Adaptive Learning In An Incomplete-Markets Model

*Andrea Giusto**

March 2012

Abstract

Recent research shows that several DSGE models provide a closer fit to the data under adaptive learning. This paper extends this research by introducing adaptive learning in the model of Krusell and Smith (1998) with uninsurable idiosyncratic risks and aggregate uncertainty. The first contribution of this paper establishes that the equilibrium of this framework is stable under least-squares learning. The second contribution consists of showing that bounded rationality enhances the ability of this model to match the distribution of wealth in the US. Learning increases significantly the Ginis of wealth and income because of the opposite effects on consumption of capital-rich and capital-poor agents. The third contribution is an empirical exercise that shows that a calibrated learning process concerning the aggregate capital stock increases the income (wealth) Gini coefficient by 8.9% (8.8%) in a period of 28 years. Overall, these findings suggest that adaptive learning has important distributional repercussions in this class of models.

*Assistant Professor of Economics, Dalhousie University. Email: andrea.giusto@dal.ca. I thank George Evans, Shankha Chakraborty, and Talan Iscan for their insightful comments, suggestions and encouragement.

1 Introduction

Expectations play a crucial role in representative-agent macroeconomic models. This paper extends the study of the role of expectations to the incomplete-markets model with uninsured idiosyncratic risk and aggregate uncertainty of Krusell and Smith (1998). In a heterogeneous-agent economy, agents must form expectations about the aggregate capital stock because this quantity has an immediate effect on their budget constraints through the determination of returns to productive factors. The main objective of this paper is to examine the effects of learning about the aggregate capital level on the endogenous distribution of wealth. I find that boundedly-rational expectations on the stock of aggregate capital significantly increase wealth and income inequality due to the opposite effects on optimal consumption levels of capital-rich and capital-poor agents. Through a calibration exercise I show that a one-standard deviation change to the expected capital stock increases the Gini coefficients of income and wealth by an average of 6.652% and 6.533% respectively in the first 4 years, and by an additional 2.266% and 2.226% in the successive 24 years. Furthermore these increases in inequality are very persistent, taking much longer to subside than to emerge.

The notion of macroeconomic fluctuations driven by expectations dates back to Pigou (1926) who suggests that expected gains in total factor productivity cause aggregate investment to rise, and therefore boost economic activity. Pigou's compelling insight hinges on the assumption that all agents in the economy respond to an expectational shock in an identical fashion – in other words, the argument uses a representative-agent logic. Yet, in the presence of a non-degenerate distribution of wealth in the economy, it is not straightforward to apply Pigou's logic to understand the consequences of an expectational shock regarding aggregate investment on economic outcomes. To see this point, consider the consumption-saving decision of an agent whose income depends on wages and interest payments. With competitive factor markets, diminishing marginal returns, and complementary factors of production, an increase in expected aggregate investment produces two distinct effects on the budget constraint of a non-representative agent. First, the complementarity of production factors implies that the increase in the stock of capital causes the expectation of wages to increase, thus expanding the agent's intertemporal budget set. This is a positive impact from the point of view of the agent. Second, higher expected investment implies a decrease in the expected rental rate of capital. Thus, there is also a *negative* impact on the agent's budget set.[1] Which one of these two effects dominates depends (among other factors) on the amount of capital owned by a particular agent: the economic significance of the second negative effect is higher for capital-rich than for capital-poor agents. This implies that although agents are otherwise identical, in response to an expected increase in aggregate investment, the capital-rich ones save more, while the capital-poor ones save less. Consequently the distribution of wealth becomes more unequal, while its first moment may either increase or decrease, depending on the initial distribution of capital. This paper develops and demonstrates this intuition using the standard heterogeneous-agent economy of Krusell and Smith (1998).

This paper contributes to three strands of the learning literature. First, Packalén (2000) and Evans and Honkapohja (2001) study the learnability of RBC-type models, and Evans and Honkapohja (2003), and Bullard and Mitra (2002) study the learnability of new-Keynesian models. These studies assume complete markets, and the learnability of equilibria of incomplete-markets DSGE models has not been established before. Second, the present work is also related to the line of research on the effects of learning on macroeconomic aggregates in general equilibrium models. Bullard and Duffy (2001) and Williams (2003) find that least-squares learning does not produce strong effects on the aggregate dynamics of various log-linearized representative-agent models. In contrast, Eusepi and Preston (2011) study the model of Beaudry and Portier (2007) under adaptive learning and find that bounded rationality yields a closer match to US data on output, consumption, and investment. Similarly, citeAevans2005monetary find that in an economy in which the monetary authority follows a Taylor rule, learning increases the volatility of the macroeconomic aggregates. Branch and Evans (2011) show that asset price bubbles may arise under least squares learning when agents estimate the risk-return tradeoff from data on past prices. Consistently with this literature, I find that adaptive learning produces closer match to the data along the unexplored dimension of the wealth and income distributions.

Third, this paper is related to the literature on heterogeneous expectations. Kurz (1994) provides a theoretical motivation to models with heterogeneous expectations, and Evans and Honkapohja (1996) give sufficient conditions for global convergence of heterogeneous-expectations economies to

[1] This argument presumes that the agent has a positive net worth.

a rational-expectations equilibrium. Brock and Hommes (1997), Brock and Hommes (1998), and Branch and McGough (2008) show that expectations heterogeneity can account for very complex (chaotic) dynamics in various asset pricing models. Branch and Evans (2006) study situations in which expectations resulting from misspecified models may coexist in equilibrium. Branch and McGough (2011) show that business cycle fluctuations are amplified in the presence of heterogeneous expectations. The approach followed in this paper is original in the sense that while expectations are shared by all the agents in the economy (and therefore expectations are not heterogeneous) the common expectational operator interacts significantly with the endogenously determined wealth distribution and therefore the *consequences* of bounded rationality are different on different agents. This modeling of the formation of expectations is motivated empirically by the substantial revisions that aggregate investment data are routinely subjected to, which imply both an objective uncertainty surrounding the real-time information on aggregate capital, as well as an economy-wide common learning process concerning these data.

In this paper, the key driver of the changes in the distribution of wealth is the inverse expected comovement of wages and capital rental rates. This negative correlation has been introduced by previous literature in at least three different ways. Kumhof and Rancière (2010) consider a shock to the bargaining power of separate categories of agents. Ríos-Rull and Santaeulalia-Llopis (2010) introduce a stochastic parameter controlling the labor share of output in the production function. Graham and Wright (2010) and Shea (2012) assume that agents learn about factor payments by using information provided by the relative markets. I adopt a fourth possibility, consisting of changing the agents' expectation of the aggregate capital stock. This strategy is parsimonious since adaptive learning affects the linear regression model that the agents of this model economy are already assumed to use to form expectations. For this reason I use the algorithm originally proposed by Krusell and Smith (1998), even though more recent alternatives have been proposed in the literature, see for example Algan, Allais, and Den Haan (2008), Reiter (2009), Kim, Kollmann, and Kim (2010), Den Haan and Rendahl (2010), Young (2010), and Maliar, Maliar, and Valli (2010).

In the previous literature on heterogeneous-agents economies there are two alternative approaches to obtain a more unequal distribution of wealth in this model. First Krusell and Smith (1998) themselves address this issue by introducing additional heterogeneity in the agents' discount factors.

Alternatively, Chang and Kim (2006) obtain high wealth inequality –without this additional layer of heterogeneity– by calibrating the agents' income process on the basis of the Panel Study of Income Dynamics. Here, I find that expectations on factor payments are an additional source of wealth and income inequality that had not been identified before.

The paper proceeds as follows: section 2 outlines the model and establishes the stability under learning of the baseline equilibrium of Krusell and Smith (1998). Section 3 studies the behavior of the model under learning and section 4 illustrates the economic intuition behind the results of section 3. Section 5 presents a calibration that assesses the empirical relevance of bounded rationality for the observed increase in income inequality in the US in the past 28 years. Section 6 concludes.

2 The Model

The model economy is the same as in Krusell and Smith (1998). There is a continuum (measure one) of infinitely-lived agents with constant relative-risk-aversion time-separable utility function $U_0 = E\sum_{t=0}^{\infty} \frac{c_t^{1-\eta}}{1-\eta}$, where η is the risk-aversion parameter and c_t^i is consumption during period t. The aggregate production function is $Y_t = z_t K_t^{\alpha} L_t^{1-\alpha}$, where K_t is time-t aggregate capital, L_t is aggregate labor, α is the capital share of output, and z_t is the aggregate productivity parameter following a two-state Markov process ($z_t \in \{z_g, z_b\}$) with known probabilities. Factors markets are competitive so that wages and real interest rates (gross of depreciation δ) are

$$w_t(\mu_t, l_t, z_t) = (1-\alpha) z_t \left(\frac{\mu_t}{l_t}\right)^{\alpha} \tag{1a}$$

$$r_t(\mu_t, l_t, z_t) = \alpha z_t \left(\frac{\mu_t}{l_t}\right)^{(\alpha-1)} \tag{1b}$$

where μ_t denotes average capital holdings and l_t denotes the employment rate at time t. Agent $i \in [0,1]$ is exposed to uninsurable idiosyncratic Markov shocks $\varepsilon_t^i \in \{0,1\}$ where $\varepsilon_t^i = 0$ means that i is unemployed at time t. The idiosyncratic shocks are correlated with aggregate productivity and the laws of large numbers imply that the measure of the employed is perfectly correlated with the aggregate shock ($l_t \in \{l_g, l_b\}$). The distribution of agents over capital holdings and employment states at time t is trivially defined over a sigma algebra containing all the possible realizations of k_t^i (individual capital holdings) and ε_t^i.

