The Business Cycles Implications of Fluctuating Long Run Expectations

Dan Tortorice, Economics Department, Brandeis University

The Business Cycles Implications of Fluctuating Long Run Expectations *

Daniel L Tortorice[†]

January 20, 2016

Abstract

I consider a real-business cycle, DSGE model where consumption is a function of the present discounted value of wage and capital income. The agent is uncertain if these income variables are stationary or non-stationary and puts positive probability on both representations. The agent uses Bayesian learning to update his probability weights on each model and these weights vary over time according to how well each model fits the data. The model exhibits an improved fit to the data relative to the rational expectations benchmark. The model requires half the level of exogenous shocks to match the volatility of output and still matches the relative volatilities of key business cycle variables. The model lowers the contemporaneous correlation of consumption and wages with output and generates positive autocorrelation in model growth rates. Impulse responses exhibit persistent responses and consistent with survey evidence forecast errors are positively serially correlated. Finally, in contrast to the existing literature, the model endogenously generates observed time varying volatility and long run predictability of business cycle variables, especially for investment.

JEL Codes: E32; E22;D83

Keywords: Business Cycles; Investment; Learning

1 Introduction

As argued by Eusepi and Preston (2011), once one departs from full information, rational expectations assumptions, consumption in standard business cycle models depends not just on one period ahead expectations but on the full present discounted value of all future wage and capital income. Building on their insight, this paper notes that these key long-run forecasts are strikingly dependent on the agent's beliefs. Specifically, these forecasts are quiet different when the agent believes that wages and capital income will return to steady state versus when he believes that there is a unit root in the income process. If data convincingly distinguished between these two possibilities then the sensitivity of long run forecasts to a unit root would not be a fundamental concern. However, as noted by several authors (e.g. Cochrane (1988); Stock (1991)) it is very difficult to distinguish between unit root processes and near unit root processes in samples sizes common in macroeconomic time series.

Motivated by these two observations I construct a real business cycle, dynamic stochastic general equilibrium model with long run uncertainty about wage and capital income. Specifically the household believes that these variables follow univariate autoregressive processes but does not know the order of integration. Instead they put positive probability on both a stationary and a non-stationary model. The agent observes the wage and rental income data generated by the model and uses Bayesian learning to update her priors on the two models. Importantly, the agent's decisions affect the equilibrium values of wages and rental rates creating important feedback between the agents beliefs and the equilibrium model outcomes.[1] Over time, depending on the realizations of income, the agent's beliefs change putting a time varying weight on the stationary models. This learning mechanism substantially affects the model's business cycle implications.

The emphasis here on a business cycle model with volatile long run expectations has an eye to address some of the key failings of business cycle models. As noted by Kocherlakota (2010) among others, the shocks embedded in business cycles models are often clearly implausible or only vague reduced form representations of real economic disturbances. Accordingly, business cycle models often generate little endogenous volatility, simply inheriting the volatility of the exogenous shocks. Changing beliefs about the long run income path can serve as an important channel to amplify productivity shocks. This need is real for business cycle models as substantial work has shown these models lack internal propagation mechanisms (e.g. Rotemberg and Woodford (1996)) and are unable to explain the positive

---

[1]The model is self-referential then in the sense of Evans and Honkapohja (2001).

autocorrelation of business cycle variables (Cogley and Nason (1995)).

I find that allowing learning about the form of the wage and rental income processes greatly improves the fit of the model over a benchmark rational expectations model. The model generates twice as much volatility of output as the benchmark model while still maintaining the model's ability to match the relative volatility of the business cycle variables. The learning model, by allowing for an increased role of expectations to determine consumption, generates a lower contemporaneous correlation between consumption and wages with output consistent with the data. Importantly, the learning model improves the propagation of shocks. The impulse response functions of the key business cycle variables exhibit persistent responses showing that productivity shocks are propagated through the system. The model also generates positive autocorrelation in variable growth rates, which is absent from the rational expectations model but a clear feature of the data. Finally, the model generates positive autocorrelation in forecast errors consistent with survey expectations data.

The model also fits some less conventional statistics on business cycle variables. First, there is evidence of negative correlation in the medium run at annual frequencies. For example, investment growth this year is negatively correlated with investment growth over the next four years. This correlation is -0.37. This statistic is matched by the learning model but the rational expectations of the model falls short predicting a correlation of only -0.18. Secondly, there is clear evidence of time varying volatility in the growth rate of business cycle variables.[2] Consider the growth rate of consmption, investment and output and the residual from an AR(1) regression of these variables. Both these growth rates and these residuals exhibit positive autocorrelation in their squared values. Simply put, large movements in the growth rates of these variables are likely to be followed by additional large movements. This fact is matched by the learning model but not the rational expectations model. Finally, the learning model generates waves of pessimism and optimism resulting in booms and busts in output over and above the more mild fluctuations in the rational expectations benchmark.

This paper stands alongside a variety of literatures related to the RBC model. It follows the spirit of the many papers critiquing and proposing mechanisms for improving the fit of these models, see for example: Burnside and Eichenbaum (1996); Christiano and Eichenbaum (1992); Hansen (1985); Schmitt-Grohe (2000). The current paper differs in that it focuses on the role of expectations in improving the fit of the RBC model and allows for a departure from strict rational expectations by the inclusion of a learning mechanism. The current

---

[2]Time varying volatility in macroeconomic data has been noted by (among others): Stock and Watson (2003); Engle (1982); Primiceri (2005).

paper also considers a larger range of data moments including the autocorrelation of forecast errors, autocorrelation of squared growth rates, and the negative correlation of growth rates at longer horizons.

The model also relates to the literature on news shocks (e.g. Beaudry and Portier (2004); Jaimovich and Rebelo (2009); Schmitt-Grohe and Uribe (2012)) which seeks to explain business cycle fluctuations with news about future productivity. The current paper differs though in substantial ways. Firstly, these models examine a different channel for business cycle fluctuations. In these models, agents receive news of future productivity and react to it. In my model, agents are using a learning mechanism to determine how long current levels of productivity will persists. Secondly, these models are rational expectations models where in this paper expectations are formed through a learning mechanism. Consequently, long run expectations are endogenously formed based on the current level of productivity and are not reliant on exogenous disturbances. This distinction matters because it is still unclear what events lead agents to anticipate changes in future productivity. Departing from the rational expectations framework is also important because it allows the model in this paper to address the positive correlation of forecast errors observed in survey data. Finally, it is worth noting that even though anticipated shocks appear to contribute substantially to business cycle fluctuations, research (e.g. Schmitt-Grohe and Uribe (2012)) demonstrates that unanticipated fluctuations still contribute strongly to the volatility of business cycle variables and therefore research in how these shocks are propagated is essential.

There is also a large literature on the role which dispersed information can have in inducing expectations driven business cycles (Woodford (2003); La'O and Angeletos (2013); Lorenzoni (2009); Nimark (2014)). A key feature of these models are heterogeneous agents who hold differing beliefs based on limited availability of information. What is different about the current paper is that fluctuations are driven by a single agent who is learning over time while fully observing all equilibrium price outcomes in the economy. While not denying the promise of the limited information approach, given the wide availability of news and economic statistics the current exercise, trying to generate booms and busts in an economy where all price information is available, seems to still be an important and relevant challenge. Additionally, these papers are less focused on matching quantitative statistics while this paper aims to match conventional business cycle statistics, the autocorrelation of forecasts errors and growth rates and statistics on time varying volatility and longer horizon mean reversion. Finally, similar to the news shocks literature, these dispersed information models rely on shocks and signals to generate variation in agent's expectations, while the current

paper generates expectations endogenously from the observable model generated data.

A variety of papers have studied models of endogenous time varying volatility in macroeconomics ostensibly as a way to model the "Great Moderation". For example, Branch and Evans (2007) study a Lucas style monetary model where agents use a time varying set of predictor variables to forecast inflation. Lansing (2009) and Milani (2014) study variants of the canonical New-Keynesian model with time variation in the learning gain used to discount past observations. Both of these mechanisms lead to time varying volatility in inflation and output. Bullard and Singh (2012) examine a model where learning generates moderation in economic activity that comes from increased uncertainty. While all these papers connect learning and time varying volatility there are key difference between the current paper and these studies. These papers focus on low frequency movements in macroeconomic volatility, i.e. they are attempts to explain the "Great Moderation." They employ Euler equation learning which leads to sub-optimal decisions given agents beliefs and often leads to substantially different conclusions than when one solves for the true optimal policy (Eusepi and Preston (2011)). Additionally, this paper, unlike papers focusing on the New-Keynesian model, employs a RBC style model allowing me to examine time varying volatility in real economic variables like consumption and investment. Finally this paper aims to be more quantitative in that I match the autocorrelation of squared growth rates as a key quantitative measure of time varying volatility.[3]

My paper is closest to two papers in the literature. The first is Eusepi and Preston (2011) (EP). I use their model and stress the importance of long run expectations. I also see improvement over the benchmark rational expectations model along many of the same dimensions as their paper. But there are several important differences. In their model agents know the true process for wages and rental rates, they only do not know the coefficients on these variables. They use a recursive least squares algorithm with constant gain to learn about these coefficients. In this case the extent of long run uncertainty is minimized. Given the real debate about the long-run effect of disturbances in the macroeconometric literature it is a natural step to extend their work to incorporate this added uncertainty. This uncertainty is quantitatively important. The model here generates 100% the output volatility of the rational expectations benchmark. In the EP model, learning generates only 15% more volatility than the rational expectations benchmark. Importantly, the learning model generates this volatility with a relative lower elasticity of labor (3.7) versus the infinitely

---

[3]In related work, Tortorice (2014) finds that learning about the permanence of shocks is important for explaining low frequency movements in consumption volatility.

elastic preferences assumed in EP. The learning model also generates lower contemporaneous correlations between output and consumption and wages while they are almost perfectly correlated in the EP model. Finally, I show that the learning mechanism in this paper is able to capture additional features of the data, e.g. time varying volatility and long run predictability.

The second closest paper to mine is Kuang and Mitra (2015). They also build off the work of Eusepi and Preston (2011) but allow learning about the growth rate of the underlying income variables. The current paper departs from their paper in substantial ways. First, in my model long run growth rates are well anchored and equal to zero for rental rates and equal to the growth rate of productivity for wage rates. In their model agents can believe that rental rates will grow at a positive rate indefinitely and that wages will grow faster than productivity indefinitely. Put a different way, agents believe there is a unit root in the growth rate of efficiency wages and the growth rate of rental rates. Since the broader macroeconomic debate is about if the level of wages is trend or difference stationary and to the extent that there is a debate, if the level of rental rates are stationary of difference stationary it seems fruitful to explore the role of these beliefs as a complement to Kuang and Mitra (2015). Secondly, the current paper tries to explain a larger variety of business cycle data. The model is successful at explaining both time varying volatility in macroeconomic data and long run mean reversion of business cycle variables like investment.

In the remaining part of the paper I outline the model and discuss its calibration and simulation. I then list the key facts the model tries to explain and examine the ability of the model to explain these facts. Next, robustness to a variety of the parameter choices is examined. The last section concludes.

2 Model

2.1 Household

The model is a standard real business cycle model with shocks to technology. However, I use a continuum of firms and households of measure one to allow for the household's use of limited information in forming expectations. Households and firms are identical but they do not know this. Additionaly, the model follows Eusepi and Preston (2011) in solving for consumption in terms of the expected future discounted value of wage and capital rental

rates. There is a continuum of households indexed by i who maximize:

$$U^i = \hat{E}^i_t \sum_{T=t}^{\infty} \beta^{T-t} \frac{C_T^{i1-\sigma} v(1 - L_T^i)}{1 - \sigma} \tag{1}$$

where C_T^i is consumption at time T, L_T^i is leisure at time T defined as $L_T^i = 1 - H_T^i$ where H_T^i is hours worked. We assume that $\sigma > 1$ and that v' and $v'' > 0$. $\hat{E}^i_t$ represents the households expectations based on its subjective beliefs described in section 2.6. Preferences are of the form analyzed in King et al. (1988). With these preferences the marginal utility of consumption rises when hours worked rises. This assumption helps the model generate co-movement between consumption, hours and output when fluctuations are driven by expectations of future income.[4] The household maximizes utility subject to the following sequence of budget constraints:

$$K_{t+1}^i = (1 - \delta(u_t^i))K_t^i + R_t^k u_t^i K_t^i + W_t(1 - L_t^i) - C_t^i \tag{2}$$

here K_t^i is capital at time t, R_t^k is the capital rental rate at time t, W_t is the wage rate at time t and u_t^i is the utilization rate of capital at time t. The first order conditions for this maximization are:

$$C_t : C_t^{i1-\sigma} v(1 - L_t^i) = \Lambda_t^i \tag{3}$$

$$L_t : \frac{-C_t^{i1-\sigma} v'(1 - L_t^i)}{1 - \sigma} = \Lambda_t W_t \tag{4}$$

$$K_{t+1} : \Lambda_t^i = \beta \hat{E}^i_t [(1 - \delta + R_{t+1}^K u_{t+1}^i)\Lambda_{t+1}^i] \tag{5}$$

$$U_t : R_t^k = \delta'(u_t^i) \tag{6}$$

note that Λ_t^i is the Lagrange multiplier on the budget constraint.

