Journal of Macroeconomics 58 (2018) 266–291

Contents lists available at ScienceDirect

Journal of Macroeconomics

journal homepage: www.elsevier.com/locate/jmacro

# The business cycle implications of fluctuating long run expectations☆

Daniel L Tortorice

*Department of Economics and Accounting, College of the Holy Cross, 1 College St. Worcester, MA 01610, USA*

ARTICLE INFO

*Keywords:*
Business cycles
Investment
Learning

*JEL classification:*
E32
E22
D83

ABSTRACT

I consider a DSGE model where consumption depends on the present discounted value of wage and capital income. The agent is uncertain if these 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 a no-learning benchmark. It requires half the standard deviation 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 growth rates. Impulse responses exhibit persistent responses and 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, without generating counterfactually high serial correlation of forecast errors.

## 1. Introduction

The last two recessions in the United States have been preceded by large swings in asset prices: equities in the 1990s and housing in the 2000s. Additionally, slow economic growth post recession has created more uncertainty about long run levels of economic activity for the US economy. These two observations have renewed interest in the role of expectations, and particularly long-run expectations, in fueling business cycles.

This exploration is important because as argued by Eusepi and Preston (2011), once one departs from full information, rational expectations assumptions, consumption in a standard business cycle model depends on subjective expectations of the present discounted value of all future wage and capital income.[1] 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 quite 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

☆ I thank Steve Cecchetti, George Hall, Blake LeBaron, Davide Pettenuzzo, three anonymous referees, the editor and seminar participants at Brandeis University, Boston College, College of the Holy Cross, Providence College, Bentley University, the Midwest Macroeconomics Meetings, the Annual Meeting of the Eastern Economics Association, the Annual Symposium of the Society for Nonlinear Dynamics and Econometrics, the 2018 Conference of Macroeconomists from Liberal Arts Colleges, and the 21st Annual Conference on Macroeconomic Analysis and International Finance for helpful comments. I thank the Tomberg and the Theodore and Jane Norman Fund at Brandeis for financial support. As always, any errors are mine.

[1] I thank an anonymous referree for suggesting this phrasing.

https://doi.org/10.1016/j.jmacro.2018.09.005
Received 7 December 2017; Received in revised form 25 September 2018; Accepted 30 September 2018
Available online 06 October 2018
0164-0704/ 

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.[2] Over time, depending on the realizations of income, the agent's beliefs change putting time varying weight on the stationary model. 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 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 as evidenced by a positive autocorrelation in variable growth rates.

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 which is matched by the learning model but not by the rational expectations model. Secondly, there is clear evidence of time varying volatility in the growth rate of business cycle variables.[3] This learning model generates time varying volatility but the rational expectations model does not.

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 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.

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. Similar models are considered by: Brock and Hommes (1997, 1998) and Evans and Ramey (1992). 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 differences 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 the Lansing and Milani papers which use a simple, three equation version of the New-Keynesian model where consumption equals output – employs an RBC style model allowing me to examine time varying volatility in a macroeconomic model where consumption does not equal output and investment is not necessarily zero. 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.[4]

Multiple papers have considered adding learning models to real business cycle models. One of the first is Williams (2003) who adds adaptive learning to the real business cycle model and finds a modest increase in volatility. Huang et al. (2009) add adaptive learning about the capital stock to a real business cycle model. Branch and McGough (2011) explore a model where agents forecast future returns using the past return. Both papers also find increased amplification.[5] The current paper differs along three important dimensions. First, it uses the Eusepi and Preston (2011) approach of forecasting only the variables exogenous to the agent's decision problem. This method allows one to solve for the true optimal policy conditional upon expectations. Secondly, this paper generates

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

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

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

[5] Branch and McGough (2011) is partially motivated by Kurz et al. (2005) who study the amplification effects of mistaken beliefs about the technology process.

more amplification and demonstrates improved fit of the real business cycle model along a larger number of dimensions, for example autocorrelation of growth rates, long run mean reversion, and time varying volatility. Finally, these above papers study simpler learning problems, for example omitting model learning (Williams and Huang et al.) or parameter learning (Branch and McGough).[6]

Learning models have also been added to business cycle models to generate persistence in macroeconomic outcomes, e.g: Milani (2007); Orphanides and Williams (2003). My paper is also an example of how adding in learning can create autocorrelation in macroeconomic outcomes. Learning has also been show to generate "animal spirits" – overoptimistic expectations about the future – by Grauwe (2011); Milani (2011) among others. This paper too generates booms and busts from overoptimistic expectations about the future but with a different learning mechanism. Finally, a more recent literature has begun to consider the role of model learning in economic fluctuations. Both Evans et al. (2012) and Cho and Kasa (2017) are important theoretical contributions in this area. Additionally, Grauwe (2012) considers a model where agents switch back and forth between fundamentalist and extrapolative forecasting rules and shows that these models generate amplification in a standard New-Keynesian model.

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. Given the real debate about the correct order of integration for macroeconomic variables in the applied time series 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% more output volatility than 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) versus the infinitely 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.

Because of the similarities of my paper to the EP paper I examine these differences in detail in 5.3. This section shows that the EP model is unable to replicate the results of this paper solely by increasing the gain parameter. Even with a high level of the gain, their model generates substantially less amplification, time varying volatility and reversals of investment growth. Additionally, increasing the gain results in counterfactually high autocorrelation of short-run forecast errors.[7]

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 sections 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 along with a detailed comparison to Eusepi and Preston (2011). The last section concludes.

## 2. Model

### 2.1. Household

The model is a standard real business cycle model with shocks to production and investment technology. As in Eusepi and Preston (2011), consumption is solved for in terms of the expected future discounted value of wage and capital rental rates. Additionally, there is a continuum of firms and households of measure one. Households and firms are identical but they do not know this which justifies the household's need to use limited information in forming expectations.

A continuum of households indexed by $i$ maximize:

[6] 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. In news shocks 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 returns to capital and wages will persist. There is also a large literature on the role which dispersed information can have in inducing expectations driven business cycles (Woodford (2003); Angeletos and La'O (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 agents who learn over time while fully observing all equilibrium price outcomes in the economy.

[7] This result is echoed in Table 4 of the EP paper.

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

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 household's 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.[8] The household maximizes utility subject to the following sequence of budget constraints:

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

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, $q_t$ is an investment specific technology shock and $u_t^i$ is the utilization rate of capital at time t. The first order conditions for this maximization are:

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

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

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

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

note that $\Lambda_t^i$ 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 households and hires labor. The firms maximize profits:

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

subject to the Cobb–Douglas production function:

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

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

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

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

### 2.3. Resource constraints and technology

Aggregate capital evolves according to:

$$K_{t+1} = (1 - \delta(u_t))K_t + q_t I_t \tag{10}$$

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

$$Y_t = C_t + I_t \tag{11}$$

where the non-indexed, aggregate variables are obtained by summing over the continuum, e.g. $Y_t = \int y_{j,t} dj = y_{j,t}$.

Finally, to close the model, technology is assumed to be stationary around a deterministic time trend so we can write:

$$\ln A_t = \ln A_0 + ln(1 + g)t + z_t \tag{12}$$

$$z_t = \rho z_{t-1} + \varepsilon_t \tag{13}$$

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

Similarly, the investment specific technology shock is also assumed to be stationary around a deterministic time trend:

$$\ln q_t = \ln q_0 + ln(1 + g_q)t + \eta_t \tag{14}$$

$$\eta_t = \rho^q \eta_{t-1} + \nu_t \tag{15}$$

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

where $g_q$ is the growth rate of the investment technology, $\nu_t$ is $\sim$ i.i.d. $N(0, \sigma_q^2)$, and $\rho^q < 1$ is the autoregressive parameter.

I choose a trend stationary productivity process versus the more common random walk assumption for two reasons. First, allowing technology to be stationary generates the ability of the model to have agents overreact to temporary changes in wages. 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 movements with new-era stories as described in Shiller (2005). Second, this productivity setup accords with the intuition for beliefs developed in section 2.7 and Fig. 1. Simply put, U.S. GDP data looks like it tends to return to a long-run trend level though there are significant departures from trends and significant doubt as to the economy's ability to return to trend. This observation leads me to consider a situation where there is a deterministic trend in GDP but allow the agent to have the same uncertainty of the macro-econometrician, i.e. they question if there is a deterministic or stochastic trend in income.

### 2.4. Model solution

I solve the model by transforming the variables to be stationary. I divide by the technology and the investment specific technology levels as described in the appendix and then linearize equations: {(3), (4), (6)–(11)} 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 derived in the appendix:

$$\begin{aligned}
\hat{c}_t + \frac{1-\sigma}{\sigma}\psi\hat{h}_t &= \frac{(1-\tilde{\beta})(1-\chi)}{\varepsilon_c}\left[\frac{1}{\tilde{\beta}}\hat{k}_t + \tilde{R}\hat{r}_t^k - \tilde{\beta}^{-1}\left(\frac{1}{1-\alpha}\hat{\gamma}_t^q + \hat{\gamma}_t\right) + \left(\varepsilon_w + \varepsilon_c\frac{\chi}{1-\chi}\right)\hat{w}_t\right] \\
&+ \frac{(1-\tilde{\beta})(1-\chi)}{\varepsilon_c}(\varepsilon_w + \varepsilon_c\frac{\chi}{1-\chi})\tilde{\beta}\hat{E}_t\sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\hat{w}_{T+1} \\
&+ \left[\frac{(1-\tilde{\beta})(1-\chi)}{\varepsilon_c} - \frac{\tilde{\beta}}{\sigma}\right]\tilde{\beta}\tilde{R}\hat{E}_t\sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\hat{r}_{T+1}^k. \\
&+ \left[\tilde{\beta} - \frac{(1-\tilde{\beta})(1-\chi)}{\varepsilon_c}\right]\hat{E}_t\sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\hat{\gamma}_{T+1} \\
&+ \left[-\frac{\tilde{\beta}\phi}{\sigma} - \frac{(1-\tilde{\beta})(1-\chi)}{\varepsilon_c(1-\alpha)}\right]\hat{E}_t\sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\hat{\gamma}^q{}_{T+1}
\end{aligned} \tag{16}$$

Here the hat notation on the variables denotes log deviation from steady state and the lower case letters represent the detrended variables. $\hat{\gamma}_t = ln(\frac{A_t}{A_{t-1}}) - ln(1+g)$ and $\hat{\gamma}_t^q = ln(\frac{q_t}{q_{t-1}}) - ln(1+g_q)$. $\tilde{\beta}, \tilde{R}, \varepsilon_w, \varepsilon_c, \phi$ and $\chi$ 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 worked increases (recall that $\sigma > 1$) because of the non-separability 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 term. Finally, consumption responds ambiguously to future rental income (the third term). There is both an income effect – after an increase 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 households have enough information to 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 these laws of motion. The required information to make this deduction includes knowledge that every agent in the model has identical preferences, technology and beliefs. An alternative to this approach is internal rationality (Adam and Marcet (2011)) where agents entertain subjective beliefs but make optimal decisions based on those beliefs. For example, Eusepi and Preston (2011) assume that households know the correct law of motion for these variables however they do not know that the productivity shock is the only disturbance in their model and they do not know the exact coefficients in the law of motion and learn about them over time. I depart even further from their assumption. I assume that the agent's subjective beliefs are given by univariate forecasting rules.

Furthermore, I follow the literature, (e.g. Eusepi and Preston (2011); Kuang and Mitra (2015)) and assume for simplicity that agents know the true model for the exogenous variables. This assumption allows me to focus on what is key for driving amplification – learning about endogenous variables. The agent then uses the univariate rules to forecast $\hat{W}_T = ln(W_T) - ln(\bar{W}_T)$ and $\hat{R}_T = ln(R_T) - ln(\bar{R}_T)$ where $\bar{W}_T$ and $\bar{R}_T$ are the balanced growth values of wages and the rental rate of capital (in levels, not

detrended[9]). However, the household does not know if the observed deviation from the balanced growth path is temporary or permanent (since the information assumptions do not enable them to map the realizations of these variables to the exogenous processes). This uncertainty leads the household to consider both a stationary process and a non-stationary process as possibilities.

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

$$x_t = \rho_0^s + \rho_1^s x_{t-1} + ...+\rho_p^s x_{t-p} + \varepsilon_t^s \tag{17}$$

with probability $p_t^s$ and that

$$\Delta x_t = \rho_0^{ns} + \rho_1^{ns} \Delta x_{t-1} + ...+\rho_p^{ns} \Delta x_{t-p} + \varepsilon_t^{ns} \tag{18}$$

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

It is useful to recall that $\hat{W}_t$ is the log deviation from balanced growth $W_t$ and that in steady state wages grow at the rate $g^w$. 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 belief for $\hat{R}_t$ is similar.

