Credit Constraints, Learning and Aggregate Consumption Volatility*

Daniel L. Tortorice†

First Version: November 2007

This Version: February 2011

Abstract

This paper documents three empirical facts. First, the volatility of consumption growth relative to income growth rose from 1947-1960 and then fell dramatically by 50 percent from the 1960s to the 1990s. Second, the correlation between consumption growth and personal income growth fell by about 50 percent over the same time period. Finally, the absolute deviation of consumption growth from its mean exhibits one break in U.S. data, and the mean of the absolute deviations has fallen by about 30 percent. First, I find that a standard dynamic, stochastic general equilibrium model is unable to explain these facts. Then, I examine the ability of two hypotheses: a fall in credit constraints and changing beliefs about the permanence of income shocks to account for these facts. I find evidence for both explanations and the beliefs explanation is more consistent with the data. Importantly, I find that estimated changes in beliefs about the permanence of income shocks have significant explanatory power for consumption changes.

JEL Codes: E21, E32, D83

Keywords: Consumption, Business Fluctuations, Learning

*A previous version of this paper was titled "Income Process Uncertainty, Learning and Aggregate Consumption Volatility". I thank Robert Barro, John Campbell, Gita Gopinath, George Hall, David Laibson, N. Gregory Mankiw, Catherine Mann, Blake LeBaron, James Stock, Aleh Tsyvinski, two anonymous referees and attendees of the Harvard macroeconomics lunch seminar, Brandeis faculty workshop, and the Society for Economic Dynamics 2010 Annual Meeting for helpful comments, suggestions, and advice. I am grateful to the Jacob K. Javits Fellowship Program at the U.S. Department of Education, Harvard University, and the Brandeis University Theodore and Jane Norman Fund for financial support. All errors are mine.

†Department of Economics and International Business School, Brandeis University, tortoric@brandeis.edu

1 Introduction

This paper establishes three facts about aggregate U.S. consumption. First, after rising from the 1950s to the 1960s the standard deviation of consumption growth *relative* to income growth fell 40% from the 1960s to the 1970s and then another 30% from the 1980s to the 1990s. Finally it rose about 25% during the most recent recession. Second the correlation between consumption growth and income growth has fallen 50% over this same period. Finally, there is a structural break in the size of the absolute deviation of consumption growth from its mean. The break is in 1977 and the total fall in the absolute deviation of consumption growth from its mean is 30%.

These facts complement the Great Moderation facts by noting that consumption volatility has fallen more than may have been expected based on the fall in income volatility alone. This paper joins a growing literature suggesting a changing relationship between consumption and income. Stock & Watson (2002) estimate structural breaks in consumption. Dynan *et al.* (2006) and Dynan *et al.* (2009) note the marginal propensity to consume has fallen substantially. This paper however is the first to document all three facts outlined above and study two different hypothesis: a reduction in credit constraints and changing beliefs about the productivity process as explanations.

I model credit constraints as in Ludvigson (1999) where impatient consumers face a binding debt limit. I estimate a fraction of credit constrained consumers that varies over time and use this model to simulate a consumption series. I examine how well this series replicates the consumption facts above.

It is not obvious that reduced credit constraints reduces the volatility of consumption. If individuals want to smooth income shocks then reducing financial constraints makes consumption smoother. However, if individuals want to borrow to respond more than one to one with income shocks, for example if shocks are permanent, then reduced financial constraints could increase consumption volatility. I therefore model uncertainty as to the permanence of income shocks.

As pointed out by Deaton (1992), if income shocks are persistent but transitory consumption is less volatile than income. However if shocks are permanent consumption may be more volatile than income. Stock (1991) and Cochrane (1988) show that distinguishing between these two models in samples the length of the U.S. macroeconomic time series is very difficult. Since these two income processes are difficult to distinguish, one might expect consumers to have uncertainty as to which is the true model. Their beliefs may change over time resulting in time varying consumption to income volatility.[1]

To examine the plausibility of this hypothesis, I study a learning model based on Cogley & Sargent (2005). My model is a standard dynamic, stochastic, general equilibrium model where output fluctuations are driven by changes in productivity, with one modification. The agent believes the productivity process is non-stationary with some probability and stationary otherwise. At each point in time the individual first updates her beliefs about the parameters of the two models, and then, using Bayes's Rule, updates her beliefs about the probability that each model is true. Based on these beliefs she chooses her optimal level of consumption. From this model I can simulate a consumption series and examine the model's ability to capture the previously outlined facts.

I find that two benchmark DSGE models with a known productivity process fail to predict the fall in consumption volatility. The learning model, however, is able to partly explain the rise and fall in consumption volatility early in the sample and also the fall and rise in consumption volatility towards the end of the sample. The credit constraints model is consistent with the early rise and fall in consumption volatility, but cannot explain the magnitude of the fall in consumption volatility. Importantly, only a model with learning about the productivity process can replicate the estimated break. Finally, the learning model's implied probability weights on the different productivity processes have significant explanatory power for consumption changes.

The learning model also has an implication for recent economic events. During the

[1] Section 6.1 provides more motivation as to why consumers share this econometric uncertainty.

current recession, the weight on the non-stationary model has increased substantially; the model implies a large drop in consumption consistent with the data. This is perhaps one of the most interesting predictions of the learning model.

This paper relates to three strands of literature. First it relates to the literature on learning in macroeconomic models (Evans & Honkapohja (2001)). It shows how relaxing the strict rational expectations assumption in favor of learning results in a significantly different optimal policy. In my learning model, the two possible productivity processes are difficult to distinguish in small samples. Hence, learning about which process is true takes a long time. However, because consumption volatility is sensitive to deviations from stationarity, the processes imply different optimal policies. Learning then introduces a very different choice of consumption relative to the no-learning model.

The paper also relates to some recent research on consumption. Guvenen (2005) finds that introducing learning into a life-cycle consumption model changes many of the model's predictions. Specifically, optimal consumption choices are very different when one learns about the trend in income than when it is known. The paper also relates to Aguiar & Gopinath (2005). They explain the differences in consumption volatility relative to income volatility across countries. In their model, different countries have different ratios of consumption volatility to income volatility because they have different productivity processes. In my paper, the variance of consumption relative to income varies over time depending on the relative likelihood of two different productivity processes.

Finally, this paper relates to the Great Moderation literature. While an extensive review would be out of place, I note three closely related papers. Cecchetti *et al.* (2005) show that the reduction in consumption volatility was accompanied by a rise in U.S. debt levels, supporting a financial innovation explanation. Dynan *et al.* (2006) show that marginal propensities to consume have fallen over time. Cecchetti *et al.* (2006) relates changes in consumption volatility across countries to changes in estimated fractions of rule-of-thumb consumers. This paper differs for three reasons. First, the question is different. I ask if financial innovation can account for the fact that consumption volatility has declined more than personal income volatility in U.S. data. Secondly, I simulate a model of credit constraints and test the model directly to see if it is consistent with the decline in consumption volatility. Finally, I solve an innovative learning model and it explains features of the data that the credit constraints model does not.

The rest of the paper proceeds as follows. Section two describes the data I use. Section three describes the empirical facts I attempt to explain. Four describes the credit constraint and learning models. Five tests these models and section six provides more intuition for the learning model's predictions. Seven examines robustness to different parameter choices. Section eight concludes.

2 Data

Data come from the National Income and Product Accounts (NIPAs). The consumption data are aggregate consumption of nondurable goods and services and income is the personal income series. Data begin in 1947 and end in 2010:1[2]. Series are transformed into a real, per-capita series by deflating with the nondurable consumption deflator, the service consumption deflator, and the GDP deflator respectively, and dividing by population.

I pay particular attention to the distinction between accrued wages and distributed wages. In general, firms distribute most of the wages accrued or earned within a quarter. However, at times firms distribute wages early or late. Occasionally this difference is large. For example, in 1992:Q4 firms distributed 63 billion dollars more in wages than employees earned, and in 1993:Q1 employees earned 72.1 billion dollars more in wages than firms distributed. These

[2] As usual in the literature, consumption analysis is carried out using only data from 1955 onwards avoiding the Korean War. Also, as Campbell & Mankiw (1990) point out the National Service Life Insurance (NSLI) benefits paid to WWII veterans in 1950:1 distort tests of the permanent income hypothesis.

In theory 1950:1 could effect learning about the productivity process. To investigate this possibility, I re-simulated the model removing personal current transfer receipts, the component of income including the insurance payment. Results were very similar. Since the 1950:1 observation does not influence the results, I use all the income data to estimate the processes. Because beliefs about the productivity process are central to this paper, I tie them to the data as much as possible.

discrepancies appear due to increase in income tax rates.[3] These distributions are large enough to affect learning about the productivity process. They appear as large shocks to income though they do not represent uncertainty. As a result, I assign to income all wages earned in a quarter regardless of how much is distributed. However, this modification does not effect the conclusions of the paper.

3 Empirical Facts

This section establishes three facts about the aggregate U.S. consumption series. First, the relative volatility of consumption to income rose early in the sample and then fell substantially over the last 45 years. Second, the correlation of consumption growth and income growth has fallen over time. Finally, there is a substantial break in the absolute deviation of consumption growth from its mean, falling about 30% over the last 60 years.

3.1 Consumption Volatility Relative to Income Volatility

In Figure 1, I plot the rolling standard deviation of consumption growth divided by the rolling standard deviation for income growth. I use a window over the next 10 years and nondurables and services consumption.[4] Results are robust to varying the window from 20 to 50 quarters. As one can clearly see, consumption variance relative to income variance rises over time reaching a peak around 1960 and then falls, quite dramatically by 40% from 1960 to 1970. It remains constant for about 20 years before beginning to fall again in 1990 by an additional 30%. During the most recent recession it increased by 25%. I obtained similar results (omitted) with nondurables alone, finding a similar, but more gradual decline

[3] I thank Kurt Kunze, an economist at the Bureau of Economic Analysis, for this insight.

[4] Standard errors are calculated using the GMM delta method formulas. See, for example, Cochrane (2001), pp. 207. In short, let $\mu = [E(\Delta \ln c_t)^2 \;\; E(\Delta \ln c_t) \;\; E(\Delta \ln y_t)^2 \;\; E(\Delta \ln y_t) \;\;]'$ and $\phi(\mu) = \frac{\sqrt{E(\Delta \ln c_t)^2 - [E(\Delta \ln c_t)]^2}}{\sqrt{E(\Delta \ln y_t)^2 - [E(\Delta \ln y_t)]^2}}$. Then $var(\phi) = \frac{1}{T}\left[\frac{d\phi}{d\mu}\right]' \sum_{j=-\infty}^{j=\infty} cov(x_t, x_{t-j}) \left[\frac{d\phi}{d\mu}\right]$ where $x_t = \left[(\Delta \ln c_t)^2 \;\; (\Delta \ln c_t) \;\; (\Delta \ln y_t)^2 \;\; (\Delta \ln y_t)\right]$. I estimate the covariance matrix using a Newey-West estimator with 5 lags.

beginning in 1960 and ending in 1980.