2.2 Firms

There is also a continuum of firms of measure one and indexed by j that rents capital from the household and hires labor. The firms maximizes profits:

$$\Pi_t^j = Y_t^j - W_t H_t^j - R_t K_t^j$$

---

[4]Similar assumptions are made in the news shocks literature. See for example, Jaimovich and Rebelo (2009).

subject to the Cobb-Douglas production function:

$$Y_t^j = (u_t^j K_t^j)^\alpha (A_t H_t^j)^{1-\alpha} \tag{7}$$

The firm's first order conditions lead to the standard factor pricing equations:

$$R_t^k = \alpha \frac{Y_t^j}{u_t^j K_t^j} \tag{8}$$

$$W_t = (1 - \alpha) \frac{Y_t^j}{H_t^j} \tag{9}$$

2.3 Technology and Resource Constraints

Here technology is assumed to be stationary around a deterministic time trend so we can write:

$$\ln A_t = \ln A_0 + (1 + g)t + z_t \tag{10}$$

$$z_t = \rho z_{t-1} + \varepsilon_t \tag{11}$$

where g is the growth rate of technology, ε_t is $\sim$ i.i.d. $N(0, \sigma_\varepsilon^2)$, and $\rho < 1$ is the autoregressive parameter.

There are several motivations for assuming that technology is a trend stationary variable versus the perhaps more common random walk assumption. Firstly, it facilitates interpretation of the model. For example, if technology were a random walk then the model would be solved in terms of normalized variables like $w_t = W_t/\tilde{A}_t$ where $\tilde{A}_t$ is a random walk. If this variable is stationary there are still permanent shocks to the wage level and there is no tendency for wages to revert to any long run deterministic trend. And if this variable is non-stationary then it means that wages are de-linked from the level of productivity in the long run. It is difficult to imagine what structural change in the economy would generate this change.[5] On the other hand if $\tilde{A}_t$ is the productivity trend: $\ln \tilde{A}_t = \ln A_0 + (1 + g)t$, the interpretation is cleaner. If w_t is stationary, this means that wages will return to their level given by the balanced growth path. If tit is non-stationary, in the long run they will be at a level above or below the long run growth path.

This productivity setup also accords with the intuition for beliefs developed in section

---

[5]A change in the share of income going to labor might. However, in a model where households own the capital as well this is not a substantial distinction when forecasting future income.

2.7 and figure 1. Simply put, U.S. GDP data looks like it tends to return to a long-run trend level thought there are significant departures from trends and significant doubt as to the economy's ability to return to trend. In fact, this is a subject of much macroeconomic debate and chosing technology to be stationary allows the agent's uncertainty to be aligned with that of the macroeconometrician. Finally, allowing technology to be stationary generates the ability of the model to have agents overreact to temporary fluctuations in wages and rental rate. This type of overreaction is consistent with the "this time is different analysis" of Reinhart and Rogoff (2009) and the tendency of agents to justify temporary movement with new-era stories as described in Shiller (2005).

To close the model note that aggregate capital evolves according to:

$$K_{t+1} = (1 - \delta(u_t))K_t + I_t \tag{12}$$

where I_t is investment and the economy's resource constraint is:

$$Y_t = C_t + I_t \tag{13}$$

where the non-indexed, aggregate variables are obtained by summing over the continuum, e.g. $Y_t = \int y_j dj = y_j$.

2.4   Model Solution

I solve the model by transforming the variables to be stationary. I divide by the balanced growth path level of technology and then linearize equations: {(3), (4), (6), (7), (8), (9), (12) and (13)} about the non-stochastic steady state. The appendix contains the linearized equations. For the Euler equation (5) I follow Eusepi and Preston (2011) and iterate forward using the linearized budget constraint to solve for consumption as a function of only current variables and future expectations of rental rates and wages. This calculation leads to the following expression for aggregate consumption:

$$\hat{c}_t + \frac{1-\sigma}{\sigma}\psi\hat{h}_t = \frac{(1-\widetilde{\beta})(1-\chi)}{\varepsilon_c}\left[\frac{1}{\widetilde{\beta}}\hat{k}_t + \widetilde{R}\hat{r}_t^k + (\varepsilon_w + \varepsilon_c\frac{\chi}{1-\chi})\hat{w}_t\right]\ldots$$

$$+\frac{(1-\widetilde{\beta})(1-\chi)}{\varepsilon_c}(\varepsilon_w + \varepsilon_c\frac{\chi}{1-\chi})\widetilde{\beta}\hat{E}_t\sum_{T=t}^{\infty}\widetilde{\beta}^{T-t}\hat{w}_{T+1}$$

$$+ \left[\frac{(1-\widetilde{\beta})(1-\chi)}{\varepsilon_c} - \frac{\widetilde{\beta}}{\sigma} \right] \widetilde{\beta}\widetilde{R}\hat{E}_t \sum_{T=t}^{\infty} \widetilde{\beta}^{T-t}\hat{r}_{T+1}^k. \tag{14}$$

Here the hat notation on the variables denotes log deviation from steady state and the lower case letters represent the detrended variables. $\widetilde{\beta}, \widetilde{R}, \varepsilon_w, \varepsilon_c$, and χ are constants defined in the appendix. $\hat{E}_t = \int E_i di$ represents the expectation averaged across consumers.

Given this equation, consumption increases as hours work increases (recall that $\sigma > 1$) because of the non-seperability assumption for household preferences. It also increases in the current level of assets, $\hat{k}_t$, and income $\hat{r}_t^k$ and $\hat{w}_t$. Consumption responds positively to the present discounted value of future labor income given by the second to last term. Finally, consumption responds ambiguously to future rental income (the last term). There is both an income effect – after an increse in $\hat{r}_t^k$, the consumer is wealthier because he owns capital which is being paid a higher rental rate – and a substitution effect– he would like to save more to take advantage of higher future capital income. The overall effect of an increase in future rental income depends on the relative magnitude of the income and substitution effects.

2.5 Expectations and Learning

In standard rational expectations, real business cycle models the households know the exact model implied laws of motion for rental rates $\hat{r}_t$ and wages $\hat{w}_t$ along with the exact coefficients in this law of motion. Eusepi and Preston (2011) assume that households know the correct law of motion for these variables however they do not know the exact coefficients and learn about them over time.[6]

I depart even further from their assumption. First the limited information implies that while the agents do observe prices, i.e. rental rates and wages, they do not necessarily observe the aggregate supply of labor and capital. As a result they forecast prices using only the past values of these prices. Additionally, they do not necessarily observe the aggregate level of technology and do not know its true functional form. Instead, they consider the case where technology is trend stationary and use this to detrend wages. This leads the household to consider a stationary process for detrended wages and rental rates (in levels) as reasonable. However, the households are concerned that this assumption may be faulty and therefore allow for the possibility that the process for wages and rental rates may be non-stationary.

---

[6]In Eusepi and Preston (2011), these variables are solely a function of the capital stock.

Hence, the household believes that for $x_t = \{\hat{r}_t, \hat{w}_t\}$:

$$x_t = \rho_0^s + \rho_1^s x_{t-1} + ... + \rho_p^s x_{t-p} + \varepsilon_t^s \tag{15}$$

with probability p_t^s and that

$$\Delta x_t = \alpha + \rho_1^{ns} \Delta x_{t-1} + ... + \rho_p^{ns} \Delta x_{t-p} + \varepsilon_t^{ns} \tag{16}$$

with probability $p_t^{ns} = 1 - p_t^s.$

It is useful to recall that $\hat{w}_t$ is the log deviation from steady state w_t and that in steady state wages grow at the rate g. Therefore, if one believes that $\hat{w}_t$ follows the stationary process then one believes that wages will return to their steady-state, balanced growth path level in the long-run. However, if one believes that it follows the non-stationary process then one believes that in the long run wages will be above or below their steady state, balanced growth path level value forever. The analog beliefs for $\hat{r}_t$ is similar, except r_t is constant in steady state. Additionally, I require that the agent believes that $\rho_0^s = 0$ for all t and that $\alpha = 0$ for all t. This ensures that long run beliefs under the stationary model are given by the balanced growth path and that long-run *growth* expectations under the non-stationary model are also given by the balanced growth path. This is a sensible restriction on beliefs given basic economic theory and resource constraints. This restriction also improves the stability of the model.

Here the agent is assumed to use a univariate forecasting equation to forecast future labor and capital income. In addition to the limited information motivations discussed previously, there are two more motivations to use this forecasting rule in the benchmark model. The first is the work of Slobodyan and Wouters (2012) who show than in a medium scale DSGE model the use for univariate forecasting (in their model an AR(2)) greatly improved the fit of the model over the full rational expectations forecasting solution. Secondly, as argued by Fuster et al. (2012) there is much psychological evidence that when faced with complicated decisions problems individuals use simplifications (i.e. heuristic as in Kahneman and Tversky (1982) and Gabaix et al. (2006)) to make their decisions. A univariate foreasting rule would be one such heuristic.

It is worth noting that the non-stationary model (16) is an extrapolation model. A variety of authors have explored the presence of extrapolative agents in economc models, mostly in models of finance. See for example, Barberis et al. (2015). In the current model agents extrapolate, but only when recent data supports extrapolation. There are at least three

reasons this is preferable to just assuming extrapolation exogenously. The first is that the model provides a justification as to why extrapolaters do not realize they are wrong: namely the recent data supports their model. Secondly, in a market we would expect extrapolaters to influence outcomes more when extrapolation seems more reasonable, both because extrapolaters will become wealthier and individuals will switch to becoming extrapolaters. The current model captures this effect albeit in a reduced form when. Finally, endogenous extrapolation is what allows this model to explain time varying volatility.

To provide one final motivation for this model of long run uncertainty, I would like to contrast this approach with a few potential other approaches. The first would be a Markov switching process where the productivity process switches between (15) and (16). I do not take this approach because my experience with these models are that they do not generate significant variation in long-run beliefs. Either the transition probabilities are high, and the initial conditions do not matter much for where you end up in the long run, or the transition probabilities are low and you do not observe many transitions in the simulation samples. The second approach would be a model where productivity has both permanent and transitory shocks and the agent has uncertainty about this. This model could be solved with the Kalman filter for example. The shortcoming of this model is that individuals react the same way to the shock at each point in time, as if it was a linear combination of a permanent and transitory shocks with the weights being the relative variances of the two shocks. This model would not generate endogenous time varying volatility.[7] The third approach would combine the two previous approaches with a Markov switching model where in one state the economy is hit by permanent shocks and one state the economy is hit by temporary shocks. However, the imperfect information version of this model is intractable as one would need to have the whole history of shocks and time varying probabilities of all past states to make forecasts. I view my approach as trying to capture the dynamics of this last approach in a straightforward tractable way.

2.6   Beliefs

I use the methods of Cogley and Sargent (2005) to calculate the parameters of each model of rental rates and wages and the probability weights on the stationary and non-stationary

---

[7]This approach is taken by Edge et al. (2007) and Boz et al. (2011) in the context of emerging markets. The current papers differs by considering learning about endogenous variables versus exogenous productivity. This paper also focuses on explaining other features of the data for example, output volatility, autocorrealtion of variables and time varying volatiltiy.

model. Their model uses Bayesian methods to recursively update the parameters on each model and then uses the likelihood of each model to calculate a probability weight on each model. For a given model (i.e. the stationary or non-stationary) indexed by $i = \{s, ns\}$, and a rental or wage history Ξ^{t-1}, we assume that agents prior beliefs about the model parameters $\Theta_{i,t-1}$ are distributed normally according to:

$$p(\Theta_{i,t-1}|\sigma_i^2, \Xi^{t-1}) = N(\Theta_{i,t-1}, \sigma_i^2 P_{t-1}^{-1})$$

and their prior beliefs concerning the model residual variance are given by:

$$p(\sigma_{i,t-1}^2|\Xi^{t-1}) = IG(s_{t-1}, v_{t-1})$$

Here N represents the normal distribution function and IG represents the inverse-gamma distribution function. P_{t-1} is the precision matrix that captures the confidence the agent has in his belief for $\Theta_{i,t-1}$, σ_i^2 is the estimate of the variance of the model residuals, s_{t-1} is an analogue to the sum of squared residuals, and v_{t-1} is a measure of the degrees of freedom to calculate the residual variance such that the point estimate of $\sigma_{i,t-1}^2$ is given by s_{t-1}/v_{t-1}. After observing the rental rate or wage the agent's posterior beliefs are given by:

$$p(\Theta_{i,t}|\sigma_i^2, \Xi^t) = N(\Theta_{i,t}, \sigma_i^2 P_t^{-1})$$
$$p(\sigma_i^2|\Xi^t) = IG(s_t, v_t)$$

Cogley and Sargent (2005) gives the following recursion to update the parameters of the beliefs:

$$P_t = P_{t-1} + x_t x_t'$$
$$\theta_t = P_t^{-1}(P_{t-1}\theta_{t-1} + x_t y_t)$$
$$s_t = s_{t-1} + y_t^2 + \theta_{t-1}' P_{t-1}\theta_{t-1} - \theta_t' P_t \theta_t$$
$$v_t = v_{t-1} + 1$$

Here x_t is the vector of right hand side variables for the model at time t and y_t is the left hand side variable for the model at time t. This recursion gives the parameters of each model. Now it is necessary to calculate the probability weight on each model.

Given a set of model parameters: $\{\Theta_i, \sigma_i\}$ we can calculate the conditional likelihood of the model as:

$$L(\Theta_i, \sigma_i^2, \Xi^t) = \prod_{s=1}^{t} p(y_s|x_s, \Theta_i, \sigma_i^2)$$

where y_s and x_s are the left and right hand side variables of the model at time s and Ξ^t is the rental and wage income history up to time t. Based on this likelihood, one can write the marginalized likelihood of the model by integrating over all possible parameters:

$$m_{it} = \iint L(\Theta_i, \sigma_i^2, \Xi^t)p(\Theta_i, \sigma_i^2)d\Theta_i d\sigma_i^2$$

Then we have the probability of the model given the observed data $p(M_i|\Xi^t) \propto m_{i,t}p(M_i) \equiv w_{i,t}$. Here we have defined the weight on model i, $w_{i,t}$ and $p(M_i)$ is the prior probability on model i.

