid according to a benefit earning rule:

$$z_t^j = \phi_t w_{t+T-j}. \tag{9}$$

where z_t^j represents the benefit paid to a retiree of age j in time t. The parameter ϕ_t is the replacement rate and shows how much of a workers' wage-indexed average period earnings (w_{t+T-j}) are replaced by social security benefits.[5]

[3] I abstract from technological growth, although it is straightforward to incorporate this into the model. If technological process was included in the model, aggregate variables would grow that the rate $(1 + n)(1 + g)$ along the balanced growth path (if g indicates the growth rate of technology). In all other regards, the model is identical.

[4] The Leeper tax given by equation (8) leaves open the possibility of a total tax rate greater than one hundred percent if bonds levels are high or if τ^1 is large (i.e., $\tau_t^0 + \tau_t^1(B_t/H_t) > 1$, for large B_t/H_t and/or large τ^1). I avoid this by imposing the restriction τ^1 is either the exogenous parameter value chosen by the economist, or the (smaller) parameter value such that $\tau_t^0 + \tau_t^1(B_t/H_t) = 1$. In the policy experiments that follow, the majority of the tax burden will come from the baseline tax rate τ^0, and the Leeper tax τ^1 will be used only to calibrate a dynamically efficient steady state with positive bonds.

[5] There is no technological growth in the model, so this roughly corresponds to the US system.

The government is not required to balance the social security budget. If social security taxes are less than social security benefits, the social security program runs a deficit. The time t deficit is defined as the time t social security benefits less the pay-roll tax revenue:

$$\text{Deficit}_t = \sum_{j=T}^{J} N_{t+1-j}\phi_t w_{t+T-j} - H_t(\tau_t^0 + \tau_t^1(B_t/H_t))w_t$$

Here H_t indicating the working age population.

The government issuance of bonds is equal to the gross interest on outstanding debt plus the social security deficit from the previous period.

$$B_{t+1} = R_t B_t + \sum_{j=T}^{J} N_{t+1-j}\phi_t w_{t+T-j} - H_t(\tau_t^0 + \tau_t^1(B_t/H_t))w_t \tag{10}$$

2.5 Markets

There are four markets: labor, capital, bonds, and goods. Prices adjust in equilibrium to ensure all markets clear.

Labor market clearing requires the total number of hours worked equal the labor input of the representative firm:

$$H_t = \sum_{j=1}^{T-1} N_{t+1-j} \tag{11}$$

Asset market clearing requires that aggregate capital and bonds next period are equal to the savings of each cohort:

$$K_{t+1} + B_{t+1} = \sum_{j=1}^{J} N_{t+1-j}a_t^j \tag{12}$$

Bond market clearing is ensured by the government's flow budget constraint, reprinted below:

$$B_{t+1} = R_t B_t + \sum_{j=T}^{J} N_{t+1-j}\phi_t w_{t+T-j} - H_t(\tau_t^0 + \tau_t^1(B_t/H_t))w_t \tag{10}$$

Goods market clearing requires that aggregate output is equal to aggregate consumption and aggregate investment. The goods market clears by Walras law. The goods market clearing equation is printed below for reference:

$$F(K_t, H_t) = \sum_{j=1}^{T} N_{t+1-j} c_t^j + K_{t+1} - (1-\delta) K_t$$

Along the balanced growth path, cohort size (N_t), labor hours worked (H_t), output (Y_t), capital (K_t), and bonds (B_t) all grow at rate n. Therefore, it will be convenient to rewrite the market clearing equations (10) and (12) in per-hours terms by defining $b_t = B_t/H_t$, and $k_t = K_t/H_t$. I will refer to variables in per-hours terms as "efficient."

The capital and bond market clearing equations can be written in efficient terms as:

$$(k_{t+1} + b_{t+1})(1+n_t) = \frac{\sum_{j=1}^{J} N_{t+1-j} a_t^j}{H_t} \tag{13}$$

$$(1+n_t) b_{t+1} = R_t b_t + \frac{\sum_{j=T}^{J} N_{t+1-j} \phi_t w_{t+T-j}}{H_t} - (\tau_t^0 + \tau_t^1 (B_t/H_t)) w_t \tag{14}$$

2.6 Rational Expectations Equilibrium

Definition 1 *Given initial conditions k_0, b_0, $a_{-1}^1, \cdots a_{-1}^{J-1}$, and an initial population $\sum_{j=1}^{J}(1+n)^{1-j} N_0$ (where N_0 is the the initial cohort of young and n is the population growth rate), a competitive equilibrium is a sequences of functions for the household savings $\left\{a_t^1, a_t^2, \cdots, a_t^J\right\}_{t=0}^{\infty}$, production plans for the firm, $\{k_t\}_{t=1}^{\infty}$, government bonds $\{b_t\}_{t=1}^{\infty}$, factor prices $\{R_t, w_t\}_{t=0}^{\infty}$, and government policy variables $\{\tau_t^0, \tau_t^1, \phi_t\}_{t=0}^{\infty}$, that satisfy the following conditions:*

1. *Given factor prices and government policy variables, individuals' decisions solve the household optimization problem (1) and (2)*

2. *Factor prices are derived competitively according to (6) and (7)*

3. *All markets clear according to (11), (13), and (14)*

There are $J+1$ equilibrium equations, which hold for time periods $t = 0, \cdots \infty$. The equilibrium equations include the capital clearing equation (13), the bond market clearing equation (14), and $J-1$ household first order conditions (15). These are reprinted below:

$$(k_{t+1} + b_{t+1})(1 + n_t) = \frac{\sum_{j=1}^{J} N_{t+1-j} a_t^j}{H_t} \tag{13}$$

$$(1 + n_t)b_{t+1} = R_t b_t + \frac{\sum_{j=T}^{J} N_{t+1-j} \phi_t w_{t+T-j}}{H_t} - (\tau_t^0 + \tau_t^1 (B_t/H_t))w_t \tag{14}$$

$$(R_t a_{t-1}^{j-1} + y_t^j - a_t^j)^{-\sigma} = \beta E_t[R_{t+1}(R_{t+1} a_t^j + y_{t+1}^{j+1} - a_{t+1}^{j+1})^{-\sigma}] \tag{15}$$
$$\text{for } j = 1, \cdots, J-1$$

Here y^j indicates period labor income during working life and social security income after retirement. Note that $a^0 = 0$, and factor prices are given by (6) and (7).

The equilibrium definition above accommodates a time-varying population growth rate n_t. In a steady-state, the growth rate of the population is constant. When the growth rate is constant, several terms can be written concisely by defining $\nu = N_t/H_t = (\sum_{j=1}^{T-1}(1+n)^{1-j})^{-1}$.

Definition 2 *The steady state is a collection $\{k, b, a^1, \cdots, a^J\}$ that solves:*

$$(k + b)(1 + n) = \nu \sum_{j=1}^{J} (1 + n)^{1-j} a^j \tag{16}$$

$$(1 + n)b = R(k)b + \nu \sum_{j=T}^{J} (1 + n)^{1-j} \phi_t w(k) - (\tau^0 + \tau^1 b)w(k) \tag{17}$$

$$(R(k)a^{j-1} + y^j - a^j)^{-\sigma} = \beta \left[R(k)(R(k)a^j + y^{j+1} - a^{j+1})^{-\sigma}\right] \tag{18}$$
$$\textit{for } j = 1, \cdots, J-1$$

with $y^j = (1 - (\tau^0 + \tau^1 b))w(k)$ for $j < T$, $y^j = \phi w(k)$ for $j \geq T$, $a^J = 0$, and factor prices $R(k)$ and $w(k)$ given by (6) and (7).

The equilibrium of the model is not unique; for many parameter combinations there are two steady states. The number of steady states depends on the model parameters and can be characterized by a saddle-node bifurcation. Chalk (2000) discusses the saddle-node bifurcation in a similar OLG model with government debt and physical capital.[6]

The intuition of the bifurcation in this model is that as a given parameter changes (like the payroll tax rate τ^0 or the population growth rate n), it impacts the social security deficit or surplus. Suppose the model is calibrated such that two equilibria exist and thus two steady states exist. As the deficit increases (endogenously in response to a parameter change), this increases government debt and crowds out

[6]See Azariadis (1993) (Chapter 8 and Chapter 20) for a discussion of saddle-node bifurcations in planar OLG models.

capital, pushing the two steady states closer together. At a critical value of the parameter, only one steady state exists; beyond that, deficits are too large, government debt is explosive, and no steady states (and no equilibria) exist.

In the exercises that follow, the model will be calibrated such that two steady states exist. I analyze the determinacy of equilibria by linearizing the model around a steady state and computing the eigenvalues of the linearized system (see Laitner (1990)). There are three predetermined variables in the model (k, b, and a^{J-1}) and J-2 free variables (a^1,...,a^{J-2}). There are three possible cases. If three eigenvalues of the linearized system are less than one in modulus and the remaining $J-2$ have modulus greater than one, the system is determinate. If more than three eigenvalues lie inside the unit circle, the system is indeterminate. If more than $J-2$ eigenvalues lie outside the unit circle, the system is explosive. I have confirmed numerically that when two steady states exist, the steady state with higher capital is stable (determinate), and the lower capital steady state is explosive. I will focus only on the determinate solutions in this paper.[7]

3 Life-cycle Horizon Learning

Under the rational expectations hypothesis (RE), households make optimal savings and consumption choices given their (rational) forecasts of future prices and policy. Households make decisions at age zero looking forward over their entire life cycle. These rational agents are fully forward looking, and consider every stage of the life-cycle when making decisions. Under RE agents know underlying equations that govern the economy and form expectations of future variables using the mathematical expected value.

This paper backs away from RE and proposes a learning model in which agents combine limited knowledge about the structure of the economy with adaptive forecasts for future macroeconomic aggregates. As a baseline, I consider agents who do not have perfect foresight of factor prices. Thus, agents do not know the future path of wages or interest rates over their life-cycle. Agents use adaptive expectations to forecast future prices, but behave optimally in all other ways. Agents update their forecasts as more information becomes available. I call this behavioral assumption Life-cycle Horizon Learning (LCH learning).

Under LCH learning, households are forward looking and make optimal savings and consumption choices given their (adaptive) forecasts of future macroeconomic aggregates. Adaptive expectations may be viewed as a special case of adaptive learning suitable for non-stochastic models. The adaptive expectations assumption in this

[7]Chalk (2000) also finds that the high capital steady state is determinate and the low capital steady state is explosive in a many-period OLG model with debt and capital.

paper is analogous to constant gain least squares learning in a model with random shocks (like a productivity shock). LCH learning is similar to infinite-horizon learning, developed in Marcet and Sargent (1989) and emphasized by Preston (2005 and 2006). The key difference is that infinite-horizon learning models are based on a representative agent who lives for an infinite number of periods. LCH learning applies the same forward looking behavior to finitely lived agents in an overlapping generations life-cycle model.[8] Throughout this paper, I assume homogenous expectations across agents.

Agents in the LCH learning model forecast wages and interest rates using adaptive expectations of the following form:

$$w^e_{t+1} = \gamma w_t + (1-\gamma)w^e_t \tag{19}$$
$$R^e_{t+1} = \gamma R_t + (1-\gamma)R^e_t \tag{20}$$

with $\gamma \in (0,1)$. Here w^e indicates expected wage, and R^e_t indicates expected interest rate. Agents also form expectations at time t of prices in period $t+j$ for $j > 1$:

$$w^e_{t+j} = w^e_{t,t+j} = w^e_{t+1} \tag{21}$$
$$R^e_{t+j} = R^e_{t,t+j} = R^e_{t+1} \tag{22}$$

When the Leeper tax is non-zero ($\tau^1 > 0$), then agents need to know future bond levels in order to estimate their after-tax pay. For simplicity, I assume that agents forecast bonds using the same adaptive learning rule they use to forecast prices.

$$b^e_{t+1} = \gamma b_t + (1-\gamma)b^e_t \tag{23}$$
$$b^e_{t,t+j} = b^e_{t,t+1} \quad \text{for } j > 1 \tag{24}$$

Agents enter the model with knowledge of the previous period's prices, bonds, and expectations. At any moment in time, all agents in the model have the same expectation for future prices and bonds. This assumption is equivalent to assuming agents inherit expectations from the previous generation. I consider alternative way for agents to forecast their income net of taxes in section 9.

LCH learning can be viewed as a small deviation from rational expectations in the sense that agents are still forward looking and only use a different rule to forecast future aggregates. Adaptive expectations plausible in a world in which the true data generating process is complex. Adaptive expectations are optimal if agents think that capital follows and IMA(1,1) process. That is, expectations of the form given in Equation (19) are *rational* if the change in capital has the following form $\Delta k_t = \epsilon_t + \theta\epsilon_{t-1}$ for shocks ϵ and parameter $\theta \in (0,1)$. The use of adaptive expectations is

[8] LCH learning is similar to the model developed in Bullard and Duffy (2001), in which agents forecast inflation over a finite life-cycle and make optimal choices based on those forecasts.

equivalent to agents believing that capital has a mixture of permanent and transitory shocks (See Muth (1960)). The gain parameter, γ, has a natural interpretation in this context. If all shocks are transitory, the optimal forecasting rule is constant, i.e., γ=0. In contrast, if all shocks are permanent, then the best forecasting rule is a random walk, i.e., $\gamma = 1$. In this particular model, all shocks are permanent, but agents are not endowed with this knowledge. I choose the gain parameter used in simulations to minimize the welfare cost of agents of inaccurately forecasting along the transition path. I explore alternative gain parameters as robustness checks.[9]

In the baseline LCH learning model, agents have full knowledge of government policy. Agents know the future path of social taxes and the benefit replacement rate if there is no policy uncertainty. If future policy is stochastic, agents form expectations of future policy parameters using *rational* expectations. Agents know the finite set of policy parameters (and/or reform dates) and the relative probability of each. I relax this assumption and have agents learn about future government policy adaptively in section 9.

A young agent in the LCH learning model chooses first period savings and consumption and plans future savings and consumption to satisfy her $J-1$ first order equations and her lifetime budget constraint. The young agent's time t plan of second period savings is denoted $a^2_{t,t+1}$. In the following period, her time $t+1$ actual choice of second period savings is given by a^2_{t+1}. Her time $t+1$ plan does not have to be consistent with her time t choice. She can update her savings decision based on the new information she receives in period $t+1$. Abstracting from policy uncertainty (assuming the path of taxes and social security benefit rates are predetermined), the young agent in time t solves:

$$(R_{t+j-1}a^{j-1}_{t+j-2} + y^{j}_{t+j-1} - a^{j}_{t+j-1})^{-\sigma} = \beta R^e_{t+j}(R^e_{t+j}a^j_{t+j-1} + y^{e,j+1}_{t+j} - a^{j+1}_{t,t+j})^{-\sigma} \quad (25)$$
$$\text{for } j = 1, \cdots, J-1$$

with $y^{e,j+1}$ indicating expected period labor income or social security income

$$y^{e,j+1}_{t+j} = (1-\tau_{t+j})w^e_{t+j} \quad \text{for } j < T$$
$$y^{e,j+1}_{t+j} = \phi_{t+j}w^e_{t+j+T-j} \quad \text{for } j \geq T$$

Similarly, an agent of age two solve $J-2$ first order equations, for the remaining $J-2$ periods of her life-cycle. In total, the J cohorts alive in any period solve $\sum_{j=1}^{J-1} j = \frac{(J-1)J}{2}$ first order conditions. Together, the decisions of households of all ages, asset market clearing (13), and government bonds (14), and the expectation equations (19), (20), and (23), create a recursive system that governs the dynamics of the economy.

[9]Evans and Ramey show that by appropriately tuning the free parameters of the forecast rule agents can obtain the best forecast rule within a given class of underparameterized learning rules (Evans and Ramey (2006)).