3.2 Time Varying Response of Consumption to Income

The data demonstrate that consumption changes appear to have moderated over time relative to income changes. To get a parametric representation of this fact I estimate a time varying coefficient in the following regression

$$\Delta \ln c_t = \alpha + \beta_t \Delta \ln y_t + \varepsilon_t \tag{1}$$

where c_t is the NIPA value of real, per-capita, nondurable and services consumption and y_t is real, per-capita, personal income. I let β_t take the form $\beta_t = \beta_0 + \beta_1 t$.[5] I found that higher order terms (quadratic and cubic) were not significant. Consumption and income growth are measured as annualized percentage changes. The results, in column 1 of table 1, indicate that there is a significant time varying component to β as β_1 is statistically significant and negative. At the beginning of the sample period $\beta_t = 0.55$; by the end of the sample period β_t is around 0.276. This result again shows a considerable moderation of consumption *relative* to personal income. While income has moderated over time, consistent with the Great Moderation facts, consumption has moderated even more.

This approach is one way of measuring a marginal propensity to consume. Dynan *et al.* (2009) estimate a marginal propensity to consume controlling for potential output, interest rates, the unemployment rate and the index of consumer sentiment. They also find a substantial fall in the response of consumption to income.

[5] Another approach, estimating β_t with the Kalman filter, results in a similar conclusion. β_t declines approximately linearly from 0.44 to 0.21. I prefer this deterministic approach because I need to make less restrictive assumptions about the nature of the error terms and statistical inference concerning β_t is straightforward.

3.3 Breaks in the Variance of Consumption

Since consumption has moderated more than income volatility it is natural to examine how much consumption has moderated. To investigate, I follow the methodology of Stock & Watson (2002) and estimate a break in the mean of the absolute value of the residual from the regression of nondurable consumption growth on a constant.[6]

$$\Delta \ln c_t = \alpha + \eta_t \tag{2}$$

$$|\eta_t| = \beta + \beta_1 \tau + \varepsilon_t \tag{3}$$

To estimate τ (the break dates) I use the methodology of Bai & Perron (1998) and the algorithms and GAUSS code available in Bai & Perron (2003). The methodology, in addition to providing a feasible way to estimate the structural break model, describes a sequential method for estimating the number of breaks (allowing for the possibility of more than one break). These methods consistently estimate the number of breaks and the proportion of the sample that occurs before the break date occurs.

To estimate the number of breaks, Bai & Perron (1998) recommend the sequential application of a test of l breaks versus the alternative of $l+1$ breaks, called the $\sup F(l+1|l)$ test. The procedure first tests one break versus the alternative of zero breaks, and then two versus one and so on until the statistic fails to reject the null hypothesis. In addition, Bai and Perron provide a test statistic, called the UD max statistic, that tests the hypothesis of zero breaks against the alternative of $k > 0$ breaks where k is unknown. The estimates of the break dates are those that minimize the sum of squared residuals in equation (3).

Table 2a presents the results from these tests[7]. Consumption growth is measured as an

[6] I chose not to include lag regressors in the regression equation below so that the interpretation of the residual is clearer. I am looking to see if deviations in consumption about its mean have changed in magnitude over time.

[7] Estimated standard errors and confidence intervals allow for heterogenity and autocorrelation in the residuals using Andrews (1991) automatic bandwidth with AR(1) approximation and a quadratic kernel. The residuals are AR(1) pre-whitened and the variance-covariance matrix is allowed to vary across segments. See Bai & Perron (2003) for details.

annualized percentage change. The consumption series has one break: at 1977:4. The means for each segment are 2.65 and 1.87. This implies a moderation in consumption volatility of 30%.

4 Explanations

This section outlines the two models I use to explain the fall in consumption volatility relative to income volatility. The first model is a model of credit constraints with a time varying fraction of credit constrained consumers. The second is a model of learning where the true productivity process is unknown and therefore agents' beliefs about the true process are changing over time.

4.1 Credit Constraints

This section of the paper outlines a model with time varying credit constraints as an attempt to explain the previously described empirical facts. Consumption may be very volatile early in the sample because individuals are credit constrained but become less volatile as more assets are accumulated or better consumption insurance becomes available. To examine this explanation I assume that aggregate consumption comes from two types of agents: an unconstrained agent and a constrained agent.

$$\Delta c_t = \lambda_t \Delta c_t^{cc} + (1 - \lambda_t)\Delta c_t^{uc} \tag{4}$$

here λ_t is the fraction of the population that is credit constrained, Δc_t^{cc} is the change in consumption of the credit constrained consumers and Δc_t^{uc} is the consumption change of the unconstrained consumer. I model the constrained consumer as in Ludvigson (1999). In this model the agent chooses consumption to maximize

$$E_t \sum_{i=0}^{\infty} \beta^i \left(\frac{C_{t+i}^{1-\gamma}}{1-\gamma} \right) \tag{5}$$

subject to $D_{t+1} = (1+r)(D_t + C_t - Y_t)$ where D_t is consumer debt, r is the real interest rate, and Y_t is labor income. Labor income evolves according to $\Delta \ln Y_t = g + \eta_t - \psi \eta_{t-1}$, where η_t is i.i.d. Consumers are credit constrained, they can borrow only up to $\overline{D}_{t+1} = \frac{1}{w} Y_t \exp(\xi_t)$ where $\xi_{t+1} = \phi \xi_t + u_{t+1.}$ Therefore, maximum consumption at time t is given by $X_t = \frac{\overline{D}_{t+1} - D_t(1+r)}{(1+r)} + Y_t$.

If we define $w_t = X_t / Y_t$ and $z_{t+1} = Y_{t+1}/Y_t$ we can solve for the policy function $\theta_t(w_t, \eta_t, \xi_t) = C_t / Y_t$, as a function of the state variables, by iterating on the Euler equation. Defining the marginal utility of consumption function $p_t(w_t, \eta_t, \xi_t) = v(\theta_t) = \theta_t^{-\gamma}$ Ludvigson (1999) shows that:

$$p_t(w_t, \eta_t, \xi_t) = \max \left[v(w_t), \beta^* E_t z_{t+1}^{-\gamma} p(w_{t+1}, \eta_{t+1}, \xi_{t+1}) \right] \tag{6}$$

where $w_{t+1} = 1 + z_{t+1}^{-1}(1+r)(w_t - v^{-1} p_t(w_t, \eta_t, \xi_t)) + \frac{\exp(\eta_{t+1} - \psi \eta_t + \xi_{t+1} - \xi_t + g) - (1+r)}{(1+r)\exp(\eta_{t+1} - \psi \eta_t + g + \ln \omega - \xi_t)}$ and $\beta^* = (1+r)\beta$. I solve for $p_t(w_t, \eta_t, \xi_t)$ by iterating on (6) using 5 grid points for η and ξ. I discretize the i.i.d. η_t and the AR(1) ξ_t using the method of Tauchen & Hussey (1991).[8] Finally, I use 100 grid points for w and linear interpolation to calculate the function between grid points of w.

To calculate the predicted aggregate consumption change I need to estimate λ_t. To do so let $\Delta c_t^{uc} = \mu + \sigma r_t + \varepsilon_t$ where r_t is the ex-post real interest rate and ε_t is i.i.d. since the unconstrained consumer does not respond to predictable variation in income and let $\Delta c_t^{cc} = \alpha + \sigma^c r_t + mpc_t \Delta y_t^p + \nu_t$ where v_t is i.i.d., Δy_t^p is the predicted change in income, and mpc_t is the credit constrained individual's marginal propensity to consume out of predictable

[8] I use the Matlab code provided by Adda & Cooper (2003) to implement this method.

income estimated from the model[9]. Then

$$\Delta c_t = \lambda_t[\alpha + \sigma^c r_t + mpc_t \Delta y_t^p + \nu_t] + (1 - \lambda_t)[\mu + \sigma r_t + \varepsilon_t] \tag{7}$$

or:

$$\Delta c_t = \mu + \sigma r_t + \lambda_t[mpc_t \Delta y_t^p + (\sigma^c - \sigma)\, r_t + \alpha - \mu] + \eta_t \tag{8}$$

To estimate this equation I first calculate Δy_t^p as the fitted value of a regression of Δy_t on the lagged income growth rate, the lagged change in the S&P 500 index, and the lag ex-post real interest rate, then I multiply by the estimated mpc_t. Since the ex-post real interest rate is also correlated with η_t, I instrument its value with the same instruments used for income growth to obtain $\widehat{r}_t$. The final equation I estimate is:

$$\Delta c_t = \mu + \sigma \widehat{r}_t + \lambda_t[mpc_t \Delta y_t^p + (\sigma^c - \sigma)\, \widehat{r}_t + \alpha - \mu] + \eta_t \tag{9}$$

I estimate the time varying λ_t using the Kalman filter (see Hamilton (1994) pp. 399 - 402). Figure 2 plots the resulting series for the estimated fraction of consumers who are credit constrained. Credit constraints rise in the 1960s with a peak of 80% credit constrained. Credit constraints fall non-monotonically to a low of 10% in 2010, though most of that fall is in the last few years. As late as 2005, 40% of the population was still estimated to be credit constrained.

I follow Campbell & Mankiw (1990) and estimate the fraction of credit constrained consumers using data from only 1955 onwards. This avoids the Korean War period and some large income swings due to government transfer payments. To estimate credit constraints pre 1955, I assume that credit constraints do not change from 1947 to 1955. This assumption is only for completeness and does not affect the results.

[9] To calculate mpc_t I simulate an series of $\eta_{i=0}^{i=t}$ and an alternative series which is the same except $\eta_{t-1} = \eta'_{t-1}$. Letting $\Delta c_t = \Delta c_t|\eta_{t-1}$, $\Delta c'_t = \Delta c_t|\eta'_{t-1}$, $\Delta y_t^p = E_t[\Delta y_t|\eta_{t-1}]$ and $\Delta y_t^{p'} = E_t[\Delta y_t|\eta'_{t-1}]$ the $mpc_t = \frac{\Delta c_t - \Delta c'_t}{\Delta y_t^p - \Delta y_t^{p'}}$.

4.2 The Learning Model

This section describes a model economy where the representative agent maximizes consumption while also investing in a capital stock. The agent combines their capital with their labor to produce output. Productivity grows at a constant rate g but is subject to stochastic shocks. The agent is learning if the correct model for the productivity shocks is one with permanent or temporary deviations from trend and also learns about parameters of each productivity model.

4.2.1 Model Description

For the learning model, the representative agent maximizes:

$$\max_{\{c_{t+j},l_{t+j},i_{t+j}\}_{j=0}^{\infty}} E_t \sum_{j=0}^{\infty} \beta^j \frac{[c_{t+j}(1-(\theta l_{t+j})^{\phi})]^{1-\gamma}}{1-\gamma} \tag{10}$$

subject to with probability $p = p_{s,t}$:

$$\begin{aligned} k_{t+1} &= (1-\delta)k_t + i_t && (11) \\ k_t^{\alpha}(A_0(1+g)^t e^{z_t} l_t)^{1-\alpha} &= c_t + i_t && (12) \\ z_t &= \theta_1^s z_{t-1} + ... + \theta_p^s z_{t-p} + \varepsilon_t^s && (13) \end{aligned}$$

and with probability $p = p_{ns,t} = 1 - p_{s,t}$:

$$\begin{aligned} k_{t+1} &= (1-\delta)k_t + i_t && (14) \\ k_t^{\alpha}(A_0(1+g)^t e^{z_t} l_t)^{1-\alpha} &= c_t + i_t && (15) \\ \Delta z_t &= \theta_1^{NS} \Delta z_{t-1} + ... + \theta_p^{NS} \Delta z_{t-p+1} + \varepsilon_t^{ns} && (16) \end{aligned}$$

Here k_t is beginning of the period capital stock, y_t is output, c_t is consumption, i_t is investment, l_t is labor supply, and $A_0(1+g)^t e^{z_t}$ is the productivity level. With probability

$p_{s,t}$ the log deviation of productivity from its trend is given by (13) and is therefore stationary around the time trend, with probability $p_{ns,t}$ the log deviation of productivity from its trend is given by the unit root process (16) and hence exhibits permanent deviations from trend.[10,11]

4.2.2 Updating of Beliefs

To choose consumption, the agent each period must update his beliefs concerning the probability that each model is true. I update beliefs as in Cogley & Sargent (2005).