Cogley and Sargent (2005) show that Bayes's rule implies

$$m_{it} = \frac{L(\Theta_i, \sigma_i^2, \Xi^t)p(\Theta_i, \sigma_i^2)}{p(\Theta_i, \sigma_i^2|\Xi_t)}$$

and therefore

$$\frac{w_{i,t+1}}{w_{i,t}} = \frac{m_{i,t+1}}{m_{i,t}} = p(y_{i,t+1}|x_{i,t}, \Theta_i, \sigma_i^2)\frac{p(\Theta_i, \sigma_i^2|\Xi_t)}{p(\Theta_i, \sigma_i^2|\Xi_{t+1})}$$

We assume that regression residuals are normally distributed allowing us to use the normal p.d.f to calculate $p(y_{i,t+1}|x_{i,t}, \Theta_i, \sigma_i^2)$. Cogley and Sargent (2005) show that $p(\Theta_i, \sigma_i^2|\Xi_t)$ is given by the normal-inverse gamma distribution and provide the analytical expressions for this probability distribution. Any choice of Θ_i, σ_i^2 will give the same ratio of weights; I use the posterior mean in my calculations.

This recursion implies the following recursion for model weights.

$$\frac{w_{s,t+1}}{w_{ns,t+1}} = \frac{m_{s,t+1}/m_{s,t}}{m_{ns,t+1}/m_{ns,t}}\frac{w_{s,t}}{w_{ns,t}}$$

To allow for an emphasized role for learning, I adapt the concept of constant gain learning from the least squares learning literature to the current setup. I introduce a gain parameter (g) that over-weights current observations.[8]

$$\frac{w_{s,t+1}}{w_{ns,t+1}} = (1-g)\frac{m_{s,t+1}/m_{s,t}}{m_{ns,t+1}/m_{ns,t}}\frac{w_{s,t}}{w_{ns,t}} + g\frac{m_{s,t+1}/m_{s,t}}{m_{ns,t+1}/m_{ns,t}}$$

---

[8]I have explored allowing for constant gain learning in the estimation of the model parameters. I have found that this dimension for learning does not quantitaively affect the resuls in this paper and therefore I omit constant gain learning of parameters.

The gain probability can be interpreted as the probability of a structural break in the economy, such that the history of the income processes no longer has any bearing on the current process generating income, hence the previous weight ratio is set to one. Thus the gain serves to overemphasize more recent observations in calculating the likelihood of each model.

In addition to a desire for agents to guard against the possibility of a structural break in the economy. There is an additional behavioral interpretation of the gain. Much psychological evidence indicates that individual's probabilistic judgments are overly influenced by more recent observations. Tversky and Kahneman (1973) refer to this tendency as the availability bias. For example, after a friend has a heart attack, an individual thinks he himself is more likely to have heart attack. This bias is also related to Rabin (2002) who calls the tendency of individuals to incorrectly infer the nature of an underlying statistical process based on a recent, small sample the "law of small numbers". In the current model, the gain functions to overweight recent observations consistent with the psychological evidence that individuals tend to overweight the most readily accessible information.

Finally, to calculate the model probabilities, the consumer normalizes the weights to one, and therefore the weight on the stationary model is given by:

$$p_{s,t} = \frac{1}{1 + w_{ns,t}/w_{s,t}}$$

Using the estimated probabilities, he can then calculate the expectational terms in the consumption equation (14)[9]

$$E_t \sum_{T=t}^{\infty} \widetilde{\beta}^{T-t} \hat{x}_{T+1} = p_{s,t} \left[E_t \sum_{T=t}^{\infty} \widetilde{\beta}^{T-t} \hat{x}_{T+1} | S \right] + (1 - p_{s,t}) \left[E_t \sum_{T=t}^{\infty} \widetilde{\beta}^{T-t} \hat{x}_{T+1} | NS \right]. \quad (17)$$

To calculate these expectations note that we can write the AR processes in matrix form: $X_t^s = \Phi^s X_{t-1}^s + \varepsilon_t^s$ where $X_t^s = \{1, X_{t-1}, ..., X_{t-p}\}'$ and $X_t^{ns} = \Phi^{ns} X_{t-1}^{ns} + \varepsilon_t^{ns}$ where $X_t^{ns} = \{X_{t-1}, 1, \Delta X_{t-1}, ..., \Delta X_{t-p}\}'$. Then the first sum is equal to second element of

$$\left[I_{p+1} - \widetilde{\beta}\Phi^s \right]^{-1} X_t^s$$

---

[9]Importantly, I make the standard assumption in the learning literature of anticipated utility Kreps (1998). This assumption is that even though individuals beliefs change in the future they take these beliefs as given when forming expectations.

and the second sum is equal to the first element of

$$\left[I_{p+2} - \widetilde{\beta} \Phi^{ns} \right]^{-1} X_t^{ns}.$$

Therefore the expectations are linear function of r_t and w_t so we can solve the consumption equation (14) simultaneously with the linearized versions of the first order conditions and resource constraints: {(3), (4), (6), (7), (8), (9), (12) and (13)}. Finally I assume that model probabilities $p_{s,t}^r$ and $p_{s,t}^w$ and coefficients are updated at the end of the period after the realization of time t variables.

2.7   Belief Motivation

To understand the specification of beliefs and motivate why there might be uncertainty concerning the ability of the economy to return to previous trend growth, examine figure 1. This figure plots annual U.S. GDP (in logs) from 1929 to 2014. In addition, I plot the linear trend from a regression of log GDP on time. What one sees is that in general U.S. GDP is fairly close to the trend line and when it is above trend it tends to return to trend and when its below trend growth tends to accelerate to return to trend. Of course, as noted by Cochrane (1988); Stock (1991), this tendency does not diminish the possibility of a unit root in the GDP process. However, it does speak to the uncertainty regarding the long run level of GDP. Examining the current situation: will GDP return to trend as it has in the past or will the level of GDP be permanently lower? That is a real question looking at current data – and the question that agents in this model address.

3   Calibration and Simulation

3.1   Calibration

Time is measured in quarters. I calibrate the model by setting the discount rate $\beta = 0.99$. I set the capital depreciation rate $\delta = 0.025$. Capital's share in production $\alpha = 1/3$. The power utility coefficient $\sigma = 1.5$ equal to the value in Eusepi and Preston (2011). The appendix examines the robustness of the results to a value of $\sigma = 1.05$ near the seperable untility case of $\sigma = 1$. The quarterly growth rate of technology equals 0.0033 consistent with the

growth rate of total factor productivity from Basu et al. (2006).[10] I set $\varepsilon_v = \frac{v''(\bar{h})\bar{h}}{v'(h)} = 0.41$ to match the volatility of hours worked. This value implies an inverse Frisch elasticity of labor supply equal to 0.27.[11] For comparison, Eusepi and Preston (2011) assume infinitely elastic labor and Kuang and Mitra (2015) set labor supply elasticity to 0.1. For the productivity process I set the autoregressive parameter $\rho = 0.975$ and the standard deviation of technology shocks $\sigma_\varepsilon = 0.004$ to match the volatility of output. Robustness to various choices for ρ is demonstrated in the appendix.

For the learning parameters I begin with a prior on the stationary model $p_0^s = 0.75$ for the wage equation and 0.95 for the rental rate equation and consider four AR lags.[12] I set $\rho^{ns} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \end{bmatrix}'$. I set $\rho^S = \begin{bmatrix} 0 & 0.95 & 0 & 0 & 0 \end{bmatrix}$.[13] I take the gain parameter to be 0.03 and demonstrate robustness to this value in the appendix.[14] This level of the gain generates forecast error autocorrealtion that are on the lowerside of the observed autocorrelation in the data. This fact is notable because in Eusepi and Preston (2011) increasing their benchmark gain level quickly leads to counterfactually high levels of autocorrelation in forecast errors.[15] I set the initial sum of squared residuals $s_0 = 0.0056$ for the wage models and $s_0 = 0.0026$ for the rental rate model. These are the median value for the standard deviation of the regression residuals on the model generated wages and rental rates across the model simulations. Finally, I set the initial precision matrix, $P_0 = 0.01 * I$. This assumption is one of a fairly diffuse prior which implies a standard deviation for the initial coefficient estimate equal to 10 times the standard deviation of the regression residuals. However, I require $P_0(1,1) = 10^{10}$ so the agent dogmatically believes that the intercepts in their expectations models are equal to zero. This restriction has two important implications for beliefs. The first is that if the agent believes the stationary model is true then he believes

---

[10]TFP data and calculations are available at http://www.frbsf.org/economic-research/indicators-data/total-factor-productivity-tfp/.

[11]The inverse Frisch elasticity of labor supply equals $\varepsilon_v - \psi\frac{(\sigma-1)^2}{\sigma}$ where ψ is defined in the appendix.

[12]The agents puts a high initial weight on the rental rate process being stationary for two reasons. One interest rates appear strongly stationary in the data. So this restriction ensures that agents beliefs are not unreasonable given the data. Secondly, this restriction helps maintain model stability during the initialization period. Similarly, there is emphasis towards stationary wages to help with stability. I choose four lags because it is common in the macroeconomic literature see for example Christiano et al. (1999); Stock and Watson (2003, 2005) but the choice is unimportant.

[13]While it may be more natural to assume a prior equal to ρ. I find that a lower initial value of ρ increases the stability of the model during the initialization phase.

[14]While the setup in this paper with model learning makes direct comparisons difficult. Least squares learning and Kalman gains in the literature range from 0.002 to 0.05. See for example: Branch and Evans (2006); Eusepi and Preston (2011); Kuang and Mitra (2015); Milani (2014)

[15]See Table 4 in Eusepi and Preston (2011)

variables will return to their steady state, balanced growth path. Secondly, if he believes in the non-stationary model he does not believe that the rental rate will grow indefinitely or that the wage rates long run growth rate will differ from that of productivity. In this sense agents expectations are tempered by economic theory. Finally, note that assumptions on priors are not key to generate results since the model is simulated for 1500 periods keeping only the last 269 data points to match the length of my data.

3.2   Simulation

To deemphasize the importance of the priors, I simulate the model for 500 trials of length 1,500 keeping only the final 269 observations. To calculate impulse responses I again simulate 500 trials of 1,500 observations, then, given the conditions and beliefs after those 1,500 observations I calculate one series assuming technology receives a one standard deviation shock at time 1501 and no shocks subsequently and one series assuming technology receives no shocks after time 1500. I calculate the impulse response as the difference between these two series and plot the median impulse responses.

   To improve stability of the model I only use updated beliefs when they lead to a stable law of motion for the variables. Given the business cycle model, conditional on the agent's weights on the stationary models, we can write the evolution of the state variables $s_{t+1} = \{z_{t+1}, \vec{r}_t^s, \vec{w}_t^s, k_t, \vec{r}_t^{ns}, \vec{w}_t^{ns}\}$ as:

$$s_{t+1} = \Psi s_t.$$

Here $\{\vec{r}_t^s, \vec{w}_t^s, \vec{r}_t^{ns}, \vec{w}_t^{ns}\}$ are the vectors of right hand side variables in the equations specifing the agents beliefs (equations (15) and (16)). The condition for stability is that all the eigenvalues of Ψ are less than one in absolute value. If this condition is not satisfied then the agent uses the beliefs from the previous period. The appendix details the impact of this stability adjustment.

4   Results

4.1   Mechanism

To provide intuition for the model results I present a random path of productivity and the resulting paths for the beliefs and model variables. Here the model is only simlulated once. In the next subsection I will present median statistics based on 500 simulations.

18

Figure 3 plots the implied path of output given a random draw of productivity. The solid line plots the prediction of the learning model and the dashed line plots the path of output for the rational expectations model for the same productivity path. We can see several notable discrepancies between the learning model and the rational expectations model. First note that there are several gaps between output under the learning model and output under the rational expectations model. In general, output is substantially more volatile under the learning model than the rational expectations model. The largest gap between the two models occurs around time 150 when there is a large increase in output of about 12% followed by a crash where it falls 8%. In contrast, output under the rational expectations model shows only a mild increase. To understand why this boom and crash occurs examine figure 4. At the same time we observe the spike in output, there is a corresponding spike in the rental rate of capital. This movement results in a substantial change in beliefs as noted in figure 5. When the real rate of return on capital spikes up agents begin to think that the process for real interest rates may be non-stationary. Not only do they observe an increase in the rental rate of capital, they begin to think that this change is more likely to be permanent. This belief leads to a large spike in investment (figure 6) which is quickly reversed when rental rates return to their steady state value. Note that these beliefs are somewhat self-fulfilling. As the rental rate increases, the household wants to save more. They need to work more hours to increase investment and output. This change leads to an increase in the rental rate because the marginal product of capital is higher. This mechanism further amplifies the boom. The boom and bust cycle lasts about 4-5 years which indicates that the model generates output movement that look more like recessions and expansions than quick quarter to quarter reversals in output. It is also worth nothing that while the boom at time 150 is the largest, the simulation also indicates investment booms around time 40 and an investment bust around time 250 among others.

The agents in the model confuse temporary movements in rental rates with permanent movements in rental rates. This mechanism is consistent with a large amount of survey evidence on investors expectations. A large number of studies have concluded that investor expected returns rise when past stock returns have been high. These beliefs hold despite the fact the high PE-ratios tend to predict lower returns in the future. See for example: Fisher and Statman (2002); Shiller (2000); Greenwood and Shleifer (2013); Vissing-Jorgensen (2004). Similarly, in the housing market, Case and Shiller (2003); Piazzesi and Schneider (2009); Shiller (2007); Case et al. (2012) all find that expectations about future returns were increasing during the housing boom of the 2000s not declining. This observation is true even

for financial industry professionals as show by Foote et al. (2012). The model here seems then to be consistent with this survey evidence and can provide one explanation as to why investors may overestimate the persistence of above average returns.