Following Krusell and Smith (1998), average capital ownership μ_t is used as a sufficient proxy for the entire distribution of wealth. The agent's dynamic problem is

$$v^i(k_t^i, \varepsilon_t^i; \mu_t, z_t) = \max_{k_{t+1}^i, c_t^i} \left\{ \frac{{c_t^i}^{1-\eta}}{1-\eta} + \beta E\left[v^i(k_{t+1}^i, \varepsilon_{t+1}^i; \mu_{t+1}, z_{t+1}) | \varepsilon_t^i, z_t \right] \right\} \tag{2}$$

subject to:

$$c_t^i + k_{t+1}^i = r(\mu_t, l_t, z_t)k_t^i + w(\mu_t, l_t, z_t)\varepsilon_t^i + (1-\delta)k_t^i \tag{2a}$$

$$\log \mu_{t+1} = \begin{cases} a_0 + a_1 \log \mu_t, & \text{if } z_t = z_g \\ c_0 + c_1 \log \mu_t, & \text{otherwise} \end{cases} \tag{2b}$$

$$k_{t+1}^i \geq -\kappa \tag{2c}$$

the transition probabilities

Equations (2a) and (2c) are the budget and credit constraints respectively ($\kappa \geq 0$ is an exogenous parameter.) Constraint (2b) is the equation used by the agents to forecast the next-period level of aggregate capital, which is information relevant for the calculation of the optimal consumption plan. In the following I refer to the free parameters of equation (2b) (a_0, a_1, c_0, and c_1) as the *expectational parameters*. Their equilibrium values (denoted a_0^*, a_1^*, c_0^*, and c_1^* respectively) are determined to deliver the best forecast of aggregate capital in a mean-squared error sense, consistent with a log-linear AR(1) econometric specification. The algorithm used to solve the model involves these steps: (1) guess initial values for the expectational parameters, (2) determine the policy function through a value-function iteration algorithm, (3) simulate a long panel with many agents and aggregate their decisions, (4) use the ordinary-least-squares estimator on the aggregated simulated data to evaluate the best-fitting expectational parameters, (5) if the initial guess is different from the OLS estimates, update the expectational parameters in the direction suggested by the estimates and repeat from step (2).

Tables 1, and 2 show the parameterization adopted by Krusell and Smith (1998) and Table 3 reports a few statistics resulting from the application of the algorithm outlined above. The average

$\beta = 0.99,$	$\delta = 0.025$
$\eta = 1,$	$\alpha = 0.36$
$z_g = 1.01,$	$z_b = 0.99$
$l_g = 0.32,$	$l_b = 0.30$

Tab. 1: Parameterization following the baseline Krusell and Smith (1998). The calibration of the parameters l_g and l_b may appear different from the values reported for unemployment rates of 4% and 10% during, respectively, expansions and recessions. Nevertheless, careful analysis of the computer code implementing the baseline model, reveals that agents are assumed to inelastically supply $\frac{1}{3}$ units of time to the labor markets in each period. In other words l_g and l_b are expressed here in units of time rather than as employment rates.

$$\begin{bmatrix} \pi_{gg00} & \pi_{gg01} & \pi_{gb00} & \pi_{gb01} \\ \pi_{gg10} & \pi_{gg11} & \pi_{gb10} & \pi_{gb11} \\ \pi_{bg00} & \pi_{bg01} & \pi_{bb00} & \pi_{bb01} \\ \pi_{bg10} & \pi_{bg11} & \pi_{bg10} & \pi_{bb11} \end{bmatrix} = \begin{bmatrix} 0.2927 & 0.5834 & 0.0938 & 0.0313 \\ 0.0243 & 0.8507 & 0.0091 & 0.1159 \\ 0.0313 & 0.0938 & 0.5250 & 0.3500 \\ 0.0021 & 0.1229 & 0.0389 & 0.8361 \end{bmatrix}$$

Tab. 2: Transition probabilities. The subscripts of each entry indicate the aggregate transition (first two characters) and the individual transition (the following two digits). For example π_{gb01} is the probability that a currently unemployed agent finds a job when the economy transitions from high productivity to low productivity.

Variable	average	st.dev.
μ	11.60	0.2692
μ_g	11.67	0.2584
μ_b	11.52	0.2591
σ	7.361	0.2244
σ_g	7.375	0.2206
σ_b	7.348	0.2273
a_0^*	0.0924	-
a_1^*	0.9633	-
c_0^*	0.0826	-
c_1^*	0.9652	-

Tab. 3: Summary statistics for the economy at the stochastic equilibrium; μ and σ denote the mean and standard deviation of the cross sectional distribution of capital averaged over 10,000 periods; the subscripts b and g denote recessions and expansions respectively. The standard errors of the expectational parameters are unreported since the precision of these estimates can be made arbitrarily small by increasing the length of the simulation.

level of capital during expansions is slightly higher than during recessions, and wealth is tightly distributed around the mean in both phases of the business cycle. The values of the expectational parameters reported in Table 3 are self-confirming: when agents adopt the log-linear forecasting model, their aggregate behavior produces data that, when used to estimate the best fitting log-linear AR(1) model, confirms the agents' initial beliefs.[2] Clearly, the standard deviation of the distribution of wealth is too small to match the Gini coefficients observed in any real economy.

2.1 E-Stability

I depart now from Krusell and Smith (1998) and I modify their baseline framework by assuming that agents do not know the equilibrium values of the aggregate capital process (i.e. a_0^*, a_1^*, c_0^*, and c_1^* from Table 3) but rather that they use regression analysis on past data to estimate them. Equation (2b) shows that the uncertainty about the equilibrium values of the expectational parameters directly translates into uncertainty about the equilibrium process governing the evolution of the aggregate stock of capital. The specification of the regression used by the agents, together with the current estimates for its parameters, are known in the learning literature as the agents' Perceived Law of Motion (PLM). Each agent's saving-consumption plan obviously depends on their PLM, and the resulting choices, once aggregated, determine the economy's Actual Law of Motion (ALM). Thus, the learning process can be conceptualized as a recursion between the PLM and the ALM: the time- t PLM determines the ALM which, in turn, produces the new data with which the new PLM at time $t+1$ is estimated. Letting $\Phi_t \equiv [\hat{a}_{0,t}\ \hat{a}_{1,t}\ \hat{c}_{0,t}\ \hat{c}_{1,t}]'$ be the vector of the time-t estimates of the expectational parameters, the learning process is represented by a sequence $\{\Phi_t\}$ in the *parameter space*, that evolves according to an unknown map denoted by T. This map depends both on the behavior of the aggregate economy as well as on the learning algorithm. Under

[2] Defining the equilibrium reported in Table 3 as a rational expectations equilibrium may be incorrect. In fact, the behavior of this model could be different when the consumption function is obtained through different representations of the state space. For this reason, the classification that best suits this equilibrium is that of a Restricted Perception Equilibrium proposed by Evans and Honkapohja (2001) and Sargent (2001). Such an equilibrium requires that the economic agents choose the best parameterization possible, within a *limited* set of expectational models and that the parameter values are self-confirming according to some metric.

least-squares learning the agents use the following recursion to update their estimates

$$\Phi_t = \Phi_{t-1} + \gamma_t R_t^{-1} y'_{t-1} [\ln \mu_t - y_{t-1} \Phi_{t-1}] \tag{3a}$$

$$R_t \equiv R_{t-1} + \gamma_t (y'_{t-1} y_{t-1} - R_{t-1}) \tag{3b}$$

where $y_t = [\mathbb{I}_g \quad \mathbb{I}_g \ln \mu_{t-1} \quad \mathbb{I}_b \quad \mathbb{I}_b \ln \mu_{t-1}]$, $\mathbb{I}_g$ is a dummy equal to 1 if $z_t = z_g$ ($\mathbb{I}_b = 1 - \mathbb{I}_g$), and $\{\gamma_t\}_{t=0}^{\infty}$ is known as "gain sequence" i.e. a positive non-increasing sequence of real numbers such that $\sum \gamma_t = \infty$ and $\sum \gamma_t^2 < \infty$.[3] Using the notation introduced in this section, an equilibrium of the learning process satisfies the condition $\Phi^* = T(\Phi^*)$.

When agents are learning about the structure of the economy, there exists an endogenous interaction between the current estimates of the expectational parameters and the process that generates the data used to update those estimates. An exogenous change to the expected future capital stock, triggers changes in the economy's aggregate saving rate, which leads in turn to changes to capital accumulation and these new data are then used by the agents to revise their expectations. Accordingly, the evolution of agents' expectations is endogenous to the model and in general, it is by no means obvious that the learning process should (a) converge, or (b) if it does converge, whether the convergence point is located at the equilibrium reported in Table 3. Precisely because such convergence criteria are not automatically satisfied, learning has been used as an equilibrium selection device – see, for example, Evans and Honkapohja (2001) – under the point of view that equilibria the are not stable under learning are less likely to be empirically relevant. Evans and Honkapohja (2001) also provide necessary and sufficient conditions for the local convergence to an equilibrium of a learning process of the kind considered here. These conditions – known as the E-stability principle – require that the matrix of first derivatives of T evaluated at an equilibrium Φ^* (denoted $DT|_{\Phi^*}$ in the following) has eigenvalues with real parts less than one, in which case the equilibrium is locally asymptotically stable under learning.

In the heterogeneous-agents economy with aggregate shocks the T-map cannot be pinned down analytically and consequently the evaluation of its matrix of first derivatives must be conducted

[3] I use dummies in the regression to keep the notation to a minimum. The matrix R is block diagonal by construction, and allowing this type of heteroskedasticity in the regressions allows the estimation of a specification of the form (2b) with a more compact notation.

numerically. This is done by initially setting the PLM at the equilibrium values reported in Table 3 and first increasing and then decreasing each parameter by a small quantity. The resulting pair of policy functions is used to perform two long Monte-Carlo simulations (a substantial initial portion of the data is discarded to ensure ergodicity). The resulting ALMs are estimated from the data generated in this way. The ratios of the changes in the estimates relative to the small initial changes in the PLM provide a numerical estimate of one of the rows of $DT|_{\Phi^*}$. A detailed step-by-step description of the algorithm used to differentiate the map T is given in the appendix. Application of this algorithm yields the following numerical evaluation of the matrix of interest

$$DT|_{\Phi^*} = \begin{bmatrix} 2.6 & -1.7 & 3.4 & -1.9 \\ 13.7 & -6.7 & 18.1 & -8.3 \\ 2.9 & -1.7 & 3.1 & -1.8 \\ 14.2 & -6.7 & 17.7 & -8.2 \end{bmatrix}.$$

The biggest eigenvalue of this matrix has a real part equal to -0.12, indicating that – according to the E-stability principle – the equilibrium Φ^* is locally asymptotically stable under learning.

3 The Dynamics Under Learning

Having established the equilibrium stability under least-squares learning, I study now the model in a neighborhood of the equilibrium expectational parameters in the parameter space. More specifically I first investigate how sensitive are the model's dynamics to small perturbations of the expectational parameters, and subsequently I study how different are the learning dynamics from the equilibrium.