There are two models of the productivity process indexed $i = s, ns$ which can be written in regression form $y_t = x_t'\theta_t + \varepsilon_t$. For the stationary model the regression equation is (13) for the non-stationary model the equation is (16).

Letting Z^t represent the joint history of Y_t and X_t up to time t, the agents prior beliefs on the parameters for each model are given by:

$$p(\theta|\sigma^2, Z^{t-1}) = N(\theta_{t-1}, \sigma^2 P_{t-1}^{-1}) \tag{17}$$

$$p(\sigma^2|Z^{t-1}) = IG(s_{t-1}, v_{t-1}) \tag{18}$$

θ_{t-1} is the parameter estimate based on $t-1$ data, $\sigma^2 P_{t-1}^{-1}$ is the estimate of the variance-covariance matrix of $\widehat{\theta}_{t-1}$, s_{t-1} is the residual sum of squares and v_{t-1} is the degrees of freedom for estimating variance of the residuals $(t-k)$. IG is the inverse gamma distribution. I follow Cogley & Sargent (2005) and use the normal and inverse gamma distributions because these allow one to update the parameters recursively and use analytical formulas to update the posterior beliefs. Stepping out of the normal-inverse gamma family would involve computationally intensive simulations at each point in time to calculate the likelihood (23) of each model.

[10] Modeling productivity as a deviation from a hp-filter trend would render productivity clearly stationary. And the agent would learn that productivity is stationary very quickly.

[11] While the productivity process is non-stationary with some probability, the stability conditions for the solution method still hold. See section 4.2.3 for details.

Maximum likelihood implies the parameters are updated recursively according to

$$
\begin{aligned}
P_t &= P_{t-1} + x_t x_t' && (19)\\
\theta_t &= P_{t-1}^{-1}(P_{t-1}\theta_{t-1} + x_t y_t) && (20)\\
s_t &= s_{t-1} + y_t' y_t + \theta_{t-1}' P_{t-1}\theta_{t-1} - \theta_t' P_t \theta_t && (21)\\
v_t &= v_{t-1} + 1 && (22)
\end{aligned}
$$

Next the agent updates the probability weights on each model. For each model the marginalized likelihood is:

$$m_{it} = \int\int \prod_{s=1}^{t} p(y_s|x_s, \theta_i, \sigma_i^2) p(\theta_i, \sigma_i^2) d\theta_i d\sigma_i^2 \tag{23}$$

and the probability weight is $w_{it} = m_{it} p_{i,0}$, where $p_{i,0}$ is the prior probability on model i.

To calculate the marginalized likelihood Cogley & Sargent (2005) note that Bayes's rule implies, for any θ_i, σ_i^2 :

$$
\begin{aligned}
p(\theta_i, \sigma_i^2|Z_t) &= \frac{\prod_{s=1}^{t} p(y_s|x_s, \theta_i, \sigma_i^2) p(\theta_i, \sigma_i^2)}{m_{it}} && (24)\\
m_{it} &= \frac{\prod_{s=1}^{t} p(y_s|x_s, \theta_i, \sigma_i^2) p(\theta_i, \sigma_i^2)}{p(\theta_i, \sigma_i^2|Z_t)} && (25)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\frac{w_{i,t+1}}{w_{i,t}} &= \frac{m_{i,t+1} p_{i,0}}{m_{i,t} p_{i,0}} && (26)\\
&= p(y_{t+1}|x_{t+1}, \theta_i, \sigma_i^2) \frac{p(\theta_i, \sigma_i^2|Z_t)}{p(\theta_i, \sigma_i^2|Z_{t+1})} && (27)
\end{aligned}
$$

As Cogley & Sargent (2005) show, this expression can be evaluated analytically since:

$$
\begin{aligned}
p(y_{t+1}|x_{t+1}, \theta_i, \sigma_i^2) &= N(y_{t+1} - x_{t+1}^{'}\theta_i, 0, \sigma_i) &(28)\\
p(\theta_i, \sigma_i^2|Z_{t+1}) &= N(\theta_i, \widehat{\theta}_{t+1}, \sigma_i P_{t+1}^{-1}) IG(\sigma_i, s_{t+1}, v_{t+1}) &(29)
\end{aligned}
$$

where N and IG indicate the normal and inverse gamma probability density functions respectively. There are analytical expressions for both distributions which are available in Cogley & Sargent (2005). Finally, to get the actual probabilities for each model one normalizes the weights to sum to one:

$$
\begin{aligned}
p_{NS,t} &= \frac{w_{NS,t}}{w_{NS,t} + w_{S,t}} &(30)\\
p_{S,t} &= 1 - p_{NS,t} &(31)
\end{aligned}
$$

4.2.3 Model Solution

Now that we can calculate the agent's beliefs at each point in time, I can solve for the optimal consumption, investment and labor supply plans. I first normalize the trending variables by the non-stochastic level of productivity and then take a linear-quadratic approximation of the utility function around the non-stochastic steady state. Subsequently, I can cast the model as a stochastic optimal linear regulator problem and solve it using standard methods (e.g. Chapter 5 of Ljungovist & Sargent (2004) and Cogley & Sargent (2005)).[12]

Letting $\widehat{x}_t = \frac{x_t}{(1+g)^t}$ (normalizing $A_0 = 1$) and substituting the resource constraint into the utility function we can rewrite the objective as:

$$
\max_{\{l_{t+j}, \widehat{i}_{t+j}\}_{j=0}^{\infty}} E_t \sum_{j=0}^{\infty} \beta^{*j} \frac{[\left(\widehat{k}_{t+j}^{\alpha}(e^{z_t} l_{t+j})^{1-\alpha} - \widehat{i}_{t+j}\right)(1 - (\theta l_{t+j})^{\phi})]^{1-\gamma}}{1-\gamma} \tag{32}
$$

[12] The solution method requires that the eigenvalues of $A - Bf$ where the optimal policy is given by $u = -fx$ (where x is the state vector), are less than $\frac{1}{\sqrt{\beta}}$ in modulus. The computer program checks this condition and finds it to hold. The matricies A and B are defined below.

where $\beta^* = \beta(1+g)^{1-\gamma}$ and rewrite the capital evolution equation as:

$$(1+g)\widehat{k}_{t+1} = (1-\delta)\widehat{k}_t + \widehat{i}_t \tag{33}$$

Next, I take a linear-quadratic approximation of the utility function. See the appendix for details.

Following Cogley & Sargent (2005) I write each sub-model (i.e. the model for each possible productivity process) as a linear regulator problem and then stack the model matrices to cast the complete model as an optimal linear regulator.

For the stationary model, the state variables are $x_t^S = [1\ z_t\ k_t\ z_{t-1}\ ...\ z_{t-p+1}]'$ and the control variables are $u_t = [l_t\ i_t]'$. For the non-stationary model, the state variables are x_t^{NS} $= [1\ z_t\ k_t\ \Delta z_t\ \Delta z_{t-1}\ ...\ \Delta z_{t-p+2}]'$ and the control variables are $u_t = [l_t\ i_t]'$. Then for model $i = \{s,\ ns\}$, I can write the problem in the linear regulator format:

$$U_t^i = \max_{\{u_{t+j}\}_{j=0}^{\infty}} E_t \sum_{j=0}^{\infty} \beta^{*j}[x_{t+j}^{i\prime} R^i x_{t+j}^i + u_{t+j}' Q u_{t+j} + 2x_{t+j}^{i\prime} W^i u_{t+j}] \tag{34}$$

subject to

$$x_{t+1}^i = A_t^i x_t^i + B^i u_t + C^i \varepsilon_{t+1}^i \tag{35}$$

the matrices R^i , Q^i, W^i, A_t^i, B^i, C^i are given in the appendix.

For the complete model the agent seeks to maximize[13]

$$\max_{\{u_{t+j}\}_{j=0}^{\infty}} p_{s,t} E_t U_t^s + (1-p_{s,t}) E_t U_t^{ns} \tag{36}$$

We can write this as an optimal control problem by stacking the matrices of the submodels.

[13] I take the approach of Kreps (1998) and Sargent (2001) of assuming that the agent treats its probability weights as constants when calculating the optimal policy. This "anticipated utility" approach is the most common in the learning literature and necessary to have a tractable solution to the problem.

$$\max_{\{u_{t+j}\}_{j=0}^{\infty}} E_t \sum_{j=0}^{\infty} \beta^{*j}[x'_{t+j}Rx'_{t+j} + u'_{t+j}Qu_{t+j} + 2x'_{t+j}Wu_{t+j}] \tag{37}$$

subject to

$$x_{t+1} = A_t x_t + Bu_t + C\varepsilon_{t+1} \tag{38}$$

where $x_t = [x_t^{s\prime}\ x_t^{ns\prime}]'$, $R = \begin{bmatrix} p_{s,t}R^S & 0 \\ 0 & p_{ns,t}R^{ns} \end{bmatrix}$, $W = \begin{bmatrix} p_{s,t}W^s \\ p_{ns,t}W^{ns} \end{bmatrix}$, $A_t = \begin{bmatrix} A_t^S & 0 \\ 0 & A_t^{NS} \end{bmatrix}$, $B = \begin{bmatrix} B^s \\ B^{NS} \end{bmatrix}$, $C = \begin{bmatrix} C^s \\ C^{NS} \end{bmatrix}$.

5 Testing the Models Empirically

5.1 Credit Model Simulation

To simulate a consumption series from the credit constraint model, I first need to parametrize the model. I use the Ludvigson (1999) calibration adjusting for quarterly data: $\psi = 0.44$, $r = 0.03/4$, $\beta = \frac{1}{1+0.15/4}$, $\phi = 0.6$, $\omega = 15$, $\gamma = 2$, $\sigma^u = 0.0025$ and $\sigma^y = 0.025$. I next need to simulate the time path for household income. I first assume an economy with 5,000 households all described by the credit constraint model. As in Ludvigson (1999), the household income process is decomposed into an aggregate shock and two idiosyncratic shocks as in: $\eta_t - \psi\eta_{t-1} = \varepsilon_{1,t} + \chi\varepsilon_{1,t-1} + \varepsilon_{2,t}^h + \varepsilon_{3,t}^h - \varepsilon_{3,t}^h$. Here ε_1 is the aggregate shock and ε_2^h and ε_3^h are household specific shocks. We need two household shocks to match observed positive autocorrelation in aggregate income and negative autocorrelation in household income. I estimate χ and the ε_1 time series from an ARIMA(0,0,1) model on the real per-capita personal income series. The ARIMA estimation also gives the growth rate of income $g = 0.0052$. I then estimate the variances of the idiosyncratic shocks by matching the variance of the household income series and its autocorrelation. I then simulate the household shocks as random draws from a normal distribution. Finally, to get a time series for ξ_t (the shock to the debt limit), I draw a random chi from the grid according to the probabilities given by the method of Tauchen & Hussey (1991).