4 Finite Horizon Life-cycle Learning

This paper proposes an additional new learning model for finitely lived agents who form expectations over a short horizon (less than or equal to the length of their lifecycle) called Finite Horizon Life-cycle Learning (FHL). This learning model is similar to N-step ahead optimal learning, developed by Branch et al. (2013). In both cases, agents make optimal choices based on their forecast of future prices (and potentially other macro variables) and their budget constraint over a finite horizon. Agents using FHL learning look forward for a fixed number of periods and make decisions based on this short-planning horizon. They update their choices each period as their planning horizon moves forward and they receive new information.

In order to solve the FHL model model, it is necessary to specify a terminal condition for asset holdings at the end of the planning horizon. When the planning horizon equals the end of the life-cycle, the terminal condition is simply that the agent dies without debt (and thus chooses optimally to die with zero assets). When the planning horizon is shorter than the life-cycle, it is less obvious what the terminal condition should be.

One simple alternative is to impose that assets are non-negative at the end of the planning horizon. Under this assumption, agents will choose to hold zero assets at the end of the planning horizon. This assumption drives the main results in the short-planning horizon literature and leads to time-inconsistent behavior (see Caliendo and Aadland (2007) and Park and Feigenbaum (2017)). This assumption is somewhat unsatisfactory, because agents treat periods beyond the planning horizon as if they did not exist. An alternative assumption is to have agents hold wealth at the end of the planning horizon. As a baseline, I assume that agents plan to hold the same amount of wealth at the end of the planning horizon that older cohorts held at the same age. Specifically, agents adaptively forecast the steady-state value of age-specific wealth at the end of their planning horizon.

As in the LCH learning model, agents using FHL learning forecast future factor prices and bonds according to (19) - (23). FHL learners also need to forecast a terminal wealth holding for the end of the planning horizon. I assume that agents forecast

$$a^{j,e}_{t,terminal} = \gamma a^{j}_{t-1} + (1-\gamma) a^{j,e}_{t-1,terminal} \quad \text{for } j = 1, \cdots, J-1 \tag{26}$$

Here, $a^{j,e}_{t,terminal}$ is amount of asset an age $j-1$ agent expects to hold at the end of the period when they are age j. This asset amount is based on the observed asset holding of age j agents from last period, and the forecast from last period.

As an example, suppose the planning horizon length is three. That is, agents looks forward three periods when making decisions. A young agent chooses first period savings and consumption, and plans saving and consumption for period two and three, using three first order equations and a terminal condition:

$$(y_t^1 - a_t^1)^{-\sigma} = \beta R_{t+1}^e (R_{t+1}^e a_t^1 + y_{t+1}^{e,2} - a_{t,t+1}^2)^{-\sigma} \tag{27}$$

$$(y_{t+1}^{e,2} - a_{t,t+1}^2)^{-\sigma} = \beta R_{t+2}^e (R_{t+2}^e a_{t,t+1}^2 + y_{t+2}^{e,3} - a_{t,t+2}^3)^{-\sigma} \tag{28}$$

$$(y_{t+2}^{e,3} - a_{t,t+2}^3)^{-\sigma} = \beta R_{t+3}^e (R_{t+3}^e a_{t,t+2}^3 + y_{t+2}^{e,4} - a_{t,terminal}^4)^{-\sigma} \tag{29}$$

Older agents follow a similar pattern. Agents who have three or fewer periods of life remaining using the standard terminal condition $a^J = 0$.

The FHL learning model can be written as a recursive system. The system includes the first order equations for the households, asset market clearing (13), and government bonds (14), and the expectation equations (19) - (23) and (26). For a planning horizon of length H, there will be $J - H$ terminal conditions and $H(J-H) + \frac{H(H-1)}{2}$ household first order equations. When the planning horizon is equal to the life-cycle, the system is identical to the Life-cycle Horizon learning model.

The shortest possible planning horizon is one. Agents who only look forward one period make their savings consumption decision based on their Euler equation, their budget constraint, and the assets they believe they need to hold at the end of the next period (the terminal condition). Branch et al. (2013) call this type of learning one-step-ahead optimal learning and compare and contrast it to Euler-equation-learning, a behavioral primitive in which agents do not explicitly consider their transversality condition. I plan to explore Euler equation learning in an OLG framework in future work.

5 Calibration

Agents enter the model at age 25 and live for six periods ($J = 6$). Each period is 10 years, and agents die with certainty at age 85. The capital share of income, discount factor, inverse elasticity of substitution, and depreciation rate are all set to standard parameter values[10]. The total factor productivity parameter A is arbitrarily set to 10 (results do not depend on this assumption). For the majority of the exercises in this paper, the population growth rate is 0.01. The parameter values are listed in Table 1.

[10] See, for example, Branch et al. (2013). My discount factor β is slightly closer to one and the depreciation rate is slightly higher. I choose the higher discount factor and depreciation rate to better match the size of tax increase necessary to ensure long-run solvency of social security

	Parameter	Value
α	Capital share of income	$1/3$
β	Discount factor	$(1/1+0.005)^{10}$
σ	Inverse elasticity of substitution	1
δ	Depreciation	$1-(1-0.10)^{10}$
n	Population growth rate	0.01
A	TFP factor	10

Table 1: Baseline parameters for 6 period model

The SSA estimates the current ratio of social security beneficiaries to workers to be 0.35 (as of the 2017 Trustee Report). This ratio is expected to increase to 0.46 by 2035, and to 0.5 by 2095 under the medium cost assumptions. The increase in the ratio of beneficiaries to workers is driven by both increasing lifespans for retirees, and declining birthrates for working generations. I abstract from these details in the model, and capture all population changes using the growth rate n. For the majority of the exercises in section 6, I will choose n such that the ratio of retirees to workers is 0.3 and then increases gradually to 0.485, as illustrated in Figure 1.

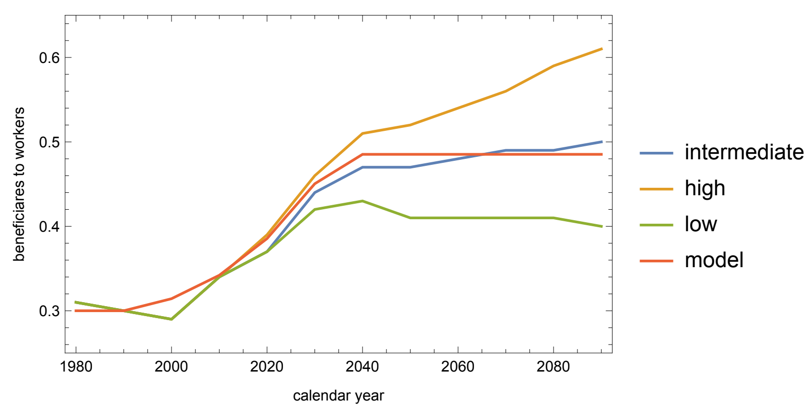

Figure 1: Ratio of beneficiaries to retirees. The red link in this graph illustrates the ratio of beneficiaries to retirees in the calibrated, six period model. The demographic change is modeled as a one-time reduction in the population growth rate n from 0.1802 to 0.01. The yellow, blue, and green lines show the high, intermediate, and low cost projections of the ratio from the 2017 SSA Trustees' Report. The model calibration closely tracks the intermediate assumptions from the SSA.

In the learning models, I set the gain parameter $\gamma = 0.93$. This gain parameter minimizes the maximum welfare cost to an agent of using adaptive forecasts in the LCH model along the transition path that includes a demographic shock and a change to social security. I construct a consumption equivalent variation that compares the utility of consumption of each cohort of agents in the LCH model to the utility of consumption for "infinitesimally rational" cohorts of agents. An infinitesimally rational agent is able to predict future prices with perfect foresight in the LCH world, but is such a small part of the market that she does not change prices. I compute the consumption equivalent variation that made an infinitesimally rational agent indifferent between her own consumption and the consumption of the Life-cycle horizon learners. I chose the gain parameter γ to minimize the maximum welfare cost. It is unsurprising that the gain parameter is close to $\gamma = 1$, since all shocks in the model are permanent. The simulation results are qualitatively sensitive to the gain parameter, and I explore alternative parameterizations as robustness checks.

6 Applications

6.1 Announced Social Security Reform

The aging of society put pressure on the social security system. In the absence of reform, the government would be required to fund social security benefits by issuing debt. There is a limit to how much the government can borrow, even in a dynamically inefficient economy. I illustrate some of these dynamics in the following example.[11] [12]

Suppose that the government runs a social security program that generates a small surplus (government assets); however, as the population ages, the operating surplus changes to a deficit. The government continues to operate the same social security system which depletes the government assets and causes the government to accrue debt. After a number of periods, the government reforms social security by cutting benefits or raising taxes.

I calibrate this example to correspond to the US system. The initial population growth rate is $n = 0.1802$ and falls to $n = 0.01$. This generates a smooth transition in the ratio of retirees to workers from 0.3 to 0.485 after six periods (which is near SSA projections for 2035-2095). The initial policy mimics the current US system: the payroll tax rate is set to $\tau^0 = 0.124$ and the benefit replacement rate is $\phi = 0.4$. Reform is calibrated as either a tax increase to $\tau^0 = 0.1516$, or a benefit cut to $\phi = 0.332$. The SSA estimates these reforms would eliminate the funding shortfall in the social security system. In order to have a dynamically efficient economy with positive government deficits, the Leeper tax rate is set to $\tau^1 = 0.045$ before and after reform.

Figure 2 plots the path of capital and bonds for this tax increase reform. The

[11] The government could also fund social security benefits using tax revenue from some other existing tax. For the sake of this paper, I will argue that diverting other federal tax dollars into the pension system is isomorphic to directly increasing social security taxes, and is thus reform.

[12] See Chalk (2000) for a discussion of maximum government deficits in a dynamically inefficient economy.

plot includes paths for an economy with Rational Expectations (in green), and for an economy with LCH Learning. Government debt increases before the reform takes place and then converges to the new higher steady state value. Debt is increasing since the government is running a deficit in the social security system. Debt would become explosive, were it not for the reform in period $t = 7$. The reform is announced, so agents adjust their behavior before the reform date. Capital increases initially as saving goes up, and then falls to the steady state.

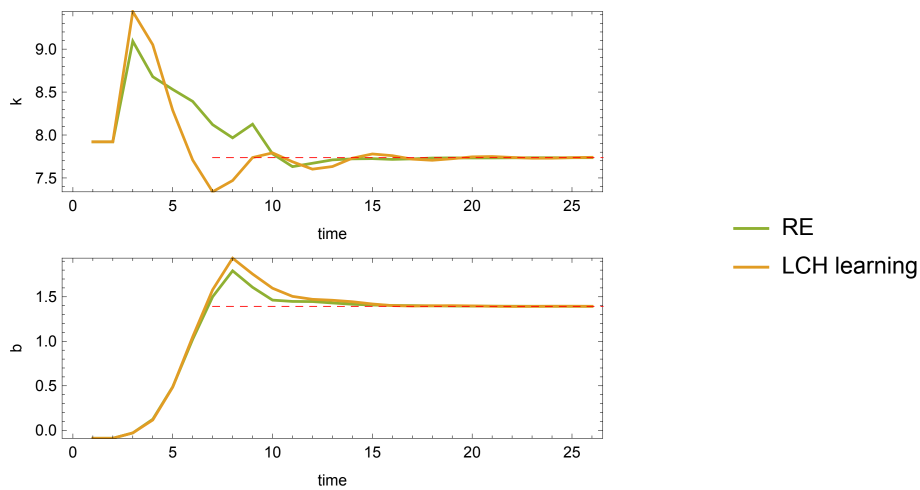

Figure 2: Time paths for capital and bonds for RE (green) and LCH (yellow) for an announced tax increase. Demographic changes drive increasing debt until tax increase in period $t = 7$. The initial population growth rate is $n = 0.1802$, and falls to $n = 0.01$ in period $t = 0$. Initial government policy is $\tau^0 = 0.124$, $\tau^1 = 0.045$, and $\phi = 0.4$. The policy change increases τ^0 to 0.1516. The red dashed lines indicate the steady state following the policy change. The gain parameter $\gamma = 0.93$ for this example. The red dashed lines indicate the steady state following the policy change.

Agents in the learning model overestimate the interest rate and underestimate the wage in the first few periods following the change in the population growth rate (which moves the economy away from the initial steady state). This is depicted in Figure 3. The expectations adjust quickly because the gain parameter γ is so large. In this example, the gain is set to 0.93.

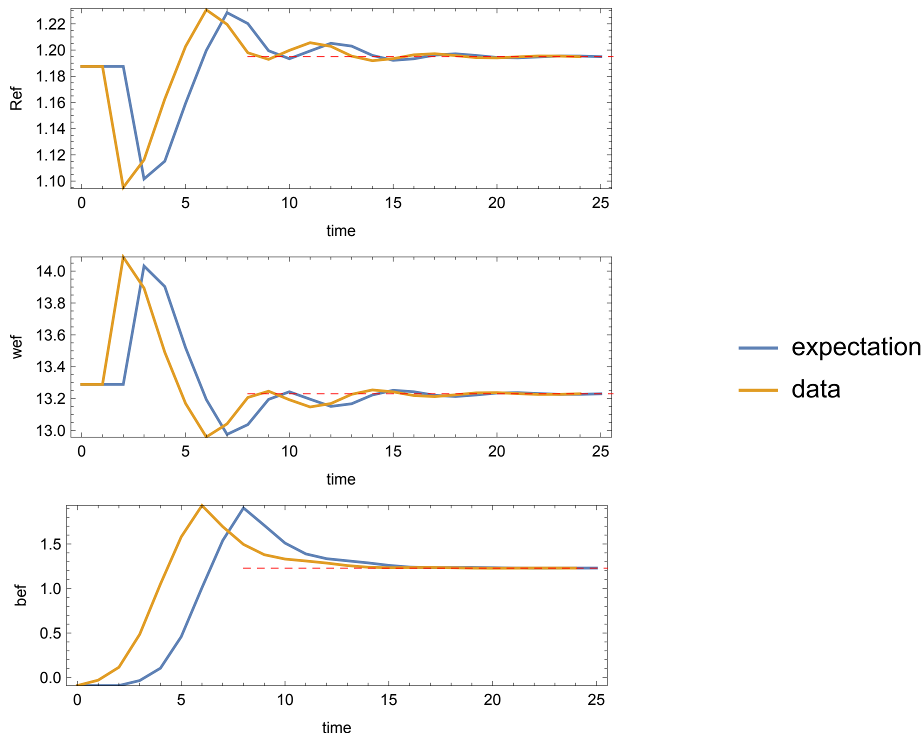

Figure 3: Time paths for expected interest rates, wages, and government debt levels (blue lines), and realized interest rates, wages, and debt (yellow lines) in the LCH model with an announced tax increase. Agents overestimate the interest rate and underestimate the wage in the initial periods as they learn that capital and debt are increasing. Demographic changes drive increasing debt until tax increase in period $t = 7$. The initial population growth rate is $n = 0.1802$, and falls to $n = 0.01$ in period $t = 0$. Initial government policy is $\tau^0 = 0.124$, $\tau^1 = 0.045$, and $\phi = 0.4$. The policy change increases τ^0 to 0.1516. The gain parameter $\gamma = 0.93$. The red dashed lines indicate the steady state following the policy change.

The agents in the learning model save more relative to the fully rational agents in the initial periods, since their estimate of prices and bonds is inaccurate. The savings choices of agents are depicted in Figure 4. The graphs of assets includes time $t-1$ which are the steady-state values that correspond to the old population growth rate. Period $t = 0$ is the first period with the lower growth rate. Agents with rational expectations do not make (large) changes in their savings behavior if they know they will not be alive for the reform. This explains the spikes in savings along the rational expectations paths leading up to reform. The initial decline in savings along the RE paths are due to the changes in the capital and bonds.