Beliefs about future wage income also drive substantial movements in business cycle variables. Note in figure 7 the path of wages. At time 210 wage growth is quite negative, with wages falling to their lowest level in the simulation. Note also the gap between wages in the rational expectations model and the learning model. This decrease in wages changes the agents beliefs about the wage process shown in figure 8. They are now more likely to think the wage process is non-stationary and, in fact, think the non-stationary wage process is equally likely. This results in a noticeable decrease in consumption under the learning model relative to the rational expectations model shown in figure 9.

Finally, note that productivity (figure 10) is clearly related to all these movement. It reaches a high point at time 150 and a low point around time 210 when these large movements in beliefs occur. The beliefs channel is then able to propagate and magnify these fluctuations in productivity leading to large changes in business cycle variables.

4.2   Simulated Results

Figure 2 gives the impulse responses to a one standard deviation shock to productivity. The rational expectations model is the dashed line, the learning model is the solid line. The learning model improves on the rational expectations model along various dimensions. First given the same productivity shock (one standard deviation) the responses under learning are substantially larger. For example the response of output is almost twice as large and the response of investment is twice as large as well. Second we see that the internal propagation mechanism of the learning model is much stronger than in the rational expectations model. The rational expectations model sees the largest effect on output, investment and hours worked immediately and the effect subsequently declines monotonically. In contrast, the effect on these variables in the learning model is much more persistent than in the rational expectations model. This results is notable because several authors, (e.g. Cogley and Nason (1995); Rotemberg and Woodford (1996)) have noted the weak internal propagation mechanisms of the real business cycle model. It is also worth nothing that for several other variables, e.g. consumption, wages and the capital stock the peak effect comes seven to eight years after the initial shock showing a very strong ability of the learning model to propagate the technology shocks to these variables.

Table 1 gives a variety of business cycles statistics for the key variables and examines the ability of the two models to match these statistics. Data are standard. Output is real gross domestic product, consumption is consumption of services and non-durables deflated with the services and non-durables deflator respectively. Investment is real gross private domestic investment. All data come from the Bureau of Economic Analysis National Income and Product Accounts. Data are transformed into per-capita terms by dividing by the US population. Hours data are total hours for the total economy calculated from Bureau Labor Statistics Data by Valerie Ramey.[16] Data begin in 1948 and end in 2014. Interest rate data are given by the 3-month T-bill rate. Wage data are the average hourly earnings of production and non-supervisory employees: total private, dollars per hour, available from the Bureau of Labor Statistics and deflated with the consumer price index. These data begin in 1964.

By calibration the learning model exactly matches the volatility of log HP-filtered output 1.7%; in contrast, given the same volatility of the productivity shock, the rational expectations benchmark model generates only 0.8% volatility in output. The learning model generates twice the amount of volatility as the rational expectations model. This is a remarkable improvement. Additionally, as argued by Burnside (1996), productivity shock volatility is much smaller than output volatility. They argue that the variance of total factor productivity shocks is at most 10% the volatility of output growth. In my model, the variance of total factor productivity shocks, $[(1 - \alpha) * 0.004]^2$ is only 7% the variance of output growth, 0.01^2, addressing the Burnside et. al critique of real business cycle models.

Consumption is half as volatile as output while investment is 4.5 times as volatile. The learning model predicts a smaller relative volatility of consumption at 0.21 while the rational expectations model comes closer to matching the volatility of consumption with a value equal to 0.46. Perhaps making it more difficult for households to smooth income shocks would improve the fit of the learning model.

The rational expectations model understates the volatility of investment predicting 2.8. The learning model comes closer to matching the data with a value equal to 4. We find that hours are 0.95 times as volatile as output which is matched by the learning model by calibration. However, the rational expectations model with the same labor supply elasticity generates hours volatility equal to only 60% the *relative* volatility of hours. The fact that the learning model is able to match the volatility of hours in noteworthy. It does so with an inverse labor supply of elasticity equal to 0.27 while many papers in the literature

---

[16]Data are available at: http://econweb.ucsd.edu/~vramey/research.html#data

need infinitely elastic labor supply to match the data. Both models understate the relative volatility of labor productivity in the data. It is 0.49 in the data. The rational expectations model predicts 0.44 the learning model predicts 0.2.

Next the table examines contemporaneous correlations between the key variables. In the data, all variables are positively correlated however the correlation is often much less than one. The rational expectations model essentially predicts a correlation of one for all the variables. The learning model does not predict the perfect correlation that the rational expectations model does for all variables. The correlation of output and consumption, $\rho(c, y) = 0.47$. While it is much higher in the data – 0.81 – in the model it is no longer perfectly correlated with output. The learning model is also able to replicate the low correlation of wages with output. The rational expectations model predicts it should be 1, the learning model predicts it should be 0.34 consistent with the value of 0.32 in the data. The learning model is able to break this extreme correlation because in introduces another channel, beliefs which drive consumption and labor supply independently from output. Results for growth rates are consistent with the results for the HP-filtered variables.

Finally, the table examines autocorrealtions of variable growth rates. In the data many variables are positively autocorrelated, for example the growth rates of output, consumption hours worked, investment and wages. As Cogley and Nason (1995) pointed out the standard real business cycle model fails to explain this positive autocorrelation. It generally generates zero or negative autocorrelation in variables and when it generates some mild positive autocorrelation it is much less than in the data. For example $\rho(\Delta y)$= -0.02 in the RE model versus 0.37 in the data. In contrast, the learning model generates positive autocorrelation in all the main variables. While admittedly the autocorrelation is smaller than what we see in the data we can see that the learning model predicts 20% of the autocorrelation in output and 50% of the autocorrelation in investment.

To examine the statistical significance of these results I examine the 90th percentile for the distribution of the model statistics across the simulations. I find that for the rational expectations model the 90th percentile of the distribution ranges from 0.07 for output to 0.1 for wages. All data statistics are outside the 90th percentile. On the other hand the 90th percentile for the learning model ranges from 0.3 for output to 0.36 for wages. And the statistics for consumption, investment and the real interest rate are within the 90% confidence interval.

This is a remarkable result given that the learning model does not contain any frictions to slow the response of consumption or investment like habit formation or investment adjust-

ment costs. Instead, the self-fulfilling expectations drive this result. When interest rates rise, agents think this might be a permanent increase. They want to invest more and they work more hours. This response leads to an increase in the marginal product of capital further raising interest rates and propagating the boom. This mechanism can then generate auto-correlation in output, hours and investment. Of course, adding investment adjustment costs could improve the autocorrelation properties of the model. However, this channel would dampen the volatility of investment, worsening the predictions of the model along that dimension. What is key about the expectations channel is that it adds both autocorrelation and amplified volatility.

Table 2 reports forecast errors from the learning model. Because the household uses an incorrect model to forecast the real interest rate and the wage they will make errors. Using data from the Survey of Professional forecasters I calculate autocorrelation in the median forecasts for a variety of variables.[17]. There is substantial autocorrelation in professional forecasts of growth rates. One quarter ahead forecast errors for real GDP growth have an autocorrelation of 0.19 while they have an autocorrelation of 0.15 for nominal GDP. For the unemployment rate forecast errors have an autocorrelation of 0.58. Looking at 4q ahead forecast errors little is done to reduce the serial correlation for real and nominal GDP. However, the unemployment rate forecast errors have a serial correlation of 0.16 now.[18] The pattern for interest rates is similar with positive autocorrelations of 1-quarter forecast errors ranging from 0.26 for the three-month Tbill to 0.43 for the ex-post real rate.[19] Finally inflation expectations errors are highly autocorrelated. The one-quarter ahead forecast errors have an autocorrelation of 0.59 for inflation based on the GDP deflator and 0.17 for CPI inflation. The autocorrelation remains for four quarter ahead inflation expectations from the GDP deflator but disappears for CPI inflation.

I find no autocorrelation of forecast errors in the rational expectations model at 1 quarter. However, the learning model generates substantial autocorrelation in wage forecast errors, the 1q ahead forecast error has an autocorrelation of 0.12, while the autocorrelation in forecast errors for rental rates is 0.13. At the 4q ahead forecast horizon the models generate

---

[17]Data are available here: https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/

[18]Forecasts errors are calculated as: $\epsilon_t^1 = x_t - E_{t-1}x_t$ for 1-quarter ahead forecast errors and the table reports $\rho(\epsilon_t^1, \epsilon_{t-1}^1)$. For 4-quarter ahead forecasts we have $\epsilon_t^4 = x_t - E_{t-4}x_t$ for 1-quarter ahead forecast errors and the table reports $\rho(\epsilon_t^4, \epsilon_{t-4}^4)$.

[19]The ex-post real rate is calculated as $i_t - \pi_{t+1}$ where i_t is the three month Tbill rate and π_{t+1} is inflation from the GDP deflator. Survey expectations correspond exactly to this definition.

small mild correlations.[20] I also examine the 90th percentile for the forecast error correlation under the rational expectations model. For the 1q ahead errors this percentile is 0.09 for both wages and rental rates. The actual data statistics are outside this confidence interval. In contrast, for the learning model this percentile is 0.47 for real interest rate forecast errors and 0.48 for wage rate forecast errors and the data all fall within this range.

Table 3 examines the ability of the model to match the long run predictability of variables in the data and the autocorrelation of squared variables and residuals. The left hand side of the first panel of table 3 examines the correlation of the sum of the HP-filtered variables with sum of the variables over the next four years. That is to say, for consumption, $\rho(c_{t-3}^{HP} + ... + c_t^{HP}, c_{t+1}^{HP} + ... + c_{t+16}^{HP})$. I find that these correlations are negative for consumption, output and investment. Both models are able to match these negative correlations and there is little difference between the rational expectations model and the learning model, though the learning model is a bit better at explaining this statistic for investment.

Next, I look at the correlation of variable growth in one year with growth over the next four years. That is to say, for consumption, $\rho(\ln c_t - \ln c_{t-4}, \ln c_{t+16} - \ln c_{t+1})$. I find that this statistic is positively correlated for consumption but negatively correlated for output and investment. For example $\rho(\ln c_t - \ln c_{t-4}, \ln c_{t+16} - \ln c_{t+1})$=0.19 and the learning model predicts 0.09 versus -0.01 in the rational expectations benchmark. For investment this correlation is -0.36. The learning model predicts a correlation of -0.3 versus -0.16 in the rational expectations model. For output, this statistics equals = -0.09. Both the learning model and the rational expectations model overshoots this correlation. The rational expectations model predicts this correlation should equal -0.12 while the learning model predict -0.25. For this statistic both models emit wide confidence intervals. However, for the rational expectations model the consumption statistic is outside the 90th percentile with a value of 0.11 versus 0.19 in the data. The 10th percentile for the investment correlation is -0.36 which is the same as the data value. In contrast, the distribution of learning model predictions is centered more inline with the data and therefore all these statistics are within the learning model confidence intervals.

Next I examine the persistence of volatility in the data and the ability of the models to account for this fact. First I examine the autocorrelation of the squared variables and the squared residuals from an AR(1) regression of the variables. I examine $\rho\left(\left[c^{HP}\right]^2\right)$ and $\rho(\varepsilon_c^2)$ where ε_c is the residual from the AR(1) regression: $c_t^{hp} = \rho c_{t-1}^{hp} + \varepsilon_{c,t}$. I find that the squares of the HP-filtered variables are positively autocorrelated and both models are able

---

[20]Small sample bias with overlapping observations prevents this correlation from being exactly zero.

to account for this fact. The learning model generates larger autocorrealtions though, more consistent with the data. For example $\rho([c^{HP}]^2)$=0.74 versus 0.57 for the learning model. The rational expectations benchmark predicts this value to be only 0.5.

These discrepancies are much larger turning out attention to the autocorrelation of squared residuals. These correlations are all positive in the data indicating time varying volatility. Large (in magnitude) residuals are likely to be followed by on average large residuals. The rational expectations model can not explain this fact. It predicts these correlations to be essentially zero. However, the learning model consistently predicts these correlation to be positive. For example, the investment residual correlation equals 0.09, versus 0.3 in the data. Similarly, the learning model generates an autocorrelation of 0.03 for squared consumption residuals compared to 0.2 in the data; and it generates 0.08 for output residuals compared to 0.17 in the data.

The last section of table 3 repeats the previous analysis but using variable growth rates. There is clear evidence that squared growth rates are positively autocorrelated in the data. The rational expectations model can not match this fact. It predicts squared growth rates should be uncorrelated over time. However, the learning model predicts a correlation of 0.09 for squared consumption growth, positive like the value of 0.24 in the data. Similarly, for output and investment we see positive autocorrelation of squared growth rates at values of 0.23 and 0.24 respectively, a fact matched by the learning model though with smaller magnitude of 0.11 and 0.14. Similarly, squared AR(1) residuals of consumption, investment and output growth are all positive. Again, the learning model is able to generate positive autocorrelation while the rational expectations model cannot. Importantly, it almost exactly matches the autocorrelation of output growth residuals with a correlation of 0.12 versus 0.13 in the data and investment residuals with an prediction of 0.17 versus 0.16 in the data. When examining 90th percentiles of the model distributions the highest values at one lag for the rational expectations model is 0.1 while it is around 0.5 for the learning model. Hence the magnitude of the autocorrelation in squared residuals is within the distribution for the learning model but outside the 90th percentile for the rational expectations model.

To investigate if this time varying volatility is only a consequence of the Great Moderation I recalculate the time varying volatility statistics from table 3 for before and after 1983.[21] The results are in Table 5. I find that there is evidence of positive autocorrelation of squared growth rates both before and after 1983. If anything, the evidence suggests that time varying volatility is stronger post the Great Moderation.

---

[21]McConnell and Perez-Quiros (2000) estimate this year as the start of the Great Moderation.