Once I have each household's consumption series I can sum across households to obtain an aggregate consumption series. I match this to the data by scaling up the mean of this series to the mean of real per-capita non-durables and services consumption in the data. Then I combine consumption changes from this scaled series with consumption changes from the learning model (described below) using equation (4) to generate the finalized consumption series from the credit constraint model. I generate two series: a credit series which combines consumption from the credit constraint model with consumption from the learning model when individuals believe that the productivity process is stationary and a credit and learning series which combines consumption from the credit constraint model with consumption from the learning model when individual's beliefs change over time.

5.2 Learning Model Simulation

The calibration of the learning model is fairly standard: $\beta = 0.99$, $\alpha = \frac{1}{3}$, $\delta = 0.025$, and $\gamma = 2$. A higher γ leads to more volatile consumption and I choose $\gamma = 2$ to better match the volatility of consumption. Finally, I choose θ and ϕ so the agent inelastically spends 1/3 of his time working. This assumption, while not necessary to generate the observed learning dynamics, allows me to choose productivity so that income in the model exactly matches income in the data as I describe below. It also allows us to interpret l_t as potential labor supply and therefore an increase in unemployment would be seen as a fall in productivity. Hence, in the current recession model productivity is below trend while some measures of labor productivity in the data are above trend. This assumption is also realistic, I believe, in the sense that consumers view increases in unemployment as the economy inefficiently using its resources.

To construct the productivity series which the agent learns about, I take the real, per-capita personal income series[14] and regress $\ln y_t = a_0 + gt + \varepsilon_t$. Growth in the model is set to g and I calculate $\widehat{y}_t = \ln y_t - a_0 - gt$. I then choose z_t so that model output $y_t^{\text{mod}} = (1+\widehat{y}_t)y^{ss}$.[15] This ensures that income in the model matches income in the data allowing me to easily combine this model with the credit constraint model.[16]

One way to set the prior beliefs would be to run a regression on the first few quarters of data and set the priors on θ_0 and P_0 based on the regression estimates. However, this approach leads to unstable estimates of the stationary model, where productivity is predicted to go off to infinity. Instead, I take a different approach and assume that the agent has some knowledge of the productivity process even though the quarterly data only begin in 1947.

I set $\theta_0^s = [0.99\ 0\ 0\ 0]$ and make this prior as diffuse as possible while still maintaining finite forecasts of productivity. This requires me to set $P_0^S = 0.15 * I$. I set $\theta_0^{ns} = [0\ 0\ 0]$ and set $P_0^{NS} = P_0^S$. I set $\sigma_0^2 = 0.01^2$ (which is the full sample estimate of the variance of the regression errors) and $v_0 = 1$. To interpret these priors remember that (17) implies that the standard deviation of the initial $\theta_0 = 0.026$. [17]

Finally, I set $k_0 = k^{ss}$, $p_{s,0} = 0.25$ and $lag = 4$. Robustness to the choice of initial lag and prior probability in shown in section 7. Finally, to compare model consumption to consumption in the data I multiply model consumption by $(1+g)^t$ and scale consumption by the mean of income in the data divided by the mean of income in the model.

From the learning model I generate three consumption series. One based on learning

[14] I use the personal income series instead of GDP because the empirical facts are based on the personal income series.

[15] $z_t = \frac{1}{1-\alpha}[\ln((1+\widehat{y}_t)y^{ss}) - \alpha \ln k_t - (1-\alpha)\ln(l^{ss})]$. Recall that k_t is chosen at time $t-1$.

[16] I have confirmed that the results for the learning model hold with more conventional measures of productivity (e.g. TFP), however the volatility over time of the income series in the model differs enough from the income series in the data that when I combine the learning model with the credit constraints model it is unclear what income series to use to calculate, for example, fact 1 the standard deviation of consumption growth to income growth. I would get different results using income from the learning model or using income from the credit constrains model which matches income in the data. Similar issues would arise in analyzing the other facts as well. The current setup removes this ambiguity ensuring that the learning model generates income exactly equaling income in the data.

[17] I have taken care to make sure the results do not depend on the prior choice. Increasing P_0^{NS} to $1.5*I$ did not effect the results nor did reducing both P_0^S and P_0^{NS} to $0.015*I$ as long as I did not update the A matrix for the first 10 periods to avoid unstable estimates of the stationary model. Also increasing $\theta^{ns} = [0.05\ 0]$, lowering $\theta^s = [0.95\ 0]$ and lowering P_0^{NS} to $0.015*I$ had no noticeable change on the results except a larger increase in p_{ns} in the last few years. The results were also insensitive to the choice of σ_0^2.

about the parameters of the two income processes and the likelihood of each model, one based on setting $p_{s,t} = 0$ for all t (non-stationary model) and one based on setting $p_{s,t} = 1$ for all t (stationary model). For the latter two models I set θ^s and θ^{ns} to the O.L.S. estimates of the complete productivity time series. They do not change over time. These last two series are my benchmark no-learning models.

5.3 Dynamics of the Relative Variances of Consumption and Income Growth

Figure 3 plots the agent's probability weight on the non-stationary productivity model implied by the learning model. Recall, the initial prior is 0.75 in 1947. As is evident from the graph the probability falls than rises reaching a peak in 1962. It remains at that level for about ten years and then falls, non-monotonically, putting more and more weight on the stationary model as time passes. In the last few years the probability on the non-stationary model has risen substantially as income has not returned to trend. This probability pattern suggests that consumption variance relative to income variance would rise early in the sample and fall later on. We saw that this was in fact true in U.S. data (figure 1), so now we examine how well the model captures this pattern.

Figure 4 plots the predictions for the relative variance of consumption to income from the models with and without learning. The no-learning, non-stationary income model predicts the ratio should have no trend. It also overestimates the level of the ratio. The no-learning, stationary model again predicts this ratio should be roughly constant and substantially underestimates its level especially early in the sample.

In contrast, the learning model, shown in the middle of Figure 4, does much better than these benchmark models. It predicts a rise and fall in consumption volatility. There is a clear rise and fall in consumption variance relative to income variance, peaking at the same time as the peak in the data. However, the model predicts too little variance at the peak of consumption volatility and from 1975 to 1990. (It is possible to raise the overall variance of consumption in the learning model by increasing the prior weight on the non-stationary model and by increasing γ, however I prefer to remain with my calibration to show that one gets good results with a very reasonable calibration.) It is important to underscore that the learning model replicates the rise and fall in the variance of consumption without using information on consumption data. It is quite remarkable that the model replicates this distinctive consumption pattern using probabilities calculated only from income data.

Another, interesting feature of the model is that it replicates the rise in consumption volatility at the end of the sample coinciding with the most recent recession. Since income has been so far below trend for so long, the model puts extra weight on the non-stationary model and this increases consumption volatility.

It may be puzzling that the learning model line and the stationary model line cross. Recall, that the learning model is updating the parameters on the productivity process every period while the stationary model uses the full sample estimates. Hence, when the learning model crosses the stationary model, it is putting most of the weight on the stationary model and its parameters for the stationary model imply less persistent deviations from trend than the no-learning, stationary model. This explains why the two lines cross. Also, volatility in the non-stationary model is quite high because the full sample estimates imply substantial autocorrelation in productivity growth.

Figure 5 examines the ability of the credit constraints model with and without learning to match this fact. The credit-constraint model combined with the non-stationary model generates a slight increase in consumption volatility over time. We see this because the volatility of consumption in the non-stationary model is higher than in the credit constraint model. As consumption becomes slightly less constrained, consumption volatility rises. When mixed with the stationary model, the credit constraints model delivers a rise and fall in consumption volatility consistent with the estimated rise and fall in credit constraints. But the magnitude is small since the estimated change in credit constraints is small. Mixing the credit constraints and the learning model (Figure 6) gives results similar to the credit constraints with stationary consumption model. The only difference is the rise and fall of consumption volatility is slightly higher with the learning model since the learning model generates a larger rise and fall in consumption. Both models though are unable to generate the low volatility of consumption at the end of the sample because estimated credit constraints have not fallen enough. Remember, these are 10 year averages so the relevant measure of credit constraints is that average over the next ten years. And while credit constraints were low in 2010, they were quite high on average from 2000-2010.

The last figure also highlights an important contribution of the learning model. If loosening of credit constraints are an important contributor to the reduction in consumption volatility it is important to have a good model of consumption choice absent credit constraints. The credit constraints model is sensitive to the choice of income process. If the income process is stationary, the credit constraints model gives the correct qualitative result. However, if the income process is non-stationary the reverse is true – consumption becomes more volatile as credit constraints loosen. One can view the learning model as completing the credit constraints model. It allows one not to impose an income process in face of uncertainty as to which process is correct and bring the model a bit closer to the data.

5.4 Time Varying Response of Consumption to Income

In section three I found that the correlation of consumption and income has fallen over time. In, $\Delta \ln c_t = \alpha + \beta_t \Delta \ln y_t + \varepsilon_t$, β_t fell over time ($\beta_t = \beta_0 + \beta_1 t$ with $\beta_0 = 0.55$ and $\beta_1 = -0.124$). Table 1 reports the ability of the different models to match this fact. We first see that the benchmark models without time varying credit constraints or learning (those with a known stationary or non-stationary productivity process) cannot match this fact (β_1 is not significantly different from zero). Both the learning model and the credit constraints model, improve over the benchmark models; these models predict a fall in β over time. For the learning model $\beta_0 = 0.7$ and $\beta_1 = -0.073$ and for the credit constraint model $\beta_0 = 0.73$ and $\beta_1 = -0.05$. Adding the time varying credit constraints to the learning model gives a similar results In this case $\beta_0 = 0.81$ and $\beta_1 = -0.06$. All models predict a smaller fall than is in the data, however the learning model comes closest to matching the data.

5.5 Breaks in the Variance of Consumption

In section three I measured breaks in the absolute deviation of consumption growth from its mean. Table 2 presents the results from this estimation. The nondurables series has one break at 1977:4. The means for each segment are 2.65 and 1.87. The simulated consumption series from the learning model is estimated to have one break occurring in the first quarter of 1982.[18] While a few years from the break estimated in the data, it is within the estimated confidence interval for the break date in the data. The respective means for each segment from the consumption model with learning are 1.5 and 0.94. The model matches both the number of breaks and the relative magnitude of the change. However, the break dates do not line up exactly. None of the consumption measures from the benchmark no-learning models show evidence of any breaks. Consumption data from the model with credit constraints shows no breaks. When the credit constraints model is combined with the learning model, there is one break in 1984, quarter 3 and the mean falls from 1.77 to 1.21.