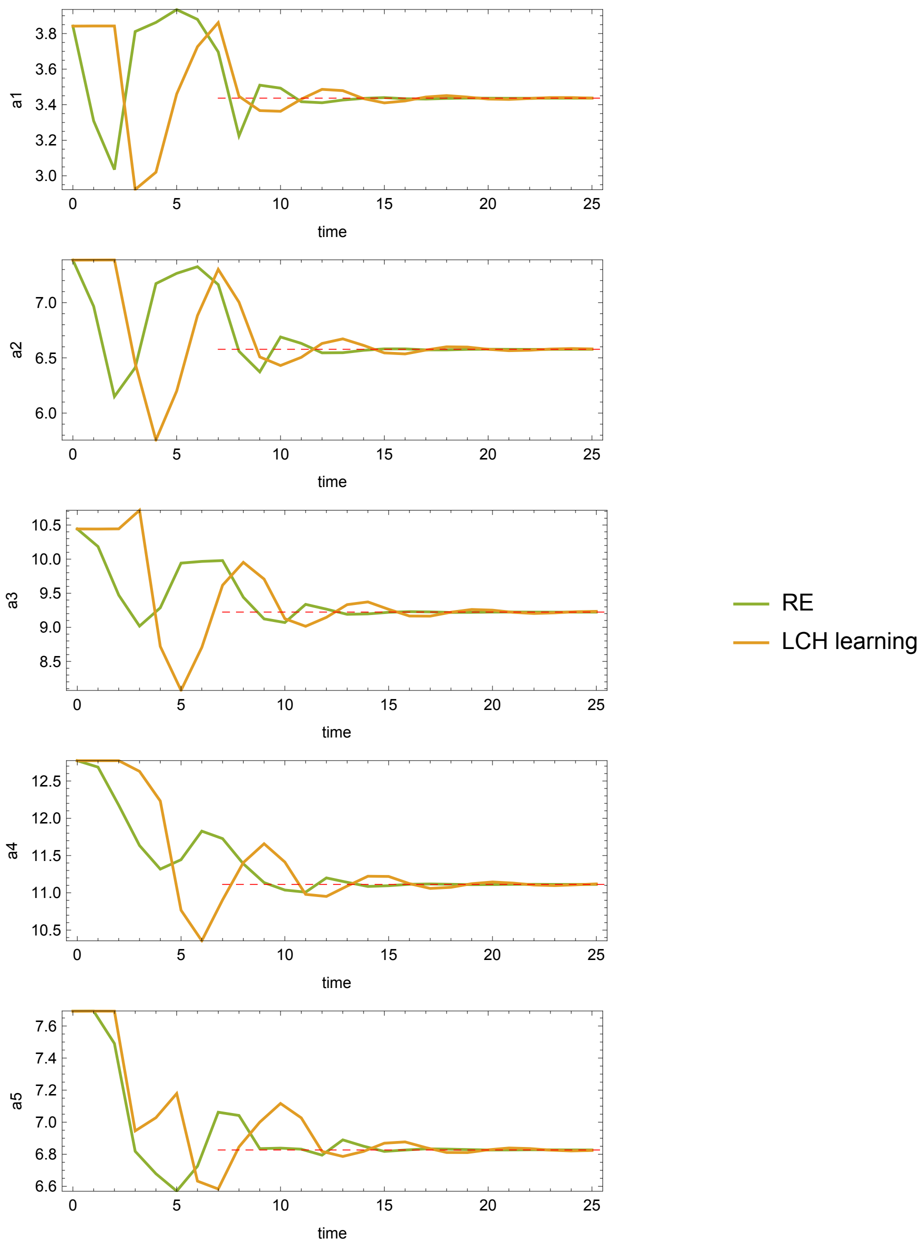

Figure 4: Time paths for the savings accounts for RE (green) and LCH (yellow) models with an announced tax increase. Demographic changes drive increasing debt until tax increase in period $t = 7$. The initial population growth rate is $n = 0.1802$, and falls to $n = 0.01$ in period $t = 0$. Initial government policy is $\tau^0 = 0.124$, $\tau^1 = 0.045$, and $\phi = 0.4$. The policy change increases τ^0 to 0.1516. The red dashed lines indicate the steady state following the policy change. The gain parameter $\gamma = 0.93$. The red dashed lines indicate the steady state following the policy change.

The agents in the model with learning over-save relative to the rational model, since they do not know future prices. Because the agents initially over-save, they are able to save less later in life relative to their plan from the first period. This is evident in Figure 5 by comparing the young (age 1) agent's planned asset holdings for future periods against the actual savings choice she makes when she reaches those periods. For the first several time periods, the agent's planned savings are above the actual savings. The planed savings and actual savings of older agents follows a similar pattern.

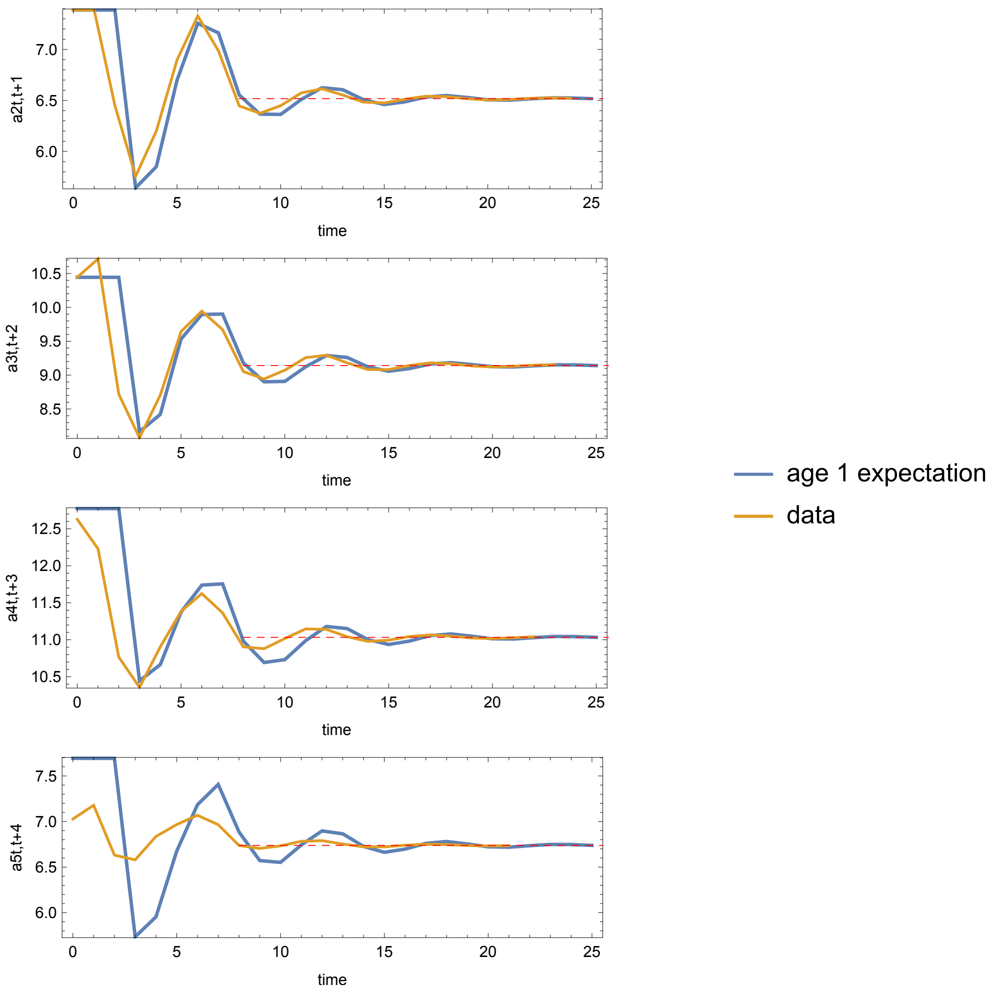

Figure 5: This graph shows planned future savings of the period in blue, and actual future savings in yellow. The expectations (or plans) are made when the agent in young (age 1) in period t. Demographic changes drive increasing debt until tax increase in period $t = 7$. The initial population growth rate is $n = 0.1802$, and falls to $n = 0.01$ in period $t = 0$. Initial government policy is $\tau^0 = 0.124$, $\tau^1 = 0.045$, and $\phi = 0.4$. The policy change increases τ^0 to 0.1516. The gain parameter $\gamma = 0.93$. The red dashed lines indicate the steady state following the policy change.

6.2 Announced Change, Finite Horizon Life-cycle Learning

Agents using FHL learning only make decisions based on their forecasts over their short planning horizon. Agents forecast prices, bonds, and the assets the expect to hold at the end of the planning horizon. Overtime, agents learn the steady state values of each of these quantities. If agents' estimates for the assets they'll need at the end of the planning horizon are close to the rational expectations (steady state) values, they make relatively better decisions using a shorter planning horizon than using a longer horizon. This is because agents using a longer horizon forecast over more periods and can make larger mistakes. This is evident by comparing the transition dynamics of different planning horizon lengths in the FLH model for the a social security tax increase in response to a demographic shock. The fluctuations are larger for longer planning horizons and smaller for shorter planning horizons, as depicted in Figure 6. The Figure depicts the time path for capital and bonds for a tax increase. The time paths for asset holdings are display cyclical behavior, with the largest cycles associated with the longest planning horizon. As in the previous section, this policy experiment replicates the tax increase suggested by the SSA. The initial population growth rate is $n = 0.1802$ and rate falls to $n = 0.01$ in period zero. Initial government policy is $\tau^0 = 0.124$, $\tau^1 = 0.045$, and $\phi = 0.4$. The policy change increases τ^0 to 0.1516.

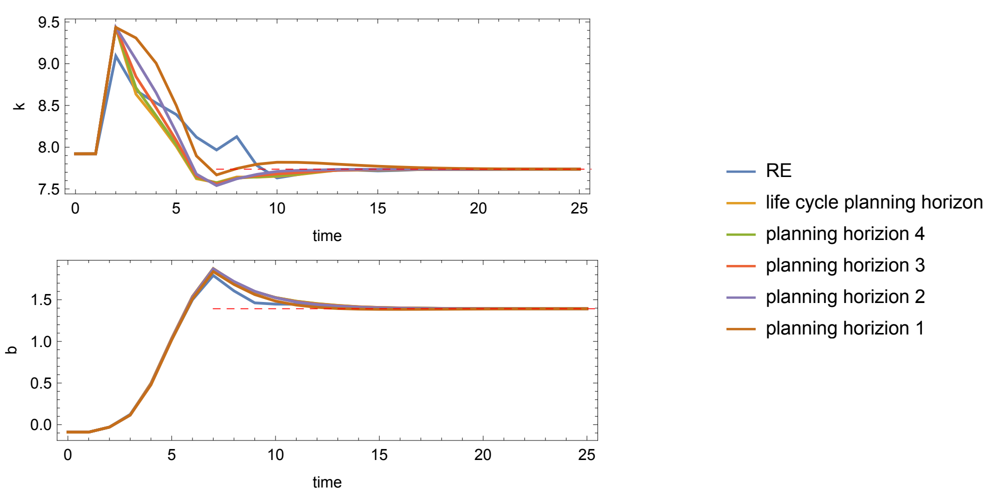

Figure 6: Time paths for capital and bonds for RE (blue), LCH (yellow), and FHL (planning horizon 4: green, 3: red, 2: purple, 1: brown) for an announced tax increase. Demographic changes drive increasing debt until tax increase in period $t = 7$. The initial population growth rate is $n = 0.1802$, and falls to $n = 0.01$ in period $t = 0$. Initial government policy is $\tau^0 = 0.124$, $\tau^1 = 0.045$, and $\phi = 0.4$. The policy change increases τ^0 to 0.1516. $\gamma = 0.93$ for all learning rules. The red dashed lines indicate the steady state following the policy change.

Agents in the FHL Learning model only respond to policy changes when the announced policy enters their planning horizon. Agents making one step ahead forecasts will not respond to announced policy until the period right before the change. Despite this fact, FHL leads to smaller aggregate fluctuations, since agents using a short planning horizon do not make inaccurate forecasts many periods into the future.

7 Welfare

7.1 Welfare Cost of Social Security Reform

The demographic changes and social security reforms proposed in the previous sections harm some generations and help others. I illustrate these effects by using a Consumption Equivalent Variation (CEV) measure. This CEV measure shows the percent of period consumption that would have to be added to an agent in the initial steady state in order to be indifferent between the steady state and being born at a particular time along a particular transition path. A negative CEV indicates that the agent would have higher utility in the initial steady state. The CEV is calculated as:

$$\sum_{j=1}^{J} \beta^{j-1} u(c^{j,ss}(1+\Delta)) = \sum_{j=1}^{J} \beta^{j-1} u(c^{j,}_{t+j-1}) \tag{30}$$

where $c^{j,ss}$ indicates age j consumption in the initial steady state, $c^{j,}_{t+j-1}$ is consumption of an agent with age j in period $t+j-1$ along a given transition path, $u(c)$ is the period utility function as defined by (5), and Δ is the Consumption Equivalent Variation.

The initial steady state is used as a baseline for comparison to capture both the effect of the demographic changes and the social security reform. Note that the initial steady state consumption is *not* possible after the initial period because of demographic changes. It is used only as a reference point to compare the relative harm of the different learning rules. The welfare cost of an announced tax increase is presented in Figure 8 and the welfare cost of an announced social security benefit cut in period $t = 6$ is presented in Figure 7 . In both examples the economy begins in a steady state with a ratio of retirees to workers of 0.3. In period zero, the population growth rate falls to 0.01, which leads the ratio of retirees to workers to increase to 0.485 over six periods to correspond with the aging of the US population. The baseline payroll tax rate is set to $\tau^0 = 0.124$ and the benefit replacement rate is $\phi = 0.4$. The benefit cut reform lowers the replacement rate to $\phi = 0.332$, which the tax increase reform increases the payroll tax to $\tau^0 = 0.1516$. Throughout both examples the Leeper tax rate is set to $\tau^1 = 0.045$ before and after reform, and the gain parameter $\gamma = 0.93$.

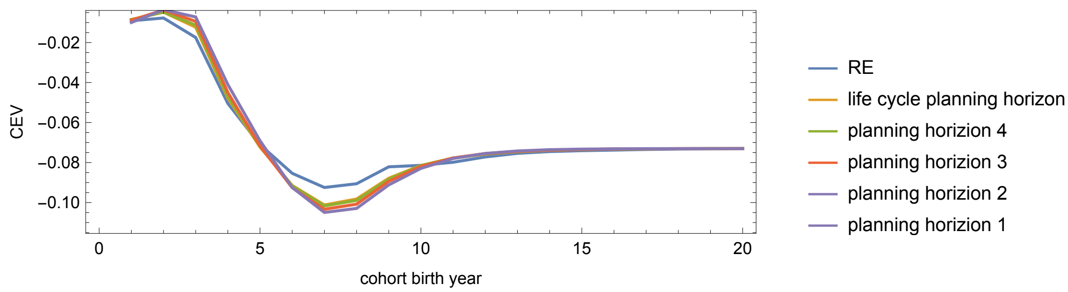

Figure 7: Consumption Equivalent variation measure of the welfare cost of an announced social security benefit cut. The population growth rate is $n = 0.1802$ in the initial steady state and falls to $n = 0.01$ in period 0. This drives the need for social security reform. The original policy parameters are $\tau^0 = 0.124$,$\tau^1 = 0.045$, $\phi = 0.4$. In period $t = 6$, the benefit replacement rate falls to $\phi = 0.332$. The gain parameter $\gamma = 0.35$. The cost of the reform is highest (lowest CEV) for agents in the LCH learning model in period 7; agents in the initial steady state would have to give up 10.1% of their period consumption to be indifferent between living in the initial steady and being born in period 7 in the LCH model.