3.1 Calibration of the learning process

Answering the two questions posed above requires the preliminary specification of what constitutes a reasonable deviation of the expectational parameters from the equilibrium values. In order to provide an answer, I exploit the link between the expectational parameters and the expected stock of capital in the next period. Since agents use an AR(1) regression, their expectation of the aggregate capital stock changes in the same direction of the change in any of the expectational parameters.

Therefore it is possible to map uncertainty about the equilibrium values of the expectational parameters to and from uncertainty about the aggregate capital stock.

To estimate a confidence interval around the aggregate capital stock I use real-time quarterly data on real gross nonresidential private investment in the US (series RINVBF from the Philadelphia Federal Reserve Bank.) The data set includes 188 vintages of data as available in each quarter from 1965:Q2 to 2012:Q2. Each vintage starts in 1947:Q1, with a only a handful of exceptions: five vintages start in 1959:Q1, five start in 1959:Q3, and one starts in 1987:Q1. To avoid normalization issues, I drop these eleven vintages and I normalize to 100 the figure for 1947:Q1 in the remaining 177 vintages. For each quarter from 1965:Q2 on, I estimate the aggregate capital by summing the investment data in the preceding quarters, net of depreciation, by using all the available vintages.[4] I repeat this calculation by using all the available vintages that include that particular quarter, and this procedure yields a time series of revisions to the aggregate capital stock available at a given point in time. I then calculate the standard deviation of these revisions and I divide this number by the initial figure. Repeating this set of calculations for each quarter, it emerges that for the median quarter, the successive revisions of the initial estimate of the capital stock are within a one standard-deviation confidence interval of $\pm 5.88\%$.[5]

3.2 Expectational Sensitivity Analysis

I study now the model's response to changes in the expectational parameters that imply estimation errors of plus/minus one and two standard deviations of the long-run equilibrium capital stock, according to the calibration discussed above.[6] The graphs in Figure 1 show the ergodic distributions of a few key statistics produced by the model under expectations that are excessive in a range of plus/minus two standard deviations from the equilibrium values. The top panel in Figure 1 shows the ergodic distributions of the capital stock. Notably, there is a negative relationship between the values of the expectational parameters and the capital stock. The bottom two panels show the

[4] This is a reasonable estimate of the capital stock. With the depreciation rate calibrated at 2.5%, the share of capital in each vintage that is accounted for by this calculation is 99.5% or more.

[5] An alternative calibration procedure based on Ríos-Rull and Santaeulalia-Llopis (2010) and yielding a comparable confidence interval is given in the appendix.

[6] Experimentation reveals that the exact way in which the expectational change is introduced – say increasing a_1 without altering a_0, c_0, or c_1, or vice versa – does not alter the results of this section in any way.

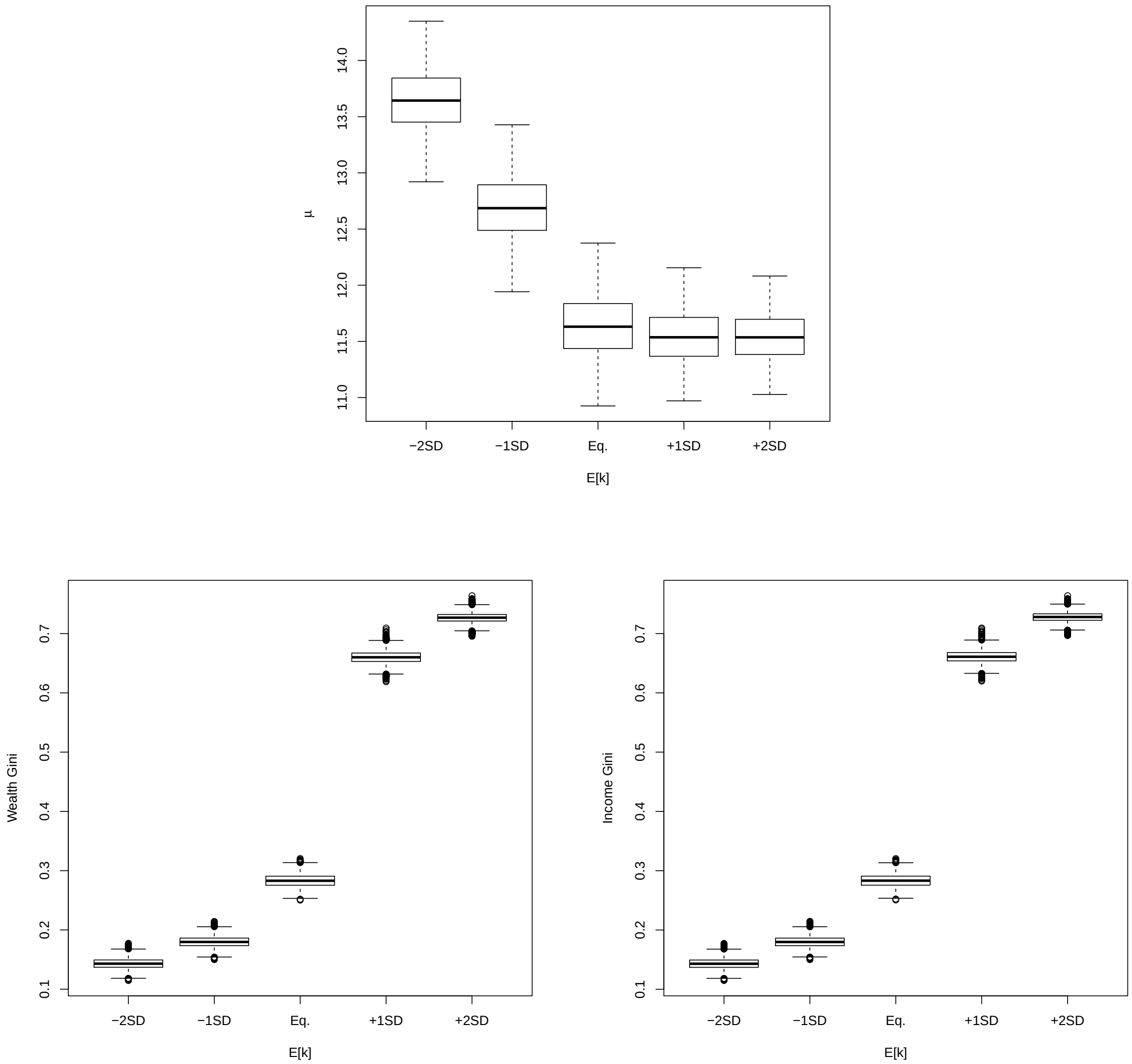

Fig. 1: Effects on the ergodic wealth distribution of changes in aggregate capital expectations. The horizontal axis indicates the various cases relative to the expected long-run capital level, for example the bars corresponding to $E[k] + 1SD$ pertain to the case in which agents expect a long-run level of capital that is one standard deviation above the equilibrium value.

ergodic distributions of the Gini coefficients of wealth and income. With respect to this latter two it is possible to notice an increasing non-linear relationship with the bias in the expected capital levels.[7] Also, the two Ginis are very similar, which is unsurprising since in this model income differences originate from idiosyncratic wage earnings and wealth ownership. Figure 1 shows that excessive expectations for the capital stock have a strong effect on the model's behavior, and in particular on the distribution of wealth. Figure 2 breaks down the changes in the distribution of wealth by quintiles. Clearly, wealth moves upwards when aggregate capital expectations are excessive: the increase in the share of the top quintile of the distribution of income is gained at the expense of the lower quintiles. Furthermore, in the last panel of Figure 2 one can see that most of the additional share of income earned by the highest quintile is indeed driven by the top percentile of the distribution of income, who experience a spectacular increase in their share of income relative to all other agents in the economy. Overall, this evidence suggests that learning has the potential to produce significant changes to the model's dynamics, provided that the learning process sojourns long enough in a region of the parameter space that causes an overestimation of the capital stock. Furthermore, the changes in the distribution of income induced by excessive capital expectations match qualitatively well the trends in the US data, where income concentrates towards the top of the distribution.

3.3 The Learning Dynamics

Figure 3 shows a long simulation of the model under learning after initializing the expectations to be one standard deviation above equilibrium. As expected from the verification of E-stability, the economy eventually converges to the equilibrium. This simulation uses the following gain sequence

$$\gamma_{t+1} = \frac{\gamma_t}{\gamma_t + 1}, \quad \text{given } \gamma_0 = 10^{-4} \tag{4}$$

which is a standard choice in the learning literature (see for example Evans and Honkapohja, 2001, Chapter 15). At the beginning of the simulation, the majority of the agents react to the the high expected level of aggregate capital by decreasing savings (this is studied thoroughly in section 4) and

[7] The asymmetry in Figure 1 depends on the fact that the autoregressive coefficients in equations (2b) are very close to one.