5.6 Relation Between Consumption Changes and Probability Changes

The last few subsections argued that a learning model with changing probability weights on the non-stationary model is consistent with the observed change in the variance of consumption over time. This subsection extends that argument. In the data large changes in consumption are associated with large changes in the estimated probability that the non-stationary model is true. I find that this fact is replicated by the learning model.

To motivate this section's analysis, imagine we wanted to investigate the hypothesis that changes in beliefs influenced consumption choice. Then to test this theory we might

[18] One can think of these breaks as "measured breaks" since the simulated data does not have discreet breaks in consumption variance.

want to identify the points in time that resulted in the largest changes in beliefs and see if consumption changed more at those times than would be predicted based on changes in income. Since my model identifies how much beliefs should change at each point in time, I can use my estimated beliefs to examine if changes in beliefs help explain consumption changes controlling for other factors that influence consumption. To this end, I estimate the following regression on the consumption data:

$$\left|\Delta \ln c_t - \overline{\Delta \ln c}\right| = \alpha + \beta \left|\Delta p_{s,t} - \overline{\Delta p_s}\right| * \left|\frac{c^s_{t-1} - c^{ns}_{t-1}}{c_{t-1}}\right| + \gamma X_t + \varepsilon_t \tag{39}$$

where the bar notation denotes mean, c_t is the NIPA measure of nondurables and service consumption, $p_{s,t}$ is the estimated probability the stationary model is true (note this is a function only of the income data and not of the consumption data), and X_t is a vector of control variables including a time variable[19] and $\left|\Delta \ln y_t - \overline{\Delta \ln y}\right|$ where y_t is personal income. $\left|\frac{c^s_{t-1} - c^{ns}_{t-1}}{c_{t-1}}\right|$ is the absolute size of the difference in the consumption choice implied by the learning model if the agent put probability one on the stationary model versus putting probability zero on the stationary model as a percent of current consumption. The model implies that consumption changes are proportional to this term (see section 6.2 for more details).

The results in table 3 indicate that the absolute value of the probability change is a significant predictor of the absolute value of the consumption deviation from its mean. The coefficient $\beta = 0.027$. In other words, large changes in the estimated probability are associated with large changes in consumption controlling for changes in income and time trends. In estimation I multiply both $\left|\Delta p_{s,t} - \overline{\Delta p_s}\right|$ and $\left|\frac{c^s_{t-1} - c^{ns}_{t-1}}{c_{t-1}}\right|$ by 100. Hence, this coefficient means that if consumption under the stationary model differs by 10% and the probability changes by 0.01 the deviation in consumption would increase by 0.27 percentage points compared to a mean of 1.4 percentage points.

Table 3 reports the ability of the different models to replicate this fact. The benchmark

[19] Higher order time variables were not significant.

models without learning or time-varying credit constraints are unable to replicate this fact. The model with learning naturally replicates this fact with a $\beta = 0.067$. However, the coefficient is too high. On the other hand, the model with credit constraints, predicts that β should be negative. Putting the credit constraints model together with the learning model gets a coefficient that is close to correct, $\beta = 0.035$.

6 Mechanism of the Learning Model

To better explain the internal mechanisms of the learning model, I discuss two implications of the model in more detail. First, I examine the largest movements in the probability the stationary model is true and show that they roughly accord with intuition. When log productivity tends to be persistently above its steady state value (0) we move toward the non-stationary model and when productivity returns towards its steady state value we move towards the stationary model. Second, I decompose the consumption variance in the learning model into variance due to directly to shocks from productivity and those due to changes in the estimated probability weights on each productivity model.

6.1 Largest Movements in the Productivity Process Probability

Recall, from (27), the dynamics of the productivity process probability are given by:

$$\frac{w_{NS,t+1}}{w_{S,t+1}} = \frac{\Phi(\varepsilon^{NS}_{t+1})w_{NS,t}}{\Phi(\varepsilon^{S}_{t+1})w_{S,t}} \tag{40}$$

$$\Phi(\varepsilon^{i}_{t+1}) = p(y^i_{t+1}|x^i_{t+1}, \theta_i, \sigma^2_i)\frac{p(\theta_i, \sigma^2_i|Z^i_t)}{p(\theta_i, \sigma^2_i|Z^i_{t+1})} \tag{41}$$

$$p_{S,t} = \frac{1}{1 + \frac{w_{NS,t}}{w_{S,t}}} \tag{42}$$

Table 4 lists the five dates when the $\frac{\Phi(\varepsilon^{NS}_{t+1})}{\Phi(\varepsilon^{S}_{t+1})}$ is lowest, and the four dates when it is the highest, (and the data from the previous quarter) in order to build intuition for the mechanisms of the model. These are the dates that provide the best evidence for the stationary model and the non-stationary model respectively. Additionally, the table lists the percentage deviation of productivity from its steady state value (1) and the non-stationary and stationary forecasts of productivity based on information from the previous period. At each date the stationary model predicts that productivity from the previous period should move toward zero in the next period. The non-stationary forecast always predicts that productivity should be farther from zero than the stationary forecast predicts.[20]

The periods, then, that give the best evidence for the stationary model are periods when productivity is far from zero and moves substantially towards zero. In contrast, the periods that give the best evidence for the non-stationary model are the periods where productivity is far from zero and moves even further away in the next period.

For example, in 2001 quarter 2, productivity moves from 5.4% above trend to 3.5% above trend. This quick reversal to the mean generates evidence for the stationary model. Similarly, during the latest recession, in quarter one of 2009 productivity falls from 9% below trend to 13% below trend providing much evidence for the non-stationary model.

This section helps illustrate, that while this model may seem based on a difficult econometric problem, too far detached from the decision making of individuals, it can capture important features of actual consumption decisions. First, imagine a situation like the 1990s where the economy begins to grow faster. Two views of the world may emerge: the first, the new era view, espouses that the gains are permanent. A second, more pessimistic view, would be that the gains are temporary and the economy will eventually return to trend. If uncertainty about these two views is important, when the recession of 2000 comes along and gives evidence against the new era view, one could imagine an additional effect on consumption from changes in beliefs, over and above the direct effect of the productivity shock. Second, note that the U.S. has done comparatively well over the last 50 years returning often to a stable trend. One could imagine that, after seeing this repeated pattern, agents will become more confident in the U.S. economy, and react less to recessions, seeing them as

[20] Whether productivity is predicted to increase or decrease in the non-stationary model depends on the lag values of Δz_t and varies by observation.

temporary deviations from trend.

6.2 Decomposition of Consumption Variance

Figure 7 decomposes the variance of consumption growth relative to income growth from the learning model into the variance due directly to income shocks and variance that comes from changes in the probability weight on each model. I confirm numerically for the learning model that

$$c_t^l = p_{s,t} c_t^s + (1 - p_{s,t}) c_t^{ns} \tag{43}$$

where c_t^l is the consumption choice of the agent at time t from the learning model, c_t^s is the optimal consumption choice if $p_{s,t} = 1$ and c_t^{ns} is the optimal consumption choice if $p_{s,t} = 0$. This equation then implies

$$\Delta c_t^l = p_{s,t} \Delta c_t^s + (1 - p_{s,t}) \Delta c_t^{ns} + (p_{s,t} - p_{s,t-1}) \left(c_{t-1}^s - c_{t-1}^{ns} \right) \tag{44}$$

Therefore we can decompose the change in consumption into changes due directly to income shocks and changes due to changes in beliefs:

$$\Delta c_t = \Delta c_t^1 + \Delta c_t^2 \tag{45}$$

$$\Delta c_t^1 = p_{s,t} \Delta c_t^s + (1 - p_{s,t}) \Delta c_t^{ns} \tag{46}$$

$$\Delta c_t^2 = (p_{s,t} - p_{s,t-1}) \left(c_{t-1}^s - c_{t-1}^{ns} \right) \tag{47}$$

A standard variance decomposition gives:

$$Var\left(\frac{\Delta c_t}{c_{t-1}} \right) = Cov\left(\frac{\Delta c_t^1}{c_{t-1}}, \frac{\Delta c_t}{c_{t-1}} \right) + Cov\left(\frac{\Delta c_t^2}{c_{t-1}}, \frac{\Delta c_t}{c_{t-1}} \right) \tag{48}$$

The first term Δc_t^1 represents the consumption response to a productivity shock today. You are hit with a shock today, you put some weight on the shock being permanent, some weight on it being temporary and adjust you consumption accordingly. But since you revised your beliefs today, you also believe that you made a mistake in your consumption choice last period. You put too much weight on one of the models. And therefore you adjust you consumption even more today based on how much your revised beliefs indicate that last period's consumption choice was in error, this effect is captured by the Δc_t^2 term.

Figure 7 contains the results from this variance decomposition, normalized by the variance of the income growth rate. Note that I report the variance of $\frac{\Delta c_t}{c_{t-1}}$ predicted by the model, divided by the variance of income.[21] There is a rise and fall in the second covariance term early in the sample (marked x). This drives about 15% of the rise in consumption volatility early in the sample. However, by 1975, as the volatility of the probability changes has fallen substantially, most of the variance of consumption is given directly by the response to changes in productivity and the continued fall in the variance of consumption is due to the increased weight put on the stationary model. Finally, during the most recent recession, the weight on the non-stationary model rises dramatically and the second covariance term contributes again about 15% of the observed volatility of consumption.

7 Robustness

7.1 Lag length and prior choice

Figure 8 explores the sensitivity of the results to different choices for the prior probability. I plot the consumption standard deviation ratio from figure 4 for the learning model with an initial prior on the stationary model of 0.25 and for a lag choice of $2, 4$, and 6. Then I plot the same figure for priors on the stationary model of $0.05, 0.25$, and 0.45 with a lag length of 4.

As seen in column one of Figure 8, the initial rise in consumption variance relative to

[21] This is not directly comparable to Figure 4 which plots the ratio of standard deviations, however the standard deviation, being non-linear, does not decompose easily.

income variance and the subsequent fall are present for all lag lengths. Moreover, the fit gets better as the lag length increases; the fit is good for lags 4 and 6. The fit deteriorates slightly for a lag length of 2. The same rise and fall pattern is evident, but the magnitude of the fall is muted. Similarly the initial rise and fall in consumption variance relative to income variance is visible for all priors. Decreasing the prior increases the mean level of the consumption variance and increasing the prior decreases the mean variance but the basic dynamic pattern of a rise and fall in consumption volatility is not changed.

8 Conclusion

This paper studies three empirical observations. First, after an increase early in the sample, the standard deviation of consumption growth relative to income growth fell by 50%. Similarly the correlation between consumption growth and personal income growth has also fallen by about 50%. Finally, the consumption series has one estimated break in the absolute deviation of consumption from its mean. The change in mean deviation is about 30%.

I examine two explanations of these facts. The first is that a fall in the fraction of credit constrained consumers has lead to a reduction in consumption volatility relative to income volatility. While this model captured some of the fall in consumption volatility it was unable to match the full magnitude of the decline in consumption volatility. It also left open the question of how an unconstrained consumer would respond to an income shock, since if productivity shocks are permanent, relaxing credit constraints could make consumption more volatile. To address these shortcomings, I studied a model with an agent who learns about whether or not income shocks are permanent or transitory. This model was consistent with the overall shape of the relative volatility of consumption to income, the declining correlation of income and consumption over time and the estimated break in the consumption growth series.