The underlying demographic changes in these examples harm agents. The model is calibrated to dynamically efficient with a social security system about the size of the US system, as the norm in this literature. Slowing population growth makes the economy even more dynamically efficient, which lowers consumption (and welfare) for all generations.

The demographic changes combined with the benefit cut lower welfare for agents in all cohorts with all learning rules (and also in the RE model). The CEV for the final steady state compared to the initial steady state is -7.29% of period consumption. That is, agents in the initial steady state would have to give up 7.29% of period consumption to be indifferent between the initial steady state and being born in the final steady state. The welfare cost is greatest for generations born right after the social security reform. The CEV reaches a minimum of -9.24% in the RE model for the cohort born in $t = 7$; and a minimum of -10.12% for the LCH model agents born in period $t = 7$. The welfare cost is decreasing in gain parameter γ; if the gain is decreased to 0.35, the minimum CEV is -11.74% and occurs in the LCH model for the cohort born in period $t = 8$. The timing of the welfare cost is also sensitive to the gain parameter, since larger gain parameters producer higher frequency cycles. The faster cycle moves the minimum capital stock a few periods ahead and thus the most harmed cohort is earlier.

Figure 8 depicts the CEV for the tax increase. The demographic changes combined with the tax increase lower welfare for all cohorts in the RE and learning models. The CEV in the final state is -10.20%, indicating agents in the initial steady state would have to give up 10.20% of their consumption to be indifferent between the initial steady state and being born in the final steady state. The welfare cost is largest in the LCH model for the cohort born in period $t = 8$. The CEV is -11.08% for this cohort. The magnitude of CEV for the learning models is decreasing in gamma. If γ is decreased to 0.35, the min CEV is -13.78% for the most harmed cohort.

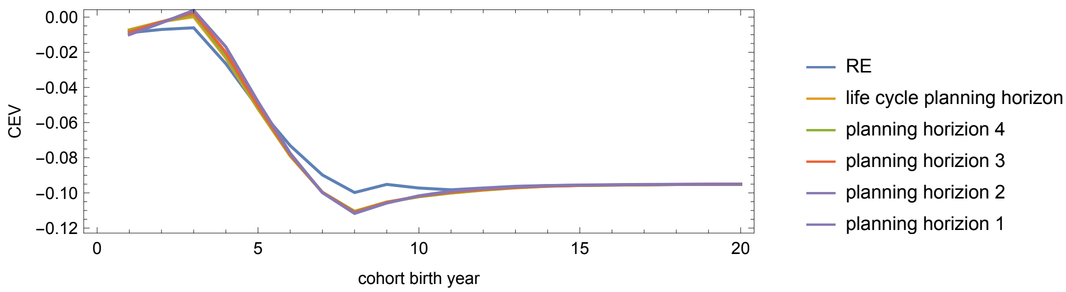

Figure 8: Consumption Equivalent variation measure of the welfare cost of an announced social security tax increase. The population growth rate is $n = 0.18$ in the initial steady state and falls to $n = 0.01$ in period 0. This drives the need for social security reform. The original policy parameters are $\tau^0 = 0.124$,$\tau^1 = 0.04$, $\phi = 0.4$. In period $t = 6$, the payroll tax rate increases to $\tau^0 = 0.1516$. $\gamma = 0.93$ for this example. The cost of the reform is highest (lowest CEV) for agents in the LCH learning model in period 8; agents in the initial steady state would have to give up 11.08% of their period consumption to be indifferent between living in the initial steady and being born in period 8 in the LCH model.

7.2 Welfare Cost of Learning

Agents using finite or life-cycle horizon learning don't fully realize the impact of a policy change or other exogenous shock. It take the agents many periods to learn the new steady state values. Converge to the new steady state is slower in the learning models than the RE model. This section measures the welfare cost of the learning models relative to the rational expectations baseline. This welfare comparison will net out the cost of changing demographics and the change in terminal steady state values due to the new social security policy. By comparing consumption in the learning model to consumption in the RE model, this measures the cost of the learning directly.

As in the previous section, welfare is measured using a Consumption Equivalent Variation measure. Here, the CEV measure shows the percent of period income that would have to be added to consumer with rational expectations (in the RE model) in order to be indifferent between living in the RE world or living in the world with learning in the same cohort. A negative CEV indicates that the agent would have higher utility in the world with rational expectations. The CEV is calculated as:

$$\sum_{j=1}^{J} \beta^{j-1} u(c_{t+j-1}^{j,RE}(1 + \Delta_{RE})) = \sum_{j=1}^{J} \beta^{j-1} u(c_{t+j-1}^{j,L}) \tag{31}$$

where $c_{t+j-1}^{j,RE}$ indicates age j consumption in period $t + j - 1$ in the model with rational expectations, $c_{t+j-1}^{j,L}$ is consumption in the model with learning, $u(c)$ is the period utility function as defined by (5), and Δ_{RE} is the Consumption Equivalent Variation.

Figures 9 and 10 show the CEV for the tax increase and benefit cut examples (parameterized as in the previous sections). In contrast to the previous section, the CEV is not always negative. That is because here the CEV measures learning compared to rational expectations. After many periods, the learning models converge to the RE model, so the CEV converges to zero.

Figure 9 illustrates the welfare cost of the tax increase in the learning models relative to the RE baseline. For the LCH model, the welfare cost is greatest for the tax increase in period $t = 8$; agents in the RE model born in period 8 would be willing to give up 1.18% of their consumption to avoid being born in period 8 in the LCH model.[13]

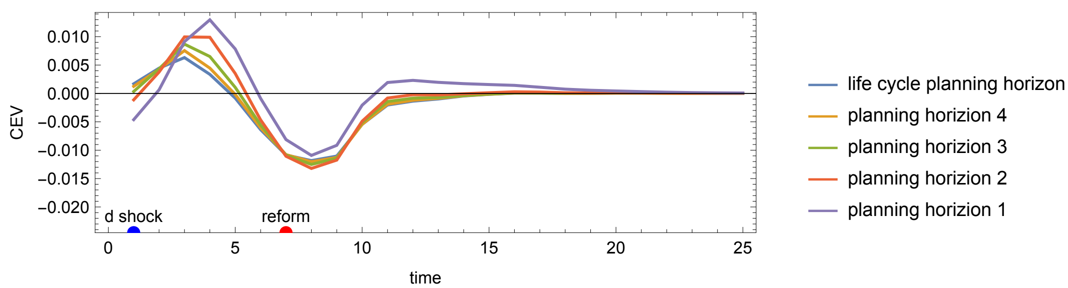

Figure 9: Consumption Equivalent variation measure of the welfare cost of an announced social security tax increase: learning models compared to RE. The population growth rate is $n = 0.18$ in the initial steady state and falls to $n = 0.01$ in period 0. This drives the need for social security reform. The original policy parameters are $\tau^0 = 0.124$,$\tau^1 = 0.045$, $\phi = 0.4$. In period $t = 6$, the payroll tax rate increases to $\tau^0 = 0.1516$. The gain parameter $\gamma = 0.93$. The cost of the reform is highest (lowest CEV) for agents in the LCH learning model in period 8; agents in the RE model would have to give up 1.18% of their period consumption to be indifferent between being born in the RE model in period 8 and being born in period 8 in the LCH model.

Figure 10 illustrates the welfare cost or the benefit cut in the learning models relative to the RE baseline. In the LCH model, the welfare cost is greatest for the cohort born in $t = 10$. Agents in the rational model would be willing to give up 0.96% of their consumption to avoid living LCH world in the same time periods.[14]

[13]When $\gamma = 0.35$ the min CEV is -4.24% and occurs for the cohort born in $t = 9$.
[14]When $\gamma = 0.35$ the min CEV is -3.06% and occurs for the cohort born in $t = 9$.

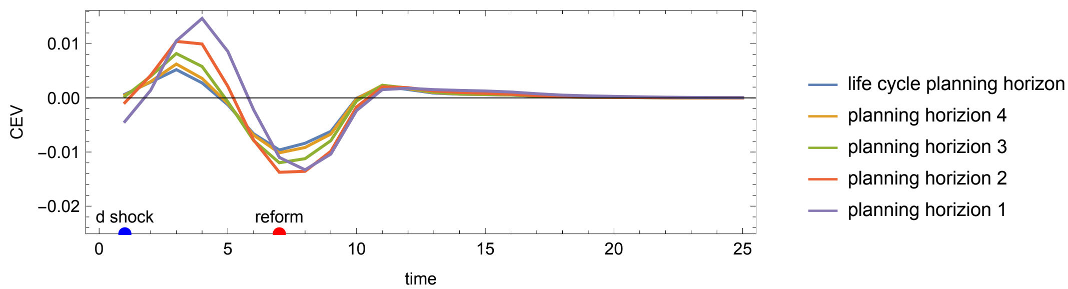

Figure 10: Consumption Equivalent variation measure of the welfare cost of an announced social security benefit cut: learning models compared to RE. The population growth rate is $n = 0.18$ in the initial steady state and falls to $n = 0.01$ in period 0. This drives the need for social security reform. The original policy parameters are $\tau^0 = 0.124, \tau^1 = 0.045$, $\phi = 0.4$. In period $t = 6$, the benefit replacement rate falls to $\phi = 0.332$. The gain parameter $\gamma = 0.93$. The cost of the reform is highest (lowest CEV) for agents in the LCH learning model in period 7; agents in the RE model would have to give up 0.9% of their period consumption to be indifferent between being born in the RE model in period 7 and being born in period 7 in the LCH model.

Many cohorts benefit from the cyclical dynamics introduced via learning. Agents benefit when the capital stock is higher because the economy is dynamically inefficient. Consider 1-step-ahead FHL learning in the context of the social security benefit cut presented in Figure 10. Cohorts born prior to the reform and many cohorts born after the reform have higher utility in the FHL model than in the RE model. Indeed, the cumulative welfare gain across cohorts can be positive. At first this may seem like a violation of the Welfare Theorems, but it is not. Bounded rationality does not present a Pareto improvement over the RE baseline. The learning dynamics improve the welfare for some agents who enjoy higher wages and higher consumption due to the increased capital stock from higher saving.

Markets exhibit a pecuniary externality in general equilibrium. Higher savings today raises capital and wages tomorrow. Agents can exploit this pecuniary externality in OLG economies, if they can coordinate behavior across generations. Feigenbaum et al. (2011) introduce the concept of "optimal irrationality" in OLG economies. They show that an irrational consumption rule *always* exists that can weakly improve upon the lifecycle/permanent-income rule in general equilibrium in a two-period model. They present a calibrated continuous time example with an escalating rule-of-thumb savings rule that increases welfare relative to rational expectations (for all but the first generation). They argue that policies that increase savings in a dynamically efficient economy can improve welfare for many (but not all) generations. The welfare gains to some generations in my model come from the same externality.

The majority of the welfare cost (or welfare gain) of living in a FHL or LCH learning model come from the general equilibrium changes in the capital stock, not from the individual forecast errors made by an agent. One way to measure this cost is to compute the life-cycle consumption for an infinitesimally rational agent who lives in a learning model.

Suppose a single rational agent is born in the LCH model. The rational agent understands that everyone else is using LCH learning, and she is also able to predict future prices with perfect foresight. She is such a small part of the market that her individual choices do not changes prices. I compute the CEV that equates the utility of the infinitesimally rational agent with the utility of the LCH agent. The minimum CEV for this comparison is -0.21% for the benefit cut reform (-0.30% for the tax increase reform). The CEV is greater than 0.1% for the first four cohorts and then quickly falls to zero. The negative sign indicates that LCH agents are worse off than the rational agent living in a LCH world. This CEV is presented below in Figure 11.

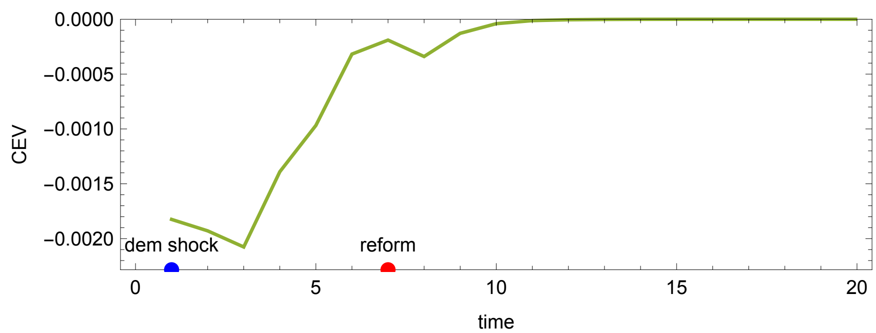

Figure 11: Consumption Equivalent Variation (CEV) measure of the welfare cost of using adaptive learning along the transition path of an announced social security benefit cut. The CEV compares an infinitesimally rational agent to a life-cycle horizon learner. The infinitesimally rational agent is effected by the general equilibrium effects of learning, but it able to predict future prices. The population growth rate is $n = 0.1802$ in the initial steady state and falls to $n = 0.01$ in period 0. This drives the need for social security reform. The original policy parameters are $\tau^0 = 0.124$,$\tau^1 = 0.045$, $\phi = 0.4$. In period $t = 6$, the benefit replacement rate falls to $\phi = 0.332$. The gain parameter $\gamma = 0.93$.

The welfare cost of learning can also be explored in the context of a recession. Agents using LCH or FHL learning will fare worse during and after a recession than fully rational agents. This is because the learning agents will not anticipate the general equilibrium effects of the initial decline in production. An example recession is presented in Appendix B.

8 Welfare Costs of Policy Uncertainty

The policy changes of the previous sections are non-stochastic and announced. However, the future of the U.S. social security system seems difficult to anticipate.

A growing body of research focuses on the welfare cost of this type of uncertainty.[15] I calculate the welfare cost of policy uncertainty under RE and under LCH learning. To do this, I model government policy as an exogenous, stochastic process.

Let $\omega_t = (\tau_t^0, \tau_t^1, \phi_t)'$ describe the government policy parameters at time t. Suppose that the social security program will be reformed at some future date. The reform date falls within a known, finite set. Suppose that the realization of ω_{t+1} depends on ω_t, and is contained in the finite set Ω. Suppose also that each possible reform converges to a steady state (the path is non-explosive).

Let the probability of realizing a particular value of ω_{t+1} given ω_t be described by $\pi(\omega_{t+1}|\omega_t)$. Using this notation, the expected value in the household first order equations (4) can be written as:

$$u'(c^j_{t+j-1}) = \beta \sum_{\omega_{t+1} \in \, \Omega} \pi(\omega_{t+1}|\omega_t) R_{t+j} u'(c^{j+1}_{t+j}) \quad \text{for } j = 1, \cdots, J-1 \tag{32}$$

Introducing aggregate uncertainty to the model raises some concerns; a steady state wealth distribution will not exist in general in an OLG model with aggregate uncertainty (see Krueger and Kubler (2004) for a discussion). I overcome this problem by modeling aggregate policy uncertainty in a simple, stylized way. Policy uncertainty only exists for a small number of periods and policy reform can only be one of a small number of policy options. These simplifications, combined with the calibration of the life-cycle as six periods, allow me to calculate the equilibrium paths for the economy with policy uncertainty.

Suppose that policy uncertainty take the following form: reform is possible in either date S, or in date $S + 1$. Within each period two reforms are possible, either a benefit cut or a tax increase. Thus, there are four possible paths for the economy. The probability of each path is $p = 0.25$. All agents in the economy know the four possible reforms and their relatively probability. I calibrate this example to correspond to the US system. The economy starts in a steady with a ratio of retirees to workers of 0.3 and with government policy: $\tau^0 = 0.124$, $\tau^1 = 0.045$, and $\phi = 0.4$. In period $t = 0$, the growth rate of the population falls to 0.01 leading the ratio of retirees to workers to grow, eventually reaching 0.485. The demographic change causes the social security system to run a deficit, which increases government debt.