Additionally, I require that the agent believes that $\rho_0^s = 0$ for all $t$ and that $\rho_0^{ns} = 0$ for all $t$. This assumption 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, i.e. we would not expect wages to grow faster than productivity forever. In addition to being a sensible restriction based on economic theory, I find that not imposing this restriction results in too many unstable paths for the model to be accurately analyzed.

Here the agent is assumed to use a univariate forecasting equation to forecast future labor and capital income. There are two motivations to consider 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 of 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 decision problems individuals use simplifications (i.e. heuristics as in Kahneman and Tversky (1982) and Gabaix et al. (2006)) to make their decisions. A univariate forecasting rule would be one such heuristic.

However, there is a downside to the choice of univariate forecasting, specifically it does not nest rational expectations and in fact has many important changes from the baseline, rational expectations model. The three are: 1. the agent ignores capital in forecasting future variables, he has unit root beliefs and he is learning about the model probabilities. In Section 5.2, I consider alternate versions of the model which decompose the results into the effects of each of these departures. Additionally, there is no explicit link between beliefs for the rental rate process and beliefs for the wage rate process, as there would be if both were forecasted with the capital stock, and these beliefs therefore could potentially be de-linked over time.

It is worth noting that the non-stationary model (18) is an extrapolation model. There is an increasingly large literature based on survey evidence that indicates agents extrapolate past returns in the future. For example, Vissing-Jorgensen (2004) and Greenwood and Shleifer (2013) both present evidence that financial market participants expect higher returns when the PE ratio rises, even though statistically returns are expected to be low. This evidence indicates agents extrapolate past returns into the future and fail to see the future mean reversion in returns. Case et al. (2012a) and Case et al. (2012b) show similar evidence of extrapolation for returns on housing. This evidence motivates my model which allows agents to partially extrapolate past returns into the future. Additionally, while it may be natural based on theory to assume that interest rates are stationary, the empirical literature does not provide unequivocal support for this hypothesis, see Rose (1988).

Motivated by this empirical evidence, a variety of authors have explored the presence of extrapolative agents in economic models, mostly in models of financial markets and asset pricing. See for example, Barberis et al. (2015) or Adam et al. (2014).[10] However, there is an important distinction between the way these models approach extrapolation and the approach in this paper. With 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 extrapolators do not realize they are wrong: namely the recent data supports their model. Secondly, in a market we would expect extrapolators to influence outcomes more when extrapolation forecast models better fit the recent data. This is for two reasons. First, extrapolators will become wealthier when their forecasting models fit the data better and therefore will have a larger impact on equilibrium prices. Secondly, as the extrapolation model forecasts better more individuals will switch to extrapolation forecasts over mean reverting or fundamentals based forecasts. The current model captures these two effects, albeit in a reduced form way. Finally, endogenous extrapolation is what allows this model to explain time varying volatility.

[9] Since the detrended variables are always stationary it is sensible to learn about the level variables which can be non-stationary in the presence of permanent shocks. Also, since the household knows the true process for the exogenous variables he can easily compute forecasts of the detrended variables from beliefs about the level variables.

[10] For an important related contribution see Winkler (2016). He uses a learning model to amplify asset prices and therefore financial frictions in a business cycle model. While related, the current paper differs because here agents learn about real wages and returns on capital in an otherwise frictionless model. Hence, the current paper achieves amplification solely through expectations. Additionally, the emphasis on time-varying volatility and long run predictability of business cycles variables is novel.

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 (17) and (18). 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.[11] 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 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 $\Xi^{t-1}$, we assume that agents prior beliefs about the model parameters $\Theta_{i,\ t}$ are distributed normally according to:

$$p(\Theta_{i,t}|\sigma^2_{i,t-1}, \Xi^{t-1}) = N(\Theta_{i,t-1}, \sigma^2_{i,t-1}P^{-1}_{t-1})$$

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

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

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 $\Theta_{i,t-1}$ , $\sigma^2_{i,t-1}$ is the estimate of the variance of the model residuals, $s_{t-1}$, the scale parameter, is an analogue to the sum of squared residuals, and $v_{t-1}$ ,the shape parameter, is a measure of the degrees of freedom to calculate the residual variance such that the point estimate of $\sigma^2_{i,t}$ is given by $s_t/v_t$. After observing the rental rate or wage the agent's posterior beliefs are given by:

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

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

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

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 $\Xi^t$ 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:

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

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

[11] This approach is taken by Edge et al. (2007), Alpanda (2017) 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, autocorrelation of variables and time varying volatility.

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

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 $\Theta_i$, $\sigma_i^2$ will give the same ratio of weights; I use the posterior mean in my calculations.

This recursion implies the following recursion for model weights.

$$\frac{w_{s,t+1}}{w_{ns,t+1}} = \frac{m_{s,t+1}/m_{s,t}}{m_{ns,t+1}/m_{ns,t}}\frac{w_{s,t}}{w_{ns,t}} \tag{19}$$

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

$$p_{s,t} = \frac{1}{1 + w_{ns,t}/w_{s,t}} \tag{20}$$

I found that in this form the learning model puts a weight of one on the stationary return process in the long run. This result eliminates long-run learning about the process for returns and severely limits the ability of the model to generate autocorrelation in growth rates, time varying volatility and long run reversals in investment growth. Therefore, I adopt the concept of constant gain learning from the least squares learning literature, see Evans and Honkapohja (2001), to the setup here to allow for perpetual learning.[12] I introduce a gain parameter ($g$) that over-weights current observations.[13] The gain probability can be interpreted as an exogenously given belief in the probability of a structural break in the economy, such that the history of the wage (rental rate) process no longer has any bearing on the current process generating wages (rental rates), 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.

With probability $1 - g$ there is no structural break and the probability is given by Eq. (20) with the weights given by Eq. (19) and with probability $g$ there is a structural break and the probability is given by Eq. (20) with the weights given by Eq. (19) but $\frac{w_{s,t}}{w_{ns,t}}$ is set to 1.[14] Therefore the model probability is given by:

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

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.[15] For example, after a friend has a heart attack, an individual thinks he himself is more likely to have a 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.

Using the estimated probabilities, the consumer can now calculate the expectational terms in the consumption Eq. (16)[16]

$$E_t\sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\hat{x}_{T+1} = p_{s,t}\left[E_t\sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\hat{x}_{T+1}|S\right] + (1 - p_{s,t})\left[E_t\sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\hat{x}_{T+1}|NS\right]. \tag{21}$$

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:

$$[I_{p+1} - \tilde{\beta}\Phi^s]^{-1}X_t^s \tag{22}$$

[12] To examine the impact of this assumption, Section 5 contains an analysis of the learning model without any gain mechanism to create perpetual learning.

[13] 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 quantitatively affect the results in this paper and therefore I omit constant gain learning of parameters.

[14] Here agents treat the probability of a structural break as exogenous. Endogenizing this parameter would most likely amplify the results in this paper because agents would use a higher gain when wages or rental rates are unusually high or low. These times are precisely when agents are more likely to believe these processes are non-stationary.

[15] Bordalo et al. (2016) explore the implications of this bias for credit cycles.

[16] 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:

$$[I_{p+2} - \tilde{\beta}\Phi^{ns}]^{-1} X_t^{ns}. \tag{23}$$

Therefore the expectations are linear functions of $\hat{R}_t$ and $\hat{W}_t$ so we can solve the consumption Eq. (16) simultaneously with the linearized versions of the first order conditions and resource constraints: {(3), (4), (6)–(11)}. 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 Fig. 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.[17]

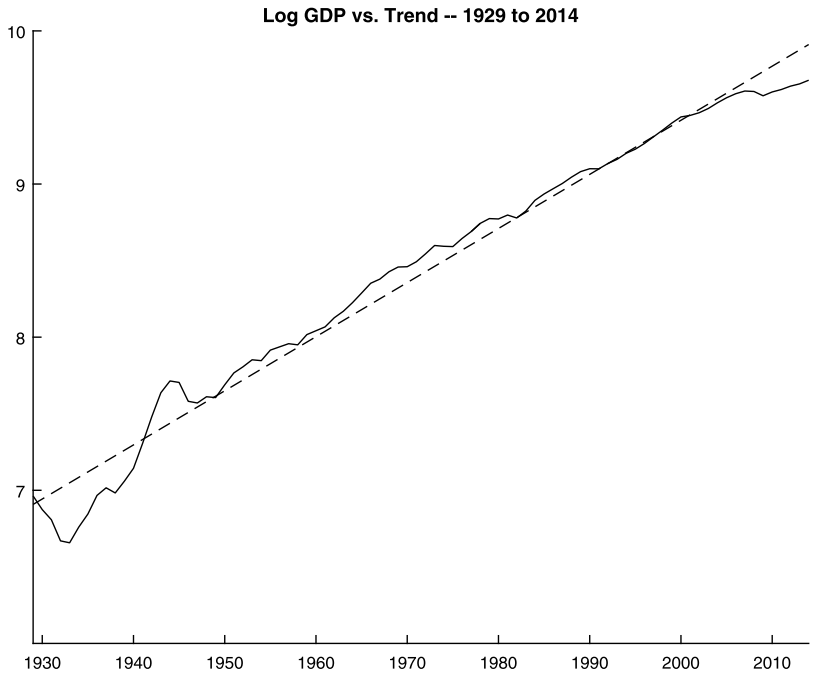


**Fig. 1.** U.S GDP relative to trend.

## 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 $\sigma = 2$ similar to Eusepi and Preston (2011) though somewhat higher because the presence of investment technology shocks reduces the co-movement between consumption and output. Section 5 examines the robustness of the results to a value of $\sigma = 1.05$ near the separable utility 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).[18] I calibrate the investment technology shock by estimating an AR(1) regression on the ratio of the GDP deflator to the investment deflator for structures as in Greenwood et al. (2000) this leads to $\rho^q = 0.986$ and $\sigma^q = 0.002$.[19] I set $\varepsilon_v = \frac{v''(\bar{h})\bar{h}}{v'(\bar{h})} = 0.7914$ to match the volatility of

[17] To find competing views on the existence of a unit root in GDP see Diebold and Senhadji (1996) and Nelson and Plosser (1982).

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

[19] I set $\gamma^q = 1$ so that the return to capital does not grow or shrink indefinitely. I also obtain $\sigma^q$ by scaling down the standard deviation of the

hours worked. This value implies an inverse Frisch elasticity of labor supply equal to 0.33.[20] For comparison, Eusepi and Preston (2011) assume infinitely elastic labor and Kuang and Mitra (2015) set the inverse 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 $\sigma_\varepsilon = 0.003598$ to match the volatility of output. Robustness to various choices for ρ is demonstrated in Section 5.

For the learning parameters I begin with a prior on the stationary model $p_0^s = 0.5$ for the wage equation and 0.75 for the rental rate equation and consider four AR lags.[21] I set $\rho^{ns} = [0\ 0\ 0\ 0\ 0]'$. I set $\rho^S = [0\ 0.975\ 0\ 0\ 0]$. I take the gain parameter to be 0.005 and demonstrate robustness to this value in the appendix.[22] This level of gain implies an expected structural break in the economy every 50 years which seems quite plausible. As an additional check on the reasonableness of this level of the gain I show that it generates forecast error autocorrelations that are on the lower side 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.[23] I set the initial sum of squared residuals $s_0 = 0.0035$ for the wage models and $s_0 = 0.006$ 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 variables will return to their steady state, balanced growth path. Secondly, if he believes the non-stationary model is true 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. Additionally, I find this assumption is necessary to have stability in the model's dynamics. Finally, note that assumptions on priors are not key to generate results since the model is simulated for 1500 periods keeping only the last 200 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 1500 keeping only the final 200 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 random shocks subsequently and one series assuming technology receives no shock at time 1500 and the same random shocks as the previous series subsequently. I calculate the impulse response as the difference between these two series and plot the median impulse responses at each horizon.

### 3.3. Stability

To ensure that the present discounted values of wage and capital income remain finite I require that the agent's estimates of both the stationary process and the non-stationary processes for wages and returns result in bounded long run forecasts by requiring that these estimated beliefs do not contain a unit root, i.e. for the stationary model the variables in levels are I(0) and for the non-stationary model the differences are I(0).[24] When this condition is not satisfied I set the agent's beliefs to the beliefs from the previous period.

In addition, to improve stability of the model, I only use updated beliefs when they lead to a stable actual law of motion for the state variables of the model. Given the model equations, conditional on the agent's beliefs, we can write the evolution of the state variables $\xi_t = \{z_t, \vec{r}_t^{\,s}, \vec{w}_t^{\,s}, k_t, \vec{r}_t^{\,ns}, \vec{w}_t^{\,ns}, q_t, \hat{\gamma}_t, \hat{\gamma}_t^q\}$ as:

$$\xi_{t+1} = \Psi\xi_t.$$

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 specifying the agents beliefs (Eqs. (17) and (18)). The condition for stability is that all the eigenvalues of Ψ are less than one in absolute value. Again, if this condition is not satisfied then the agent uses the beliefs from the previous period. These stability adjustments are not important for the median statistics, in part because they occur early in the simulation in data that is not used to calculate the median statistics. For example, more than half

(*footnote continued*)
regression residuals by 1/3, the share of structures in total investment. An alternative would be to calibrate q with the total investment defaltor but this leads to an estimate of $\sigma^q$ high enough that the learning model can match the volatility of output without any shocks to productivity.