References

ADDA, JEROME, & COOPER, RUSSELL W. 2003. *Dynamic economics: Quantitative methods and applications*. MIT Press.

AGUIAR, MARK, & GOPINATH, GITA. 2005. Emerging market business cycles: The cycle is the trend. *Mimeo*, April.

ANDREWS, DONALD W K. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, **59**(3), 817–58.

BAI, JUSHAN, & PERRON, PIERRE. 1998. Estimating and testing linear models with multiple structural changes. *Econometrica*, **66**(1), 47–78.

BAI, JUSHAN, & PERRON, PIERRE. 2003. Computation and analysis of multiple structural change models. *Journal of applied econometrics*, **18**(1), 1–22.

CAMPBELL, JOHN Y, & MANKIW, N GREGORY. 1990. Permanent income, current income, and consumption. *Journal of business & economic statistics*, **8**(3), 265–79.

CECCHETTI, STEPHEN G, FLORES-LAGUNES, ALFONSO, & KRAUSE, STEFAN. 2005. Assessing the sources of changes in the volatility of real growth. *In:* KENT, CHRISTOPHER, & NORMAN, DAVID (eds), *The changing nature of the business cycle*. RBA Annual Conference Volume. Reserve Bank of Australia.

CECCHETTI, STEPHEN G, FLORES-LAGUNES, ALFONSO, & KRAUSE, STEFAN. 2006. Financial development, consumption smoothing, and the reduced volatility of real growth. *mimeo*, December.

COCHRANE, JOHN. 2001. *Asset pricing*. Princeton University Press.

COCHRANE, JOHN H. 1988. How big is the random walk in gnp? *Journal of political economy*, **96**(5), 893–920.

COGLEY, TIMOTHY, & SARGENT, THOMAS J. 2005. The conquest of us inflation: Learning and robustness to model uncertainty. *Review of economic dynamics*, **8**(2), 528–563.

DEATON, ANGUS. 1992. *Understanding consumption*. Oxford University Press.

DYNAN, KAREN E., ELMENDORF, DOUGLAS W., & SICHEL, DANIEL E. 2006. Can financial innovation help to explain the reduced volatility of economic activity? *Journal of monetary economics*, **53**(1), 123–150.

DYNAN, KAREN E., EDELBERG, WENDY, & PALUMBO, MICHAEL. 2009. The effects of population ageing on the relationship among aggregate consumption, saving and income. *American economic review: Papers and procedings*, May.

EVANS, GEORGE, & HONKAPOHJA, SEPPO. 2001. *Learning and expectations in macroeconomics*. Princeton University Press.

GUVENEN, FATIH. 2005. Learning your earning: Are labor income shocks really very persistent? *Mimeo*, July.

HAMILTON, JAMES. 1994. *Time series analysis.* Princeton University Press.

KREPS, DAVID. 1998. *Anticipated utility and dynamic choice, a 1997 schwartz lecture.* Frontiers of Research in Economic Theory.

LJUNGOVIST, LARS, & SARGENT, THOMAS. 2004. *Recursive macroeconomic theory.* The MIT Press.

LUDVIGSON, SYDNEY. 1999. Consumpiton and credit: A model of time-varying liquidity constraints. *Review of economics and statistics*, **81(3)**, 434–447.

NAKAJIMA, MAKOTO. 2007. Solving rbc models with l-q approximation. *mimeo uiuc.*

SARGENT, THOMAS. 2001. *The conquest of american inflation.* Princeton University Press.

STOCK, JAMES H. 1991. Confidence intervals for the largest autoregressive root in u.s. macroeconomic time series. *Journal of monetary economics*, **28**(3), 435–459.

STOCK, JAMES H., & WATSON, MARK W. 2002 (Aug.). *Has the business cycle changed and why?* NBER Working Papers 9127. National Bureau of Economic Research, Inc.

TAUCHEN, GEORGE, & HUSSEY, ROBERT. 1991. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica*, **59**(2), 371–96.

Table 1: Consumption Growth Rate on the Income Growth Rate

		Benchmark Models				
	Data	Non-Stationary	Stationary	Learning	Credit	Credit and Learning
Δln(income)	0.55***	1.03***	0.50***	0.70***	0.73***	0.81***
	[0.08]	[0.04]	[0.03]	[0.04]	[0.02]	[0.02]
Δln(income)*(time/100)	-0.124***	0.021	-0.009	-0.073***	-0.050**	-0.061***
	[0.043]	[0.021]	[0.017]	[0.025]	[0.024]	[0.012]
Constant	1.17***	-0.11	1.17***	0.90***	0.76***	0.60***
	[0.19]	[0.08]	[0.05]	[0.08]	[0.07]	[0.04]
Observations	221	221	221	221	221	221

Newey-West Standard Errors with Five Lags in Brackets (* significant at 10%; ** significant at 5%; *** significant at 1%)

This table reports the result of regressing anuualized consumption growth from the data and the predictions of the different models on the annualized income growth rate, alowing the coeficient to vary by time.

Table 2a: Breaks in Consumption Volatility

		Benchmark Models				
	Data	**Non-Stationary**	**Stationary**	**Learning**	**Credit**	**Credit and Learning**
supF(1|0)	7.38*	2.32	3.2	9.47*	5.26	7.29
supF(2|1)	3.44	5.24	4.64	3.71	4.33	3.86
UD Max	7.38	6.07	4.4	10.61*	5.26	7.29
Break Date 1	1977:4	--	--	1982:1	--	1984:3
90% CI	1957:2-1993:1	--	--	1973:2-1995:3	--	1971:2-2000:2
δ1	2.65	--	--	1.5	--	1.77
	(0.20)	--	--	(0.11)	--	(0.12)
δ2	1.87	--	--	0.94	--	1.21
	(0.17)	--	--	(0.11)	--	(0.13)

This Table reports the results from testing for a break in the residual of the annualized consumption change regressed on a constant. The means of the residual are given by δ, for before and after the breaks. supF and UDMax are tests for L+1 breaks vs. L and any breaks respectively and the critical values for the tests are given in Table 2b. Star denotes >= 10% significance.

Table 2b: Critical Values for the Tests

	10%	**5%**	**1%**
supF(1|0)	7.04	8.58	12.29
supF(2|1)	8.51	10.13	13.89
supF(3|2)	9.41	11.14	12.66
UD Max	7.46	8.88	12.37

Critical values come from Bai and Perron (2003).

Table 3: Absolute Change in Consumption Growth on Absolute Change in Normalized Beliefs

		Benchmark Models				
	Data	**Non-Stationary**	**Stationary**	**Learning**	**Credit**	**Credit and Learning**
$100*\|\Delta(ps_t)-E(\Delta(ps))\|*100*\|(cs_{t-1} - cns_{t-1})/c_{t-1}\|$	0.027*	0.004	-0.004	0.067***	-0.019**	0.035***
	[0.015]	[0.011]	[0.006]	[0.007]	[0.009]	[0.006]
$\|\Delta\ln(y_t)-E(\Delta\ln(y))\|$	0.24***	1.01***	0.46***	0.50***	0.65***	0.66***
	[0.06]	[0.04]	[0.02]	[0.03]	[0.03]	[0.03]
t/100	-0.09	0.02	-0.02	-0.18***	-0.14***	-0.16***
	[0.14]	[0.07]	[0.04]	[0.04]	[0.04]	[0.04]
Constant	0.96***	0.15	0.14**	0.29***	0.33***	0.31***
	[0.21]	[0.11]	[0.06]	[0.08]	[0.09]	[0.09]
Observations	221	221	221	221	221	221

Newey-West Standard Errors with Five Lags in Brackets (* significant at 10%; ** significant at 5%; *** significant at 1%)

Column 1 reports the results of regressing the absolute value of annualized consumption growth minus its mean on the estimated change in probability that the stationary model is true times the difference in consumption predicted by the stationary and non-stationary model, controlling for the changes income (y) and time trends . Columns 2-6 redoes the analysis of column 1 using simulated data from the non-stationary, no-learning model from the stationary, no-learning model, from the learning model, from the credit constraint model, and from the credit constraint with learning model.

Table 4: Largest Movements in the Non-Stationary Probability

Movements towards the Stationary Model

Year	Quarter	PNS	$\Phi(\varepsilon^{NS})/\Phi(\varepsilon^{S})$	Productivity	NS Forecast	S Forecast
2001	1	0.07	1.10	5.44	5.20	4.82
2001	2	0.05	0.66	3.50	5.42	5.02
1974	3	0.50	0.85	3.96	5.01	4.60
1974	4	0.41	0.70	1.51	3.90	3.49
1949	4	0.76	1.04	-8.83	-7.83	-7.87
1950	1	0.70	0.70	0.74	-8.82	-7.76
2001	2	0.05	0.66	3.50	5.42	5.02
2001	3	0.03	0.72	1.72	3.42	3.08
1959	1	0.55	0.89	-6.84	-7.91	-7.03
1959	2	0.47	0.73	-4.88	-6.83	-5.89

Movements towards the Non-Stationary Model

Year	Quarter	PNS	$\Phi(\varepsilon^{NS})/\Phi(\varepsilon^{S})$	Productivity	NS Forecast	S Forecast
2008	4	0.08	1.56	-9.02	-7.08	-6.80
2009	1	0.19	2.74	-13.45	-9.21	-8.96
2008	2	0.03	1.09	-3.83	-3.35	-3.10
2008	3	0.05	2.05	-6.95	-3.83	-3.58
2009	2	0.18	0.93	-13.28	-13.73	-13.37
2009	3	0.28	1.72	-15.14	-13.45	-13.06
2004	4	0.02	0.81	-2.81	-4.18	-3.92
2005	1	0.03	1.59	-4.95	-2.75	-2.51
2008	3	0.05	2.05	-6.95	-3.83	-3.58
2008	4	0.08	1.56	-9.02	-7.08	-6.80

This Table lists the largest movements in the non-stationary probability along with percentage deviation of productivity from its steady state and its predictions.

Figure 1: Ratio of the Standard Deviation of Consumption Growth to Income Growth

Note: Dashed lines represent the 95% confidence interval.