Reform is calibrated as either a tax increase to $\tau^0 = 0.1516$, or a benefit cut to $\phi = 0.32$.The reform takes place in either date $t = 6$ or $t = 7$. The SSA estimates these reforms would eliminate the funding shortfall in the social security system. In order to have a dynamically efficient economy with positive government debt, the

[15]I follow Caliendo et al. (2016) and use the words “uncertainty” and “risk” interchangeably in this paper. All of the examples I will consider have known probabilities, and thus might be called “risky.”

Leeper tax rate is set to $\tau^1 = 0.045$ before and after reform. The gain parameter for the learning model is set to $\gamma = 0.93$.

Before considering the welfare cost of reform, First consider the transition dynamics of the four possible equilibrium paths in the RE model. The four possible paths for the economy are identical up to period five. Agents in (or before) period five face the same uncertainty about future social security policy. This uncertainty is either fully or partially resolved in period six. Along the first two paths of the economy, agents observe that policy was reformed in period six (either as a tax increase or benefit cut). Uncertainty is resolved for these two paths. Along the other two paths, agents observe that policy was not reformed in period six, so they know reform will take place in period seven, but they do not which reform until the policy is realized. In period seven, all uncertainty is resolved, and all four paths are (potentially) different. There are two possible steady states for this example, the tax increase steady state and the benefit cut steady state. Both tax increase paths converge to the tax increase steady state; similarly, both benefit cut paths converge to the benefit cut steady state. The asset paths for this example are presented in Figure 12.

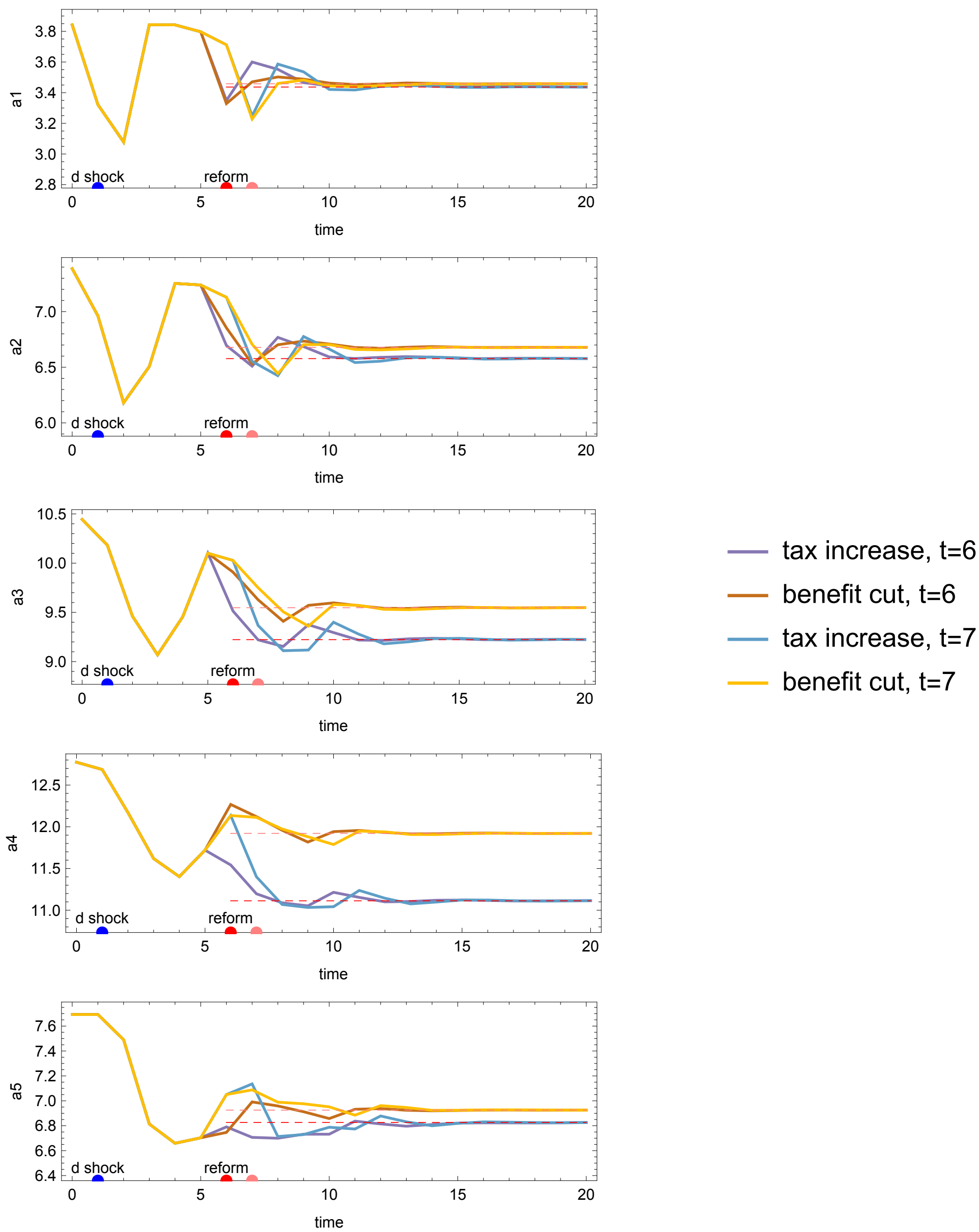

Figure 12: Transition paths for asset holdings in the RE model with policy uncertainty. Four reforms are possible. Either taxes are increased in period 6 (purple line), benefits are cut in period 6 (red/brown line), taxes are increased in period 7 (light blue line), or benefits are cut in period 7 (gold line). All four paths are identical until period 5. This is because all agents face the same policy uncertainty for the first five periods. The two possible steady states are indicated with red dashed lines and two possible reform dates are indicated with red and pink dots. The demographic shock is indicated with a blue dot.

This same policy uncertainty experiment is possible in the learning models; I will present results for the LCH model. Agents in the learning model do not know future prices, they make decisions based on their adaptive expectations. In the initial few periods, following the change in the population growth rate, the LCH agents over-save compared to the rational agents because they are overestimating the interest rate. This drives up capital in the LCH model compared to the RE model. The LCH agents benefit from this higher capital stock (since the economy is dynamically efficient) and don't have to save as much in the periods right before the policy reform. Following a policy reform, the paths of savings, capital and bonds are cyclical in the LCH model since agents are slowly updating their forecasts of prices and bonds.

Note, that all of the agents in the RE and LCH model know the policy process. All agents know the possible dates of reform, the new policy parameters, and the probability of each reform. All agents are forward looking (to the end of their life-cycle). The only difference is that the LCH agents forecast future prices and bonds adaptively, while the rational agents make fully rational forecasts. The paths of assets and capital for the LCH and RE models are presented together in Figures 13 and 14.

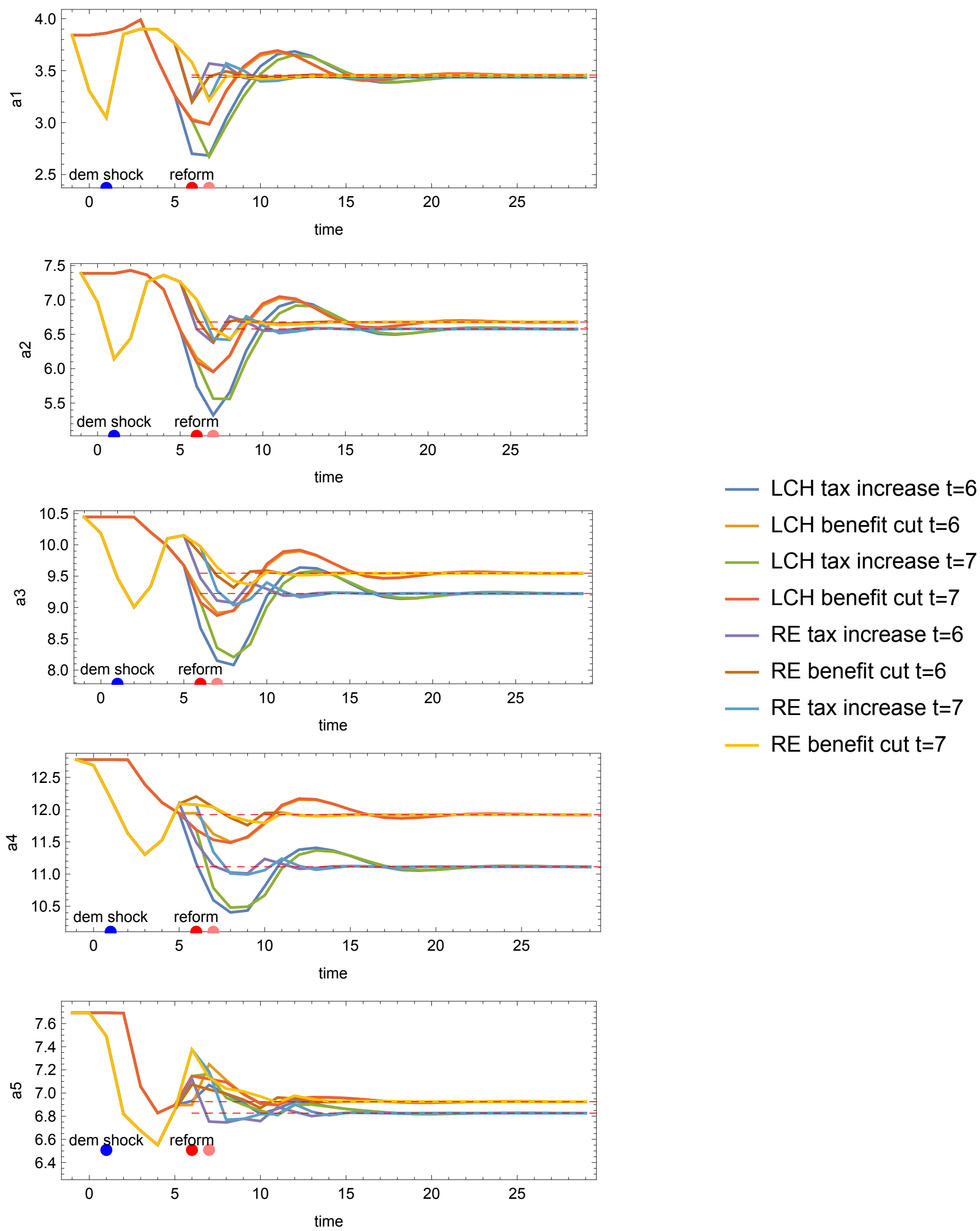

Figure 13: Transition paths for asset holdings in the LCH model and RE model with policy uncertainty. Four reforms are possible in each model. Either taxes are increased in period 6 (LCH: blue, RE: purple line), benefits are cut in period 6 (LCH: yellow, RE: red/brown line), taxes are increased in period 7 (LCH:green, RE: light blue line), or benefits are cut in period 7 (LCH: red, RE: gold line). The two possible steady states are indicated with red dashed lines. The benefit cut paths converge to a steady state with lower age 1 and 2 savings (a^1 and a^2) and higher age 3, 4 and 5 savings (a^3, a^4, a^5). The LCH paths are cyclical, while the RE paths converge to the new steady states quickly. The dynamics before the reform are driven by demographic change from period zero, and the forward looking behavior of agents anticipating reform.

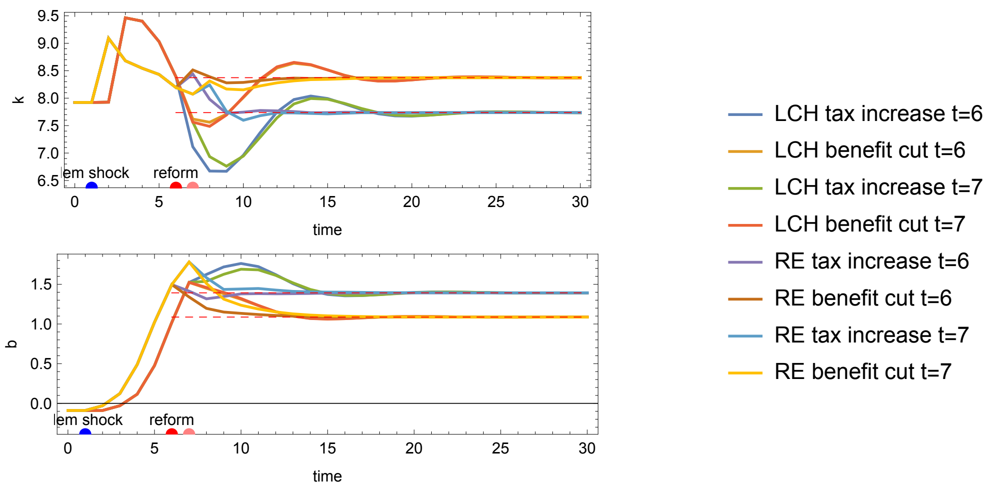

Figure 14: Transition paths for capital and bonds in the LCH model and RE model with policy uncertainty. Four reforms are possible in each model. Either taxes are increased in period 6 (LCH: blue, RE: purple line), benefits are cut in period 6 (LCH: yellow, RE: red/brown line), taxes are increased in period 7 (LCH:green, RE: light blue line), or benefits are cut in period 7 (LCH: red, RE: gold line). The LCH paths are cyclical, while the RE paths converge to the new steady states quickly. The dynamics before the reform are driven by demographic change from period zero, and the forward looking behavior of agents anticipating reform. The two possible steady states are indicated with red dashed lines. The benefit cut paths converge to a steady state with higher capital.

8.1 Welfare Cost

The welfare cost of policy uncertainty is calculated in this section using two Consumption Equivalent Variation techniques: one for the rational expectations model, and one for the learning model. In each, the utility of an agent who experiences a given realization of the uncertain policy process (that is, they experience a vector of policy parameters over their lifecycle $\omega_t, \dots, \omega_{t+J}$) is compared to the utility of an agent who experience the same policy realization with certainty (that is, the policy parameters $\omega_t, \dots, \omega_{t+J}$ are anticipated). Kitao (2017) constructs a similar welfare measure to examine the welfare cost of policy uncertainty (particularly when considering uncertainty over the timing and type of policy reform). Bütler (1999) uses a similar metric (although her model contains agent-level policy misperceptions rather than aggregate policy uncertainty). The CEV for policy uncertainty in the RE model is calculated as:

$$\sum_{j=1}^{J} \beta^{j-1} u(c_{t+j-1}^{j,A,RE}(1+\Delta_u)) = \sum_{j=1}^{J} \beta^{j-1} u(c_{t+j-1}^{j,A,u}) \tag{33}$$

where $c_{t+j-1}^{j,A,RE}$ indicates age j consumption in period $t+j-1$ in the model with policy path A and rational expectations and $c_{t+j-1}^{j,A,u}$ is consumption in the model with RE and policy uncertainty where policy path A was realized. As before, $u(c)$ indicates the period utility function, and Δ_u is the CEV for policy uncertainty.

The welfare cost of social security policy uncertainty is illustrated below in Figure 15. The figure depict the CEV for each of the four possible paths in the RE model. The calibration of this example is identical to example 14. Recall, the economy begins in a steady state with a high population growth rate. In period zero, the growth rate falls and the social security system begins to run a deficit. Bonds increase until reform is enacted and the economy converges to a new steady state associated with the particular reform. The four possible reforms are a benefit cut in period 6 or 7, or a tax increase in period six or seven.