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

[21] 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.

[22] 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)

[23] See Table 4 in Eusepi and Preston (2011) and Section 5.3 in this paper.

[24] The exact requirement for the present discounted values to be finite are that the eigenvalues of the matrices $\Phi^s$ and $\Phi^{NS}$ are all less than $\frac{1}{\hat{\beta}}$ in absolute value, see Hamilton (1994) pp. 19–20 which holds when the beliefs meet the unit root requirement.

**Table 1**
Mechanism.

| | | | |
|---|---|---|---|
| Correlation of beliefs and volatility | | | |
| | $\rho(\lvert R_t\rvert, p^r_{s,t})$ | $\rho[\sigma_t(I_t), \mu_t(p^r_{s,t})]$ | $\rho[\sigma_t(Y_t), \mu_t(p^r_{s,t})]$ |
| | −0.68 | −0.35 | −0.32 |
| | $\rho(\lvert W_t\rvert, p^w_{s,t})$ | $\rho[\sigma_t(Wt), \mu_t(p^w_{s,t})]$ | $\rho[\sigma_t(Y_t), \mu_t(p^w{}_{s,\ t})]$ |
| | −0.79 | −0.05 | −0.03 |
| Regression of belief changes on future outcomes | | | |
| | $\Delta R_{t+1}$ | $\Delta h_{t+1}$ | $\Delta p^r{}_{ns,t+1}$ |
| Constant | −0.000002 | 0.00003 | −0.00002 |
| | (0.00004) | (0.00008) | (0.0002) |
| $\Delta p^r_{ns,t}$ | −0.006 | −0.011 | 0.340*** |
| | (0.003) | (0.010) | (0.033) |
| $\Delta p^r_{ns}$,t * $R_t$ | 1.133*** | 3.221*** | −0.366 |
| | (0.044) | (0.127) | (0.380) |
| $R_t$ | −0.012*** | 0.000 | −0.011 |
| | (0.002) | (0.005) | (0.027) |
| | $\Delta W_{t+1}$ | $\Delta h_{t+1}$ | $\Delta p^w_{ns,t+1}$ |
| Constant | −0.00002 | −0.00003 | −0.00001 |
| | (0.00005) | (0.0001) | (0.0002) |
| $\Delta p^w_{ns,t}$ | 0.005 | −0.033 | 0.259*** |
| | (0.007) | (0.041) | (0.040) |
| $\Delta p^w_{ns,t}$ * $W_t$ | 0.928*** | −5.237*** | −0.576 |
| | (0.104) | (0.601) | (0.717) |
| $W_t$ | −0.016*** | −0.013** | 0.005 |
| | (0.002) | (0.006) | (0.012) |

This table reports the results from correlating the volatility of output and investment with probability weights on the stationary model. It also reports the results for regressing changes in returns to capital, wages hours worked and beliefs on changes in beliefs and their interaction with wages and returns on capital. Standard errors are reported in parentheses and are calculated with a Newey–West estimator with 12 lags. *** $p<0.01$, ** $p<0.05$

the simulations do not have a single stability adjustment in the sample period used to calculate the median statistics. Among all simulations only 1.6% had adjustments more than 10% of the time throughout the entire simulation. And only four of the 500 simulations had adjustments more than 10% of the time in the sample used to calculate the statistics. For more details on the stability adjustments see the appendix.

## 4. Results

### 4.1. Mechanism

This section illustrates how the model generates time varying volatility with changing weights on the stationary and non-stationary models and also how the model generates self-fulfilling beliefs. First, I examine the mechanism that determines the probability weights on the models and how this impacts the volatility of economic outcomes like output and investment. Consider the weight on the stationary model ($p^r_{s,t}$) for the log deviation of the return on capital from its balanced growth path value ($\hat{R}_t$). This weight should be small when $|\hat{R}_t|$ is persistently large. This is when $\hat{R}_t$ is far from its steady state value of 0 and is not showing much tendency to revert to its mean. To demonstrate this I report the median correlation between $|\hat{R}_t|$ and ($p^r_{s,t}$) across the 500 simulations. The results are in Table 1. I find that this correlation is strongly negative equal to −0.68 demonstrating that times when $|\hat{R}_t|$ is large are the times when the weight on the stationary model is small.

Next, I show that these times are when the volatility of outcomes like output and investment are larger. To show this I report the correlation between $\mu_t(p^r_{s,t})$ and $\sigma_t(\hat{I}_t)$ where the t subscript denotes a 40 quarter moving average centered around time *t*, $\mu$ represents the moving mean and $\sigma$ represents the moving standard deviation. The median correlations are given in Table 1. This correlation is negative, equal to -0.35, demonstrating the negative relationship between the volatility of investment and the weight on the stationary model for the return on capital. The table also reports the correlation for the standard deviation of output which is negative as well.

Now, I repeat this analysis for the wage beliefs. The first correlation is negative, so that the weight on the stationary model for wages is low when wages are far from their steady state. The other correlations while negative are small. This results suggests that the time varying volatility the model generates is driven mostly by changes in beliefs about the stationarity of the return on capital and less by changes in beliefs about the stationarity of wages.

Next, I examine the ability of the model to generate self-fulfilling beliefs. Specifically, if $\hat{R}_t$ is large and positive, and the individual revises his belief on the non-stationary model ($p^r_{ns,t}$) upwards, does this tend to raise next period's return on capital, $\hat{R}_{t+1}$. To examine

this result I simulate the model for 10,000 quarters and then run the following regression on the data:

$$\Delta\hat{R}_{t+1} = \alpha + \beta\Delta p^r_{ns,t+1} + \gamma\Delta p^r_{ns,t+1} * \hat{R}_t + \delta\hat{R}_t + \varepsilon_t$$

In interpreting this regression it is important to recall that beliefs are updated at the end of the period so that $p^r_{s,t+1}$ is determined using data at time $t$. The coefficient of interest is γ; it tells us the impact on next period's return to capital of revising your belief on the non-stationary model upward when $\hat{R}_t$ is positive. I find that γ is positive and significantly different than zero. This means that when $\hat{R}_t$ is positive, and $\Delta p^r_{ns,t+1}$ is positive, next period's return to capital increases on average and therefore the return on capital will move father from its steady state of zero. Conversely, when $\hat{R}_t$ is negative, revising your belief upward leads to an even more negative $\hat{R}_t$ next period. To understand the magnitude of this coefficient, $\gamma = 1.1$, note that if $\Delta p^r_{ns,t+1} = 0.1$ and $\hat{R}_t = 0.05$ (5% above the steady state value), then next period's $\hat{R}_t$ would, all else being equal, increase to 0.0555.

I then repeat the regression replacing the dependent variable ($\Delta\hat{R}_{t+1}$) with $\Delta\hat{h}_{t+1}$ and then $\Delta p^r_{ns,t+2}$. Again, when $\Delta\hat{h}_{t+1}$ is the dependent variable γ is positive showing that when $\Delta\hat{h}_t$ is positive, and the agent increases his belief on the non-stationary model, this will lead to increased labor supply next period. Note that in neither the rental rate nor hours regression is $\beta$ significantly different than zero. This is because the effect of revising your beliefs has a different sign if you are above or below the steady state and therefore the average effect is roughly zero. Finally, when $\Delta p^r_{ns,t+2}$ is the dependent variable we see that $\beta$ is positive and significant. This result demonstrates that changes in beliefs are positively autocorrelated. If we start out above steady state, then revising your belief upward leads to an even larger value for the rental rate, which leads to another upward revision of beliefs. On the other hand, if we start out below steady state, then if we revise up our weight on the non-stationary model, the rental rate becomes more negative, and we further revise up our weight on the non-stationary model. The positive autocorrelation occurs if rental rates are above and below the steady state and so here $\beta$ is positive. γ is not significantly different because the effect does not depend on the level of the rental rate.

This self-fulfilling belief channel is an important source of amplification, a point made by Adam et al. (2017). If one were learning about an exogenous variables (for example fundamentals like productivity) then there is no way for beliefs to influence that variable reducing the ability of the model to amplify shocks.

I then repeat the above regression using $\hat{W}_t$ the deviation of the wage from the balanced growth path. The results are similar. When wages are above the steady state, an increase in the belief that wages are non stationary leads the wage to increase further next period. This comes from a decrease in labor supply, presumably because of the income effect. Again changes in beliefs about the stationarity of the wage process are positively autocorrelated.

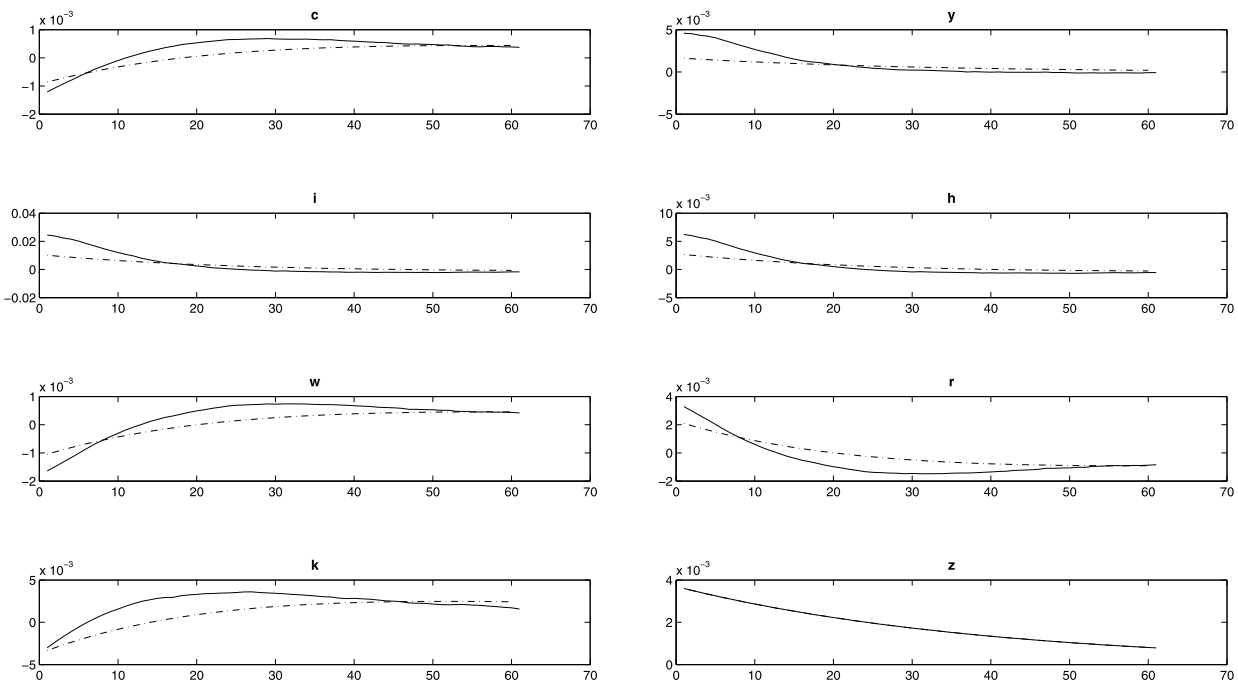


**Fig. 2.** Impulse responses to one standard deviation productivity shock.
Note: Learning model (solid), rational expectations benchmark (dashed).

### 4.2. Simulated results

#### 4.2.1. Business cycle statistics

Fig. 2 gives the impulse responses to several of the detrended model variables to a one standard deviation shock to productivity.[25] 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 substantially larger and the response of investment is more than 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 2 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.[26] Data begin in 1964 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.

By calibration the learning model exactly matches the standard deviation of log HP-filtered output 1.55%; in contrast, given the same volatility of the productivity shock, the rational expectations benchmark model generates only a 0.7% standard deviation 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, $[(1-\alpha)*0.0036]^2$ is only 9% the variance of output growth ($0.008^2$) addressing the Burnside et. al critique of real business cycle models.

Consumption is half as volatile as output while investment is 4.4 times as volatile. The learning model predicts a smaller relative volatility of consumption at 0.22 while the rational expectations model comes closer to matching the volatility of consumption with a value equal to 0.53. 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 2.6. The learning model comes closer to matching the data with a value equal to 4.1. We find that hours are 0.97 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 53% 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 0.33 while many papers in the literature need infinitely elastic labor supply to match the data. While the rational expectations model does match the relative volatility of labor productivity in the data, the learning model tends to understate it. This value is 0.5 in the data, the learning model predicts 0.21.