Figure 2: Estimated Fraction of Credit Constrained Consumers

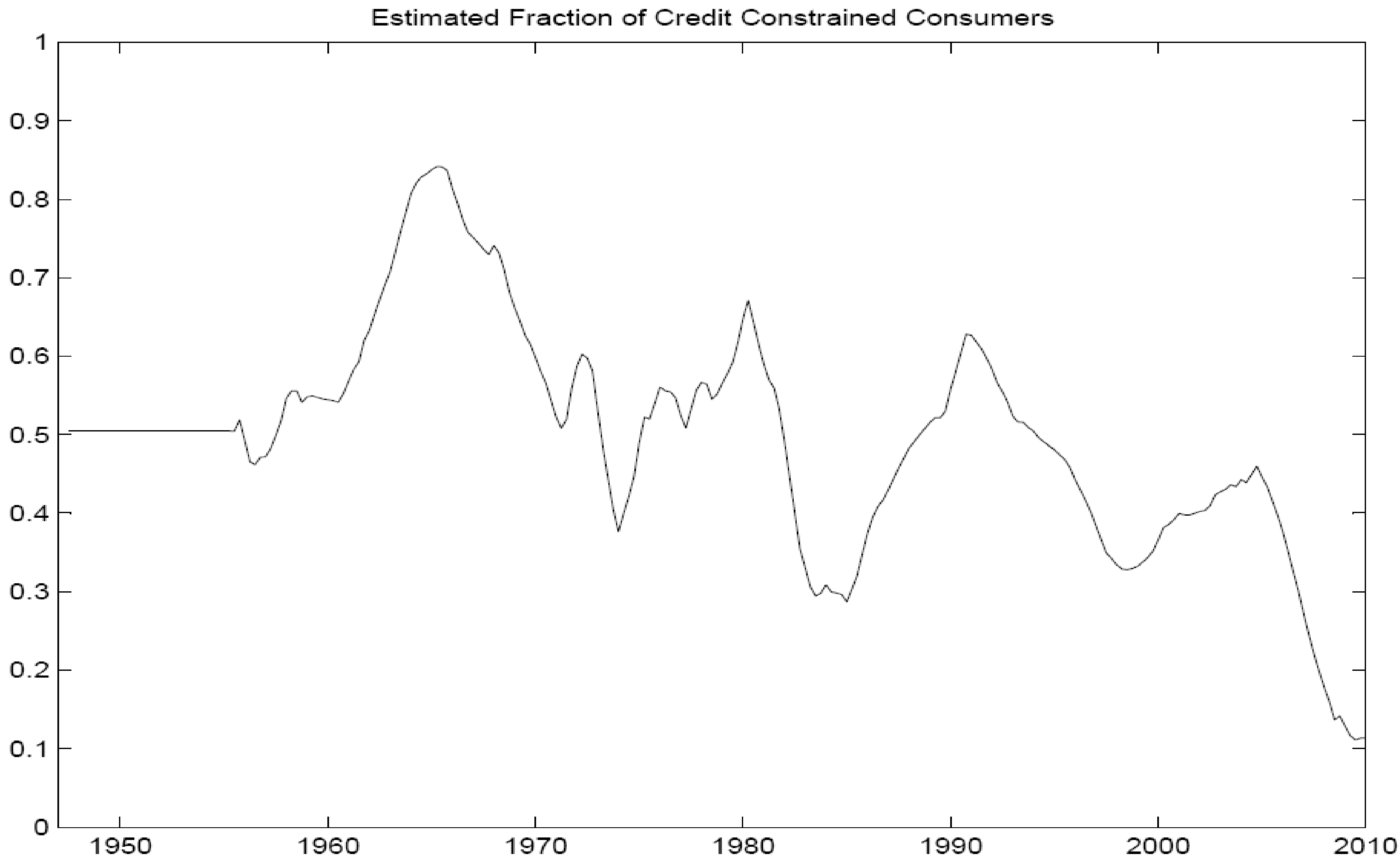


Figure 3: Probability Weight on the Non-Stationary Model

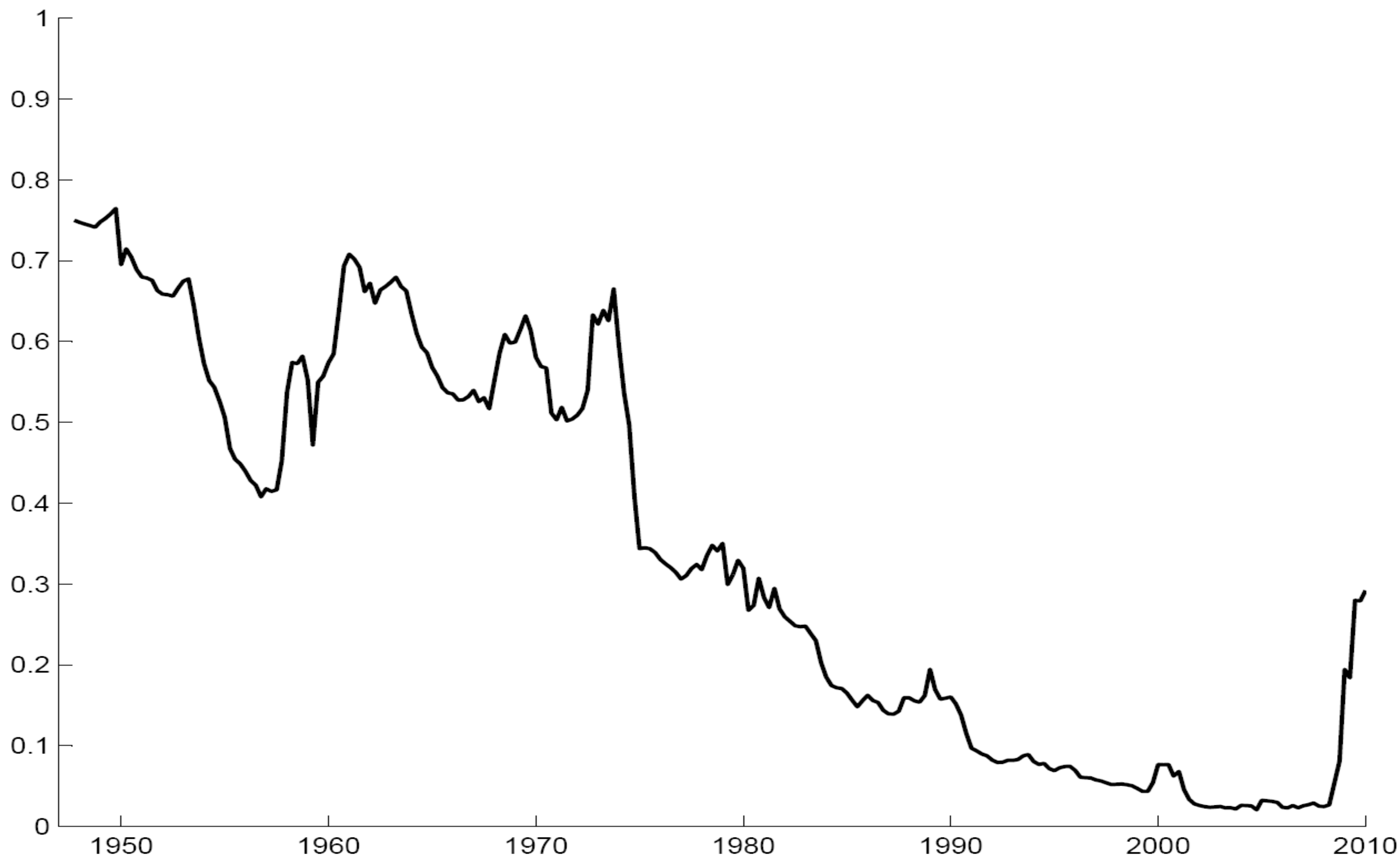


Figure 4: Ratio of Standard Deviation of Consumption Growth to Income Growth (Models)

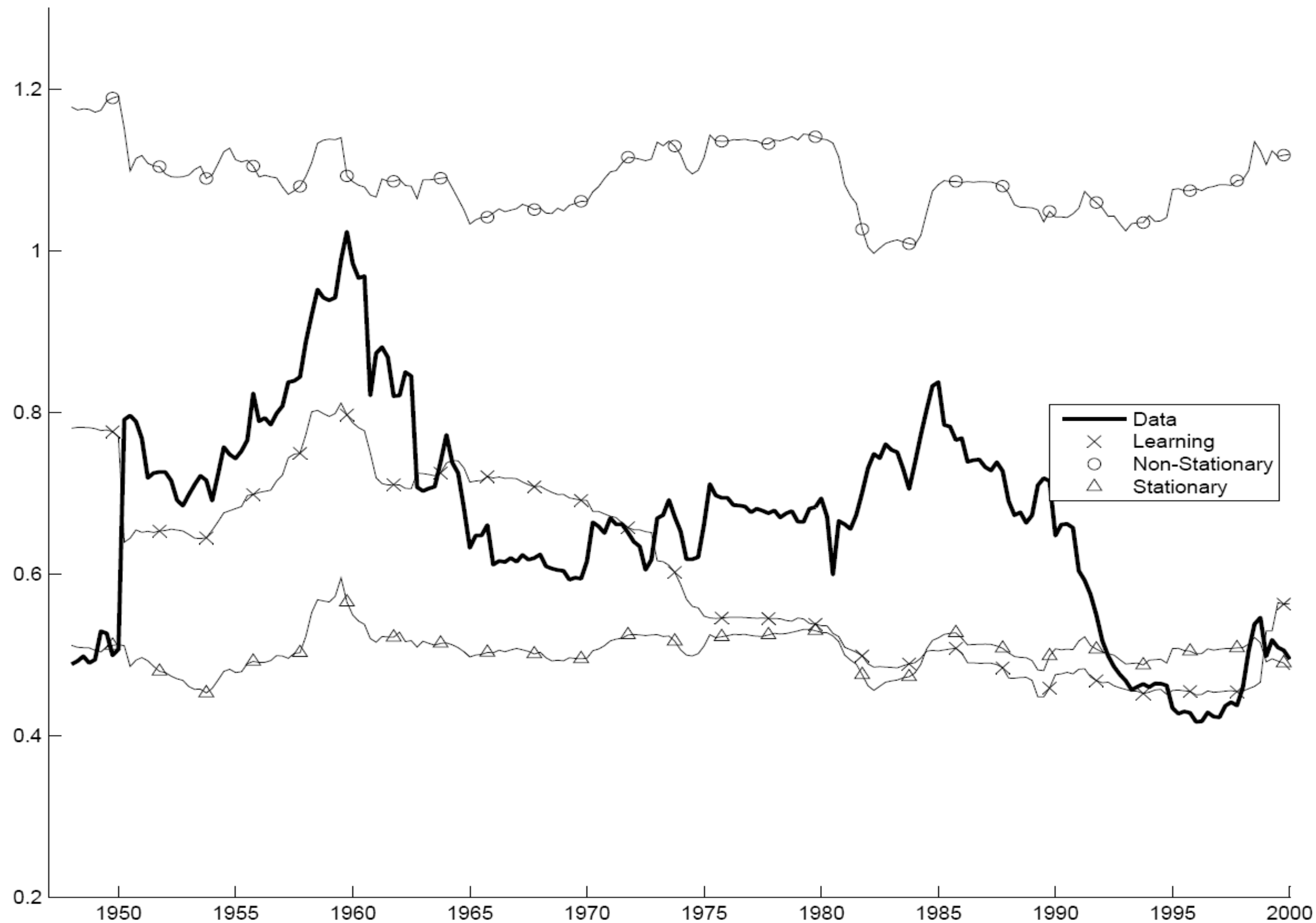


Figure 5: Ratio of Standard Deviation of Normalized Consumption Growth to Income Growth Credit Model