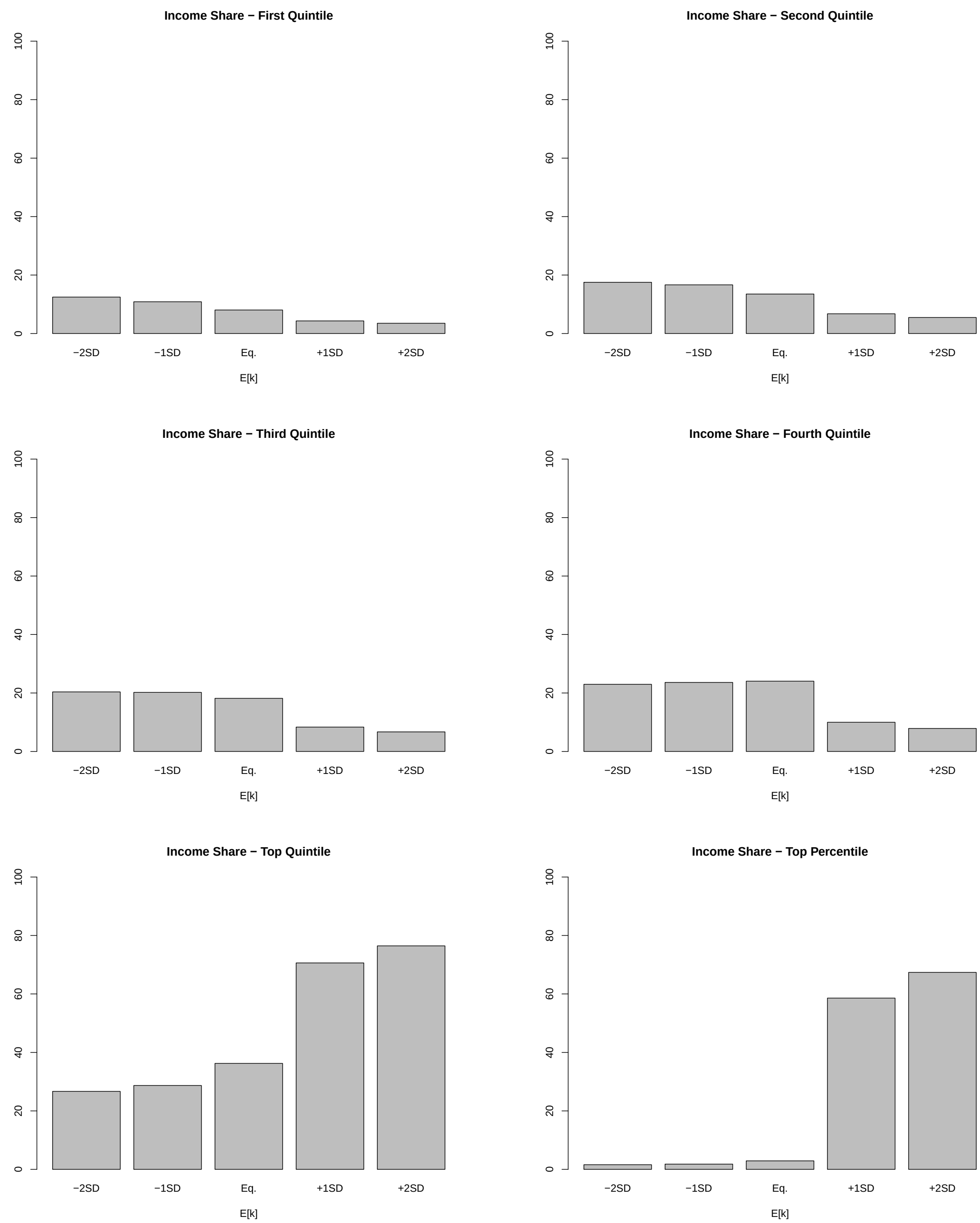

Fig. 2: Average changes in the shares of income earned by each quintile of the distribution of income corresponding to near-equilibrium expectations. The unit of measure of the vertical axes are percentage points.

increasing consumption immediately. This causes in turn a drop in the capital accumulation process which increases the interest rate. This significantly increases the income of a minority of wealthy agents and they respond by accumulating more capital. In the mean time, the learning dynamics converge back to the equilibrium dynamics, and once the expectational parameters are close enough to the equilibrium values, the wealthier (poorer) agents revert towards more (less) consumption, and the wealth distribution converges back to the equilibrium process. Overall, figure 3 shows that learning has important effects on the distribution of wealth (the Gini coefficient increases by 50% in this example) and that the inequality produced in this manner is *extremely* persistent. While the expectational parameters converge to their equilibrium values at the expected rate of $\sqrt{t}$, the distribution of wealth remains more unequal than at the equilibrium for a long time. One crude measure of this persistence is given by comparing the 254 periods it takes for capital to first return within the two-standard-deviation range of the ergodic process to the 10,335 periods for the Gini of wealth to do the same – about 40 times longer.

4 The Economic Intuition

I argue in this section that the behavior of the model illustrated in section 3 can be rationalized by the opposite effects on expected wages and rental rates, resulting from an overestimation of the capital stock. To facilitate exposition, I will denote with E^o an expectational operator that yields an overestimation of future capital levels. The application of E^o to equations (1) implies that the *expected* wage rate is higher, and the return rate to capital lower than under equilibrium expectations (denoted with E^*). These changes in expected returns to productive factors can be usefully decomposed into substitution and wealth effects. First, higher expected wages have a positive wealth effect on current consumption. Second, the expected decrease in the interest rates further increases consumption through the substitution effect. Third, the expected decrease in rental rates produces a wealth effect that may either increase or decrease consumption depending on each agent's net worth: an agent with positive net assets faces a negative wealth effect, while one with negative net assets faces a positive wealth effect.[8] All three effects work in the same direction

[8] Logarithmic preferences and market completeness imply that the income and substitution effects offset each other exactly for every agent. However, under incomplete markets, Aiyagari (1994) and Huggett (1993) show that the interest rate is lower than the rate of time preference.

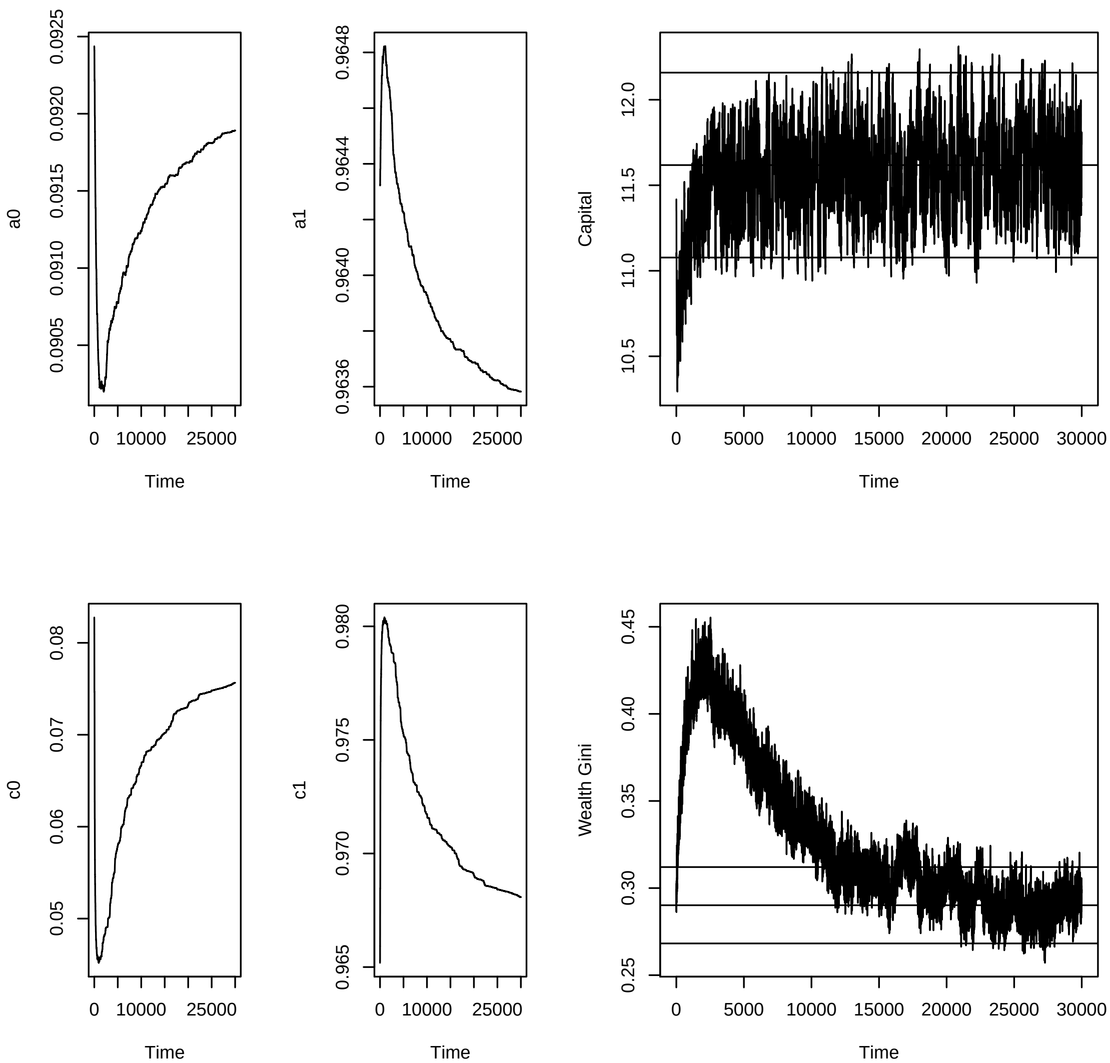

Fig. 3: Time paths under learning of the expectational parameters and of aggregate capital and its Gini coefficient. The initial parameter values are set so that the initial aggregate capital is overestimated by one standard deviation.

for agents with negative net worth and therefore their current consumption increases when E^o is used instead of E^*. The net effect of E^o on the consumption of agents with positive net worth is ambiguous since the first wealth and substitution effects are balanced by the third wealth effect. As one considers richer and richer agents, the amount of income earned by renting capital becomes more and more significant, relative to the income earned by offering labor services. Accordingly, it is intuitive to speculate that there must exist a critical threshold of wealth above (below) which E^o has a *negative* (positive) net effect on the agent's expected permanent income. The following Proposition refprop:incomeEffects proves that this threshold exists, and furthermore that it is positive and finite. The proof, as given below, is somewhat simplified by the assumption that E^o is a point-expectational operator.[9] A proof for the general case is given in the appendix.

Proposition 1. *An agent with point expectational operator E^o will anticipate higher permanent income than an identical agent with equilibrium expectations E^* if and only if*

$$k^i_{t+1} < \psi E^o \mu_{t+1} \tag{5}$$

where $\psi = E_t\left[\frac{z_{t+1}\varepsilon^i_{t+1}}{l^\alpha_{t+1}}\right] / E_t\left[\frac{z_{t+1}}{l^{\alpha-1}_{t+1}}\right]$, and E_t is the operator consistent with the probabilities reported in Table 2.