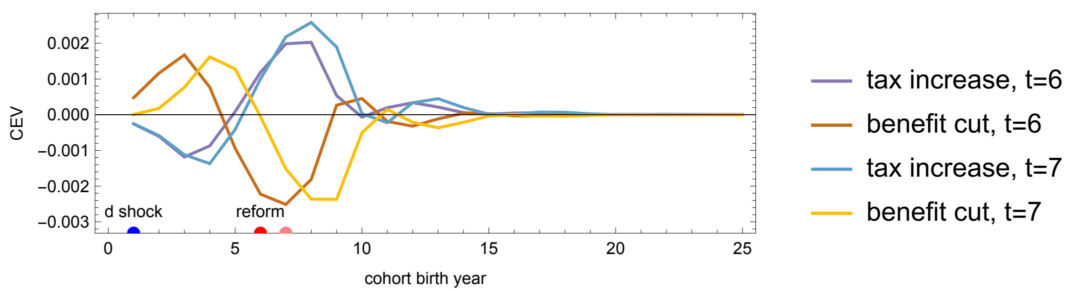

Figure 15: Consumption Equivalent variation measure of the welfare cost of policy uncertainty in the RE model. The welfare cost is constructed by comparing the utility of realized policy parameters of a given path with the utility of that same policy change in a perfect foresight, RE model. The CEV is show pictured for all eight possible paths: taxes are increased in period 6 (LCH: blue, RE: purple line), benefits are cut in period 6 (LCH: yellow, RE: red/brown line), taxes are increased in period 7 (LCH:green, RE: light blue line), or benefits are cut in period 7 (LCH: red, RE: gold line). A negative CEV indicates that an agent would have higher utility in a RE model without policy uncertainty. A positive CEV indicates that utility is higher in the uncertain model.

The welfare cost of social security policy uncertainty for the RE model is small. The cost is the less than 0.3% of period consumption for cohorts along all four paths. The initial cohorts benefit if the reform is a benefit cut, since they do not have to pay higher taxes and do not experience the benefit reduction in their lifetime. The agents born a few periods before and after the benefit cut are harmed because they receive lower benefits, but are not alive to experience the feedback effect of higher wages that result in the new steady state.

The relatively small welfare cost of policy uncertainty in the rational expectations framework that I find is consistent with work by other authors. Kitao (2017) finds the welfare cost of pension reform uncertainty in a model calibrated to the Japanese economy to be 0.8%-1.5% of period consumption.[16] Caliendo et al. (2015) find the

[16]Kitao constructs a CEV measure similar to the CEV used in this paper which compares utility of agents who experience a particular reform under uncertain with utility of agents who experience the same reform with certainty. She also compares reforms realized in later periods to the baseline

welfare cost of social security policy uncertainty to be 0.01% of period consumption for an agent with average earnings.[17] Part of the reason Kitao and I find larger welfare costs than Caliendo et al. is because Kitao and I both use a general equilibrium framework. Caliendo et al. use a partial equilibrium model, which misses the feedback effect of savings on future wages and interest rates. Bütler (1999) finds that the welfare cost of misperceived social security benefits is between 0-3.46% of period consumption in a partial equilibrium model. Bütler's results are not easily comparable to other papers, because agent beliefs are model consistent in her paper.[18]

To assess the cost of policy uncertainty in the learning model the consumption equivalent variation compares the lifetime utility of an agent in a LCH model who experiences a given realization of the uncertain policy process (that is, they experience a vector of policy parameters over their lifecycle $\omega_t, \ldots, \omega_{t+J}$) to the utility of an agent in a LCH model who experience the same *announced* policy realization (that is, the policy parameters $\omega_t, \ldots, \omega_{t+J}$ are anticipated).

The learning CEV is calculated as:

$$\sum_{j=1}^{J} \beta^{j-1} u(c_{t+j-1}^{j,A,L}(1+\Delta_{uL})) = \sum_{j=1}^{J} \beta^{j-1} u(c_{t+j-1}^{j,A,u}) \tag{34}$$

where $c_{t+j-1}^{j,A}$ indicates age j consumption in period $t+j-1$ in the model with policy path A and LCH learning, and $c_{t+j-1}^{j,A,u}$ is consumption in the model with learning and policy uncertainty where policy path A was realized. As before, $u(c)$ indicates the period utility function, and Δ_{uL} is the CEV for policy uncertainty comparing learning to learning.

This learning CEV measures the change in consumption that is the result of policy uncertainty in the LCH framework. The learning CEV is presented in below in Figure 16.

of earlier reform (and finds CEV in the range of 2-3.5% of period consumption). Finally, she also considers a CEV that compares the utility of agents who experience a particular reform with the utility of agents who experience the *expected* reform. Kitao finds similar magnitudes using both CEV measures.

[17] The welfare metric developed in Caliendo et al. (2015) is not exactly the same as the CEV developed in this paper. Caliendo et al. consider a continuous distribution of policy reform dates and options. The CEV metric they construct compares the utility of agent endowed with expected wealth over all possible reforms who experiences the average (or expected) reform to the utility of an agent who faces uncertainty. The also perform heterogeneity analysis and find the welfare cost of social security policy uncertainty to be 2.39 times higher for the lowest earners when the policy is benefit cut, and 1.7 times higher for tax increase. The welfare cost for highest earners is 0.6 times the welfare cost for an average earning when the policy is a benefit cut, and 2.25 times higher for tax increase. These costs are still quite small, since $2.39 \times 0.01 = 0.0239$.

[18] See Phelan (1999) for a discussion of Bütler (1999).

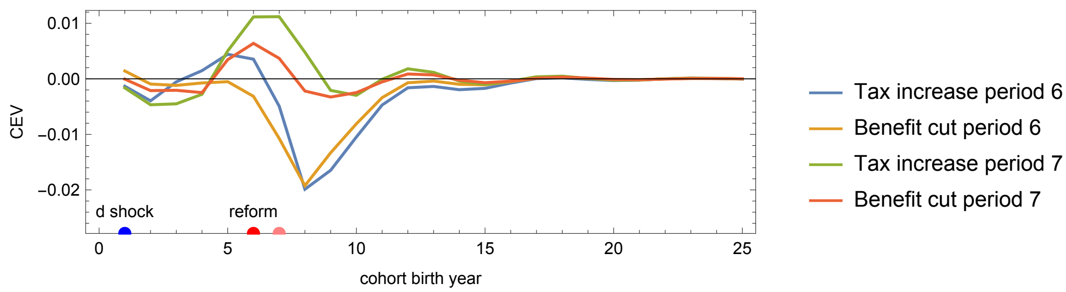

Figure 16: Consumption Equivalent variation measure of the welfare cost of policy uncertainty in the LCH model. The welfare cost is constructed by comparing the utility of realized policy parameters of a given path with the utility of that same policy change in a LCH model. The CEV is pictured for all four possible paths: tax increased in period 6 (blue), benefit cut in period 6 (yellow), tax increase in period 7 (green), or benefit cut in period 7 (red). A negative CEV indicates that an agent would have higher utility in a LCH model with announced policy (no uncertainty). A positive CEV indicates that utility is higher in the uncertain model. The gain parameter $\gamma = 0.93$ in this example.

The learning uncertainty CEV is positive for several cohorts if taxes are increased in period 7. This is because the possibility of facing reform in period 6 increases agent savings and drives up the capital stock in the model with policy uncertainty, relative to the model with announced tax increases. The accumulation of extra capital raises agent welfare, relative to the baseline. The maximum (positive) CEV is 1.11% and occurs in period 9. The cohort of agents born in period 9 in the LCH model with a tax increase in period 7 would need to have 1.11% added to their period consumption in order to be indifferent between their consumption, and the consumption of agents born in period 7 in a LCH model that had a 25% chance of a tax increase in period 6, a 25% chance of a tax increase in period 7, a 25% chance of a benefit cut in period 6, and a 25% chance of a benefit cut in period 7 (where the tax increase in period 7 was realized).

The learning uncertainty CEV is negative for a few early cohorts who are harmed by increased precautionary saving, and for a few cohorts following the reform who experience the bottom of the swing in the capital stock. The swings in state variables are larger in the uncertain model, and thus the welfare cost is larger. The min CEV is -1.98% occurs in period $t = 8$ for agents who are born after the tax increase in period $t = 6$. The minimum CEV along the other three paths is between -0.32% and -1.92%. As in the previous examples, the CEV is depends partially on the gain parameter. The min CEV is -1.94% when $\gamma = 0.35$.

My analysis suggests that rational expectations models may understate the welfare cost of social security policy reform *and* the welfare cost of policy uncertainty. The life-cycle horizon learning model I propose is a small deviation from rational expectations in which agents maintain model consistent beliefs Agents in the LCH model are still optimizing and they are still forward looking, yet the welfare cost of policy uncertainty is much larger. The welfare cost is driven mainly by the cyclical changes in capital stock that are introduced into the model by adaptive learning. The most harmed agents in the LCH model would be willing to give up nearly 2% of period consumption to avoid living in a world of policy uncertainty. In the RE model, the most harmed agents would only be willing to give up 0.3% of period consumption to avoid the policy uncertainty. The welfare cost in the learning model is nearly an order of magnitude larger than in the rational case.

9 Robustness

9.1 Forecast tax burden

In main specification, agents fully understand the Leeper-tax and forecast government debt levels in order to estimate their upcoming taxes. I back away from that assumption in this section and assume that agents do not fully understand the tax system. Agents observe the total tax burden (rate) they face in the current period, and they use adaptive learning to forecast their future tax burden. They do not incorporate knowledge about the structure of the tax system. Agents continue to forecast social security benefits as the replacement rate ϕ times the expected wage at the time of retirement.

Agents in this framework understand that a fraction of their wage is taken by the government every period, but they do not distinguish between the Leeper tax and the payroll tax. Agents are not forward looking with regard to tax changes. They are, however, forward looking with regard to social security benefit changes. This assumption will be relaxed in the next section.

Under this specification, agents forecast

$$w^e_{t+1} = \gamma w_t + (1-\gamma) w^e_t \tag{35}$$

$$R^e_{t+1} = \gamma R_t + (1-\gamma) R^e_t \tag{36}$$

$$\tau^e_{t+1} = \gamma(\tau^0_t + \tau^1_t b_t) + (1-\gamma)\tau^e_t \tag{37}$$

with $\gamma \in (0, 1)$. Forecasts many periods ahead are equal to the one-step-ahead forecasts: $x^e_{t+j} = x^e_{t,t+j} = x^e_{t+1}$, for $x = R, w, \tau$ and $j > 1$.

Agents do not respond to an announced tax change until after the change has been is implemented. Savings of working-age cohorts increase as the policy is implemented, instead of before. Thus the welfare cost of learning (compared to the RE baseline) is larger in this specification than in the baseline learning model. For example, consider the same set-up as the announced policy change in section 6.1. The initial ratio of retirees to workers is 0.3. The ratio increases over six periods to 0.485 which causes the social security system to run deficits; capital falls and bonds increase until reform is enacted. Initial policy is $\tau^0 = 0.124$, $\tau^1 = 0.045$ and $\phi = 0.4$. Reform is calibrated as a payroll tax increase to $\tau^0 = 0.1516$. This policy change is depicted in Figure 17 and the CEV is presented in figure 18.

Agents in the learning models do not respond to the tax increase until the period of the policy change. The swings in capital stock that follow the policy change are larger than in the baseline learning model when agents anticipate tax changes. The minimum CEV in the tax-forecasting learning model is in period $t = 8$. An agent born in period $t = 8$ in the RE model would be willing to give up 1.42% of period consumption to avoid being born in the same period in the LCH tax-forecasting model. The minimum CEV in the bond-forecasting model is -1.18%. As with pervious examples, the CEV depends on the learning gain and is smaller for larger gains. The gain for these examples was set to $\gamma = 0.93$.

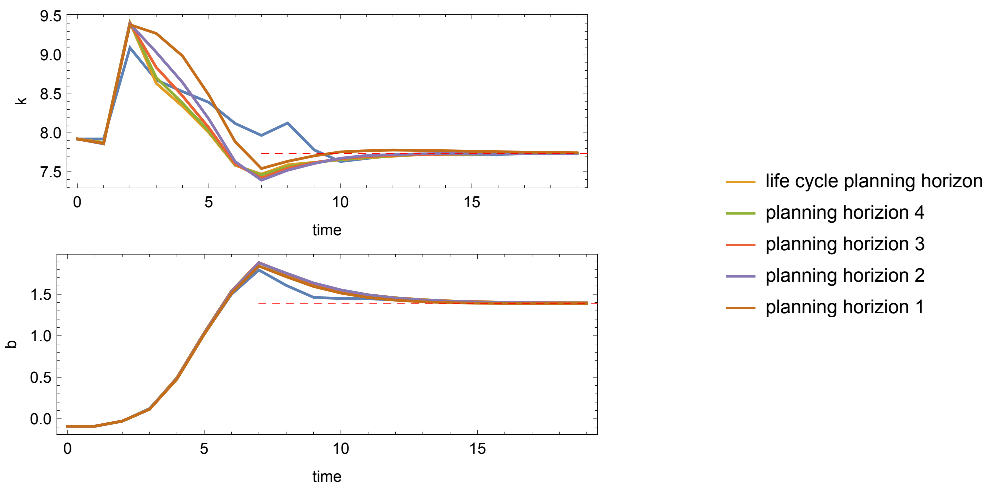

Figure 17: Equilibrium paths for the rational expectations economy and learning economies that forecast the tax burden according to equation 37. The economy begins with a ratio of retirees to workers of 0.3; this falls to 0.485 over six periods. Social security deficits lead to an accumulation of government debt until taxes are raised in period $t = 6$. The RE path is shown in blue, LCH learning in yellow, finite horizon learning in green (for 4 period planning horizon), red (3), purple (2), and orange/brown (1-step-ahead).

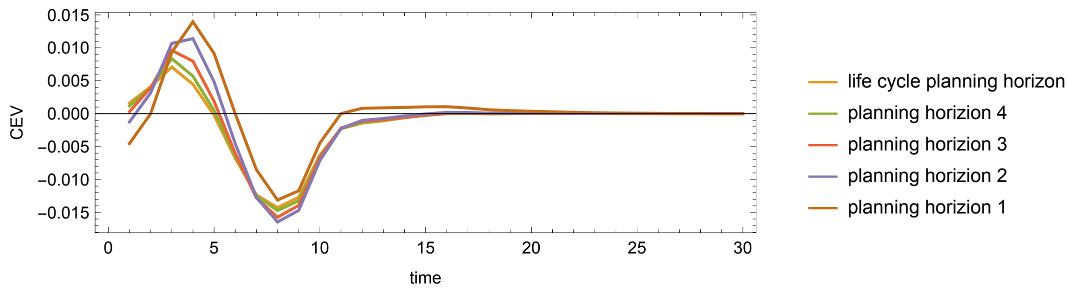

Figure 18: Consumption equivalent variation paths learning economies that forecast the tax burden according to equation 37. The CEV is calculated according to 31. The economy begins with a ratio of retirees to workers of 0.3; this falls to 0.485 over six periods. Social security deficits lead to an accumulation of government debt until taxes are raised in period $t = 6$. The CEV for LCH learning is sown in yellow, finite horizon learning in green (for 4 period planning horizon), red (3), purple (2), and orange/brown (1-step-ahead). The welfare cost is greatest (smallest CEV) for LCH learning.

9.2 Forecast policy adaptively

As a final robustness check, suppose that agents do not anticipate any policy changes, but rather agents learn all policy adaptively. Agents forecast their tax burden $(1 - \tau_t)w_t$ and social security benefit, z_t adaptively. Agents expect to receive the same social security benefit in both periods of retirement (which is consistent with the actual policy process). They forecast social security benefits for the period in which they retire.