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, $\rho(c, y) = 0.58$. While it is much higher in the data – 0.85 – 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 0.29 consistent with the value of 0.31 in the data. The learning model is able to break this extreme correlation because it 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 autocorrelations 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

[25] 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 random shocks subsequently and one series assuming technology receives no shock at time 1500 and then the same random shocks as the previous series subsequently. I calculate the impulse response as the difference between these two series and plot the median impulse responses at each horizon. Both series are assumed to always have the same shock to investment technology.

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

**Table 2**
Business cycle statistics.

| | Data | RE | Learning | | Data | RE | Learning | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Standard deviation of HP-filtered output | | | | Standard deviation of output growth | | | | | | |
| $\sigma(y)$ | 0.0155 | 0.007 | 0.0155 | $\sigma(\Delta y)$ | 0.008 | 0.005 | 0.01 | | | |
| Relative volatility of HP-filtered variables | | | | Relative volatility of variable growth rates | | | | | | |
| $\sigma(c)/\sigma(y)$ | 0.55 | 0.53 | 0.22 | $\sigma(\Delta c)/\sigma(\Delta y)$ | 0.54 | 0.52 | 0.23 | | | |
| $\sigma(i)/\sigma(y)$ | 4.4 | 2.6 | 4.1 | $\sigma(\Delta i)/\sigma(\Delta y)$ | 4.8 | 2.6 | 3.9 | | | |
| $\sigma(h)/\sigma(y)$ | 0.97 | 0.51 | 0.97 | $\sigma(\Delta h)/\sigma(\Delta y)$ | 0.83 | 0.51 | 0.92 | | | |
| $\sigma(y/h)/\sigma(y)$ | 0.5 | 0.5 | 0.21 | $\sigma(\Delta y/h)/\sigma(\Delta y)$ | 0.74 | 0.49 | 0.21 | | | |
| Contemporaneous correlations of HP-filtered variables | | | | Contemporaneous correlations of variable growth rates | | | | | | |
| $\rho(c, y)$ | 0.85 | 0.99 | 0.58 | $\rho(\Delta c, \Delta y)$ | 0.55 | 0.99 | 0.68 | | | |
| $\rho(h, y)$ | 0.87 | 0.99 | 0.98 | $\rho(\Delta h, \Delta y)$ | 0.69 | 0.99 | 0.98 | | | |
| $\rho(i, y)$ | 0.91 | 0.99 | 0.99 | $\rho(\Delta i, \Delta y)$ | 0.81 | 0.99 | 0.99 | | | |
| $\rho(r, y)$ | 0.42 | 0.84 | 0.92 | $\rho(\Delta r, \Delta y)$ | 0.33 | 0.85 | 0.94 | | | |
| $\rho(w, y)$ | 0.31 | 0.99 | 0.29 | $\rho(\Delta w, \Delta y)$ | 0.04 | 0.99 | 0.44 | | | |
| Autocorrelations of variable growth rates | | | | | | | | | | |
| | Data | RE | Learning | Data | RE | Learning | Data | RE | Learning | |
| | | $\Delta y$ | | | $\Delta c$ | | | $\Delta h$ | | |
| lag 1 | 0.33 | −0.02 | 0.06 | 0.53 | −0.007 | 0.08 | 0.48 | −0.03 | 0.09 | |
| lag 2 | 0.25 | −0.02 | −0.01 | 0.39 | 0 | 0.04 | 0.4 | −0.02 | −0.004 | |
| lag 3 | 0.12 | −0.01 | −0.03 | 0.42 | 0 | 0.04 | 0.26 | −0.02 | −0.04 | |
| | | $\Delta i$ | | | $\Delta r$ | | | $\Delta w$ | | |
| lag 1 | 0.23 | −0.03 | 0.08 | 0.22 | −0.02 | 0.1 | 0.8 | −0.005 | 0.12 | |
| lag 2 | 0.1 | −0.02 | −0.004 | −0.17 | −0.01 | 0.02 | 0.79 | 0 | 0.05 | |
| lag 3 | 0.07 | −0.02 | −0.04 | 0.16 | −0.02 | −0.01 | 0.78 | 0.002 | 0.05 | |

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 correlations. The same statistics are reported for the growth rates of these variables along with four lags of their autocorrelations.

autocorrelation in variables. 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 18% of the autocorrelation in output and 35% 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 0.09 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 0.25 for consumption to 0.39 for hours. And the statistic for investment is 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 adjustment costs. Instead, the self-fulfilling expectations drive this result. When rental 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 autocorrelation 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.

### *4.2.2. Forecast errors*

Table 3 reports forecast errors from the learning model. Because the household uses an incorrect model to forecast the real interest rate and the wage agents will make errors. Using data from the Survey of Professional Forecasters I calculate autocorrelation in the median forecasts for a variety of variables.[27]. 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.[28] The pattern for interest rates is similar with positive autocorrelations of 1-quarter forecast errors ranging from 0.26 for the three-month T-bill to 0.43 for the ex-post real rate.[29] 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.

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

[28] 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 forecast errors we have $\epsilon_t^4 = x_t - E_{t-4}x_t$ and the table reports $\rho(\epsilon_t^4, \epsilon_{t-4}^4)$.

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

**Table 3**
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 | |
| | T-bill | 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 | 0 | 0.09 | 0.12 |
| 4Q | 0 | 0 | 0.07 | 0.19 |

Note: This table reports the autocorrelation of forecast errors 1 quarter and 4 quarters ahead. Results for the data are in the top panel for the listed variables. Results for the return on capital and wage variables for the rational expectations model (based on a simulation of length 100,000) and the learning model are in the bottom panel.

**Table 4**
Long run predictability and time varying volatility.

| Long run predictability | Data | RE | Learning | | Data | RE | Learning |
|---|---|---|---|---|---|---|---|
| $\rho(c_{t-3} + \ldots + c_t, c_{t+1} + \ldots + c_{t+16})$ | −0.4 | −0.33 | −0.29 | $\rho(\ln c_t - \ln c_{t-4}, \ln c_{t+16} - \ln c_t)$ | 0.22 | −0.06 | 0.07 |
| $\rho(y_{t-3} + \ldots + y_t, y_{t+1} + \ldots + y_{t+16})$ | −0.43 | −0.35 | −0.4 | $\rho(\ln y_t - \ln y_{t-4}, \ln y_{t+16} - \ln y_t)$ | −0.12 | −0.12 | −0.25 |
| $\rho(i_{t-3} + \ldots + i_t, i_{t+1} + \ldots + i_{t+16})$ | −0.51 | −0.34 | −0.4 | $\rho(\ln it - \ln i_{t-4}, \ln i_{t+16} - \ln i_t)$ | −0.33 | −0.15 | −0.3 |
| Squared Autocorrelations | | $c^2$ | | | | $(\varepsilon_c)^2$ | |
| lag 1 | 0.8 | 0.47 | 0.53 | lag 1 | 0.09 | −0.02 | 0.02 |
| lag 2 | 0.52 | 0.18 | 0.26 | lag 2 | 0.06 | −0.01 | 0.02 |
| lag 3 | 0.25 | 0.04 | 0.11 | lag 3 | 0.06 | −0.01 | 0.01 |
| | | $y^2$ | | | | $(\varepsilon_y)^2$ | |
| lag 1 | 0.76 | 0.47 | 0.54 | lag 1 | 0.14 | −0.01 | 0.06 |
| lag 2 | 0.49 | 0.18 | 0.2 | lag 2 | 0.08 | −0.01 | 0.06 |
| lag 3 | 0.21 | 0.04 | 0.05 | lag 3 | 0.11 | −0.01 | 0.02 |
| | | $i^2$ | | | | $(\varepsilon_i)^2$ | |
| lag 1 | 0.71 | 0.47 | 0.54 | lag 1 | 0.15 | −0.01 | 0.07 |
| lag 2 | 0.35 | 0.18 | 0.22 | lag 2 | 0.04 | −0.01 | 0.07 |
| lag 3 | 0.09 | 0.04 | 0.06 | lag 3 | 0.04 | −0.01 | 0.01 |
| | | $(\Delta c)^2$ | | | | $(\varepsilon_{\Delta c})^2$ | |
| lag 1 | 0.32 | −0.01 | 0.03 | lag 1 | 0.1 | −0.02 | 0.01 |
| lag 2 | 0.28 | −0.01 | 0.02 | lag 2 | 0.08 | −0.01 | 0.02 |
| lag 3 | 0.15 | 0 | 0 | lag 3 | 0.1 | 0 | 0 |
| | | $(\Delta y)^2$ | | | | $(\varepsilon_{\Delta y})^2$ | |
| lag 1 | 0.12 | −0.01 | 0.06 | lag 1 | 0.09 | −0.02 | 0.07 |
| lag 2 | 0.04 | −0.01 | 0.05 | lag 2 | 0.127 | −0.01 | 0.05 |
| lag 3 | 0.06 | 0 | 0.02 | lag 3 | 0.03 | 0 | 0.02 |
| | | $(\Delta i)^2$ | | | | $(\varepsilon_{\Delta i})^2$ | |
| lag 1 | 0.11 | −0.01 | 0.09 | lag 1 | 0.04 | −0.02 | 0.09 |
| lag 2 | 0.15 | −0.01 | 0.06 | lag 2 | 0.17 | −0.01 | 0.05 |
| lag 3 | 0.03 | 0 | 0.02 | lag 3 | 0.01 | −0.004 | 0.02 |

Note: This table contains the correlation of the sum of 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 contains 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.

I find no autocorrelation of forecast errors in the rational expectations model at 1 or 4 quarters.[30] 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 0.09. At the 4q ahead forecast horizon the learning models generate autocorrelation of forecast errors of 0.07 for rental rates and 0.19 for wages. I also examine the 90th percentile for the forecast error correlation in the learning model. For the learning model this percentile is 0.41 for one-quarter real interest rate forecast errors and 0.38 for one-quarter wage rate forecast errors and the data all fall within this range.

[30] This result is of course expected for a rational expectations model.

#### 4.2.3. Long run predictability and time varying volatility

Table 4 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 4 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 $\rho(\ln c_t - \ln c_{t-4}, \ln c_{t+16} - \ln c_{t+1}) = 0.22$ and the learning model predicts 0.07 versus $-0.06$ in the rational expectations benchmark. For investment this correlation is $-0.33$. The learning model predicts a correlation of $-0.3$ versus $-0.15$ in the rational expectations model. For output, this statistic equals $= -0.12$. The rational expectations model matches this statistic while the learning model overshoots it predicting a correlation of $-0.25$.

For this statistic both models emit wide confidence intervals. However, for the rational expectations model the consumption statistic is larger than the 90th percentile with a value of 0.19 versus 0.22 in the data. The 10th percentile for the investment correlation is $-0.32$ which is higher than 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 $\rho([c^{HP}]^2)$ and $\rho(\varepsilon_c^2)$ where $\varepsilon_c$ 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 to account for this fact. The learning model generates larger autocorrelations though, more consistent with the data. For example $\rho([y^{HP}]^2) = 0.76$ versus 0.54 for the learning model. The rational expectations benchmark predicts this value to be only 0.47.

These discrepancies are much larger turning our 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 0.07, versus 0.15 in the data. Similarly, the learning model generates an autocorrelation of 0.02 for squared consumption residuals compared to 0.09 in the data; and it generates 0.06 for output residuals compared to 0.14 in the data.

The last section of Table 4 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 cannot match this fact. It predicts squared growth rates should be uncorrelated over time. However, the learning model predicts a correlation of 0.03 for squared consumption growth, positive like the value of 0.32 in the data. Similarly, for output and investment we see positive autocorrelation of squared growth rates at values of 0.12 and 0.11 respectively, a fact matched by the learning model though with smaller magnitude of 0.06 and 0.09. Similarly, squared AR(1) residuals of consumption, investment and output growth are all positive. Again, the learning model is able to generate positive squared autocorrelations while the rational expectations model cannot. Importantly, it replicates the autocorrelation of squared output growth residuals with a correlation of 0.07 versus 0.09 in the data and investment residuals with a prediction of 0.09 versus 0.04 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.[31]

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 so as 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 presumably 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

### 5.1. Different parameter values

Table 5 considers robustness of the results to varying some of the calibrated parameters. For these exercises I keep all the

[31] Though I have omitted the results due to space constraints, I have checked that time varying volatility exists in both the pre and post Great Moderation samples to ensure that time varying volatility is not due solely to the Great Moderation.