Proof: Household i's income at $t+1$ is given by

$$y^i_{t+1} \equiv z_{t+1}\alpha\left[\frac{\mu_{t+1}}{l_{t+1}}\right]^{\alpha-1} k^i_{t+1} + z_{t+1}(1-\alpha)\left[\frac{\mu_{t+1}}{l_{t+1}}\right]^{\alpha}\varepsilon^i_{t+1}$$

the assumption of point expectations implies that $E^o[\mu^\alpha_{t+1}] = [E^o\mu_{t+1}]^\alpha$. Therefore, the time-$t$ expectation for y^i_{t+1}, conditional on the current aggregate and individual states, can be written as

$$E_tE^o[y^i_{t+1}] = \sum_{s\in\{b,g\}} P_{z_tz_s}\left(z_s\alpha\left[\frac{E^o\mu_{t+1}}{l_s}\right]^{\alpha-1} k^i_{t+1} + z_s(1-\alpha)\left[\frac{E^o\mu_{t+1}}{l_s}\right]^{\alpha}\pi_{z_ts\varepsilon^i_t1}\right)$$

[9] A point expectational operator pertains to a degenerate probability distribution that places a unit mass of probability on a single event. Point expectations correspond to the situation in which agents believe to know with certainty the relevant variables. This is an accurate approximation for the purpose of applying this proposition to the simulations summarized in Figure 1 since the standard errors of the regressions run by the agents are extremely small, consistently with Krusell and Smith (1998).

where $P_{z_t z_g}$ and $P_{z_t z_b}$ are the probabilities that the aggregate state transitions from z_t to z_g and z_b respectively, and and $\pi_{z_t s \varepsilon_t 1}$ are the probabilities that agent i is employed at $t+1$ conditional on ε_t^i and the aggregate state transition. These probabilities are found in Table 2, and they are assumed to be known by the agents. Now, calculate the effect of a change expected level of average capital at $t+1$ on expected income

$$\frac{\partial EtE^o[y_{t+1}^i]}{\partial E^o \mu_{t+1}} = \alpha(1-\alpha)E^o \mu_{t+1}^{\alpha-2} \left\{ -k_{t+1}^i E_t \left[\frac{z_{t+1}}{l_{t+1}^{\alpha-1}} \right] + E^o \mu_{t+1} E_t \left[\frac{z_{t+1}\varepsilon_{t+1}^i}{l_{t+1}^{\alpha}} \right] \right\}$$

This expression shows that the necessary and sufficient condition for a marginal increase in the expected aggregate capital level (μ_{t+1}) to increase expected personal income ($\partial E_t E^o[y_{t+1}^i]/\partial E^o \mu_{t+1} > 0$) is $k_{t+1}^i < \psi E^o \mu_{t+1}$.

□

Proposition 1 implies that a high expectation of future capital affects the expected income of agents asymmetrically across a critical wealth threshold. Those agents *above* the threshold expect *lower* income in the future, and since they are dynamic optimizers that seek to smooth consumption, they *decrease* current consumption thus saving more and accumulating *more wealth.* On the contrary, those agents *below* the critical threshold expect *higher* income in the future, hence they consume *more* presently and end up owning *less* wealth. The polarization of wealth and income then arises as a direct consequence of these considerations.[10]

I conclude this section by explaining the decreasing relationship between the stock of capital and the changes of expectations as shown in the top panel of Figure 1. The decrease in average capital corresponding to the increase in expected capital takes place because the equilibrium ergodic distribution of wealth does not feature very rich agents: at the equilibrium, the wealthiest percentile owns about 3 times the average capital, no agent receives as much income from interest payments as from wages, and the top 0.1 percentile earns five times more from labor than from capital. These figures are consistent with the standard deviations reported in Table 3. Accordingly, the vast majority of agents are likely to be located below the threshold identified in Proposition 1. This

[10] Notice that *both* classes of agents behave sub-optimally, and therefore it is not obvious to state which group suffers the most in terms of foregone welfare.

implies that expectation that overestimate aggregate capital, have a negative effect on the savings of most agents, which explains the decreasing relation between the expectational parameters and the mean of the distribution of wealth.

4.1 Self-Insurance and Precautionary Savings

In this model, the agents are exposed to uninsurable shocks and therefore they accumulate extra capital (precautionary savings) as a form of self-insurance against prolonged spells of unemployment. However, the demand for precautionary savings is unlikely to drive the increased dispersion of the distribution of wealth. There are two reasons for not considering precautionary savings in this discussion. First, Aiyagari (1994) and Dıaz, Pijoan-Mas, and Rıos-Rull (2003) argue that precautionary savings account only for a small proportion of total wealth in this type of models, and therefore consumption smoothing and intertemporal substitution motives dominate the dynamics of individual savings decisions. Second, changes to expectations of the kind considered in this paper, will push precautionary savings in the same direction independently of wealth. When agents have high expectations for aggregate capital, the stochastic environment appears more favorable because the positive payoff (the wage in case of employment) is expected to increase, and the probabilities do not change. Furthermore, as interest payments decrease, capital accumulation becomes a less effective form of self-insurance, and therefore the net effect on precautionary savings is unambiguously negative, independent of the agent's individual level of wealth. Opposite and symmetric considerations apply to the case in which agents underestimate aggregate capital.

5 An Empirical Application

This section presents a calibrated learning process yielding an estimate of the effect of boundedly rational expectations for aggregate capital on income inequality over a number of time periods that correspond to 28 years of real time. The time window is chosen to allow for a direct comparison with the Report by the Congressional Budget Office (2011), documenting significant increases in income inequality in the US from 1979 to 2007. Since the evolution of the expectational parameters is endogenous it is possible that the learning dynamics bring the expectational parameters to regions of the parameter space that are not acceptable on the basis of prior restrictions. For example, values for a_1 or c_1 exceeding or equal to one are not consistent with the existence of a stationary equilibrium.[11] To overcome this problem, it is not uncommon in the learning literature to invoke a "projection facility," i.e. a mechanism according to which the agents ignore estimates that they know to be impossible. It turns out that if the initial expectations are not too far from the equilibrium values it is possible to avoid using a projection facility, and for this reason I initiate the simulations of this section at parameter values implying an overestimation of the capital stock by by one standard deviation (i.e. +5.88% of the ergodic equilibrium value).

The gain sequence is given by the recursion of equation (4) and the initial value of this sequence (γ_0) must be calibrated. This free parameter controls the speed with which agents learn the equilibrium law of motion of the economy or, equivalently, the speed with which they reach equilibrium expectations. As the empirical counterpart to the changes in expectations are the measurement errors in the aggregate stock of capital, my strategy for calibrating γ_0 consists of requiring that its value is chosen so that the biases in the expected level of capital evolve in a way that is similar to the evolution of the revisions to aggregate capital stock, as calculated with the procedure of section 3. This strategy may be implemented by either looking at one-period-ahead conditional expectations biases, or by looking at long-run unconditional biases. Focusing on long-run expectations is clearly preferable since agents are dynamic optimizers. Accordingly, I define the variable

$$\log bias_t = \frac{1}{2}\left[\frac{\hat{a}_{0,t}}{1-\hat{a}_{1,t}} - \frac{a_0^*}{1-a_1^*}\right] + \frac{1}{2}\left[\frac{\hat{c}_{0,t}}{1-\hat{c}_{1,t}} - \frac{c_0^*}{1-c_1^*}\right]$$

to measure the magnitude of the deviation of the unconditional expectations of long-run capital from the equilibrium value. (The factors of $\frac{1}{2}$ are due to the fact that the probabilities of Table 2 imply that the unconditional probability of a positive productivity shock is the same as that of a negative shock.) The learning schedule is then required to be such that the partial autocorrelation diagram of the series $\{bias_t\}$ is similar to the partial autocorrelation diagram of the revisions to estimates of the capital stock. Figure 4 reports the partial autocorrelation function of the variable $bias_t$ averaged over one hundred simulations with $\gamma_0 = 2.25 \times 10^{-2}$. The second panel of Figure 4 shows the partial autocorrelation function of the revisions of the estimated capital stock for quarter

[11] The special case in which a_1 and c_1 are simultaneously equal to one and $a_0 = -c_0$ is instead acceptable as it does not imply explosive aggregate capital. It is straightforward to argue that the probability of the learning algorithm to reach *exactly* these estimates is infinitesimal.

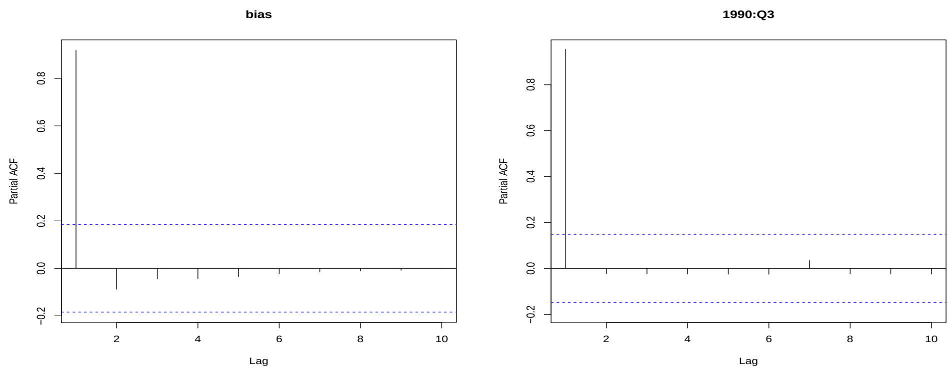

Fig. 4: Left: partial autocorrelation function for the average bias across the 100 Monte Carlo replications. Right: partial autocorrelation function for the time-series of the revisions of the estimated aggregate capital stock available in quarter 3 of 1990. This autocorrelation function is typical.

3 of 1990. This quarter is typical and the picture would be nearly identical with most other quarters or their average.

How much income inequality can be generated by the expectational channel identified in this paper, assuming that agents overestimate the stock of capital by one-standard deviation, and learn according to the calibrated learning schedule, over a period of time of the 28 years? A Monte Carlo strategy with 100 repetitions is adopted to answer this question, and the results can be viewed as estimates of the impulse response functions following a one-time-period one-standard-deviation change to the estimated stock of aggregate capital. The results of these simulations are illustrated in Figure 5 and in Table 4. With the notable exception of the fourth quintile, the model successfully matches the direction (although not the magnitude) of the change for each quintile of the income distribution. Furthermore, the model seems to better capture the dynamics of the income distribution at the lower three quintiles, while it overpredicts the change in the share of the top 80^{th} to 99^{th} percentiles, and it grossly underpredicts the increase in the share of income of the top percentile. One likely factor accounting for this failure is the US tax system, whose progressivity decreases sharply at the higher end of the income distribution. As mentioned above, the results shown in Table 4 do not capture fully the changes in the data and this can be explained on several grounds. First, these results pertain to a one-time, one-standard-deviation change to

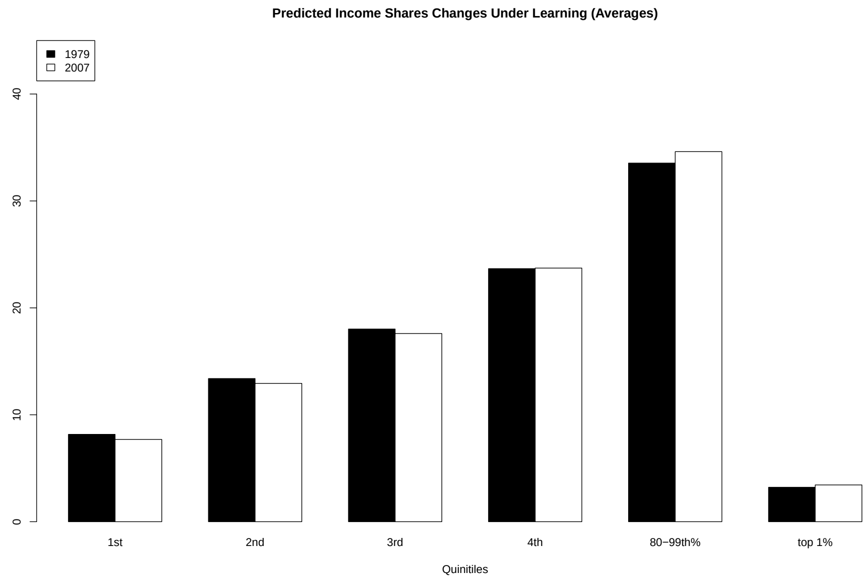

Fig. 5: Changes in the shares of income earned by each quintile in 112 periods in the model following a one standard-deviation increase in the expected capital stock.

expected capital and implementing the more canonical two-standard-deviations case would surely help the model to predict more extreme changes, at an unfeasibly high computational cost.[12] And second, no one factor in isolation is likely to explain by itself the trends in these data, as indicated by the existence of a large literature on the distributions of wealth and income.