Agents forecast

$$w^e_{t+1} = \gamma w_t + (1 - \gamma) w^e_t \tag{38}$$
$$R^e_{t+1} = \gamma R_t + (1 - \gamma) R^e_t \tag{39}$$
$$\tau^e_{t+1} = \gamma(\tau^0_t + \tau^1_t b_t) + (1 - \gamma)\tau^e_t \tag{40}$$
$$z^e_{t+1} = \gamma \phi_t w_t + (1 - \gamma) z^e_t \tag{41}$$

with $\gamma \in (0, 1)$. Forecasts many periods ahead are equal to the one-step-ahead forecasts: $x^e_{t+j} = x^e_{t,t+j} = x^e_{t+1}$, for $x = R, w, \tau, z$ and $j > 1$.

Under this specification, agents in learning models do not respond to announced policy changes. They only respond to tax or benefit changes when they have been implemented. The welfare cost of tax changes is identical to the previous section. However, the welfare cost of social security benefit changes is larger. Using the set-up from the previous section, suppose the population growth rate is initial high, and then falls in period $t = 0$, leading to an accumulation of government debt until benefits are cut in period $t = 6$. The CEV is lowest in period $t = 7$ and equals -1.84%. This means an agent born in the rational model in period $t = 7$ would be willing to give up 1.84% of her period consumption in order to avoid being born in the same period in a model where agents adaptively learn policy with a life-cycle planning horizon. In contrast, the minimum CEV in the baseline LCH model (where agents anticipate policy change and adaptively forecast prices and bonds) is -0.9%. The gain parameter for these examples was $\gamma = 0.93$. The welfare costs are larger for a smaller gain.

9.3 Sensitivity analysis

The qualitative results of this paper are sensitive to parameter choice. As a robustness check, I calibrated the model to be dynamically *inefficient* and ran similar experiments (raising taxes or lowering social security benefits). I parameterize a dynamically inefficient model by setting the discount factor β greater than one. This increases agents desire to consume in old age and drives up savings, and thus the capital stock. I use the parameterization of Bullard and Russell (1999): $\sigma = 4.2$, annual $\beta = (\frac{1}{1-0.041})$, $\alpha = 0.25$, and annual $\delta = 0.1$.[19] Under this specification, agents in learning models still suffer relative to RE agents when facing demographic changes, policy changes, and policy uncertainty. However, the welfare cost of learning is much smaller. For example, the welfare cost that compares LCH to RE of an announced benefit cut in response to the demographic shock is less than 1% of period consumption.

10 Conclusion

Demographic changes in the United States make future social security reform likely, as more beneficiaries are supported by each worker paying taxes. If the program is left unchanged, benefits will exceed tax receipts and the Social Security Trust Fund will be depleted by around the year 2034. The welfare cost to agents of social security reform is not limited to agents alive during the policy change. Using two new models of bounded rationality, I show that the welfare cost of announced policy changes can be quite large if there are general equilibrium feedback effects along the transition path.

I relax the rational expectations assumption and model agents who forecast future interest rates and wages adaptively. I model this in two main ways. First, I develop a model of Life-cycle Horizon Learning, in which agents make optimal decisions over their lifecycle, given their (potentially imperfect) forecasts of prices. Second, I model Finite Horizon Life-cycle Learning, in which agents plan over a shorter planning

[19] My model does not compare directly to Bullard and Russell (1999) in two key ways; first I do not have stochastic survival probability, nor do I have endogenous labor choice. Thus, this calibration should be viewed as a robustness exercise only, and not an attempt to replicate Bullard and Russell (1999).

horizon and make optimal choices conditional on their forecasts and the assets they plan to hold at the end of their planning horizon.

The maximum welfare cost occurs along the transition path for the tax increase in the LCH learning model; agents born after the reform while the capital stock is depressed are worse off by up to about 1% of period consumption compared to agents in a fully rational model. The welfare costs of learning are not always negative; agents benefit when the capital stock is increased as the result of over-saving by the learning agents. This increase in welfare in a boundedly rational model is an extension of the "optimal irrationality" introduced by Feigenbaum et al. (2011).

The uncertainty regarding the future of social security is also economically interesting. Social security policy uncertainty is not terribly costly to agents in a rational expectations framework. Forward looking, rational agents are able to save and partially self-insure against the aggregate risk of unfavorable policy change. The greatest welfare cost of policy uncertainty in the rational model is equivalent to less than 0.3% reduction in period consumption. I find that the maximum welfare cost to agents in a model with social security policy uncertainty is larger in a model with adaptive learning. The maximum welfare cost of policy uncertainty in the LCH model (compared to the LCH model with announced policy) is 1.98%. The CEV are significantly larger in the learning model than the rational expectations model.

Policy makers who use rational expectations models to predict the welfare effects of social security policy change will understate the cost of reform. To the extent that the learning dynamics are a realistic depiction of agent level behavior, the welfare costs of announced or uncertain social security reforms might be quite large. If agents are unable to predict the general equilibrium effects of a change in government policy, they will not be able to respond optimally. The results of this paper suggest that policy makers can help agents by announcing policy and also by explaining how the policy will impact wages and interest rates.

The learning models I develop in the paper include demographic changes and endogenous government debt. Several interesting questions that are beyond the scope of this paper can be addressed in this framework. I plan to explore the relationship between delaying social security reform, growing deficits, and explosive government debt in future work. The learning models developed in this paper provide an excellent framework to examine explosive debt that is not possible in a standard, rational expectations framework. The model could also be used to explore the relationship between recessions, public pensions, and lifecycle savings.

A Stability of REE under Learning

Determinacy is often used as an equilibrium selection device in rational expectations models with multiple equilibria. A complementary approach to selecting equilibria is to conduct stability analysis under learning. If agents' (non-rational) expectations and forecasts converge to a rational expectations equilibrium (REE) in a model with learning, the REE is said to be stable under learning (see Evans and Honkapohja (2001)). Branch et al. (2013) show that the unique REE in an infinite-horizon Ramsey model is stable under N-step optimal learning. Similarly, I numerically verify that determinate REE in my model are stable under LCH learning and FHL Learning.

Given constant (potentially incorrect) expectations $p^e = (R^e, w^e, b^e, a^{j,e}_{terminal})'$, the learning dynamics of the FHL model asymptotically converge to $p = (R, w, b, a^j)'$. Similarly, the dynamics of the LCH model converge from beliefs $p^e = (R^e, w^e, b^e)'$ to $p = (R, w, b)'$. This converge from beliefs to actual prices is called a T-map.

$$T : \mathbb{R}^{J-H+3} \rightarrow \mathbb{R}^{J-H+3} \tag{42}$$

A fixed point of the T map is E-stable if it locally stable under the ordinary differential equation $\frac{dp}{d\tau} = T(p) - p$. E-stability requires the real parts the eigenvalues of the derivative matrix $dT < 1$. I have numerically verified all determinate steady states in the paper are E-stable under LCH learning and FHL learning (at all horizons).

The stability under learning of the determinate REE in this model can be illustrated graphically. The dynamics of the learning model converge to the REE given arbitrary initial conditions near the steady state. Figure 19 illustrates this convergence in the LCH and FHL learning models, calibrated with $J = 6$ (six period lives) and parameters calibrated as detailed in section 5. The example begins with capital, bonds, assets (not pictured) and expectations of all of these things below the steady state values. Agents update their expectations as they receive more information and eventually learn the steady state values, which are indicated with a dashed red line. Agents overshoot the steady state initially. This pattern is observed in Evans et al. (2009).

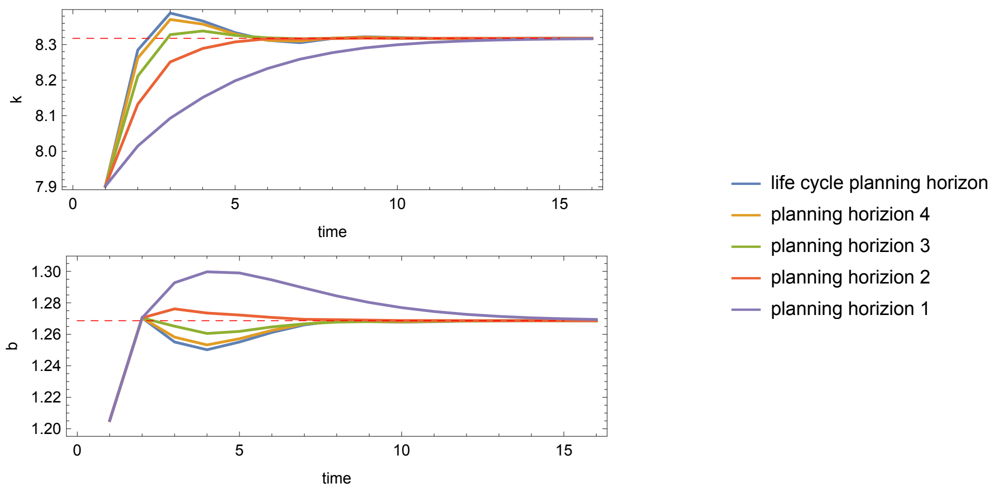

Figure 19: This graphs shows convergence to a steady state for the Finite Horizon Lifecycle Learning model (FHL) and Life-cycle Horizon Learning model (LCH) with different planning horizons. The red dashed line indicates the steady-state. The initial values were chosen to be near the steady-state. The initial expectations were all set equal to the initial values. The top graph shows the path of capital; the bottom graph shows the path of bonds. The converge of the asset choices $(a^1, \ldots, a^5)$ follow a similar oscillating pattern.

Branch et al. (2013) find that longer planning horizons converge to the REE more quickly; and that agents make larger errors (and there are larger aggregate fluctuations) when the planning horizon is longer. I find a similar result; the fluctuations in the economy are greatest for life-cycle horizon learning and decrease as the planning horizon decreases.

B Recessions and Life-cycle Learning

The examples in the main text illustrate the welfare cost of life-cycle and finite-horizon learning in the face of social security policy changes. An alternative framework to view the welfare cost of learning is to simulate a recession and compare the lifecycle utility of agents in the RE model with the utility of agents in a learning model. I do this using the same CEV as defined in equation 31.

I simulate as recession as a surprise decrease in the TFP factor A. The recession is depicted in Figure 20 below. I model the recession as a one period reduction in the TFP factor from $A = 10$ to $A = 9.5$. The recession reduces output per worker by 2.9% and aggregate consumption by 7.53%.[20] Agents in the RE model know

[20]During the Great Recession (2007-2009) output fell by 7.2% and consumption fell by 5.4%. The average post-war recession saw a decline in output and consumption of 4.4% and 2.1%, respectively.

that the recession is only one period. Agents in the learning models do not know the recession will end; they continue to forecast prices and bonds adaptively. The recession generates cycles in capital and bonds, as agents overshoot the steady state before converging. Note that converge back to the steady state is non-monotonic in the rational and learning models (see Azariadis et al. (2004) for a discussion of cycles in rational, multi-period OLG economies). The oscillatory dynamics are larger and last longer in the learning dynamics. The decrease in output and consumption are also larger in the learning models relative to the RE baseline.[21] The welfare cost of experiencing the recession in a model with learning compared to the same recession in the RE model is depicted in Figure 21. The welfare cost varies by cohort. The cost is greatest (smallest CEV) for agents born the period after recession in the LCH model. Agents born in period $t = 11$, following a recession in period $t = 10$, are worse off in the learning model by 1.10% of period consumption. The largest welfare cost occurs after the recession due to the depressed capital stock in the learning model relative to the RE baseline.

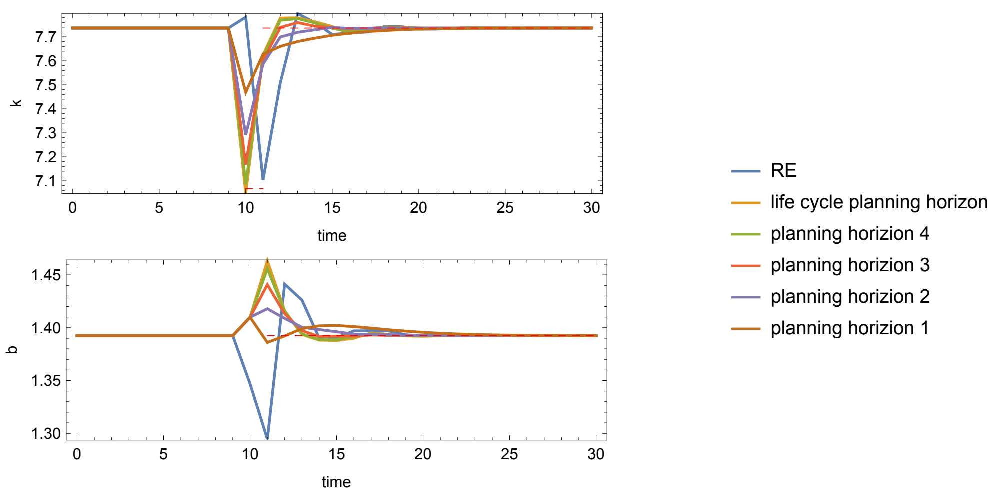

Figure 20: Equilibrium paths for capital and bonds in an economy with a recession in period $t = 10$ that lasts for 1 period. The population growth rate is constant at $n = 0.01$, and social security policy is constant $\tau^0 = 0.1516$,$\tau^1 = 0.045$, $\phi = 0.4$. The recession is a surprise reduction in the TFP factor from $A = 10$ to $A = 9.5$ in period $t = 10$. The TPF factor is depressed for a single period, and the returns to the pre-recession level. In the learning models, the gain parameter $\gamma = 0.93$.

See Christiano (2017).

[21]Eusepi and Preston (2011) show that adaptive learning in an RBC model creates dynamics that more closely match that data than a standard, RE model. They also show that the shocks required to generate a realistic recession are smaller in the model with learning.

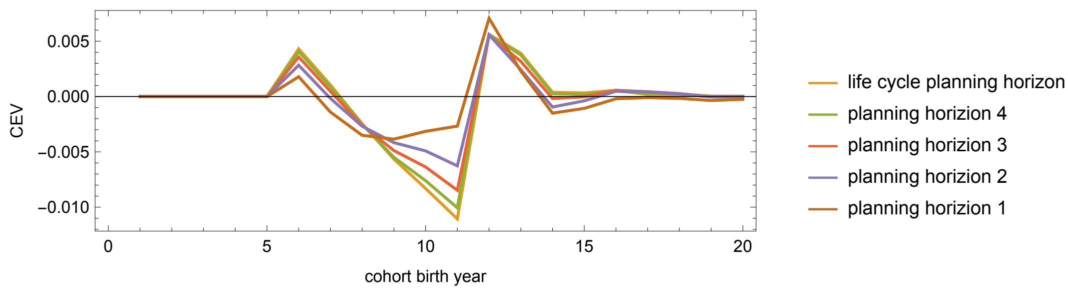

Figure 21: Consumption Equivalent variation measure of the welfare cost of a recession. The population growth rate is constant at $n = 0.01$, and social security policy is constant $\tau^0 = 0.1516$,$\tau^1 = 0.045$, $\phi = 0.4$. The recession is a surprise reduction in the TFP factor from $A = 10$ to $A = 9.5$ in period $t = 10$. The TPF factor is depressed for a single period, and the returns to the pre-recession level. In the learning models, the gain parameter $\gamma = 0.93$.

The welfare cost is greatest to the cohorts born the period after the recession in the LCH model since they experience lower consumption relative to the other generations. Consumption falls during the recession (period $t = 10$) and during the following for rational agents. The dip in consumption after the recession reflect the decision of agents to save as the economy recovers. Consumption also falls for learning agents during the recession. Since the learning agents adaptively forecast prices, they anticipate depressed wages and interest rates the following period, so they continue to consume less for several periods. This is depicted in Figure 22.