**Table 5**
Robustness.

| | Benchmark | prod ar(1) 0.9 | 0.99 | gain 0.015 | 0 | PS = 1 | PS = 0.65 | σ = 1.05 | VAR |
|---|---|---|---|---|---|---|---|---|---|
| Volatility | | | | | | | | | |
| $\sigma(y)$ | 0.0155 | 0.015 | 0.0157 | 0.017 | 0.013 | 0.014 | 0.0155 | 0.015 | 0.009 |
| $\sigma(c)/\sigma(y)$ | 0.24 | 0.2 | 0.23 | 0.2 | 0.24 | 0.22 | 0.2 | 0.77 | 0.42 |
| $\sigma(i)/\sigma(y)$ | 4.1 | 4 | 4.1 | 4.2 | 3.7 | 3.9 | 4 6.2 | 3.4 | |
| Correlations | | | | | | | | | |
| $\rho(c, y)$ | 0.55 | 0.7 | 0.48 | 0.4 | 0.74 | 0.76 | 0.66 | −0.83 | 0.81 |
| $\rho(h, y)$ | 0.98 | 0.99 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.97 | 0.94 |
| Autocorrelations of growth rates | | | | | | | | | |
| $\Delta y$ | 0.06 | 0.03 | 0.08 | 0.11 | −0.02 | −0.03 | −0.04 | 0.1 | 0.23 |
| $\Delta c$ | 0.08 | 0.01 | 0.08 | 0.09 | 0.06 | 0.07 | 0.1 | 0.18 | −0.07 |
| $\Delta i$ | 0.09 | 0.07 | 0.1 | 0.14 | −0.02 | −0.03 | −0.05 | 0.12 | 0.22 |
| Autocorrelation of forecast errors | | | | | | | | | |
| r | 0.09 | 0.09 | 0.1 | 0.11 | 0.03 | −0.01 | −0.01 | 0.15 | 0.02 |
| w | 0.12 | 0.1 | 0.1 | 0.14 | 0.09 | 0.1 | 0.08 | 0.1 | −0.02 |
| Long run predictability | | | | | | | | | |
| $\rho(lnc_t - \ln c_{t-4}, \ln c_{t+16} - lnc_t)\rho(lnc_t - \ln c_{t-4}, \ln c_{t+16} - lnc_t)$ | 0.07 | −0.09 | 0.07 | 0.05 | 0.09 | 0.13 | 0.11 | −0.31 | -0.03 |
| $\rho(lny_t - \ln y_{t-4}, \ln y_{t+16} - lny_t)$ | −0.25 | −0.31 | −0.28 | −0.3 | −0.21 | −0.19 | −0.28 | −0.34 | -0.2 |
| $\rho(lni_t - \ln i_{t-4}, \ln i_{t+16} - lni_t)$ | −0.3 | −0.31 | −0.32 | −0.33 | −0.24 | −0.23 | −0.31 | −0.39 | -0.26 |
| Time varying volatility | | | | | | | | | |
| $(\varepsilon_c)^2$ | 0.02 | 0.01 | 0.1 | 0.03 | −0.02 | −0.01 | −0.01 | 0.09 | −0.01 |
| $(\varepsilon_y)^2$ | 0.06 | 0.04 | 0.06 | 0.08 | 0 | −0.01 | −0.01 | 0.06 | 0.18 |
| $(\varepsilon_i)^2$ | 0.07 | 0.05 | 0.07 | 0.1 | 0 | −0.01 | −0.01 | 0.07 | 0.2 |
| $(\varepsilon_{\Delta c})^2$ | 0.01 | 0.02 | 0.01 | 0.03 | −0.01 | −0.02 | −0.02 | 0.21 | −0.01 |
| $(\varepsilon_{\Delta y})$ 2 | 0.07 | 0.05 | 0.1 | 0.11 | 0 | −0.01 | −0.02 | 0.1 | 0.15 |
| $(\varepsilon_{\Delta i})2$ | 0.09 | 0.05 | 0.13 | 0.16 | 0 | −0.02 | −0.02 | 0.15 | 0.19 |

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

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 $\rho = 0.9$. 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 almost the same volatility. The standard deviation of output $\sigma(y) = 0.015$ versus 0.0155 in the benchmark case. The relative standard deviation of consumption is a bit lower and the relative standard deviation of investment is about the same. The contemporaneous correlations of consumption and hours worked are similar, with an elevated correlation between consumption and output. 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 0.07 when $\rho = 0.9$ versus 0.09 in the benchmark case. The model still generates long run predictability in growth rates. For example, $\rho(\ln i_t - \ln i_{t-4},\ \ln i_{t+16} - \ln i_t) = -0.31$ which almost exactly matches the benchmark value. Finally, the model still is able to generate time varying volatility with just a slight decrease for the benchmark values. Similarly, when $\rho = 0.99$ the model generates essentially the same volatility, $\sigma(y) = 0.0157$ versus 0.0155 in the benchmark case, and maintains the same relative volatilities and contemporaneous correlations. The model also generates slightly higher autocorrelations for growth rates. For example output growth has an autocorrelation $\rho(\Delta y) = 0.08$ versus 0.06 as the benchmark value. The model also generates similar autocorrelation of forecast errors. The autocorrelation of the forecast errors for returns now equals 0.1 for wages versus a value of 0.09 for the benchmark case. Values for long run predictability are stronger, $\rho(\ln i_t - \ln i_{t-4},\ \ln i_{t+16} - \ln i_t) = -0.32$ 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 $g = 0.005$. I consider increasing $g = 0.015$. The first effect of increasing the gain variable is to increase the volatility of output. $\sigma(y) =$ 1.55% versus 1.7% when the gain variable equals 0.015 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 0.4 versus 0.55 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 0.14 versus 0.09 in the benchmark case. The correlations representing long run predictability are more negative. For example $\rho(\ln y_t - \ln y_{t-4},\ \ln y_{t+16} - \ln y_t) = -0.3$ versus -0.25 for the benchmark case and $\rho(\ln i_t - \ln i_{t-4},\ \ln i_{t+16} - \ln i_t) = -0.33$ versus -0.3 for the baseline case. Finally the model generates more time varying volatility than the baseline case does. For example, the autocorrelation of squared investment growth rate residuals equals 0.16 versus 0.09 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 1.3%. The relative volatilities of consumption and output however show little change. Consumption now is more correlated with output: $\rho(c, y) = 0.74$ versus 0.55 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. The model also generates different autocorrelations of long run growth rates: $\rho(\ln c_t - \ln c_{t-4},\ \ln c_{t+16} - \ln c_t) = 0.09$ versus 0.07 for the benchmark case and $\rho(\ln i_t - \ln i_{t-4},\ \ln i_{t+16} - \ln i_t) = -0.24$ 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 $\rho(\varepsilon^2_{\Delta i}) = 0$ when the gain variable equals zero versus 0.09 in the baseline case. Given the model's degraded fit when the gain value equals zero, I find it necessary to include a non-zero gain in the benchmark calculation.

### 5.2. Drivers of the results

In Section 2.5, I noted that the model has three main deviations from textbook rational expectations, RBC models. First, I have model learning, second I have unit root beliefs, and third I ignore forecasting with the capital stock. In this subsection of the paper I examine the consequences of each of these assumptions.

Now, in the robustness table, I consider the importance of model learning. I re-simulate the learning model but set the weights on the stationary model $p^r_{s,t}$ and $p^w_{s,t} = 1$ for all $t$. Output volatility decreases 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 generate 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. It also highlights the role of neglecting the capital stock in forecasting future wages and returns. Even though the agent is not wrong to think both wages and rental rates are stationary, volatility is still increased over the benchmark RE case.

To examine the role that beliefs about a unit root in wages and rental rates play in the results I re-simulate the learning model but set $p^r_{s,t}$ and $p^w_{s,t} = 0.65$ for all $t$.[32] The model does not generate more volatility than in the benchmark case and the relative volatilities and correlations are similar. However, the model loses its ability to explain autocorrelation in growth rates, rental rate forecast errors, and autocorrelation of squared residuals. It appears, therefore, that while unit root beliefs not surprisingly help with amplification, learning is necessary to generate autocorrelation 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 of output is unchanged 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 in magnitude. However, for consumption these statistics are amplified. For example, the autocorrelation in consumption growth rates is 0.18 versus 0.08 in the benchmark case. Time varying volatility for consumption growth rate residuals equals 0.21 versus 0.01 in the benchmark case.

Finally, I consider including additional information in the forecasting of wages and returns. Specifically, I include the variables $k_t$, $z_{t-1}$, and $q_{t-1}$ in the stationary factor price equations (17) and $\Delta k_t$, $\Delta z_{t-1}$, and $\Delta q_{t-1}$ in the non-stationary factor price equations (18). 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 $r_t$ , $z_{t-1}$, and $q_{t-1}$. 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. $k_{t+1} = \rho^{k,s}_0 + \rho^{k,s}_1 k_t + \rho^{k,s}_2 z_{t-1} + \rho^{s}_3 q_{t-1} + \varepsilon^{k,s}_t$ with probability $p^{ks}_t$ and that it is non-stationary as, i.e. $\Delta k_{t+1} = \rho^{k,ns}_0 + \rho^{k,ns}_1 \Delta k_t + \rho^{k,ns}_2 \Delta z_{t-1} + \rho^{k,ns}_3 \Delta q_{t-1} + \varepsilon^{k,ns}_t$ with probability $1 - p^{ks}_t$. 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 5. There is a fall in the volatility of output to 0.9% and the relative volatility of consumption rises more in line with the data. The model generates quite similar contemporaneous correlations of hours with output and a higher correlation of consumption with output. The model generates larger autocorrelation in growth rates but no autocorrelation in forecast errors. The model generates slightly less negative correlation of growth rates with long-run growth rates and the model's prediction for time varying volatility is higher. This exercise also suggests that omitting capital in the forecasting equation is an important source of amplification in the learning model.

### 5.3. Comparison to Eusepi and Preston (2011)

The business cycle model in this paper is quite similar to the model of Eusepi and Preston (2011) (EP) (the two exceptions are that in their model technology is a random walk and there are no investment technology shocks). However, the learning mechanism in their paper versus this paper is quite different. In this paper agents are learning about the process for returns to capital and wages, while in EP they are learning about the parameters of a known process. In addition, in this paper agents are using a univariate model of lagged values of wages and capital income to forecast future values, while in EP they use the minimum state variable solution and forecast using expectations of future levels of the capital stock. In this section I compare the results in this paper to the results in EP and argue that the different assumption about what agents are learning about are key to generating the results. In short, one can not replicate the results in this paper using the EP model or the EP model with a higher gain value.

Table 6 shows the key moments for the model in this paper and then the EP model for their gain calibration and a higher gain

[32] I find this is the lowest constant weight I can put on the stationary model and still maintain stability.

**Table 6**
Comparison with Eusepi and Preston (2011).

| | Benchmark | EP my calibration | | EP benchmark calibration | benchmark cal |
|---|---|---|---|---|---|
| | | gain | | gain | |
| | | 0.002 | 0.015 | 0.002 | 0.015 |
| Volatility | | | | | |
| $\sigma$(y) | 0.0155 | 0.006 | 0.008 | 0.01 | 0.021 |
| $\sigma$(c)/$\sigma$(y) | 0.24 | 0.71 | 0.53 | 0.42 | 0.18 |
| $\sigma$(i)/$\sigma$(y) | 4.1 | 2 | 2.9 | 2.9 | 4 |
| Correlations | | | | | |
| ρ(c,y) | 0.55 | 0.99 | 0.89 | 0.97 | 0.28 |
| ρ(h,y) | 4.4 | 0.98 | 0.91 | 0.99 | 0.98 |
| Autocorrelations of growth rates | | | | | |
| Δy | 0.06 | 0.07 | 0.34 | 0.17 | 0.27 |
| Δc | 0.08 | −0.01 | −0.09 | −0.05 | −0.39 |
| Δi | 0.09 | 0.19 | 0.27 | 0.36 | 0.13 |
| Autocorrelation of forecast errors | | | | | |
| r | 0.09 | 0.06 | 0.54 | 0.44 | 0.9 |
| w | 0.12 | 0.06 | 0.54 | 0.44 | 0.9 |
| Long Run Predictability | | | | | |
| $\rho(lnc_t - \ln ct_4, \ln c_{t+16} - lnc_t)$ | 0.07 | 0 | 0.06 | 0.08 | 0.2 |
| $\rho(lny_t - \ln y_{t-4}, \ln y_{t+16} - lny_t)$ | −0.25 | −0.05 | −0.07 | −0.14 | −0.32 |
| $\rho(lni_t - \ln i_{t-4}, \ln i_{t+16} - lni_t)$ | −0.3 | −0.08 | −0.14 | −0.2 | −0.36 |
| Time Varying Volatility | | | | | |
| $(\varepsilon_c)^2$ | 0.02 | −0.01 | −0.01 | −0.02 | 0.03 |
| $(\varepsilon_y)^2$ | 0.06 | 0 | 0.22 | 0.04 | 0.24 |
| $(\varepsilon_i)^2$ | 0.07 | 0.04 | 0.31 | 0.1 | 0.2 |
| $(\varepsilon_{\Delta c})^2$ | 0.01 | −0.01 | −0.01 | −0.01 | 0.04 |
| $(\varepsilon_{\Delta y})^2$ | 0.07 | −0.01 | 0.1 | −0.01 | 0.18 |
| $(\varepsilon_{\Delta i})^2$ | 0.09 | 0 | 0.27 | 0 | 0.19 |

Note: This table compares the main results to the results of Eusepi and Preston (2011). The first set of EP results uses a calibration of the model in this paper. The second set uses the benchmark EP calibration with $\rho = 1$ and infinitely elastic labor supply. Results are given for EP's choice of constant gain parameter 0.002 and a higher gain parameter 0.015.

value. The first set of results uses a calibration of the model in this paper.[33] The second set of results uses the EP benchmark calibration with $\sigma = 1$ and infinitely elastic labor supply.