	Quintiles					
	1^{st}	2^{nd}	3^{rd}	4^{th}	80-99%	top 1%
Data	-2%	-3%	-2.2%	-2.1%	0.5%	9.4%
Model	-0.4735%	-0.4528%	-0.4216%	0.0503%	1.0705%	0.2204%
Model/Data	0.2368	0.1509	0.1916		2.1410	0.0234

Tab. 4: Comparison to the US data of the changes in the shares of income predicted by the model after a one-standard-deviation increase in expected aggregate capital.

Figure 6 shows the responses of the expectational parameters with the two-standard deviations bounds. The resulting series for the variable $bias_t$ is shown on the right of Figure 6. Figure 7

[12] At the two standard deviation level most simulations need to be discarded as the expectational parameters often drift to unacceptable regions of the parameter space.

	Dependent Variable							
	Income Gini	Income Gini	Wealth Gini	Wealth Gini	Income Share 0-20%	Income Share 0-20%	Income Share 20-40%	Income Share 20-40%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept					7.7673	8.0046	13.0798	13.2850
					0.0064	0.0483	0.0041	0.0297
Time(1-112)	0.0763		0.0750		-0.0010		-0.0016	
	0.0005		0.0005		0.0001		0.0001	
Time(1-16)		0.4033		0.3964		-0.0265		-0.0196
		0.0217		0.0218		0.0050		0.0031
N	11300	1700	11300	1700	11300	1700	11300	1700
R^2	0.6600	0.7570	0.6525	0.7481	0.0092	0.2038	0.0526	0.2698

	Dependent Variable							
	Income Share 40-60%	Income Share 40-60%	Income Share 60-80%	Income Share 60-80%	Income Share 80-99%	Income Share 80-99%	Income Share Top 1%	Income Share Top 1%
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Intercept	17.8603	17.9903	23.7768	23.7182	34.1851	33.7530	3.3309	3.2530
	0.0020	0.0118	0.0022	0.0155	0.0086	0.0623	0.0014	0.0096
Time(1-112)	-0.0028		-0.0005		0.0046		0.0012	
	0.0000		0.0000		0.0001		0.0000	
Time(1-16)		-0.0125		0.0063		0.0443		0.0076
		0.0012		0.0016		0.0064		0.0010
N	11300	1700	11300	1700	11300	1700	11300	1700
R^2	0.4159	0.4865	0.0161	0.1231	0.0998	0.3004	0.1948	0.3437

Tab. 5: Standard errors reported under the estimates. All figures are significant at the 1% level. Income and Wealth Gini are expressed as percent deviations from ergodic equilibrium values.

reports the responses of the income shares and the income and wealth Ginis, and it shows that the changes produced by adaptive learning persist long after the expectational parameters are close to equilibrium values. Jointly, these graphs are consistent with an overestimation of the capital stock and with the intuition presented in section 3.

Figure 7 also shows that, except for the share of income earned by the third quintile (which displays a steady and prolonged decline), most of the effects of the excessive expectations are produced quickly. For example, 75% of the peak effect on the income Gini is produced in the first

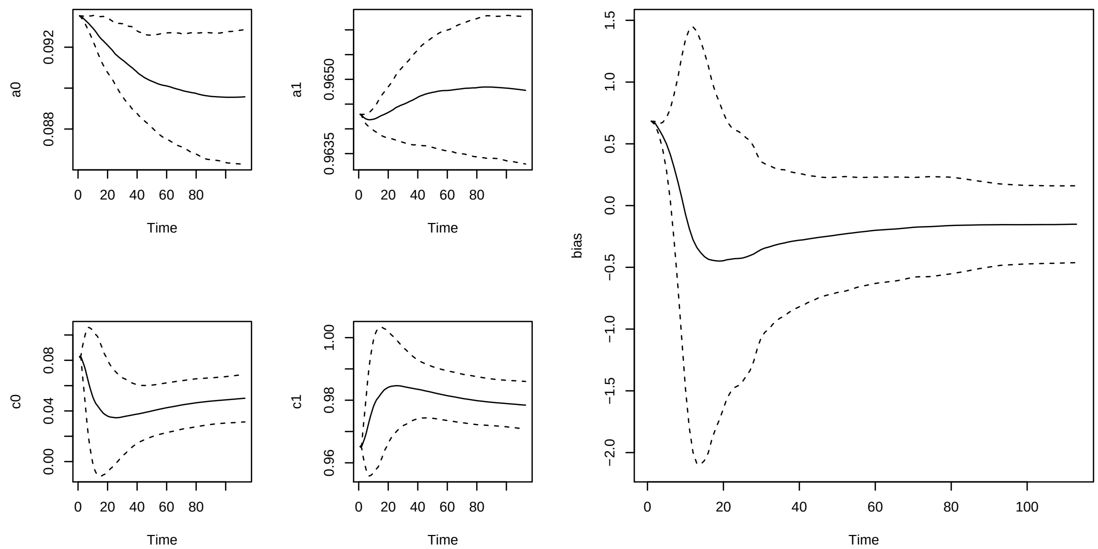

Fig. 6: Time paths of the expectational parameters and of the implied bias in unconditional capital expectations, with the two-standard deviation bounds.

16 time periods. For this reason, in order to assess the statistical significance and magnitude of the effect of a one-standard deviation change in expected aggregate capital, I use regression analysis both over the entire sample of 112 time periods (28 years) and over a shorter 16-periods (4 years) window. Table 5 shows the results of regressing the changes of wealth and income Ginis (regressions 1 through 4) and income shares by quintile on a deterministic trend (regressions 5 through 16). This analysis uncovers that the effects documented are statistically significant in all cases considered in the sample produced via Monte Carlo replications. The model is at odds with the data only in relation to the change in the fourth quintile of the income distribution (-2.1% according to CBO estimates), where it predicts a small but significant decrease in the full time period considered (regression 11), while also predicting an initial increase (regression 12). Columns (1) through (4) of Table 5 show that the average effect of a one-standard deviation overestimation of the capital stock is to increase the Gini of income (wealth) by .0763% (.075%) per quarter, for a cumulative average increase of 8.918% (8.759%) over 28 years. In the first 4 years instead, the average increase in the Ginis are respectively .4033% and .3964% per quarter, totaling 6.652% and 6.533%.

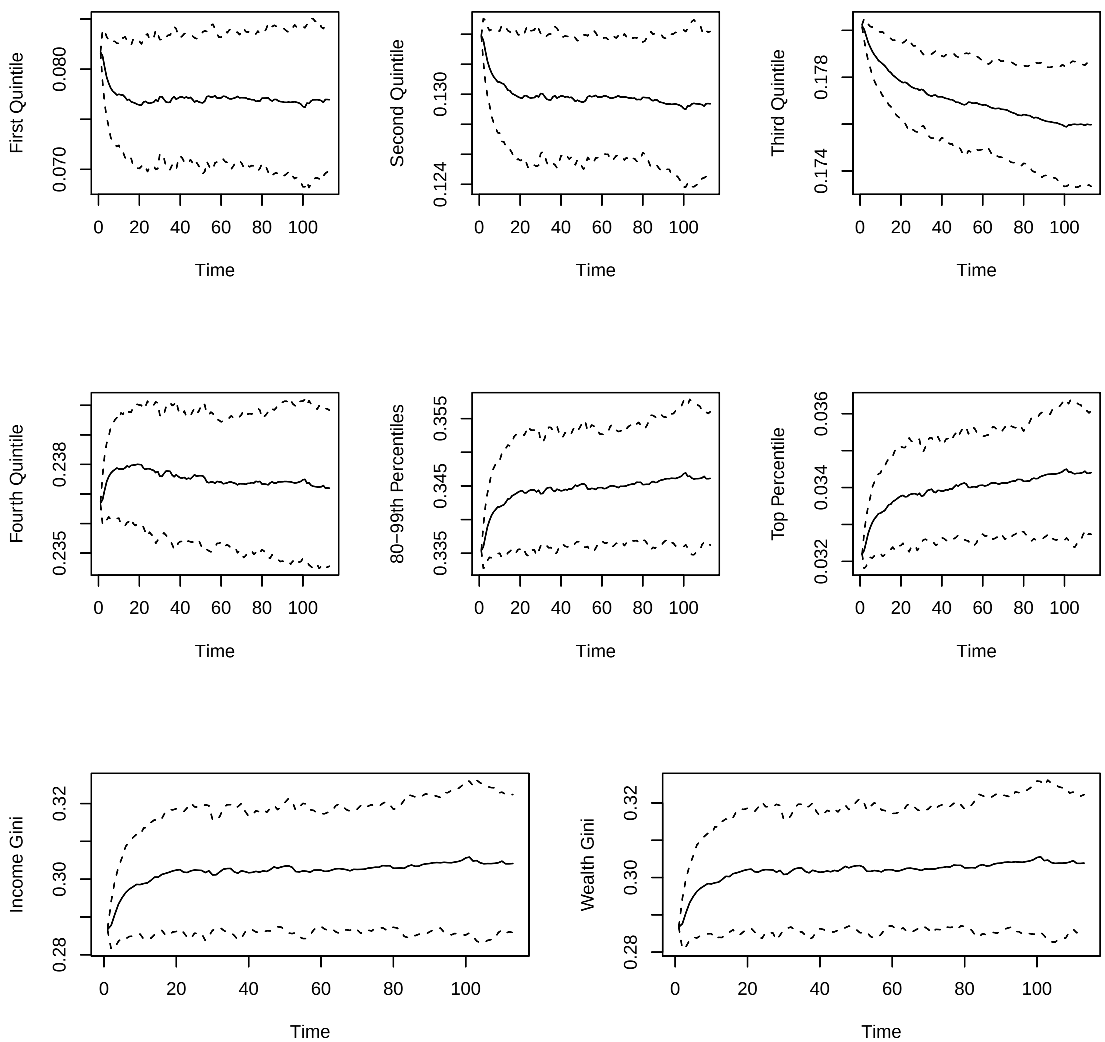

Fig. 7: Responses of the income shares of each quintile with the two-standard deviations bounds. At the bottom, the response of the income and wealth Gini coefficients.