Now of course the rational expectations model can match time varying volatility by assuming the *exogenous* disturbances have time varying volatility. I find this unsatisfying for a multitude of reasons. First, it is an additional assumption that must be added into the baseline model to help it match this one fact. While here the expectations channel improves the model on many dimensions. Secondly, a goal of this paper is to minimize the need to rely on exogenous disturbances to match the data, in trying to address the Kocherlakota (2010) critique that many shocks are difficult to match up with economic experience. Making these shocks not only large, but time varying in their volatility, seems, in this spirit, a step in the wrong direction. Thirdly, very few shocks are truly exogenous. Even productivity is presumable the outcome of a research and development process that depends on economic incentives. Therefore to the extent that we wish to understand time varying volatility it is important to endogenize it. This model gives one explanation for time varying volatility: that given their beliefs agents will respond differently to unexpected changes in economic conditions.

5   Robustness

Table 4 considers robustness of the results to varying some of the calibrated parameters. For these exercises I keep all the parameters at their benchmark values and vary one parameter. I then report a sampling of statistics used to evaluate the main model to examine robustness of the results to different parameter choices. I present only a sampling of statistics for clarity, however results are quite similar for the omitted statistics.

To consider robustness to the choice of the AR(1) parameter for the productivity process, I rerun the simulation setting $\rho = 0.95$. Results are given in column 2 of table 5. All in all, the results are quite similar to the benchmark case (whose values are presented for reference in column 1) . The model generates slightly less volatility, the standard deviation of output $\sigma(y) = 0.016$ versus 0.017 in the benchmark case. However the relative standard deviation of consumption and investment are the same as are the contemporaneous correlations of consumption and hours worked. The model still generates autocorrelation in forecast errors and autocorrelation in variable growth rates, though these correlations are slightly smaller, For example the autocorrelation of investment growth is 0.06 when $\rho = 0.95$ versus 0.1 in the benchmark case. The model still generates long run predictability in growth rates. For example, $\rho(\ln i_t - \ln i_{t-1} - 4, \ \ln i_{t+16} - \ln i_t) = -0.3$ which exactly matches the benchmark value. Finally, the model still is able to generate time varying volatility with almost no change

for the benchmark values. Similarly, when $\rho = 0.99$ the model generates the same volatility, $\sigma(y) = 0.017$ versus 0.017 in the benchmark case, and maintains the same relative volatilities and contemporaneous correlations. The model also generates slightly higher autocorrealtions for growth rates. For example output growth has an autocorrelation $\rho(\Delta y) = 0.08$ versus 0.07 as the benchmark value. The model also generates stronger autocorrelation of forecast errors. The autocorrelation of the forecast errors for wages now equals 0.15 versus a value of 0.13 for the benchmark case. Values for long run predictability are stronger, $\rho(\ln i_t - \ln i_{t-1} - 4, \ \ln i_{t+16} - \ln i_t) = -0.32$ versus the benchmark value of -0.3. Finally, the model again can generate time varying volatility, and the statistics almost exactly match the baseline case.

Next I consider varying the gain parameter. In the benchmark case $g = 0.03$. I consider increasing $g = 0.05$. The first effect of increasing the gain variable is to increase the volatility of output. $\sigma(y) = 2.2\%$ versus 1.7% when the gain variable equals 0.03 however it does not significantly impact the relative volatilities of consumption and investment. Increasing the gain however, lowers the contemporaneous correlation of consumption with output to 0.07 versus 0.47 but does not affect the contemporaneous correlations of hours with output. The model with a higher gain value exhibits larger autocorrelation of growth rates, for example investment growth rate autocorrelation equals 0.16 versus 0.1 in the benchmark case. The correlations representing long run predictability are all more negative. For example $\rho(\ln c_t - \ln c_{t-1} - 4, \ \ln c_{t+16} - \ln c_t) = -0.29$ versus 0.09 for the benchmark case and $\rho(\ln i_t - \ln i_{t-1} - 4, \ \ln i_{t+16} - \ln i_t) = -0.41$ versus -0.3 for the baseline case. Finally the model generates more time varying volatility that the baseline case does. For example, the autocorrelation of squared investment growth rate residuals equals 0.24 versus 0.17 in the benchmark case. Next I consider the model with learning but no gain so that $g = 0$. In this case volatility of output falls to 1.4%. The relative volatilities of consumption and output however show little change. Consumption now is slightly more correlated with output: $\rho(c, y) = 0.9$ versus 0.47 in the baseline case. One important distinction is that the model needs some gain to match the autocorrelation of growth rates. The model generates no positive autocorrelation of variable growth rates of output and investment when $g = 0$. The model still generates autocorrelation of forecast errors for wages, though the model is unable to generate considerable autocorrelation of forecast errors for rental rates. Forecast errors for the rental rate have an autocorrelation of 0.02 when $g = 0$ versus 0.13 in the benchmark case. The model also generates more positive autocorrelation of long run growth rates: $\rho(\ln c_t - \ln c_{t-1} - 4, \ \ln c_{t+16} - \ln c_t) = 0.15$ versus 0.09 for the benchmark case and $\rho(\ln i_t - \ln i_{t-1} - 4, \ \ln i_{t+16} - \ln i_t) = -0.21$ versus -0.3 for the benchmark case. Finally, the

gain is also important in generating time varying volatility. When $g = 0$ the model generates no time varying volatility. For example, the autocorrelation of squared residuals from the AR(1) regression on investment growth $\varepsilon_{\Delta i}^2 = 0.01$ when the gain variable equals zero versus 0.12 in the baseline case.

Next the robustness table considers the importance of model learning. I re-simulate the learning model but set $p_{s,t}^r$ and $p_{s,t}^w = 1$ for all t. Output volatility increases though the relative volatilities of investment and consumption show little change. Consumption becomes more correlated with output, similar to the case where the gain equals zero. As in that case the model no longer generates autocorrelation in output growth and investment growth, nor does it generates any endogenous time varying volatility. This robustness check highlights the importance of model learning, as opposed to parameter learning, in generating autocorrelation of growth rates and time varying volatility.

I also consider a version of the model close to the separable utility case ($\sigma = 1$) with a value of $\sigma = 1.05$. Volatility falls slightly to 1.4% and the relative volatility of consumption and investment rise. Consumption is now negatively correlated with output as the substitution effect of increased rental rates become stronger. For investment and output variables there is very little difference in the statistics. We still see autocorrelation of growth rates, long run predictability, and time varying volatility of quite similar magnitude. However, for consumption these statistics are amplified. For example, the autocorrelation in consumption growth rates is 0.32 versus 0.08 in the benchmark case. Time varying volatility for consumption growth rate residuals equals 0.33 versus 0.08 in the benchmark case.

Finally, I consider including information about the capital stock in the forecasting of wages and returns. Specifically, I include the variable k_t in the stationary factor price equations (15) and Δk_t in the non-stationary factor price equations (16). For comparison, Eusepi and Preston (2011) use the functional form $r_t = \alpha + \beta k_t + \varepsilon_t$. I allow this as a candidate model for the rental rate while also including lags of r_t. Additionally, since the agent must now forecast the future capital stock I assume that the agent believes that the capital stock is stationary, i.e. $k_t = \rho_0^s + \rho_1^s k_{t-1} + \varepsilon_t^s$ with probability p_t^{ks} and that it is non-stationary as, i.e. $\Delta k_t = \rho_0^{ns} + \rho_1^{ns} \Delta k_{t-1} + \varepsilon_t^{ns}$ with probability $1 - p_t^{ks}$. Agents learn in the same manner as they do about the wage and rental rate equations as described in section 2.6.

Results are in the column labeled VAR in table 4. There is a fall in the volatility of output to 1% and the relative volatility of consumption rises more in line with the data. The model generates quite similar contemporaneous correlations of consumption and hours

with output. The model no longer generates autocorrelation in growth rates for output but continues to generate positive autocorrelation in consumption and investment. The model continues to generate autocorrelation in forecast errors for rental rates however there is only a small autocorrelation in wage forecast errors. The model generates slightly more negative correlation of growth rates with long-run growth rates and the model's predictions for the autocorrelation for rental rate forecasts and time varying volatility are unchanged.

6   Conclusion

In this paper I considered a real business cycle model where consumption depends on the present discounted value of all future capital and wage income. To this model, I added long run uncertainty. The household is unsure about the stationarity of wage and capital income and puts some probability on the possibility that these variables are non-stationary. The household learns about the true model using Bayesian learning and therefore has time varying beliefs about the nature of the income processes.

I found that relative to a rational expectations benchmark, the model amplified the volatility of output and improved upon the model's prediction for the contemporaneous correlation of variables. The model exhibited persistent impulse responses and generated positive autocorrelation of variable growth rates. The model also generated positively autocorrelated forecast errors, consistent with evidence from the Survey of Professional Forecasters. Finally, the model also better fit some less conventional business cycle statistics. The model matched the medium frequency reversals in key variables like the growth rate in investment and the model generated time varying volatility consistent with the data.

References

Nicholas Barberis, Robin Greenwood, Lawrence Jin, and Andrei Shleifer. X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics*, 115(1):1–24, 2015.

Susanto Basu, John G. Fernald, and Miles S. Kimball. Are Technology Improvements Contractionary? *American Economic Review*, 96(5):1418–1448, December 2006.

Paul Beaudry and Franck Portier. An exploration into Pigou's theory of cycles. *Journal of Monetary Economics*, 51(6):1183–1216, September 2004.

Emine Boz, Christian Daude, and C. Bora Durdu. Emerging market business cycles: Learning about the trend. *Journal of Monetary Economics*, 58(6):616–631, 2011.

William Branch and George W. Evans. Model Uncertainty and Endogenous Volatility. *Review of Economic Dynamics*, 10(2):207–237, April 2007.

William A. Branch and George W. Evans. A simple recursive forecasting model. *Economics Letters*, 91(2):158–166, May 2006.

James Bullard and Aarti Singh. Learning And The Great Moderation. *International Economic Review*, 53(2):375–397, 05 2012.

Craig Burnside. Production function regressions, returns to scale, and externalities. *Journal of Monetary Economics*, 37(1):177–201, 1996.

Craig Burnside and Martin Eichenbaum. Factor-Hoarding and the Propagation of Business-Cycle Shocks. *American Economic Review*, 86(5):1154–74, December 1996.

Karl E. Case and Robert J. Shiller. Is There a Bubble in the Housing Market? *Brookings Papers on Economic Activity*, 34(2):299–362, 2003.

Karl E. Case, Robert J. Shiller, and Anne K. Thompson. What Have They Been Thinking? Homebuyer Behavior in Hot and Cold Markets. *Brookings Papers on Economic Activity*, 45(2 (Fall)):265–315, 2012.

Lawrence J Christiano and Martin Eichenbaum. Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations. *American Economic Review*, 82(3):430–50, June 1992.

Lawrence J. Christiano, Martin Eichenbaum, and Charles L. Evans. Monetary policy shocks: What have we learned and to what end? In J. B. Taylor and M. Woodford, editors, *Handbook of Macroeconomics*, volume 1 of *Handbook of Macroeconomics*, chapter 2, pages 65–148. Elsevier, 1999.

John H Cochrane. How big is the random walk in gnp? *Journal of Political Economy*, 96 (5):893–920, October 1988.

Timothy Cogley and James M Nason. Output Dynamics in Real-Business-Cycle Models. *American Economic Review*, 85(3):492–511, June 1995.

Timothy Cogley and Thomas J. Sargent. The conquest of us inflation: Learning and robustness to model uncertainty. *Review of Economic Dynamics*, 8(2):528–563, April 2005.

Rochelle M. Edge, Thomas Laubach, and John C. Williams. Learning and shifts in long-run productivity growth. *Journal of Monetary Economics*, 54(8):2421–2438, November 2007.

Robert F Engle. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1007, July 1982.

Stefano Eusepi and Bruce Preston. Expectations, learning, and business cycle fluctuations. *American Economic Review*, 101(6):2844–72, October 2011.

George Evans and Seppo Honkapohja. *Learning and Expectations in Macroeconomics*. Princeton University Press, 2001.

K.L. Fisher and M. Statman. Blowing Bubbles. *Journal of Psychology and Financial Markets*, 3:53–65, 2002.

Christopher L. Foote, Kristopher S. Gerardi, and Paul S. Willen. Why Did So Many People Make So Many Ex Post Bad Decisions? The Causes of the Foreclosure Crisis. NBER Working Papers 18082, National Bureau of Economic Research, Inc, May 2012.

Andreas Fuster, Benjamin Hebert, and David Laibson. Natural expectations, macroeconomic dynamics, and asset pricing. *NBER Macroeconomics Annual*, 26(1):1 – 48, 2012.

Xavier Gabaix, David Laibson, Guillermo Moloche, and Stephen Weinberg. Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model. *American Economic Review*, 96(4):1043–1068, September 2006.

Robin Greenwood and Andrei Shleifer. Expectations of Returns and Expected Returns. Nber working papers, National Bureau of Economic Research, Inc, January 2013.

Gary D. Hansen. Indivisible labor and the business cycle. *Journal of Monetary Economics*, 16(3):309–327, November 1985.

Nir Jaimovich and Sergio Rebelo. Can News about the Future Drive the Business Cycle? *American Economic Review*, 99(4):1097–1118, September 2009.

Daniel Kahneman and Amos Tversky. On the psychology of prediction. In Daniel Kahneman, Paul Slovic, and Amos Tversky, editors, *Judgment under uncertainty*, pages 48–68. Cambridge University Press, 1982.

Robert G. King, Charles I. Plosser, and Sergio T. Rebelo. Production, growth and business cycles : I. The basic neoclassical model. *Journal of Monetary Economics*, 21(2-3):195–232, 1988.

Narayana R. Kocherlakota. Modern macroeconomic models as tools for economic policy. *The Region*, (May):5–21, 2010.

David Kreps. *Anticipated Utility and Dynamic Choice, a 1997 Schwartz Lecture*. Frontiers of Research in Economic Theory, 1998.

Pei Kuang and Kaushik Mitra. Long-Run Growth Uncertainty. Discussion Papers 15-07, Department of Economics, University of Birmingham, 2015.