In the first column of the EP model results, we see that the EP model generates substantially less volatility than the model in this paper: about one-third of the volatility. The model is also unable to match the new moments introduced in this paper, there is much less long run predictability of output growth and investment growth. There is also no real evidence of consistent time varying volatility.

I then look to see if these shortcomings can be addressed by increasing the gain in the EP model. I increase the gain value to 0.015 over their benchmark value of 0.002. The EP model now generates about 1/2 the volatility of the model in this paper and generates time varying volatility in output and investment, but not consumption. This level of volatility is similar to the level of volatility generated in this paper's model when the agent forecasts with the capital stock suggesting that is a key difference between this model and EP. Additionally, the model's ability to generate long run predictability in output growth and investment growth is improved through still outperformed by the model in this paper. However, there is a significant problem with increasing the gain in the EP model. Looking at the autocorrelation of forecast errors the EP model generates an autocorrelation of 0.54 versus 0.15 and 0.09 for the benchmark model in this paper. This level of autocorrelation is far above mostly all the data moments in table 3 and suggests an unreasonably high value of the gain. To sum up, the benchmark model in this paper generates more amplification, reversals in long run growth rates, and time varying volatility than the EP model. And while increasing the gain can narrow this gap, it comes at the cost of unreasonably high autocorrelation of forecast errors.

Next I consider the EP model with their benchmark calibration of $\sigma = 1$ and infinitely elastic labor supply. This model naturally generates more amplification with more elastic labor supply. However the level of amplification is still around 50% below the model in this paper with the low level of the gain. With a high gain, the EP benchmark calibration of the model is still able to generate more amplification, time varying volatility and reversals of output and investment growth however this comes at the expense of unreasonably high forecast error autocorrelation. The model generates a forecast error autocorrelation of 0.9, far above any values in Table 3.

To understand why this paper's learning model and the EP model make different predictions for the autocorrelation of short run forecast errors it is important to understand the source of model misspecification in the EP model. Following the notation in Eusepi and Preston (2011), let $z_t = [r_t, w_t, k_{t+1}]'$ and $q_t = [1, k_t]'$. Additionally, let the coefficients in the perceived law of motion be

[33] I use $\sigma = 1.75$ , $\varepsilon_v = 0.6125$ and $\sigma^g = 0.0045$ which generates similar moments in this paper's model without investment technology shocks.

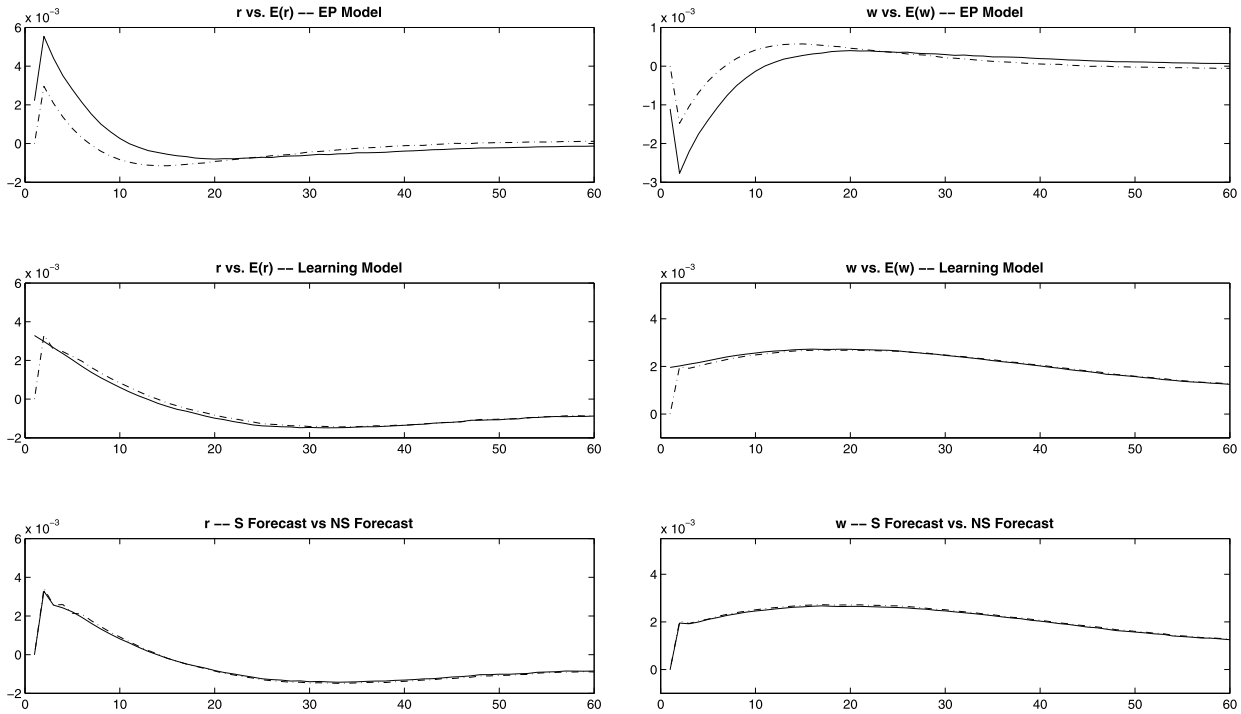


**Fig. 3.** One period ahead forecasts (dashed) vs. outcomes (solid): EP vs. learning.

given by $\omega^r = [\omega_0^r \omega_1^r]$, $\omega^w = [\omega_0^w \omega_1^w]$, $\omega^k = [\omega_0^k \omega_1^k]$. In the EP model (similarly to this paper's model) beliefs are updated at the end of period $t-1$ and then those beliefs are used at time $t$ to forecast the future values of $z_t$. That is to say that agents forecast one period ahead variables using $z_{t+1} = \Omega_{t-1} q_{t+1}$ where $\Omega_{t-1} = [\omega_{t-1}^r \omega_{t-1}^w \omega_{t-1}^k]'$ are the coefficients updated at the end of period $t-1$.[34] However the actual law of motion for $z_{t+1} = T_1(\Omega_t) q_t + T_2(\Omega_t) \hat{\gamma}_{t+1}$ where $\hat{\gamma}_{t+1}$ is technology growth unforecastable at time $t$. Here $T_1(\Omega_t)$ and $T_2(\Omega_t)$, the so-called T-maps, are non-linear functions of the previous periods estimates of $\Omega$.[35] Therefore agents will make forecast errors, the source of the forecast errors being that $\Omega_{t-1} \neq T_1(\Omega_t)$.

To understand the direction of this misspecification, examine Fig. 3. This figure plots the impulse responses of the one-period ahead forecasts and outcomes for the EP model and this paper's learning model. The top panel contains the impulse response for the EP model, the solid line is the actual outcome of the variable and the dashed line is the forecast of the variable from the previous period. What one sees is that the agent consistently underestimates the rental rate and overestimates the efficiency wage for the first 20 periods after the technology shock. The reason agents do this is that their misspecified model neglects the impact that updating their beliefs will have on the rental rate and wage. Specifically, the updating of beliefs after the technology shock leads to an increase in labor supply which drives up the return on capital and lowers the wage. The magnitude of this effect depends on how much beliefs change, which depends directly on the gain variable. When the gain is high, we get large persistent mis-estimation which results in a high autocorrelation of forecast errors.

On the other hand, in this paper's learning model, as one sees in panel 2 of Fig. 3, the forecast errors for r and w, while in the same direction are much smaller and are more easily swamped out by future shocks. The reason for this result is shown in the last panel of Fig. 3. Since both the rental rate and wage are highly persistent series the stationary and non-stationary forecasts given in the last panel are both quite close and accurate. Additionally neglected updating of beliefs is not a major factor in the evolution of these variables. First, because there is decreasing gain in learning about the parameters of each model they are close to stable over time. Additionally, the shocks do not move the weight put on the stationary or non-stationary model in a predictable direction. If the economy is above steady state a positive shock will move it towards the non-stationary model, but if it is below trend it will move it towards the stationary model. Therefore, unlike in the EP model, there is no systematic updating of beliefs that is neglected in the forecasting model.

This is not to say though that the learning model in this paper does not have important forecast errors. Examining Fig. 4 I plot, as a dashed line, the forecast for the rental rate one period after the technology shock (after beliefs have been updated) for a horizon of 1 to 40 quarters and as a solid line the actual evolution of r. Note this is not an impulse response function but the forecasted path of these variables based on information one period after the technology shocks and the actual evolution of the variables absent any additional technology shocks. One sees that the short run forecasts are smaller for the learning model in this paper versus the EP

[34] Note that $k_{t+1}$ is known at time t.

[35] See Eusepi and Preston (2011) and Evans and Honkapohja (2001) for more detail.

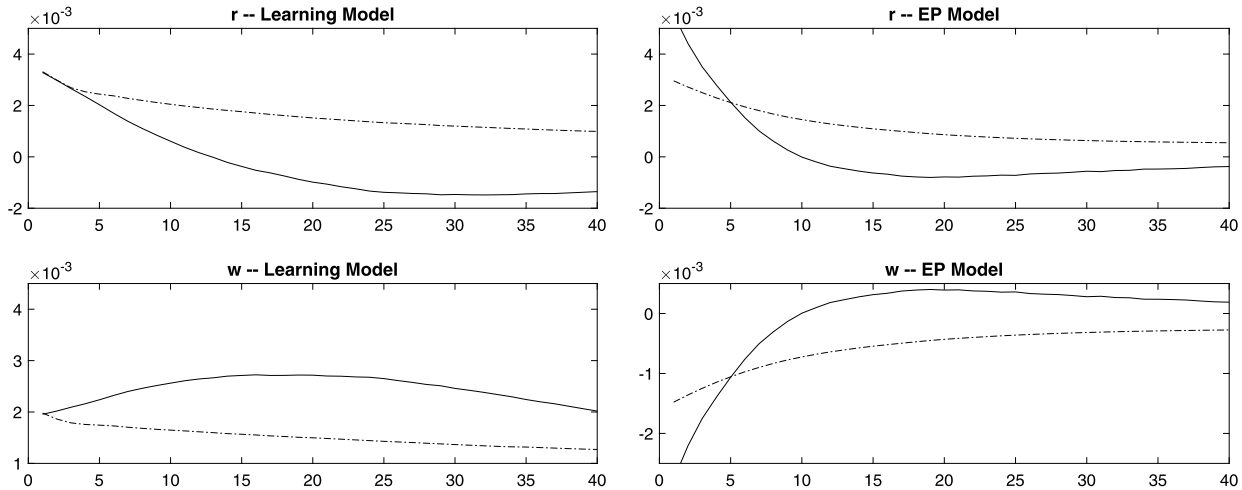


**Fig. 4.** Forecasts period after shock (dashed) vs. outcomes (solid): EP vs learning.

model. However, the long run forecast errors are larger in the learning model in this paper.

## 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.

## Appendix. Appendices - Not for publication

~

### 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 $Z_t = q_t^{\frac{\alpha}{1-\alpha}} A_t$ for $C_t$, $Y_t$, $I_t$, $W_t$ . $k_t = K_t/(Z_{t-1}q_{t-1})$, $\lambda_t = \Lambda_t/(Z_t^{-\sigma}q^{-1})$, $r_t = q_t R_t^k$ and $h_t = H_t$.