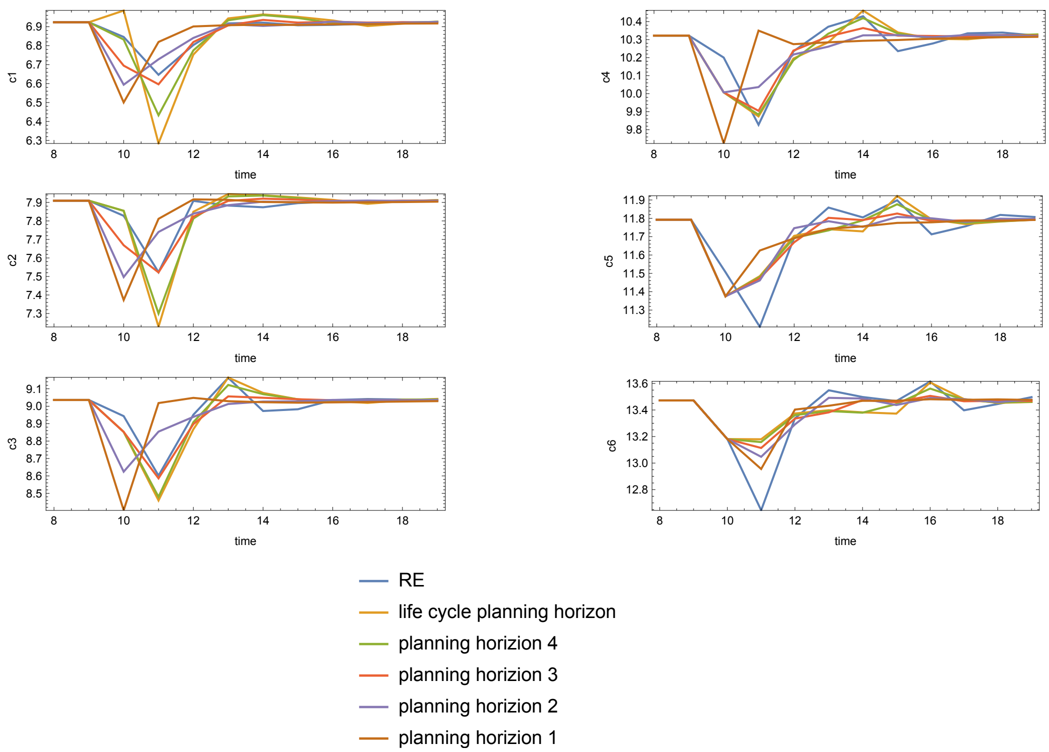

Figure 22: Path of consumption for the rational and learning models for a recession in period $t = 10$. Consumption falls the most during the recession for the 1-period ahead finite horizon learners (dark orange line). Consumption is depressed for many periods in the LCH learning model (yellow line, often very close to the green line). The population growth rate is constant at $n = 0.01$, and social security policy is constant $\tau^0 = 0.1516$,$\tau^1 = 0.045$, $\phi = 0.4$. The recession is a surprise reduction in the TFP factor from $A = 10$ to $A = 9.5$ in period $t = 10$. The TPF factor is depressed for a single period, and the returns to the pre-recession level. In the learning models, the gain parameter $\gamma = 0.93$.

The welfare cost of the learning depends on both the length and depth of the recession, as well as on the gain parameter γ. If the recession were three periods rather than one, the greatest welfare cost would be 1.98% of period consumption (compared to 1.10% for a one period recession). Similarly, if the recession is deeper and the TFP factor falls from $A = 10$ to $A = 9.1$ instead of $A = 9.5$, the largest welfare cost of a one period recession is 2.26% of period consumption. In all three cases, the largest welfare cost occurs under LCH learning for the cohort of agents born after the recession. The welfare cost also depends on the gain parameter γ. When the gain parameter is larger, the welfare cost is also larger. This is because agents place more weight on recent data when the gain parameter is high, and so they expect prices to be more depressed following a one period recession than agents using a smaller gain. The minimum CEV for LCH agents experiencing a one period recession is 0.7% of period consumption for a gain parameter of 0.01, and is equal to 1.10% of period consumption for a gain of 1.

C Response to announced policy, FHL Learning

Agents in the FHL Learning model only respond to policy changes when the announced policy enters their planning horizon. Agents making one step ahead forecasts will not respond to announced policy until the period right before the change. This is illustrated by a simple example in Figure 23. The economy starts off in a dynamically efficient steady state with government debt and a social security system that runs a small surplus. Payroll taxes are increased and benefits are cut in period $t = 6$. This policy change burdens all generations, as take-home pay and social security benefits are lower. The response of young agents is visible in the path of age one savings a^1 on the top of Figure 23. Young agents with a life-cycle planning horizon (that is, they look forward five periods to the end of the life-cycle), increase their savings in period $t = 1$, since they will experience the benefit cut in their old age. Agents with a planning horizon of four periods do not respond until period $t = 2$, when the policy change is four periods away. You agents with a planning horizon of one don't increase their savings until period $t = 5$, despite the fact that the policy change was announced prior. The savings choices of agents at model ages two-five are also included in the figure. Agents of model age five (which is calibrated to be age 75) do not respond to the announced policy until one period ahead, regardless of the planning horizon. These agents do not change their old age savings earlier because they will not be alive to be bothered by the policy change.

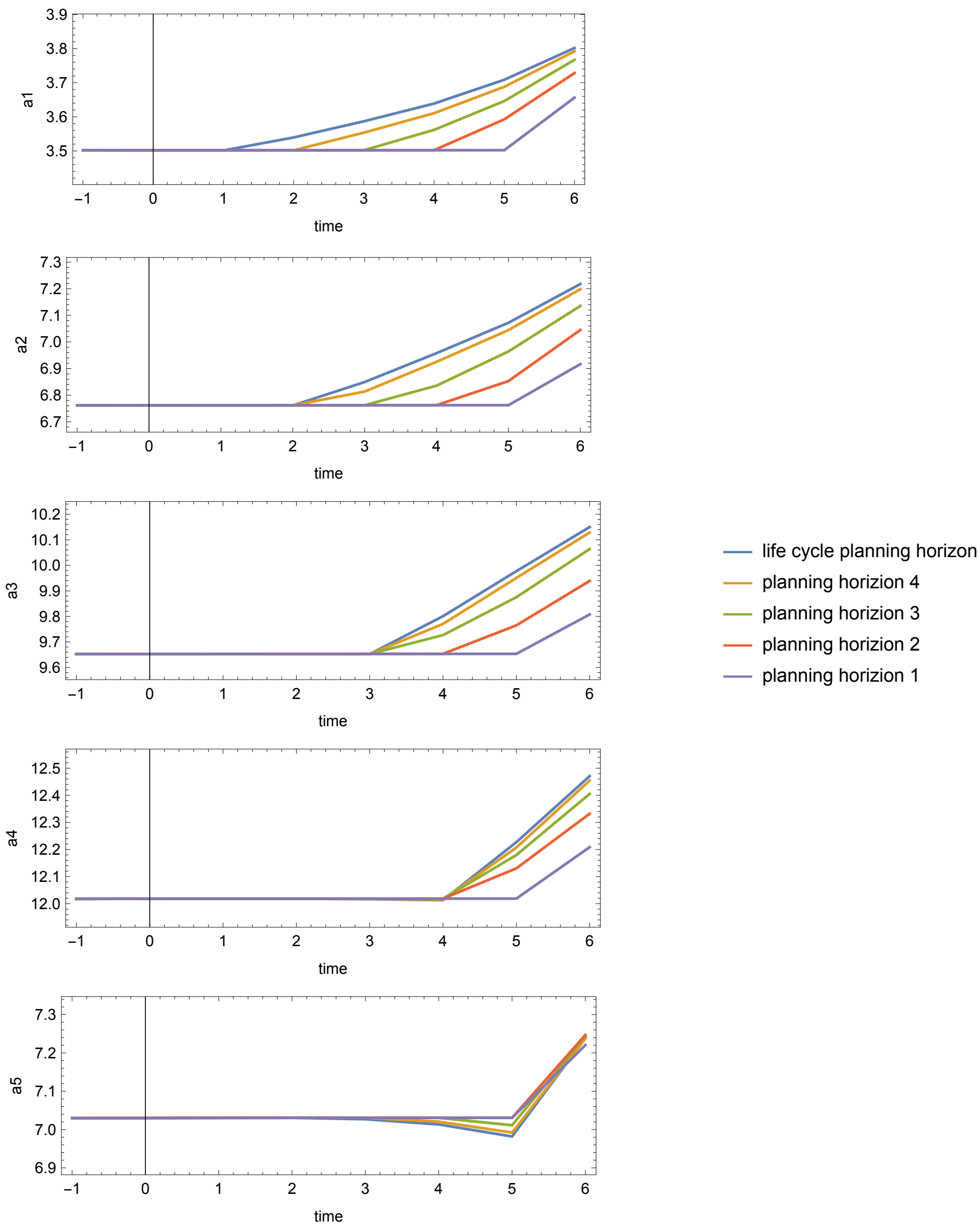

Figure 23: Transition paths for asset holdings for the FHL learning model with different panning horizons. The economy starts off in a steady state; in period $t = 6$ taxes are raised *and* benefits are cut. Agents do not respond to the announced policy change until it enters their planning horizon. The figure only includes the first few time periods, to show the earlier responses of longer planning horizons. The age 5 savings choice a^5 does not change until one period before the policy change. This is because agents do not choose a^5 until the second to last period of their life when their planning horizon is only one period ahead.

D Note on golden rule

The golden rule level of capital maximizes consumption of all generations and occurs when $R = (1 + n)$ in this model.

$$k_{gr} = \left(\frac{\alpha}{n + \delta}\right)^{\frac{1}{1-\alpha}}$$

The consumption profile is smooth over the lifecycle if $R = \beta^{-1}$. This is not the case in general. $R = \beta^{-1}$ if

$$k_{smooth} = \left(\frac{\alpha}{\beta^{-1} + \delta - 1}\right)^{\frac{1}{1-\alpha}}$$

Thus, the golden rule level of capital corresponds to the level of capital such that $R = \beta^{-1}$ (and the lifecycle consumption profile is constant) when $n = \frac{1-\beta}{\beta}$.

The social security program crowds out capital and can be set at a level to ensure $k = k_{gr}$, or $k = k_{smooth}$; it's less clear why a planner would want the latter.

References

Adam, K. (2007). Experimental evidence on the persistence of output and inflation*. *The Economic Journal*, 117(520):603–636.

Adam, K., Marcet, A., and Nicolini, J. P. (2016). Stock market volatility and learning. *Journal of Finance*, 71(1):33–82.

Azariadis, C. (1993). *Intertemporal macroeconomics*. Oxford, UK ; Cambridge, USA : Blackwell, Oxford, UK ; Cambridge, USA.

Azariadis, C., Bullard, J., and Ohanian, L. (2004). Trend-reverting fluctuations in the life-cycle model. *Journal of Economic Theory*, 119(2):334 – 356.

Branch, W., Evans, G. W., and McGough, B. (2013). Finite horizon learning. In Sargent, T. and Vilmunen, J., editors, *Macroeconomics at the Service of Public Policy*, chapter 8, pages 141–163. Oxford University Press, Oxford.

Branch, W. A. (2004). The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations*. *Economic Journal*, 114(497):592–621.

Branch, W. A. and Mcgough, B. (2011). Business cycle amplification with heterogeneous expectations. *Economic Theory*, 47(2/3):395–421.

Bullard, J. and Duffy, J. (2001). Learning and excess volatility. *Macroeconomic Dynamics*, 5(2):272–302.

Bullard, J. and Russell, S. (1999). An empirically plausible model of low real interest rates and unbacked government debt. *Journal of Monetary Economics*, 44(3):477–508.

Bütler, M. (1999). Anticipation effects of looming public-pension reforms. *Carnegie-Rochester Conference Series on Public Policy*, 50:119 – 159.

Caliendo, F. and Aadland, D. (2007). Short-term planning and the life-cycle consumption puzzle. *Journal of Economic Dynamics and Control*, 31(4):1392–1415.

Caliendo, F., Casanova, M., Gorry, A., and Slavov, S. (2016). The welfare cost of retirement uncertainty. Working paper.

Caliendo, F., Gorry, A., and Slavov, S. (2015). The cost of uncertainty about the timing of social security reform. NBER Working Papers 21585, National Bureau of Economic Research, Inc.

Caprioli, F. (2015). Optimal fiscal policy under learning. *Journal of Economic Dynamics and Control*, 58(Supplement C):101 – 124.

Chalk, N. A. (2000). The sustainability of bond-financed deficits: An overlapping generations approach. *Journal of Monetary Economics*, 45(2):293–328.

Christiano, L. J. (2017). The Great Recession: A Macroeconomic Earthquake. Technical report, Federal Reserve Bank of Minneapolis.

Davig, T., Leeper, E. M., and Walker, T. B. (2010). “unfunded liabilities” and uncertain fiscal financing. *Journal of Monetary Economics*, 57(5):600–619.

Eusepi, S. and Preston, B. (2011). Expectations, learning, and business cycle fluctuations. *The American Economic Review*, 101(6):2844–2872.

Evans, G. and Ramey, G. (2006). Adaptive expectations, underparameterization and the lucas critique. *Journal of Monetary Economics*, 53(2):249–264.

Evans, G. W. and Honkapohja, S. (2001). *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton.

Evans, G. W. and Honkapohja, S. (2009). Learning and macroeconomics. *Annual Review of Economics*, 1(1):421–449.

Evans, G. W., Honkapohja, S., and Mitra, K. (2009). Anticipated fiscal policy and adaptive learning. *Journal of Monetary Economics*, 56(7):930 – 953.

Feigenbaum, J., Caliendo, F. N., and Gahramanov, E. (2011). Optimal irrational behavior. *Journal of Economic Behavior and Organization*, 77(3):285 – 303.

Gasteiger, E. and Zhang, S. (2014). Anticipation, learning and welfare: the case of distortionary taxation. *Journal of Economic Dynamics and Control*, 39(Supplement C):113 – 126.

Kitao, S. (2017). Policy uncertainty and the cost of delaying reform: A case of aging japan. forthcoming, Review of Economic Dynamics.

Krueger, D. and Kubler, F. (2004). Computing equilibrium in olg models with stochastic production. *Journal of Economic Dynamics and Control*, 28(7):1411–1436.

Laitner, J. (1990). Tax changes and phase-diagrams for an overlapping generations model. *Journal of Political Economy*, 98(1):193–220.

Marcet, A. and Sargent, T. (1989). Convergence of least squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic Theory*, 48(2):337–368.

Mitra, K., Evans, G. W., and Honkapohja, S. (2013). Policy change and learning in the rbc model. *Journal of Economic Dynamics and Control*, 37(10):1947 – 1971.

Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association*, 55(290):299–306.

Park, H. and Feigenbaum, J. (2017). Bounded rationality, lifecycle consumption, and social security. Working papers.

Pfajfar, D. and Santoro, E. (2010). Heterogeneity, learning and information stickiness in inflation expectations. *Journal of Economic Behavior and Organization*, 75(3):426–444.

Phelan, C. (1999). Anticipation effects of looming public-pension reforms. *Carnegie-Rochester Conference Series on Public Policy*, 50:161 – 164.

Preston, B. (2005). Learning about monetary policy rules when long-horizon expectations matter. Mpra paper, University Library of Munich, Germany.

Preston, B. (2006). Adaptive learning, forecast-based instrument rules and monetary policy. *Journal of Monetary Economics*, 53(3):507–535.

Sargent, T. (1993). *Bounded Rationality in Macroeconomics*. Oxford University Press.

SSA, B. o. T. (2017). The 2017 annual report of the board of trustees of the federal old-age and survivors insurance and federal disability insurance trust funds. House Documents 114-145, Social Security Administration.