Kevin Lansing. Time Varying U.S. Inflation Dynamics and the New Keynesian Phillips Curve. *Review of Economic Dynamics*, 12(2):304–326, April 2009.

Jennifer La'O and George-Marios Angeletos. Sentiments. *Econometrica*, 81:739–779, June 2013.

Guido Lorenzoni. A Theory of Demand Shocks. *American Economic Review*, 99(5):2050–84, December 2009.

Margaret M. McConnell and Gabriel Perez-Quiros. Output fluctuations in the united states: What has changed since the early 1980's? *American Economic Review*, 90(5):1464–1476, December 2000.

Fabio Milani. Learning and time-varying macroeconomic volatility. *Journal of Economic Dynamics and Control*, 47(C):94–114, 2014.

Kristoffer P. Nimark. Man-Bites-Dog Business Cycles. *American Economic Review*, 104(8): 2320–67, August 2014.

Monika Piazzesi and Martin Schneider. Momentum Traders in the Housing Market: Survey Evidence and a Search Model. *American Economic Review*, 99(2):406–11, May 2009.

Giorgio E. Primiceri. Time Varying Structural Vector Autoregressions and Monetary Policy. *Review of Economic Studies*, 72(3):821–852, 2005.

Matthew Rabin. Inference By Believers In The Law Of Small Numbers. *The Quarterly Journal of Economics*, 117(3):775–816, August 2002.

Carmen Reinhart and Kenneth Rogoff. *This Time is Different: Eight Centuries of Financial Folly.* Princeton University Press, 2009.

Julio Rotemberg and Michael Woodford. Real business cycle models and the forcastable movements in output, hours and consumption. *The American Economic Review*, 86(1): 71–89, 1996.

Stephanie Schmitt-Grohe. Endogenous business cycles and the dynamics of output, hours, and consumption. *American Economic Review*, 90(3):1136–1159, 2000.

Stephanie Schmitt-Grohe and Martin Uribe. What's News in Business Cycles. *Econometrica*, 80:2733–2764, 2012.

R.J. Shiller. Measuring Bubble Expectatations and Investor Confidence. *Journal of Psychology and Financial Markets*, 1:49–60, 2000.

Robert Shiller. *Irrational Exuberence.* Princeton University Press, 2005.

Robert J. Shiller. Understanding recent trends in house prices and homeownership. *Proceedings - Economic Policy Symposium - Jackson Hole*, pages 89–123, 2007.

Sergey Slobodyan and Raf Wouters. Learning in a Medium-Scale DSGE Model with Expectations Based on Small Forecasting Models. *American Economic Journal: Macroeconomics*, 4(2):65–101, April 2012.

James H. Stock. Confidence intervals for the largest autoregressive root in u.s. macroeconomic time series. *Journal of Monetary Economics*, 28(3):435–459, December 1991.

James H. Stock and Mark W. Watson. Has the business cycle changed and why? In *NBER Macroeconomics Annual 2002, Volume 17*, NBER Chapters, pages 159–230. National Bureau of Economic Research, Inc, 2003.

James H. Stock and Mark W. Watson. Understanding Changes In International Business Cycle Dynamics. *Journal of the European Economic Association*, 3(5):968–1006, 09 2005.

Daniel L. Tortorice. Credit Constraints, Learning, And Aggregate Consumption Volatility. *Macroeconomic Dynamics*, 18(02):338–368, March 2014.

Amos Tversky and Daniel Kahneman. Availability: A Heuristic for Judging Frequency and Probabilities. *Cognitive Psychology*, 5:207–232, 1973.

Annette Vissing-Jorgensen. Perspectives on Behavioral Finance: Does Irrationality Disappear with Wealth? Evidence from Expectations and Actions. In *NBER Macroeconomics Annual 2003, Volume 18*, NBER Chapters, pages 139–208. National Bureau of Economic Research, Inc, 2004.

Michael Woodford. Imperfect Common Knowledge and the Effects of Monetary Policy. In Joseph Stiglitz Philippe Aghion, Roman Frydman and Michael Woodford, editors, *Knowledge, Information and Expectations in Modern Macroeconomics:In Honor of Edmund S. Phelps*, pages 25–58. Princeton University Press, 2003.

Figure 1: U.S GDP Relative to Trend
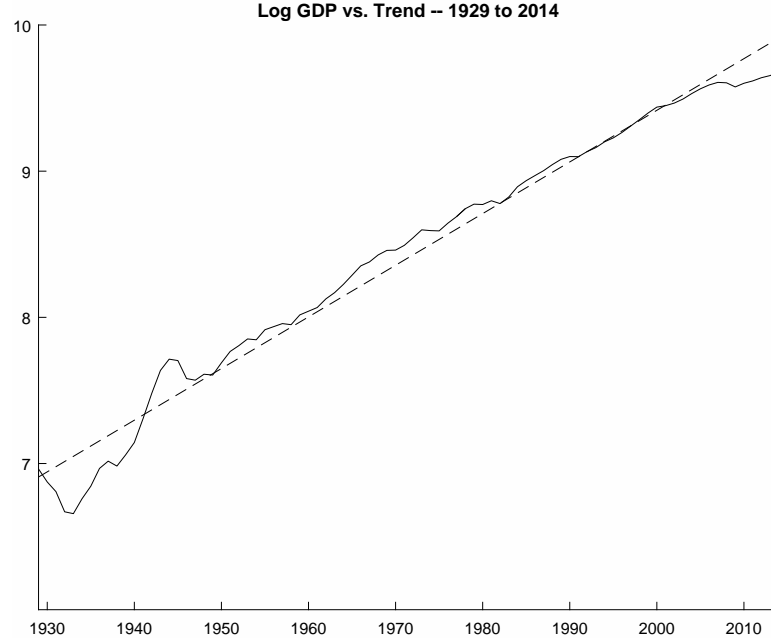

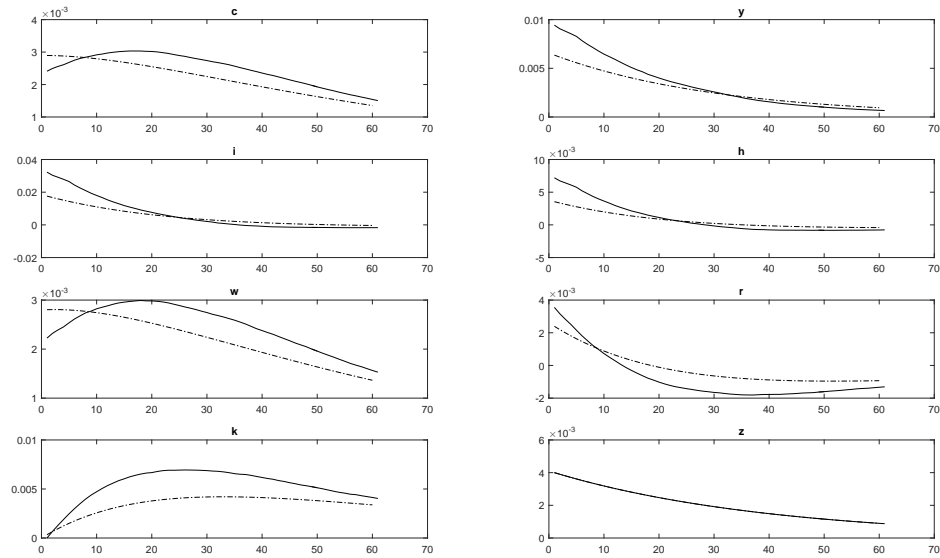**Log GDP vs. Trend -- 1929 to 2014**

Figure 2: Impulse Responses to One Standard Deviation Productivity Shock

Note: Learning model (solid), Rational Expectations Benchmark (Dashed)

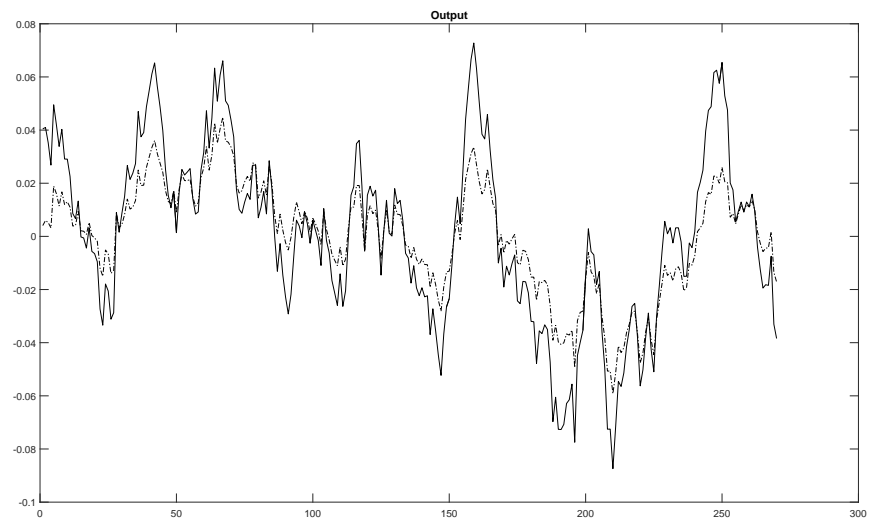Figure 3: Simulated Path of Output
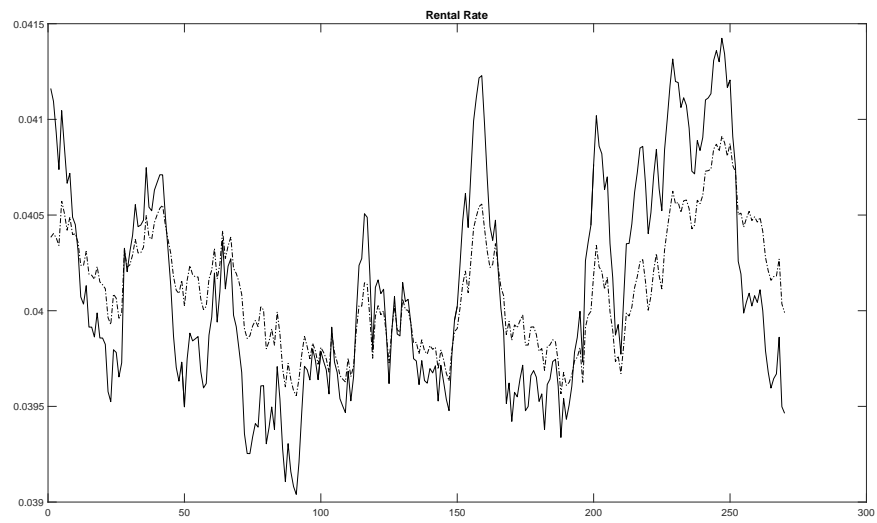

Figure 4: Rental Rate for Capital

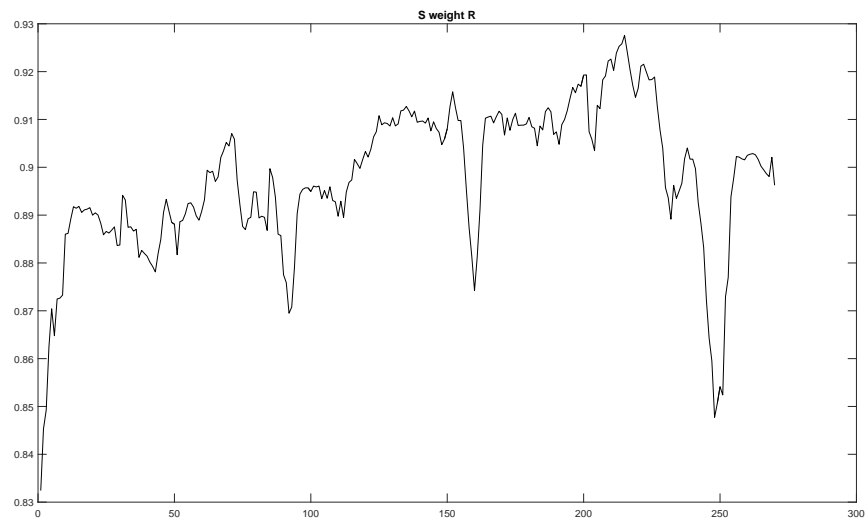Figure 5: Weight on Stationary Rental Rate Model
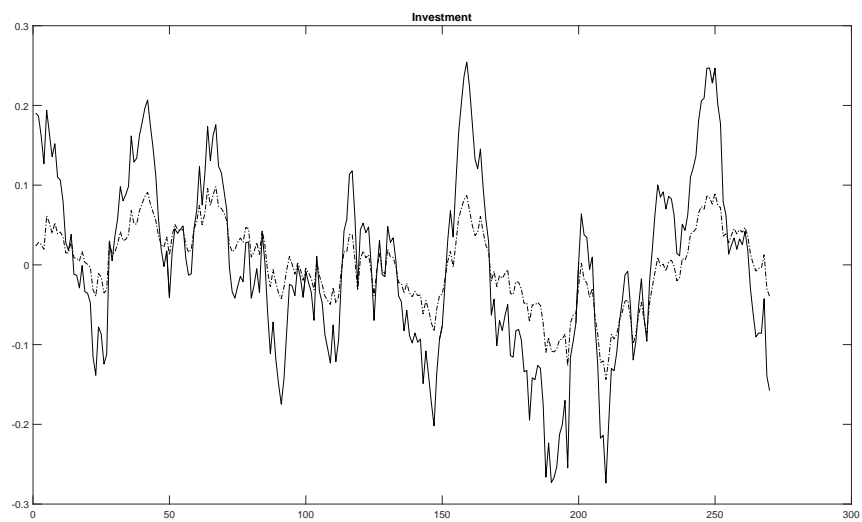