The consumption first order condition linearizes as:

$$-\sigma\hat{c}_t - \psi(1-\sigma)\hat{h}_t = \hat{\lambda}_t \tag{24}$$

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:

$$(1-\sigma)\hat{c}_t + \varepsilon_v\hat{h}_t = \hat{\lambda}_t + \hat{w}_t \tag{25}$$

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

$$\hat{r}_t^k = \hat{y}_t - \hat{u}_t - \hat{k}_t + \hat{\gamma}_t + \frac{1}{1-\alpha}\hat{\gamma}_t^q \tag{26}$$

$$\hat{w}_t = \hat{y}_t - \hat{h}_t \tag{27}$$

here $\hat{\gamma}_t = ln(\frac{A_t}{A_{t-1}}) - ln(1+g)$ and $\hat{\gamma}_t^q = ln(\frac{q_t}{q_{t-1}}) - ln(1+g^q)$

The production function linearizes as

$$\hat{y}_t = \alpha\hat{k}_t + \alpha\hat{u}_t + (1-\alpha)\hat{h}_t - \alpha\hat{\gamma}_t - \frac{\alpha}{1-\alpha}\hat{\gamma}_t^q \tag{28}$$

The capital evolution equation linearizes as

$$\hat{k}_{t+1} = (1-\delta)\tilde{g}\left(\hat{k}_t - \hat{\gamma}_t - \frac{1}{1-\alpha}\hat{\gamma}_t^q\right) + \frac{i}{k}\hat{i}_t - \theta\delta\tilde{g}\hat{u}_t \tag{29}$$

here $\frac{i}{k}$ is the steady state investment capital ratio and $\tilde{g} = (1+g)^{-1}(1+g^q)^{\frac{1}{\alpha-1}}$.

Finally the resource constraint linearizes as

$$\hat{y}_t = \frac{c}{y}\hat{c}_t + \frac{i}{y}\hat{i}_t \tag{30}$$

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:

$$ru = \frac{1}{\beta(1+g_\lambda)} - (1-\delta)$$

here $g_\lambda$ the growth rate of marginal utility is equal to $(1+g^q)^\phi(1+g)^{-\sigma}$ where $\phi = \frac{\alpha(1-\sigma)-1}{1-\alpha}$.

The factor price equation (8) yields:

$$\frac{Y}{K/q} = \frac{ru}{\alpha}$$

And the capital evolution Eq. (10) gives:

$$\frac{I}{K/q} = (1+\bar{g})(1+g^q) - (1-\delta)$$

where $(1+\bar{g}) = (1+g^q)^{\frac{\alpha}{1-\alpha}}(1+g)$ is the growth rate of Y, I and C. from the resource constraint

$$\frac{C}{K/q} = \frac{Y}{K/q} - \frac{I}{K/q}.$$

Then $\frac{c}{k} = \frac{C}{K/q}(1+g^q)^{\frac{1}{1-\alpha}}(1+g)$ , $\frac{i}{k} = \frac{I}{K/q}(1+g^q)^{\frac{1}{1-\alpha}}(1+g)$ and $\frac{y}{k} = \frac{c}{k} + \frac{i}{k}$ . Similarly, $\frac{y}{i} = \frac{Y/(K/q)}{I/(K/q)}$, $\frac{c}{y} = 1 - \frac{i}{y}$ and $\frac{c}{i} = \frac{c/y}{i/y}$.

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

$$\psi = \frac{wh}{c} = \frac{wh}{y}\frac{y}{c} = \frac{(1-\alpha)}{\alpha}r\left(\frac{C}{K/q}\right)^{-1}$$

To determine $\theta$ we use the fact that

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

## C Consumption equation

The derivation of the consumption equation follows Eusepi and Preston (2011) allowing for investment specific technology shocks. We begin with the log-linearized budget constraint:

$$\hat{k}_t = \tilde{\beta}(\frac{c}{k}\hat{c}_t + \hat{k}_{t+1} + \tilde{\beta}^{-1}\left(\frac{1}{1-\alpha}\hat{\gamma}_t^q + \hat{\gamma}_t\right) - \tilde{R}(\hat{r}_t + \frac{1-\alpha}{\alpha}(\hat{h}_t + \hat{w}_t)$$

$$\tilde{\beta} = \beta(1+g^q)^{\phi+\frac{1}{1-\alpha}}(1+g)^{1-\sigma}$$

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

Substituting for $\hat{h}_t$ in the budget constraint using the labor supply equation we arrive at:

$$\hat{k}_t = \tilde{\beta}(\varepsilon_c \hat{c}_t + \hat{k}_{t+1} + \tilde{\beta}^{-1}\left(\frac{1}{1-\alpha}\hat{\gamma}_t^q + \hat{\gamma}_t\right) - \tilde{R}\hat{r}_t - \varepsilon_w \hat{w}_t)$$

where

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

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

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

Iterating the budget constraint forward we get:

$$\varepsilon_c \hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t}\hat{c}_T = \tilde{\beta}^{-1}\hat{k}_t + \hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t}\left[\varepsilon_w \hat{w}_T + \tilde{R}\hat{r}_T - \tilde{\beta}^{-1}\left(\frac{1}{1-\alpha}\hat{\gamma}_T^q + \hat{\gamma}_T\right)\right]$$

Next combining the expressions for the log-linearized Euler equation and labor supply we get that:

$$\hat{E}_t \hat{c}_T = \frac{1}{1-\chi}[\hat{c}_t + \psi\frac{1-\sigma}{\sigma}\hat{h}_t + \hat{E}_t \sum_{k=t}^{T-1} \tilde{\beta}^{T-t}(\frac{\tilde{\beta}\tilde{R}}{\sigma}\hat{r}_{k+1} - \hat{\gamma}_{k+1} + \frac{\phi}{\sigma}\hat{\gamma}_{k+1}^q) - \chi\hat{w}_t)$$

where

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

Substituting this expression into the iterated budget constraint above and simplifying we arrive at:

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

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

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

$$+ \left[\tilde{\beta} - \frac{(1-\tilde{\beta})(1-\chi)}{\varepsilon_c}\right]\hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t}\hat{\gamma}_{T+1}$$

$$+ \left[-\frac{\tilde{\beta}\phi}{\sigma} - \frac{(1-\tilde{\beta})(1-\chi)}{\varepsilon_c(1-\alpha)}\right]\hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t}\hat{\gamma}^q{}_{T+1}$$

## 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 actual law of motion, or agents beliefs contain a unit root, i.e. for the stationary model the level of the variables is not I(0) *Not* I and for the non-stationary model the differences are not I(0) *Not* I, it is assumed that agents use the previous period's beliefs. Fig. 5 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.[36]

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 where adjustments happen more than 10% of the time is rare, occurring in about 1.6% of the sample. In only a few 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

[36] 53% to be exact.

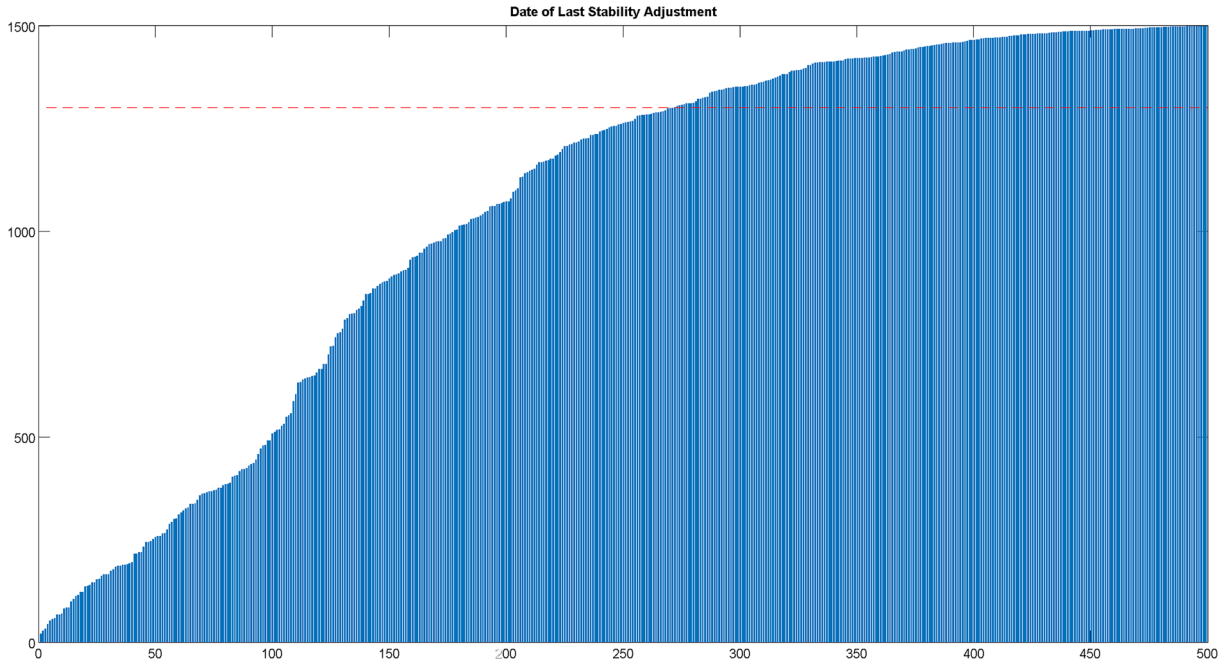


**Fig. 5.** Date of last stability adjustment.

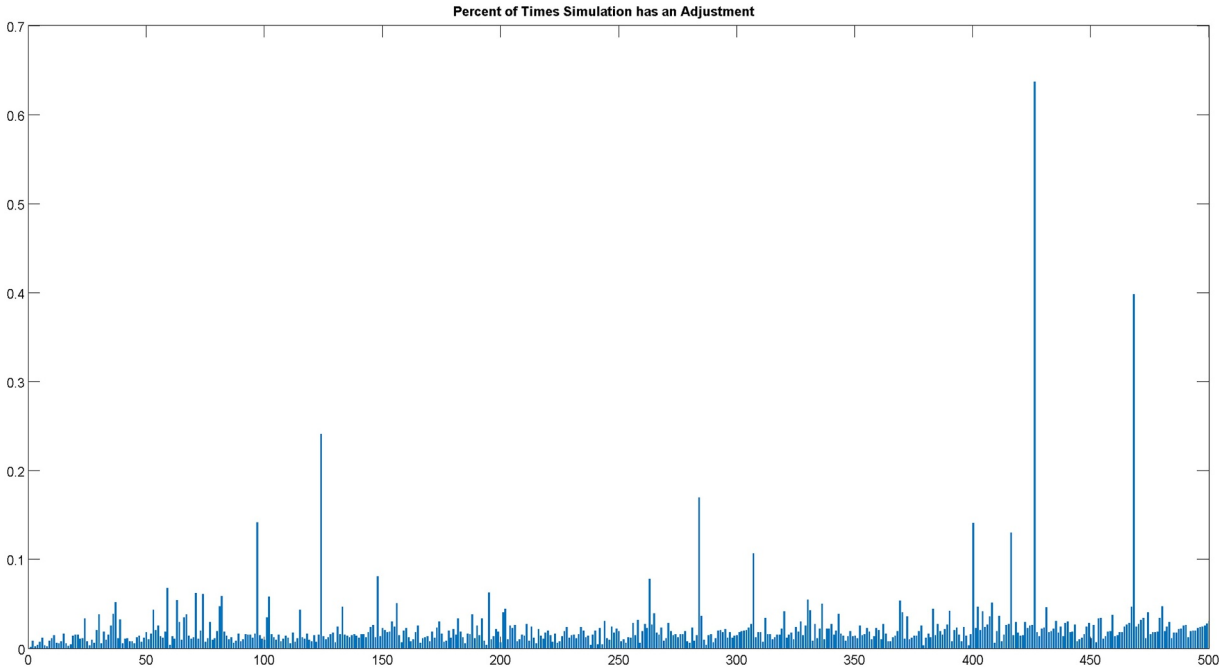


**Fig. 6.** Percent of times simulation has adjustment.

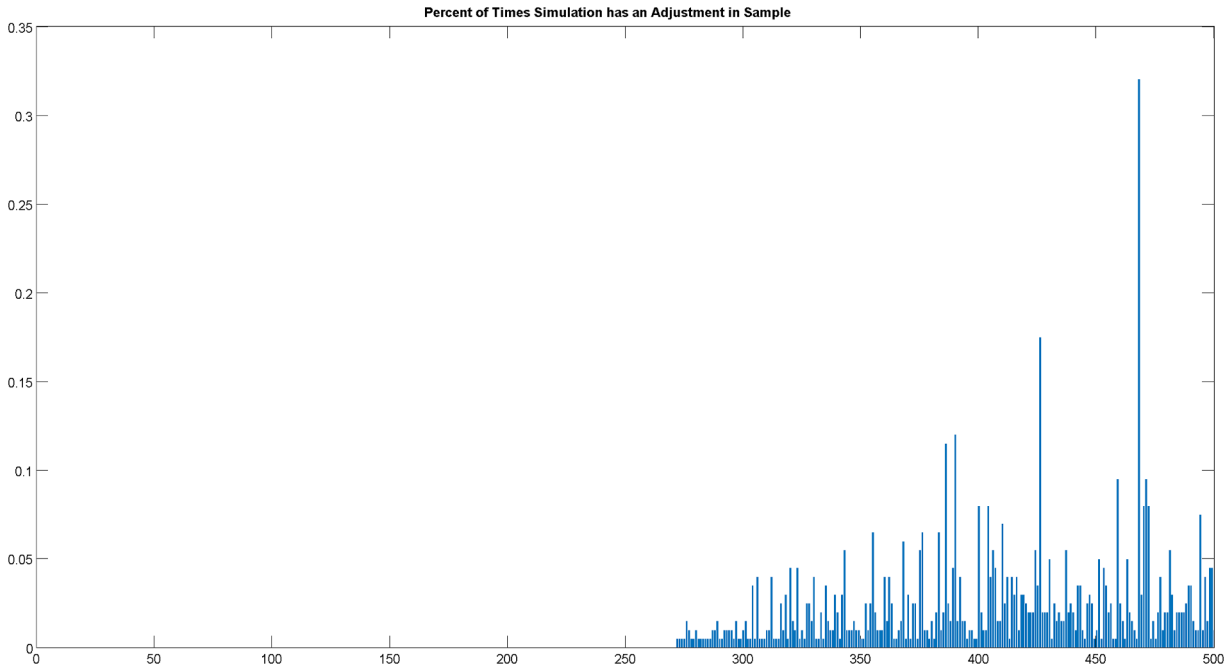


**Fig. 7.** Percent of times simulation has adjustment in sample.

statistics for each of the 500 simulations. Here the influence of the stability adjustment is small, only 4 out of the 500 samples used to calculate the statistics has a stability adjustment more than 10% of the time.