6 Conclusion

The study of dynamic stochastic general equilibrium models under learning has proved to be a productive exercise towards the goal of capturing several features of the data produced by real economies, both qualitatively and quantitatively. Until now, the effects of adaptive learning in DSGE models featuring incomplete markets have not been studied. From a theoretical point of view this gap is noticeable because one of the reasons why markets may be incomplete is precisely the limit to the computational ability of the agents engaging in economic activity. From the empirical point of view, this work shows that – consistently with the existing literature on the effects of learning in a variety of different models – learning gives the model studied here a better match to the empirical distributions of wealth and income.

References

Aiyagari, S. (1994). Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, *109*(3), 659–684.

Algan, Y., Allais, O., & Den Haan, W. (2008). Solving heterogeneous-agent models with parameterized cross-sectional distributions. *Journal of Economic Dynamics and Control*, *32*(3), 875–908.

Beaudry, P., & Portier, F. (2007). When can changes in expectations cause business cycle fluctuations in neo-classical settings? *Journal of Economic Theory*, *135*(1), 458–477.

Branch, W., & Evans, G. (2006). Intrinsic heterogeneity in expectation formation. *Journal of Economic theory*, *127*(1), 264–295.

Branch, W., & Evans, G. (2011). Learning about risk and return: A simple model of bubbles and crashes. *American Economic Journal: Macroeconomics*, *3*(3), 159–191.

Branch, W., & McGough, B. (2008). Replicator dynamics in a cobweb model with rationally heterogeneous expectations. *Journal of Economic Behavior and Organization*, *65*(2), 224–244.

Branch, W., & McGough, B. (2011). Business cycle amplification with heterogeneous expectations. *Economic Theory*, *47*(2), 395–421.

Brock, W., & Hommes, C. (1997). A rational route to randomness. *Econometrica: Journal of the Econometric Society*, *65*(5), 1059–1095.

Brock, W., & Hommes, C. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic dynamics and Control*, *22*(8-9), 1235–1274.

Bullard, J., & Duffy, J. (2001). Learning and excess volatility. *Macroeconomic Dynamics*, *5*(02), 272–302.

Bullard, J., & Mitra, K. (2002). Learning about monetary policy rules. *Journal of Monetary Economics*, *49*(6), 1105–1129.

Chang, Y., & Kim, S. (2006). From individual to aggregate labor supply: a quantitative analysis based on a heterogeneous agent macroeconomy. *International Economic Review*, *47*(1), 1–27.

Congressional Budget Office. (2011). *Trends in the Distribution of Household Income Between 1979 and 2007.* Congress of the United States, Congressional Budget Office. October 2011 release.

Den Haan, W., & Rendahl, P. (2010). Solving the incomplete markets model with aggregate uncertainty using explicit aggregation. *Journal of Economic Dynamics and Control*, *34*(1), 69–78.

Dıaz, A., Pijoan-Mas, J., & Rıos-Rull, J. (2003). Precautionary savings and wealth distribution under habit formation preferences. *Journal of Monetary Economics*, *50*(6), 1257–91.

Eusepi, S., & Preston, B. (2011). Expectations, learning, and business cycle fluctuations. *American Economic Review*, *101*, 284472.

Evans, G., & Honkapohja, S. (1996). Least squares learning with heterogeneous expectations. *Economics Letters*, *53*(2), 197–201.

Evans, G., & Honkapohja, S. (2001). *Learning and expectations in macroeconomics.* Princeton Univ Pr.

Evans, G., & Honkapohja, S. (2003). Expectations and the stability problem for optimal monetary policies. *Review of Economic Studies*, *70*(4), 807–824.

Graham, L., & Wright, S. (2010). Information, heterogeneity and market incompleteness. *Journal of Monetary Economics*, *57*(2), 164–174.

Huggett, M. (1993). The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control*, *17*(5-6), 953–969.

Kim, S., Kollmann, R., & Kim, J. (2010). Solving the incomplete market model with aggregate uncertainty using a perturbation method. *Journal of Economic Dynamics and Control*, *34*(1), 50–58.

Krusell, P., & Smith, A. (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, *106*(5), 867–896.

Kumhof, M., & Rancière, R. (2010). Inequality, Leverage and Crises. *IMF Woking Paper*(10/268), 28.

Kurz, M. (1994). On the structure and diversity of rational beliefs. *Economic Theory*, *4*(6), 877–900.

Maliar, L., Maliar, S., & Valli, F. (2010). Solving the incomplete markets model with aggregate uncertainty using the Krusell-Smith algorithm. *Journal of Economic Dynamics and Control*, *34*(1), 42–49.

Packalén, M. (2000). On the learnability of rational expectations equilibria in three business cycle models. *University of Helsinki, Department of Economics*.

Pigou, A. (1926). *Industrial fluctuations.* London: MacMillan.

Reiter, M. (2009). Solving heterogeneous-agent models by projection and perturbation. *Journal of Economic Dynamics and Control*, *33*(3), 649–665.

Ríos-Rull, J., & Santaeulalia-Llopis, R. (2010). Redistributive shocks and productivity shocks. *Journal of Monetary Economics*, *57*(8), 931–948.

Sargent, T. (2001). *The conquest of American inflation.* Princeton Univ Pr.

Shea, P. (2012). Learning by doing, short-sightedness, and indeterminacy*. *The Economic Journal*.

Williams, N. (2003). Adaptive learning and business cycles. *Manuscript, Princeton University*.

Young, E. (2010). Solving the incomplete markets model with aggregate uncertainty using the Krusell-Smith algorithm and non-stochastic simulations. *Journal of Economic Dynamics and Control*, *34*(1), 36–41.

Appendices

A General Proof of Proposition 1

Proposition 2. *An agent with expectational operator E^o will anticipate higher personal income in the future if and only if*

$$k_{t+1}^i < -\frac{(1-\alpha)E_t\left[\frac{z_{t+1}\varepsilon_{t+1}^i}{l_{t+1}^{\alpha}}\right]Cov^*[\mu_{t+1},\mu_{t+1}^{\alpha}]}{\alpha E_t\left[\frac{z_{t+1}}{l_{t+1}^{\alpha-1}}\right]Cov^*[\mu_{t+1},\mu_{t+1}^{\alpha-1}]} \tag{6}$$

Proof: Household i's expected income is

$$E_tE^o[y_{t+1}^i] = P_1z_g\alpha E^o\left[\frac{\mu_{t+1}}{l_g}\right]^{(\alpha-1)}k_{t+1}^i + P_1z_g(1-\alpha)E^o\left[\frac{\mu_{t+1}}{l_g}\right]^{(\alpha)}P_1^1$$

$$+P_2z_b\alpha E^o\left[\frac{\mu_{t+1}}{l_b}\right]^{(\alpha-1)}k_{t+1}^i + P_2z_b(1-\alpha)E^o\left[\frac{\mu_{t+1}}{l_b}\right]^{(\alpha)}P_2^1$$

where $P_1 = P_{z_t z_g}$, P_1^1 is prob. i is employed *conditional* on the aggregate transition being $z_t \to z_g$. $P_2 = 1 - P_1 = P_{z_t z_b}$.

$$E_t E^o[y_{t+1}^i] = P_1 z_g \alpha l_g^{1-\alpha} k_{t+1}^i E^o\left[\mu_{t+1}^{\alpha-1}\right] + P_1 z_g (1-\alpha) l_g^{-\alpha} P_1^1 E^o[\mu_{t+1}^{\alpha}]$$

$$+P_2 z_b \alpha l_b^{1-\alpha} k_{t+1}^i E^o\left[\mu_{t+1}^{\alpha-1}\right] + P_2 z_b (1-\alpha) l_b^{-\alpha} P_2^1 E^o[\mu_{t+1}^{\alpha}]$$

$$= \alpha k_{t+1}^i E^o[\mu_{t+1}^{\alpha-1}] \underbrace{E_t\left[\frac{z_{t+1}}{l_{t+1}^{\alpha}}\right]}_{X_1} + (1-\alpha) E^o[\mu_{t+1}^{\alpha}] \underbrace{E_t\left[\frac{z_{t+1}\varepsilon_{t+1}^i}{l_{t+1}^{\alpha}}\right]}_{X_2}$$

A change in the agent's regression parameters will change the mean of the normal probability distribution function used by the agents to perform forecasts. Conditional on time t information I will denote the conditional mean with μ^*. That is

$$\mu^* = \begin{cases} e^{a_0}\mu_t^{a_1} & \text{if } z_t = z_g \\ e^{c_0}\mu_t^{c_1} & \text{if } z_t = z_b \end{cases}$$

The effect of optimistic expectations is singled out by calculating

$$\frac{\partial E_t E^o[y_{t+1}^i]}{\partial \mu^*} = \alpha k_{t+1}^i X_1 \frac{\partial E^o[\mu_{t+1}^{\alpha-1}]}{\partial \mu^*} + (1-\alpha) X_2 \frac{\partial E^o[\mu_{t+1}^{\alpha}]}{\partial \mu^*}$$