Figure 6: Simulated Path of Investment
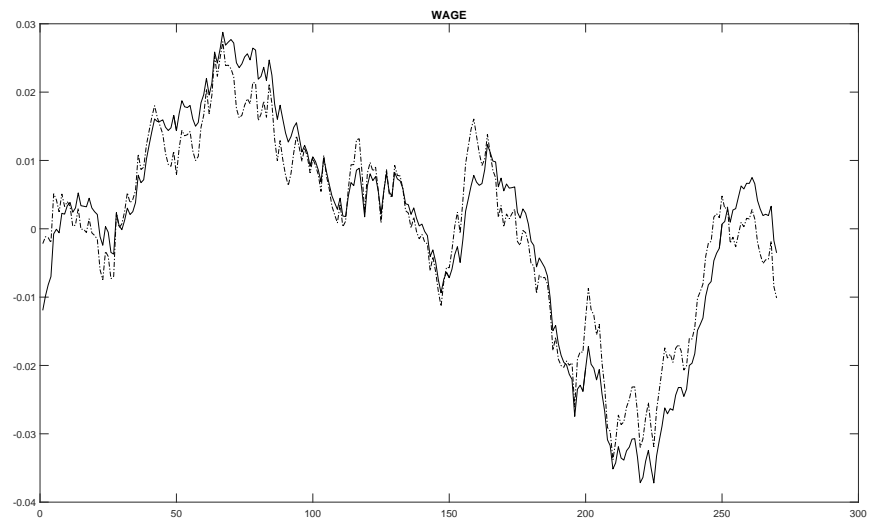
Figure 7: Simulated Path of Wages

Figure 8: Simulated Probability W follows stationary model
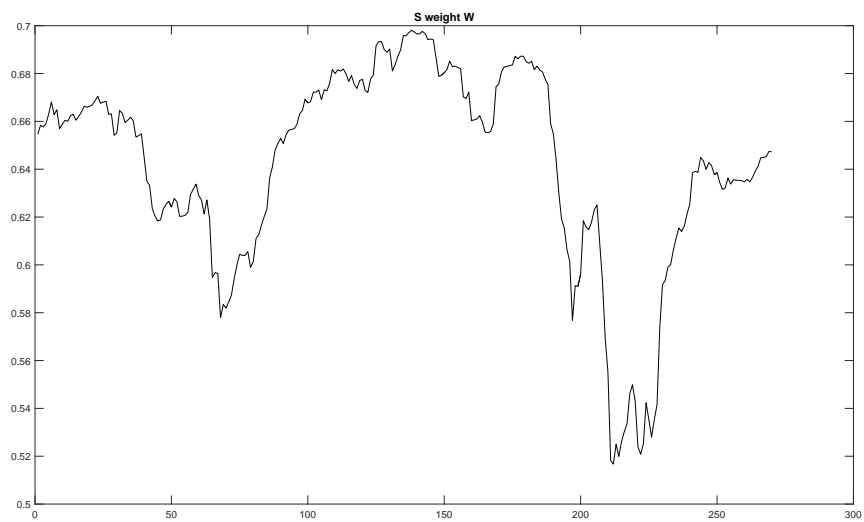
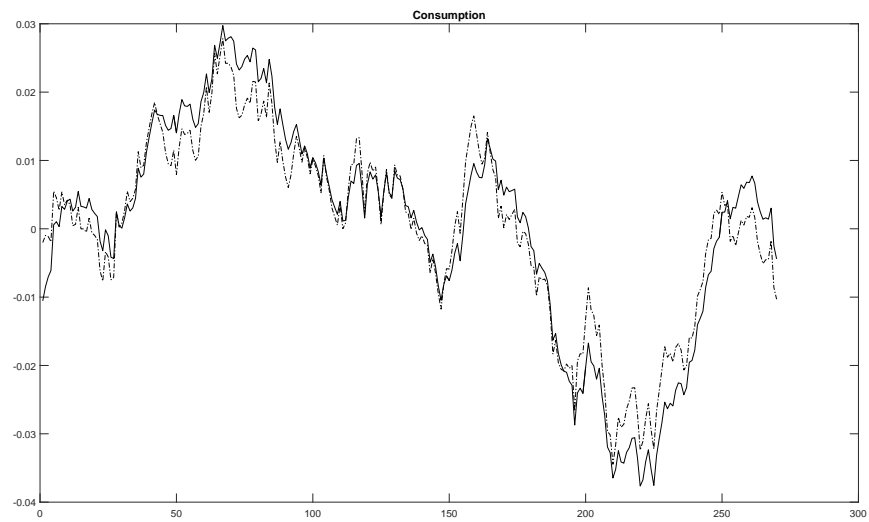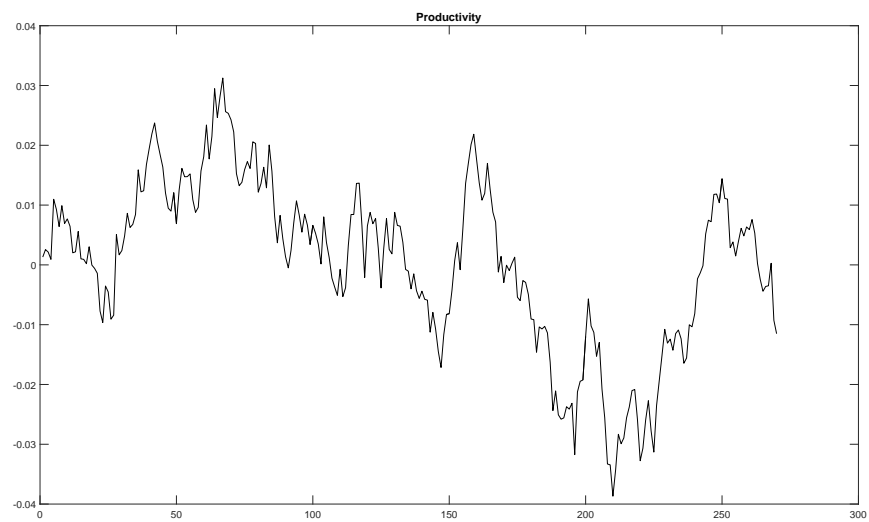Figure 9: Simulated Consumption

Figure 10: Simulated Path of Productivity

Appendices

A   Linearized Model

For this section I use the hat notation where $\hat{x}_t = \ln x_t - \ln x^{ss}$. I also use the lower case letter notation $x_t = X_t/Z_t$ where $Z_t = A_0(1+g)^t$ for C_t, Y_t, I_t, W_t and K_t, $\lambda_t = \Lambda_t/Z_t^{-\sigma}$, $r_t = R_t^k$ and $h_t = H_t$.

The consumption first order condition linearizes as:

$$-\sigma \hat{c}_t - \psi(1-\sigma)\hat{h}_t = \hat{\lambda}_t \tag{18}$$

where $\psi = \frac{v'(\bar{h})\bar{h}}{v(\bar{h})}\frac{1}{\sigma-1}$ and the bar denotes the steady state value.

The labor supply first order condition linearizes as:

$$(1-\sigma)\hat{c}_t + \varepsilon_v \hat{h}_t = \hat{\lambda}_t + \hat{w}_t \tag{19}$$

where $\varepsilon_v = \frac{v''(\bar{h})\bar{h}}{v'(\bar{h})}$. We assume the depreciation function $\delta(u_t) = \frac{1}{\theta}u_t^\theta$. which gives the linearized condition for the optimal choice of capital utilization

$$\frac{1}{\theta-1}\hat{r}_t^k = \hat{u}_t$$

The factor price equations linearizes as

$$\hat{r}_t^k = \hat{y}_t - \hat{u}_t - \hat{k}_t \tag{20}$$

$$\hat{w}_t = \hat{y}_t - \hat{h}_t \tag{21}$$

The production function linearizes as

$$\hat{y}_t = \alpha\hat{k}_t + \alpha\hat{u}_t + (1-\alpha)(\hat{h}_t + \hat{z}_t) \tag{22}$$

The capital evolution equation linearizes as

$$(1+g)\hat{k}_{t+1} = (1-\delta)\hat{k}_t + \frac{i}{k}\hat{i}_t - \theta\delta\hat{u}_t \tag{23}$$

here $\frac{i}{k}$ is the steady state investment capital ratio.

Finally the resource constraint linearizes as

$$\hat{y}_t = \frac{c}{y}\hat{c}_t + \frac{i}{y}\hat{i}_t \tag{24}$$

where $\frac{c}{y}$ and $\frac{i}{y}$ are the steady state ratios.

B   Steady State

To get the steady state return on capital I use the Euler equation (5) which gives:

$$R = \frac{(1+g)^\sigma}{\beta} - (1-\delta)$$

The factor price equation (8) yields:

$$\frac{y}{k} = \frac{R}{\alpha}$$

And the capital evolution equation (12) gives:

$$\frac{i}{k} = \delta + g$$

from the resource constraint

$$\frac{c}{k} = \frac{y}{k} - \frac{i}{k}$$

which can be used to calculate $\frac{c}{y}$ and $\frac{i}{y}$.

Finally combining the consumption (3) and labor supply (4) first order conditions gives:

$$\psi = \frac{wh}{c} = \frac{wh}{y}\frac{y}{c} = (1-\alpha)\frac{y}{c}$$

To determine θ we set steady state utilization to 1, i.e. $\bar{u} = 1$. The implication then is

$$
\begin{aligned}
r^k &= u^{\theta-1} \\
r^k u &= u^{\theta} \\
r^k &= \theta\delta \\
\theta &= \frac{r^k}{\delta}
\end{aligned}
$$

41

C Consumption Equation

The derivation of the consumption equation follows Eusepi and Preston (2011) allowing for stationary technology shocks. We have the following constants:

$$\widetilde{\beta} = \beta(1+g)^{1-\sigma}$$

$$\widetilde{R} = \frac{R}{(1+g)}$$

$$\varepsilon_c = \left[\frac{c}{k(1+g)} + \left[\varepsilon_h - \frac{\sigma-1}{\sigma}\psi\right]^{-1}\right]\widetilde{R}\frac{1-\alpha}{\alpha}$$

$$\varepsilon_h = \varepsilon_v - \frac{(\sigma-1)^2}{\sigma}\psi$$

$$\varepsilon_w = \left[1 + \left[\varepsilon_h - \frac{\sigma-1}{\sigma}\psi\right]^{-1}\right]\widetilde{R}\frac{1-\alpha}{\alpha}$$

$$\chi = \frac{(1-\sigma)\psi}{\sigma\varepsilon_h + (1-\sigma)\psi}$$

D Stability Statistics

In this section I provide some statistics on the stability adjustments made in the model as described in section 3.2. When the beliefs in the model would create an unstable law of motion, it is assumed that agents use the previous period's beliefs. Figure 11 shows the date of the last stability adjustment across the 500 simulations used to calculate median statistics. The bar represents the date of the last adjustment and the dashed line represents the first date used in the sample for calculating statistics. Simulations are sorted by the last date there is an adjustment. What this figure shows is that in more than half the simulations there is no stability adjustments in the sample period suggesting that median statistics are not influenced much by this adjustment.[22]

Next I look at the percent of time there is a stability adjustment in each of the 500 simulations of the model. Simulations are sorted as above. We can see that a simulation

---

[22]54% to be exact.

Figure 11: Date of Last Stability Adjustment
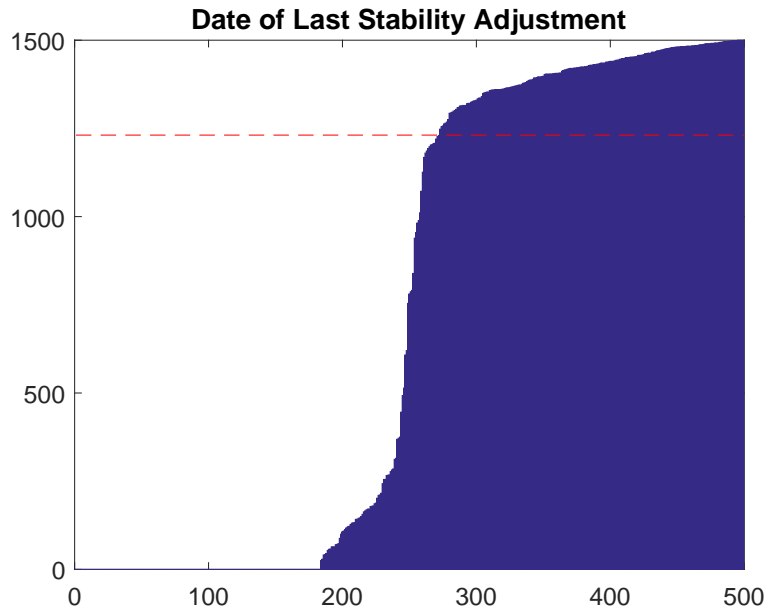
where adjustments happen more than 5% of the time is rare, occurring in about 2% of the sample. In only a handful of the simulations is a stability adjustment a common occurrence.

Finally, I look at the percent of times there is a stability adjustment in the sample that is used for calculating the simulation statistics for each of the 500 simulations. Here the influence of the stability adjustment is small, only about 3% of the sample has a stability adjustment more than 5% of the time in the sample used to calculate statistics.