## References


Adam, Klaus, Marcet, Albert, Beutel, Johannes, 2017. Stock Price Booms and Expected Capital Gains. American Economic Review, American Economic Association 107 (8), 2352–2408 August.

Adam, K., Marcet, A., 2011. Internal rationality, imperfect market knowledge and asset prices. J. Econ. Theory 146 (3), 1224–1252.

Adam, K., Marcet, A., Beutel, J., 2017. Stock price booms and expected capital gains. Am. Econ. Rev. 107 (8), 2352–2408.

Alpanda, S., 2017. Regime-Switching productivity growth and Bayesian learning in real business cycles. Mimeo. University of Central Florida.

Angeletos, G.-M., La'O, J., 2013. Sentiments. Econometrica 81, 739–779.

Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2015. X-CAPM: an extrapolative capital asset pricing model. J. Financ. Econ. 115 (1), 1–24.

Basu, S., Fernald, J.G., Kimball, M.S., 2006. Are technology improvements contractionary? Am. Econ. Rev. 96 (5), 1418–1448.

Beaudry, P., Portier, F., 2004. An exploration into Pigou's theory of cycles. J. Monet. Econ. 51 (6), 1183–1216.

Bordalo, Pedro, Gennaioli, Nicola, Shleifer, Andrei, 2018. Diagnostic Expectations and Credit Cycles. Journal of Finance, American Finance Association 73 (1), 199–227 February.

Boz, E., Daude, C., Bora Durdu, C., 2011. Emerging market business cycles: learning about the trend. J. Monet. Econ. 58 (6), 616–631.

Branch, W., Evans, G.W., 2007. Model uncertainty and endogenous volatility. Rev. Econ. Dyn. 10 (2), 207–237.

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

Branch, W.A., Evans, G.W., 2006. A simple recursive forecasting model. Econ. Lett. 91 (2), 158–166.

Brock, W., Hommes, C., 1997. A rational route to randomness. Econometrica 65 (5), 1059–1096.

Brock, W.A., Hommes, C.H., 1998. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. J. Econ. Dyn. Control 22 (8–9), 1235–1274.

Bullard, J., Singh, A., 2012. Learning and the great moderation. Int. Econ. Rev. 53 (2), 375–397.

Burnside, C., 1996. Production function regressions, returns to scale, and externalities. J. Monet. Econ. 37 (1), 177–201.

Burnside, C., Eichenbaum, M., 1996. Factor-hoarding and the propagation of business-cycle shocks. Am. Econ. Rev. 86 (5), 1154–1174.

Case, K.E., Shiller, R.J., Thompson, A.K., 2012. What have they been thinking? Homebuyer behavior in hot and cold markets. Brook. Pap. Econ. Act. 45 (2 (Fall)), 265–315.

Case, K.E., Shiller, R.J., Thompson, A.K., 2012. What have they been thinking? homebuyer behavior in hot and cold markets. Brook. Pap. Econ. Act. 45 (2 (Fall)), 265–315.

Cho, I.-K., Kasa, K., 2017. Gresham'S law of model averaging. Am. Econ. Rev. 107 (11), 3589–3616.

Christiano, L.J., Eichenbaum, M., 1992. Current real-business-cycle theories and aggregate labor-market fluctuations. Am. Econ. Rev. 82 (3), 430–450.

Christiano, L.J., Eichenbaum, M., Evans, C.L., 1999. Monetary policy shocks: what have we learned and to what end? In: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics. Handbook of Macroeconomics 1. Elsevier, pp. 65–148.

Cochrane, J.H., 1988. How big is the random walk in GNP? J. Polit. Econ. 96 (5), 893–920.

Cogley, T., Nason, J.M., 1995. Output dynamics in real-business-cycle models. Am. Econ. Rev. 85 (3), 492–511.

Cogley, T., Sargent, T.J., 2005. The conquest of us inflation: learning and robustness to model uncertainty. Rev. Econ. Dyn. 8 (2), 528–563.

Diebold, F., Senhadji, A., 1996. Deterministic vs. stochastic trend in u.s. GNP, yet again. Am. Econ. Rev. 86 (5), 1291–1298.

Edge, R.M., Laubach, T., Williams, J.C., 2007. Learning and shifts in long-run productivity growth. J. Monet. Econ. 54 (8), 2421–2438.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica 50 (4), 987–1007.

Eusepi, S., Preston, B., 2011. Expectations, learning, and business cycle fluctuations. Am. Econ. Rev. 101 (6), 2844–2872.

Evans, G., Honkapohja, S., 2001. Learning and Expectations in Macroeconomics. Princeton University Press.

Evans, G.W., Honkapohja, S., Sargent, T.J., Williams, N., 2012. Bayesian model averaging, learning and model selection. CEPR Discussion Papers. C.E.P.R.
Evans, G.W., Ramey, G., 1992. Expectation calculation and macroeconomic dynamics. Am. Econ. Rev. 82 (1), 207–224.
Fuster, A., Hebert, B., Laibson, D., 2012. Natural expectations, macroeconomic dynamics, and asset pricing. NBER Macroecon. Annu. 26 (1), 1–48.
Gabaix, X., Laibson, D., Moloche, G., Weinberg, S., 2006. Costly information acquisition: experimental analysis of a boundedly rational model. Am. Econ. Rev. 96 (4), 1043–1068.
Grauwe, P.D., 2011. Animal spirits and monetary policy. Econ. Theory 47 (2), 423–457.
Grauwe, P.D., 2012. Lectures of Behavioral Macroeconomics. Princeton University Press, Princeton, New Jersey.
Greenwood, J., Hercowitz, Z., Krusell, P., 2000. The role of investment-specific technological change in the business cycle. Eur. Econ. Rev. 44 (1), 91–115.
Greenwood, Robin, Shleifer, Andrei, 2014. Expectations of Returns and Expected Returns. Review of Financial Studies, Society for Financial Studies 27 (3), 714–746.
Hamilton, J., 1994. Time Series Analysis. Princeton University Press.
Hansen, G.D., 1985. Indivisible labor and the business cycle. J. Monet. Econ. 16 (3), 309–327.
Huang, K., Liu, Z., Zha, T., 2009. Learning, adaptive expectations and technology shocks. Econ. J. 119 (536), 377–405.
Jaimovich, N., Rebelo, S., 2009. Can news about the future drive the business cycle? Am. Econ. Rev. 99 (4), 1097–1118.
Kahneman, D., Tversky, A., 1982. On the psychology of prediction. In: Kahneman, D., Slovic, P., Tversky, A. (Eds.), Judgment Under Uncertainty. Cambridge University Press, pp. 48–68.
King, R.G., Plosser, C.I., Rebelo, S.T., 1988. Production, growth and business cycles : I. the basic neoclassical model. J. Monet. Econ. 21 (2–3), 195–232.
Kocherlakota, Narayana R., 2010. Modern macroeconomic models as tools for economic policy. The Region, Federal Reserve Bank of Minneapolis, issue May, pp. 5–21.
Kreps, D., 1998. Anticipated Utility and Dynamic Choice, A 1997 Schwartz Lecture. Frontiers of Research in Economic Theory, Cambridge University Press.
Kuang, Pei, Mitra, Kaushik, 2016. Long-run growth uncertainty. Journal of Monetary Economics 79 (C), 67–80 Elsevier.
Kurz, M., Jin, H., Motolese, M., 2005. The role of expectations in economic fluctuations and the efficacy of monetary policy. J. Econ. Dyn. Control 29 (11), 2017–2065.
Lansing, K., 2009. Time varying u.s. inflation dynamics and the new Keynesian phillips curve. Rev. Econ. Dyn. 12 (2), 304–326.
Lorenzoni, G., 2009. A theory of demand shocks. Am. Econ. Rev. 99 (5), 2050–2084.
Milani, F., 2007. Expectations, learning and macroeconomic persistence. J. Monet. Econ. 54 (7), 2065–2082.
Milani, F., 2011. Expectation shocks and learning as drivers of the business cycle. Econ. J. 121 (552), 379–401.
Milani, F., 2014. Learning and time-varying macroeconomic volatility. J. Econ. Dyn. Control 47 (C), 94–114.
Nelson, C.R., Plosser, C.I., 1982. Trends and random walks in macroeconmic time series : some evidence and implications. J. Monet. Econ. 10 (2), 139–162.
Nimark, K.P., 2014. Man-bites-dog business cycles. Am. Econ. Rev. 104 (8), 2320–2367.
Orphanides, A., Williams, J.C., 2003. Imperfect knowledge, inflation expectations, and monetary policy. In: Bernake, B., Woodford, M. (Eds.), Inflation Targeting. Chicago University Press.
Primiceri, G.E., 2005. Time varying structural vector autoregressions and monetary policy. Rev. Econ. Stud. 72 (3), 821–852.
Rabin, M., 2002. Inference by believers in the law of small numbers. Q. J. Econ. 117 (3), 775–816.
Reinhart, C., Rogoff, K., 2009. This Time is Different: Eight Centuries of Financial Folly. Princeton University Press.
Rose, A.K., 1988. Is the real interest rate stable? J. Financ. 43 (5), 1095–1112.
Rotemberg, J., Woodford, M., 1996. Real business cycle models and the forcastable movements in output, hours and consumption. Am. Econ. Rev. 86 (1), 71–89.
Schmitt-Grohe, S., 2000. Endogenous business cycles and the dynamics of output, hours, and consumption. Am. Econ. Rev. 90 (3), 1136–1159.
Schmitt-Grohe, S., Uribe, M., 2012. What's news in business cycles. Econometrica 80, 2733–2764.
Shiller, R., 2005. Irrational Exuberence. Princeton University Press.
Slobodyan, S., Wouters, R., 2012. Learning in a medium-scale DSGE model with expectations based on small forecasting models. Am. Econ. J. Macroecon. 4 (2), 65–101.
Stock, J.H., 1991. Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series. J. Monet. Econ. 28 (3), 435–459.
Stock, J.H., Watson, M.W., 2003. Has the business cycle changed and why? NBER Macroeconomics Annual 2002. 17. National Bureau of Economic Research, Inc, pp. 159–230. NBER Chapters
Stock, J.H., Watson, M.W., 2005. Understanding changes in international business cycle dynamics. J. Eur. Econ. Assoc. 3 (5), 968–1006.
Tortorice, D.L., 2014. Credit constraints, learning, and aggregate consumption volatility. Macroecon. Dyn. 18 (02), 338–368.
Tversky, A., Kahneman, D., 1973. Availability: a heuristic for judging frequency and probabilities. Cognit. Psychol. 5, 207–232.
Vissing-Jorgensen, A., 2004. Perspectives on behavioral finance: does irrationality disappear with wealth? evidence from expectations and actions. NBER Macroeconomics Annual 2003. 18. National Bureau of Economic Research, Inc, pp. 139–208. NBER Chapters
Williams, N., 2003. Adaptive learning and business cycles. Technical Report. University of Wisconsin.
Winkler, F., 2016. The role of learning for asset prices and business cycles. Finance and Economics Discussion Series. Board of Governors of the Federal Reserve System (U.S.).
Woodford, M., 2003. Imperfect common knowledge and the effects of monetary policy. In: Philippe Aghion, J.S., Frydman, R., Woodford, M. (Eds.), Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps. Princeton University Press, pp. 25–58.