So one needs to calculate $\frac{\partial E^o[\mu_{t+1}^{\alpha}]}{\partial \mu^*}$

$$\frac{\partial E^o[\mu_{t+1}^{\alpha}]}{\partial \mu^*} = \frac{\partial}{\partial \mu^*} \int x^{\alpha} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu^*)^2}{2\sigma^2}\right] dx$$

$$= \int x^{\alpha} \frac{(x-\mu^*)}{\sigma^2} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu^*)^2}{2\sigma^2}\right] dx$$

$$= \frac{1}{\sigma^2}\left[\int x^{\alpha+1}\phi(x)dx - \mu^* \int x^{\alpha}\phi(x)dx\right]$$

$$= \frac{1}{\sigma^2}\left(E^o[\mu_{t+1}^{\alpha+1}] - \mu^* E^o[\mu_{t+1}^{\alpha}]\right)$$

And similarly

$$\frac{\partial E^o[\mu_{t+1}^{\alpha-1}]}{\partial \mu^*} = \frac{1}{\sigma^2}\left(E^o[\mu_{t+1}^{\alpha}] - \mu^* E^o[\mu_{t+1}^{\alpha-1}]\right)$$

Now consider the two random variables $X = \mu_{t+1}$ and $Y = X^{\alpha-1}$. Obviously X and Y are not independent and exploiting the fact that $Cov[X,Y] = E[XY] - E[X]E[Y]$ one sees that $\frac{\partial E^o[\mu_{t+1}^{\alpha-1}]}{\partial \mu^*} = Cov^*[\mu_{t+1}, \mu_{t+1}^{\alpha-1}]/\sigma^2$. With a similar line of reasoning it is possible to show that setting $Y = X^{\alpha}$ the relationship $Cov[X,Y] = E[XY] - E[X]E[Y]$ yields

$$\frac{\partial E^o[\mu_{t+1}^{\alpha}]}{\partial \mu^*} = \frac{Cov^*[\mu_{t+1}, \mu_{t+1}^{\alpha}]}{\sigma^2}$$

Finally the condition $\frac{\partial E_t E^o[y_{t+1}^i]}{\partial \mu^*} > 0$ requires

$$\alpha k_{t+1}^i X_1 Cov^*[\mu_{t+1}, \mu_{t+1}^{\alpha-1}] > -(1-\alpha) X_2 Cov^*[\mu_{t+1}, \mu_{t+1}^{\alpha}]$$

$$k_{t+1}^i < -\frac{(1-\alpha) X_2 Cov^*[\mu_{t+1}, \mu_{t+1}^{\alpha}]}{\alpha X_1 Cov^*[\mu_{t+1}, \mu_{t+1}^{\alpha-1}]}$$

where the reversal of the direction of the inequality sign is necessary because $Cov^*[\mu_{t+1}, \mu_{t+1}^{\alpha-1}] < 0$. This is because $\mu_{t+1}^{\alpha-1}$ is a decreasing function of μ_{t+1}. Finally, since μ_{t+1}^{α} is an increasing function of μ_{t+1}, the quantity on the RHS is positive. This proves that optimism about the capital accumulation process implies a higher personal income for agent i if and only if i is not "too wealthy."

$\square$

The Algorithm for Calculating DT

The matrix of first derivatives of the T map is derived numerically, by changing the PLM and deriving the implied ALM. The following steps describe the algorithm to derive the matrix of first derivatives of the T map at the equilibrium.

1. Set $\iota = 1$ and pick a small constant ϑ.
2. Set $\Phi = [a_0^* \ a_1^* \ c_0^* \ c_1^*]'$, and let $\Phi[\iota]$ denote the ι-th element of Φ. Set $\Phi[\iota] = \Phi[\iota] + \vartheta$.
3. Using a value function iteration algorithm derive the policy function $\hat{f}_{\iota}^{+}$.

4. Using 5000 agents, run a 210,000-periods long Monte Carlo experiment using $\hat{f}_i^+$ to simulate the economy.

5. Drop the first 200,000 periods of the simulation and use OLS on the remaining sample to obtain an estimate of the ALM: $(a_0)_\iota^+$, $(a_1)_\iota^+$, $(c_0)_\iota^+$, and $(c_1)_\iota^+$

6. Set $\Phi = [a_0^* \; a_1^* \; c_0^* \; c_1^*]'$ and $\Phi[\iota] = \Phi[\iota] - \vartheta$. Repeat Steps 3, 4, and 5 but this time denote the implied ALM as $(a_0)_\iota^-$, $(a_1)_\iota^-$, $(c_0)_\iota^-$, and $(c_1)_\iota^-$

7. Populate the ι-th row of DT as follows

$$DT_\iota = \left[\frac{(a_0)_\iota^+ - (a_0)_\iota^-}{2\vartheta} \quad \frac{(a_1)_\iota^+ - (a_1)_\iota^-}{2\vartheta} \quad \frac{(c_0)_\iota^+ - (c_0)_\iota^-}{2\vartheta} \quad \frac{(c_1)_\iota^+ - (c_1)_\iota^-}{2\vartheta} \right]$$

8. If $\iota < 4$ set $\iota = \iota + 1$ and repeat from Step 2.

The matrix reported in the main text is derived by using $\vartheta = 0.0001$, which is appropriate when numbers are have a double precision representation in the computer.

Least Squares Learning Algorithm

The simulation of the model under least-squares learning is implemented in these steps.

1. Choose an in initial point in the parameter space Φ_0 and initial distribution of wealth Γ_0, a simulation length T, and the number of agents N.

2. Choose a constant γ_0 and calculate the gain $\gamma_t = \frac{\gamma_{t-1}}{1+\gamma_{t-1}}$ for $t = 1, 2, \ldots, T$.

3. Set $t = 0$

4. Solve the dynamic program (2).

5. Draw an aggregate shock and N idiosyncratic shocks. Use the policy obtained in step 4 to determine the individual saving decisions, and aggregate them to obtain μ_{t+1}.

6. Use equations (3) to update Φ_t.

7. Set $t = t + 1$ and repeat from step 4 until $t = T$.

The initial distribution of wealth Γ_0 is the equilibrium distribution corresponding to Table 3. The initial values for the expectational parameters imply an overestimation of the long run capital of +5.88%. Also, this algorithm requires the initialization of the matrix R_0 for the application of the recursive least squares algorithm, as it can be seen from equations (3). I use the variance-covariance matrix of a regression specified as in (2b) run on equilibrium data.

An alternative calibration procedure

Deviations from equilibrium values of the expectational parameters in equation (2b) imply that the agents have a bias in their expectations of future stock of capital. An alternative way to think of shocks to capital levels as a source of randomness, is through shocks to the elasticity of output to labor (or share of labor's output, in competitive factor pricing models). I now show that the assumed production function allows to consider expectations about aggregate capital and investment as expectations on a stochastic labor share of output, with the same effect on factors payments.

To facilitate the exposition I suppose now that agents overestimate the capital stock. Such an expectational shock enters in the model *exclusively* through the intertemporal consumption decisions of the agents. More specifically, only the budget constraints (2a) are affected through the expected payments to productive factors i.e. through the functions w and r. Because w is an increasing function of capital, a higher expected future capital stock implies a higher expected future wage rate. Conversely, since r is a decreasing function of the capital stock, high expectations for it imply low expected future capital rental rates. Such a divergent effect on interest rates and wages can also be produced through a stochastic shock to the labor share of output of the kind considered by Ríos-Rull and Santaeulalia-Llopis (2010) who postulate a production function of the following form

$$Y_t = z_t K_t^{\alpha-\zeta_t} L_t^{1-\alpha+\zeta_t} \tag{7}$$

where ζ_t is a covariance-stationary zero-mean random variable.[13] A production function such as (7), and competitive factor markets can be used to show that a positive value of ζ_t increases wages and

[13] Ríos-Rull and Santaeulalia-Llopis (2010) consider a production function that allows for population and labor-augmenting technology growth.

decreases interest rates. The equivalence with a stochastic capital shock can also be seen directly by rewriting the production function (7) in the following equivalent form

$$Y_t = z_t (x_t K_t)^{\alpha} L_t^{1-\alpha} \tag{8}$$

where x_t is a random variable defined as follows

$$x_t \equiv \left(\frac{\mu_t}{l_t} \right)^{\zeta_t/\alpha} \tag{9}$$

where the timing of the shocks is such that the time-t capital-to-labor ratio (μ_t/l_t) is a predetermined variable when ζ_t is realized. Equation (8) shows that from the viewpoint of a non-representative agent, a stochastic capital stock has the same effects on the budget constraints as a stochastic labor share of output. Ríos-Rull and Santaeulalia-Llopis (2010) use US data to estimate a VAR in the Solow residual and the random variable ζ introduced above. They report the following estimates (standard errors below the coefficients)

$$\begin{array}{rcclcl} \ln z_t & = & 0.952 & \ln z_{t-1} & -0.004 & \zeta_{t-1} \\ & & (0.023) & & (0.043) & \\ \zeta_t & = & 0.050 & \ln z_{t-1} & +0.931 & \zeta_{t-1} \\ & & (0.011) & & (0.019) & \end{array} \tag{10}$$

and the variance-covariance matrix Σ of the residuals is reported to be

$$\Sigma = \begin{bmatrix} 0.00675^2 & -0.0001 \\ - & 0.00304^2 \end{bmatrix}$$

These estimates show that the two types of shocks have weak cross-effects on each other. The effect of ζ_{t-1} on the current value of the aggregate productivity shock is not statistically significant, while z_{t-1} has a significant impact on ζ_t. Since the magnitude of this effect is trivial compared to the autoregressive component, I consider these two shocks as (approximately) linearly independent. The variance-covariance matrix reported above shows that the normal range of variation of innovations to ζ is about 45% of that of innovations to $\ln z$. According to the parameterization of Tables 1 and 2 the (unconditional) standard deviation of $\ln z$ is 0.01. Using the definition of the random variable x in equation (9), and the unconditional averages reported in Table 3, turns out that the one-standard-deviation interval $(-0.0045, 0.0045)$ for ζ maps to an (average) one-standard-deviation confidence interval for x of $(0.9557, 1.0463)$. In other words, according to estimates reported by Ríos-Rull and Santaeulalia-Llopis (2010), on average the effect on wages and rental rates produced by a one-standard-deviation shock to the labor share of output, are equivalently produced by changes in the levels of capital of around $\pm 4.5\%$.

By using directly the stochastic labor share of output that corresponds to the shocks in expected capital, it is possible to give the core intuition analyzed in this paper most simply: when redistribution from capital to labor is *expected* –as opposed to it actually taking place– the households that rely on labor as their primary source of income expect a windfall which they start spending presently. Instead those households that earn the larger share of income from interest payments expect smaller disposable income, and therefore they cut their consumption. Again, inequality is the consequence of such expectations.