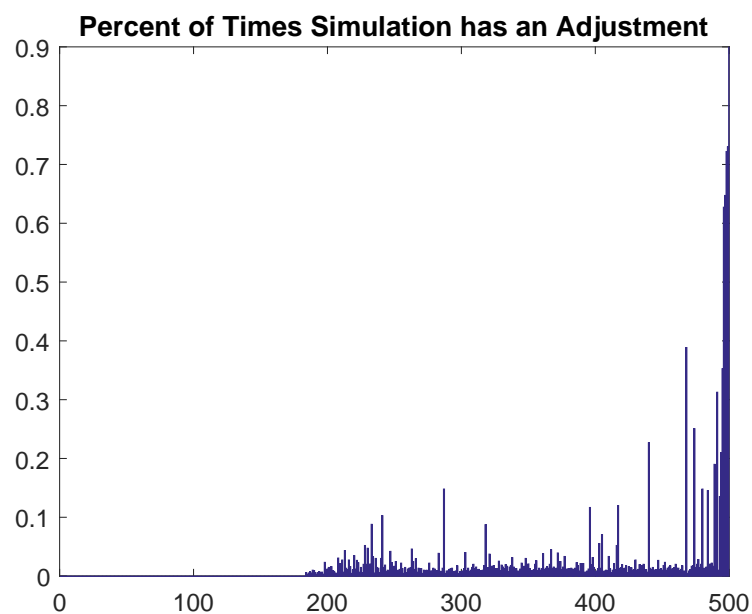Figure 12: Percent of Times Simulation Has Adjustment
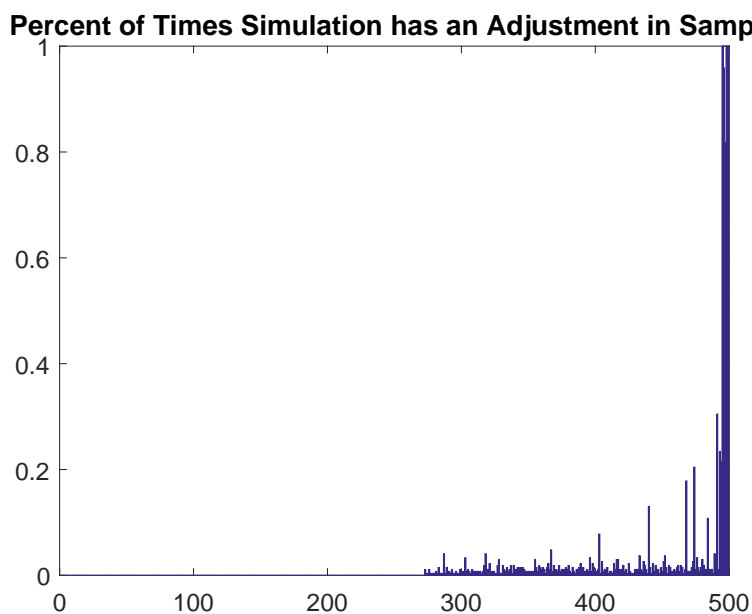

Percent of Times Simulation has an Adjustment

Figure 13: Percent of Times Simulation Has Adjustment in Sample

Percent of Times Simulation has an Adjustment in Sample

Table 1: Business Cycle Statistics

	Data	RE	Learning		Data	RE	Learning
Standard deviation of HP-filtered output				**Standard deviation of output growth**			
$\sigma(y)$	0.017	0.008	0.017	$\sigma(\Delta y)$	0.01	0.006	0.012
Relative volatility of HP-filtered variables				**Relative volatility of variable growth rates**			
$\sigma(c)/\sigma(y)$	0.5	0.46	0.21	$\sigma(\Delta c)/\sigma(\Delta y)$	0.52	0.45	0.21
$\sigma(i)/\sigma(y)$	4.5	2.8	4	$\sigma(\Delta i)/\sigma(\Delta y)$	5	2.8	4
$\sigma(h)/\sigma(y)$	0.95	0.56	0.95	$\sigma(\Delta h)/\sigma(\Delta y)$	0.81	0.56	0.94
$\sigma(y/h)/\sigma(y)$	0.49	0.44	0.2	$\sigma(\Delta y/h)/\sigma(\Delta y)$	0.72	0.44	0.2
Contemporaneous correlations of HP-filtered variables				**Contemporaneous correlations of variable growth rates**			
$\rho(c,y)$	0.81	0.99	0.47	$\rho(\Delta c,\Delta y)$	0.51	0.99	0.55
$\rho(h,y)$	0.87	0.99	0.99	$\rho(\Delta h,\Delta y)$	0.71	0.99	0.99
$\rho(i,y)$	0.87	0.99	0.99	$\rho(\Delta i,\Delta y)$	0.79	0.99	0.99
$\rho(r,y)$	0.4	0.98	0.96	$\rho(\Delta r,\Delta y)$	0.32	0.99	0.99
$\rho(w,y)$	0.32	0.99	0.34	$\rho(\Delta w,\Delta y)$	0.04	0.99	0.45

Autocorrelations of variable growth rates

	Data	RE	Learning		Data	RE	Learning		Data	RE	Learning
		Δy				Δc				Δh	
lag 1	0.37	-0.02	0.07		0.3	0	0.08		0.5	-0.03	0.11
lag 2	0.21	-0.02	-0.04		0.3	0	0.02		0.31	-0.03	-0.05
lag 3	0.02	-0.02	-0.04		0.22	0	0.03		0.12	-0.02	-0.06
		Δi				Δr				Δw	
lag 1	0.21	-0.03	0.1		0.24	-0.02	0.08		0.8	0	0.1
lag 2	0.14	-0.02	-0.05		-0.17	-0.02	-0.04		0.79	0.01	0.02
lag 3	-0.03	-0.02	-0.06		0.13	-0.02	-0.05		0.78	0	0.02

Note: This table gives the standard deviation of log HP-filtered output, the relative volatilities of log HP-filtered consumption, investment, hours worked, and labor productivity, and their contemporaneous correlation. The sames statistics are reported for the growth rates of these variables along with four lags of their autocorrealtions.

Table 2: Autocorrelation of Forecast Errors

Data

	Real GDP	Nominal GDP	Unemployment Rate
1Q	0.19	0.15	0.58
4Q	0.1	0.12	0.16

	Tbill	10-year Treasury	Corporate Bond	Real Rate
1Q	0.26	0.29	0.38	0.43
4Q	0.13	-0.31	-0.2	

	Inflation (GDP Deflator)	Inflation (CPI)
1Q	0.59	0.17
4Q	0.59	0.01

Model

	RE		Learning	
	r	w	r	w
1Q	-0.01	-0.01	0.13	0.12
4Q	0.06	-0.01	0.05	-0.03

Note: This table reports the autocorrelation of forecast errors for the return on capital and wage variables in the learning model.

Table 3: Long Run Predictability and Time Varying Volatility

Long Run Predictability	Data	RE	Learning		Data	RE	Learning
$\rho(c_{t-3}+..+c_t, c_{t+1}+…c_{t+16})$	-0.34	-0.33	-0.28	$\rho(\ln c_t - \ln c_{t-4}, \ln c_{t+16} -\ln c_t)$	0.19	-0.01	0.09
$\rho(y_{t-3}+..+y_t, y_{t+1}+…y_{t+16})$	-0.36	-0.34	-0.38	$\rho(\ln y_t - \ln y_{t-4}, \ln y_{t+16} -\ln y_t)$	-0.09	-0.12	-0.25
$\rho(i_{t-3}+..+i_t, i_{t+1}+…i_{t+16})$	-0.5	-0.35	-0.4	$\rho(\ln i_t - \ln i_{t-4}, \ln i_{t+16} -\ln i_t)$	-0.36	-0.16	-0.3
Squared Autocorrelations		c^2				$(\varepsilon_c)^2$	
lag 1	0.75	0.49	0.53	lag 1	0.24	-0.01	0.03
lag 2	0.45	0.2	0.27	lag 2	0.11	-0.01	0.06
lag 3	0.18	0.05	0.13	lag 3	0.08	-0.01	0.02
		y^2				$(\varepsilon_y)^2$	
lag 1	0.67	0.48	0.54	lag 1	0.17	-0.01	0.08
lag 2	0.35	0.18	0.2	lag 2	0.08	-0.01	0.12
lag 3	0.1	0.04	0.04	lag 3	0.2	-0.02	0.04
		i^2				$(\varepsilon_i)^2$	
lag 1	0.58	0.48	0.54	lag 1	0.3	-0.01	0.09
lag 2	0.32	0.18	0.2	lag 2	0.12	-0.01	0.14
lag 3	0.08	0.04	0.06	lag 3	0.32	-0.02	0.05
		$(\Delta c)^2$				$(\varepsilon_{\Delta c})^2$	
lag 1	0.24	-0.01	0.09	lag 1	0.4	-0.01	0.08
lag 2	0.26	-0.01	0.06	lag 2	0.21	-0.01	0.05
lag 3	0.17	-0.01	0.01	lag 3	0.13	-0.01	0.01
		$(\Delta y)^2$				$(\varepsilon_{\Delta y})^2$	
lag 1	0.23	-0.01	0.11	lag 1	0.13	-0.01	0.12
lag 2	0.14	-0.01	0.1	lag 2	0.13	-0.01	0.1
lag 3	0.06	-0.01	0.04	lag 3	0.03	-0.01	0.03
		$(\Delta i)^2$				$(\varepsilon_{\Delta i})^2$	
lag 1	0.24	-0.01	0.14	lag 1	0.16	-0.01	0.17
lag 2	0.18	-0.01	0.11	lag 2	0.22	-0.01	0.12
lag 3	0.31	-0.01	0.04	lag 3	0.22	-0.01	0.05

Note: This table contains the correlation of the sum four lags of the HP-filtered log variables with the next 16 HP-filtered log variables, and the correlation of the four quarter growth rate of the variables with growth over the next 16 quarters. It also containes the autocorrelation of the squared HP-filtered variables and growth rates and the autocorrelations of the squared residuals from an AR(1) regression on the HP-filtered variables and the growth rates.

Table 4: Robustness

	Benchmark	prod. ar(1)		gain		PS = 1	σ = 1.05	VAR
		0.95	0.99	0.05	0			
Volatility								
$\sigma(y)$	0.017	0.016	0.017	0.022	0.014	0.016	0.014	0.01
$\sigma(c)/\sigma(y)$	0.21	0.21	0.21	0.18	0.22	0.18	0.35	0.4
$\sigma(i)/\sigma(y)$	4	3.9	4	4.2	3.7	3.8	4.7	3.8
Correlations								
$\rho(c,y)$	0.47	0.55	0.47	0.07	0.9	0.88	-0.52	0.48
$\rho(h,y)$	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.93
Autocorrelations of growth rates								
Δy	0.07	0.04	0.08	0.13	-0.02	-0.02	0.07	-0.01
Δc	0.08	0.04	0.1	0.1	0.08	0.13	0.32	0.03
Δi	0.1	0.06	0.12	0.16	-0.03	-0.03	0.11	0.04
Autocorrelation of forecast errors								
r	0.13	0.1	0.15	0.16	0.02	0.01	0.16	0.11
w	0.12	0.09	0.14	0.13	0.12	0.19	0.07	0.02
Long Run Predictability								
$\rho(\ln c_t - \ln c_{t-4}, \ln c_{t+16} - \ln c_t)$	0.09	0	0.14	-0.29	0.15	0.23	-0.1	-0.07
$\rho(\ln y_t - \ln y_{t-4}, \ln y_{t+16} - \ln y_t)$	-0.25	-0.26	-0.26	-0.4	-0.18	-0.18	-0.26	-0.3
$\rho(\ln i_t - \ln i_{t-4}, \ln i_{t+16} - \ln i_t)$	-0.3	-0.3	-0.32	-0.41	-0.21	-0.21	-0.33	-0.41
Time Varying Volatility								
$(\varepsilon_c)^2$	0.03	0.04	0.03	0.07	0	-0.01	0.12	0.04
$(\varepsilon_y)^2$	0.08	0.08	0.08	0.1	0	-0.01	0.08	0.08
$(\varepsilon_i)^2$	0.09	0.09	0.08	0.1	0.01	-0.01	0.09	0.09
$(\varepsilon_{\Delta c})^2$	0.08	0.08	0.06	0.19	0	-0.02	0.33	0.04
$(\varepsilon_{\Delta y})^2$	0.12	0.12	0.1	0.19	0	-0.01	0.12	0.11
$(\varepsilon_{\Delta i})^2$	0.17	0.15	0.14	0.24	0.01	-0.01	0.19	0.15

Note: This table reports robustness of the main results to varying the model parameters.

Table 5: Autocorrelation of Squared Changes Before and After Great Moderation

Before			After		
	c^2	$(\varepsilon_c)^2$		c^2	$(\varepsilon_c)^2$
lag 1	0.72	0.18	lag 1	0.81	0.27
lag 2	0.37	0.04	lag 2	0.57	0.13
lag 3	0.1	0.02	lag 3	0.3	0.07
	y^2	$(\varepsilon_y)^2$		y^2	$(\varepsilon_y)^2$
lag 1	0.62	0.07	lag 1	0.76	0.34
lag 2	0.26	-0.02	lag 2	0.53	0.02
lag 3	-0.02	0.12	lag 3	0.31	0
	i^2	$(\varepsilon_i)^2$		i^2	$(\varepsilon_i)^2$
lag 1	0.51	0.23	lag 1	0.78	0.48
lag 2	0.27	0.05	lag 2	0.41	0.09
lag 3	0.02	0.3	lag 3	0.12	0
	$(\Delta c)^2$	$(\varepsilon_{\Delta c})^2$		$(\Delta c)^2$	$(\varepsilon_{\Delta c})^2$
lag 1	0.18	0.4	lag 1	0.22	0.17
lag 2	0.18	0.17	lag 2	0.34	0.17
lag 3	0.07	0.07	lag 3	0.28	0.21
	$(\Delta y)^2$	$(\varepsilon_{\Delta y})^2$		$(\Delta y)^2$	$(\varepsilon_{\Delta y})^2$
lag 1	0.15	0.03	lag 1	0.18	0.22
lag 2	0.06	0.05	lag 2	0	0.01
lag 3	-0.04	-0.08	lag 3	-0.06	0.1
	$(\Delta i)^2$	$(\varepsilon_{\Delta i})^2$		$(\Delta i)^2$	$(\varepsilon_{\Delta i})^2$
lag 1	0.17	0.08	lag 1	0.36	0.33
lag 2	0.12	0.18	lag 2	0.04	-0.02
lag 3	0.27	0.17	lag 3	0.13	0.11