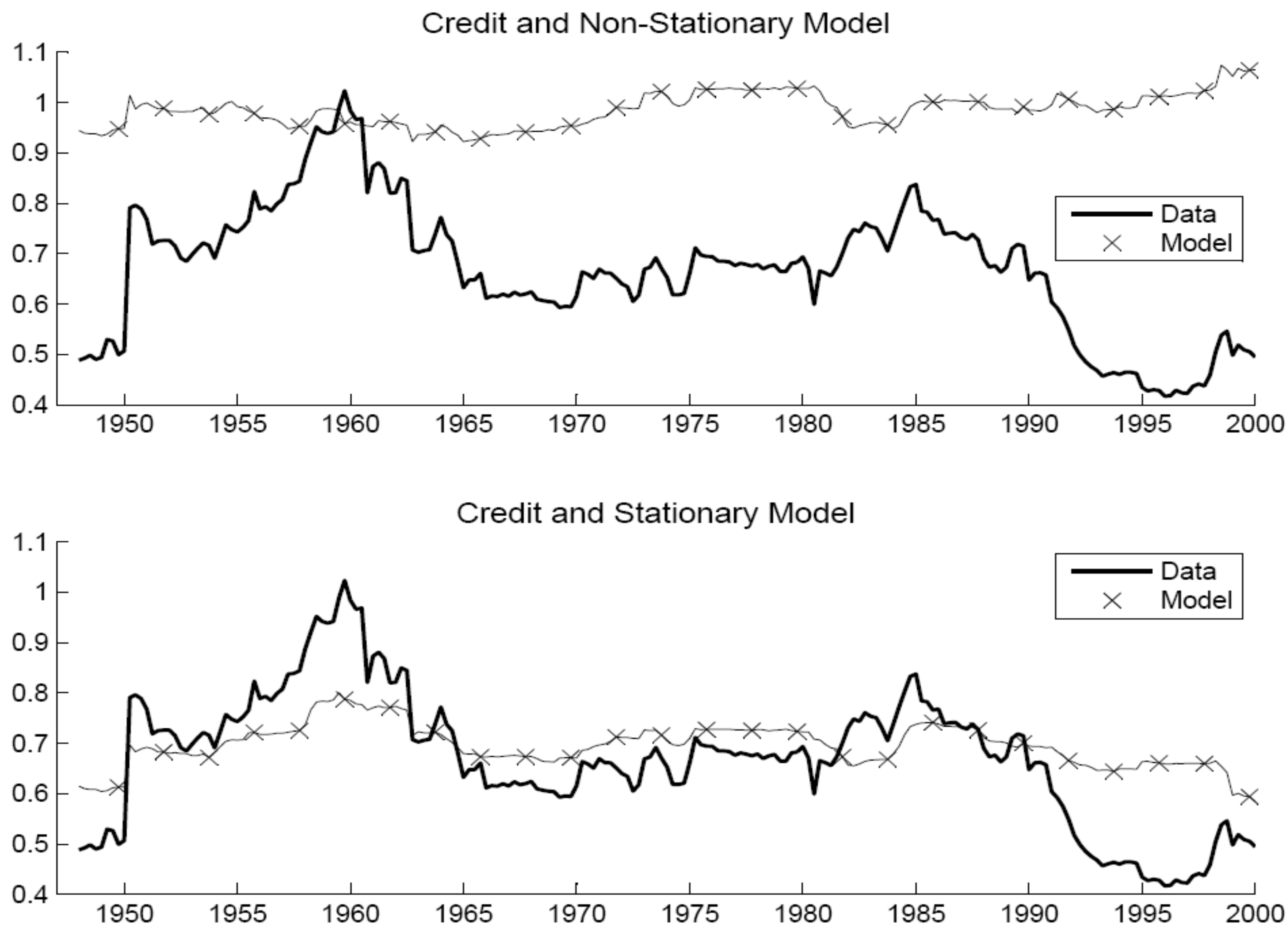


Figure 6: Ratio of Standard Deviation of Consumption Growth to Income Growth Credit Model with Learning

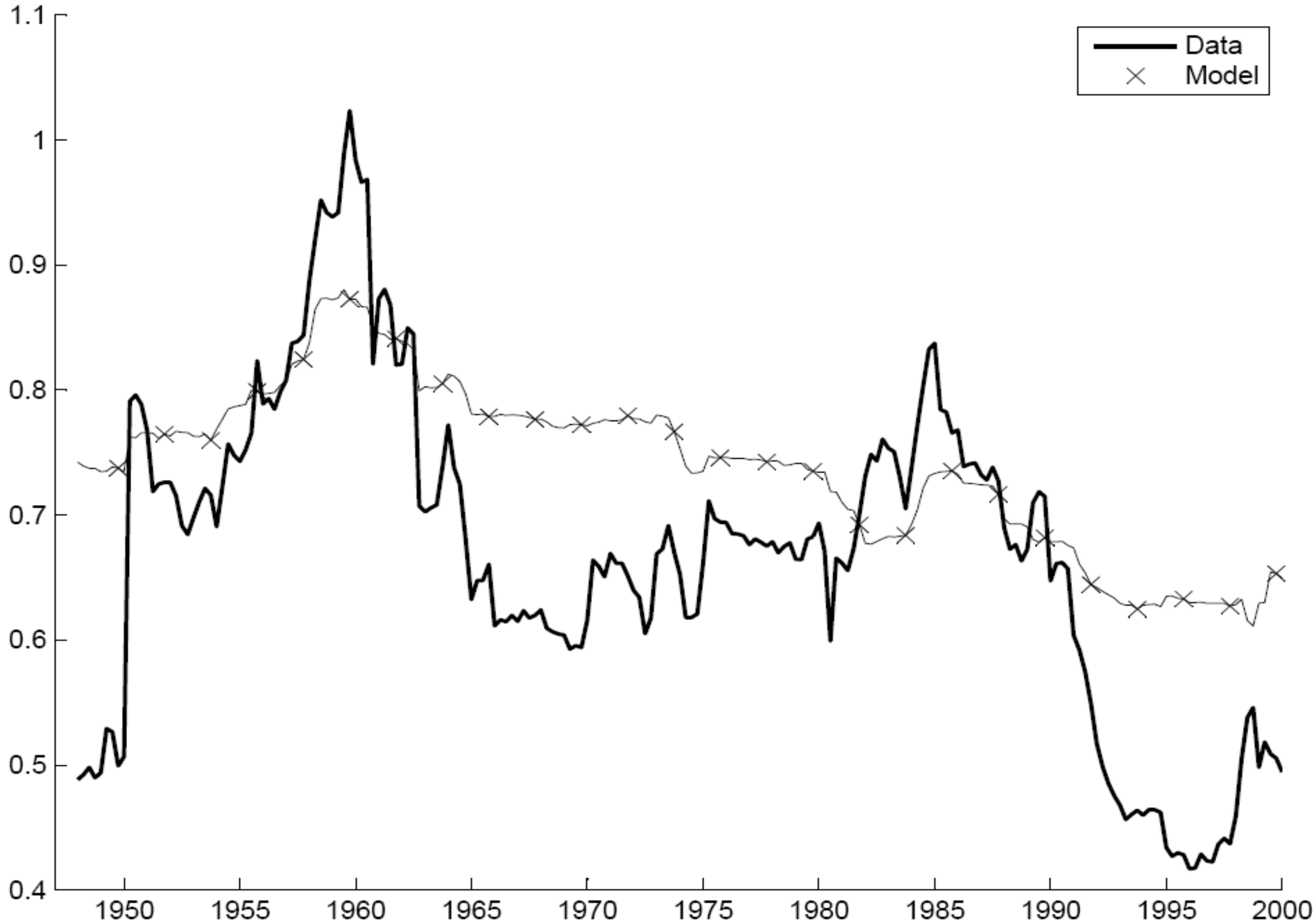
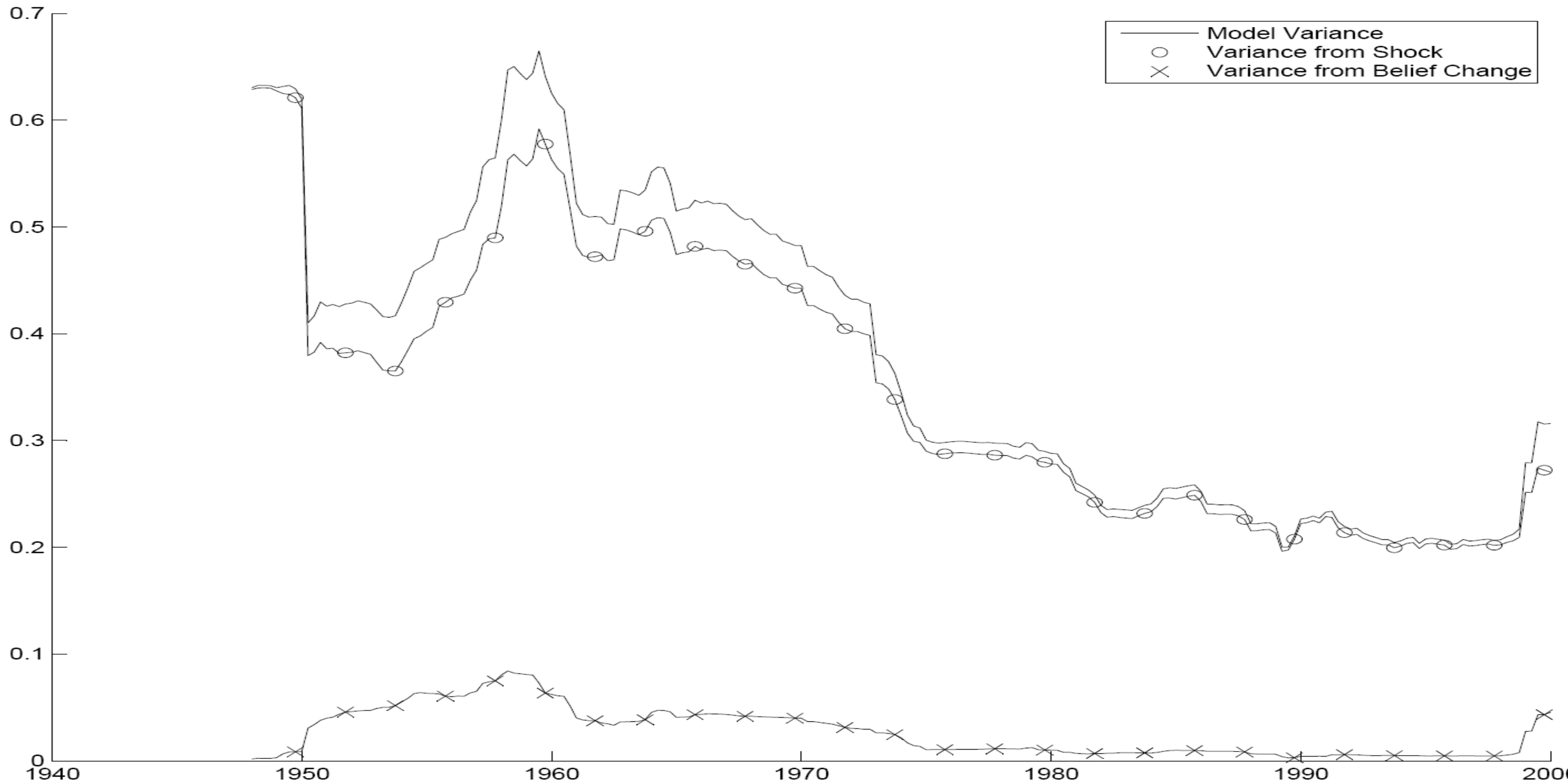
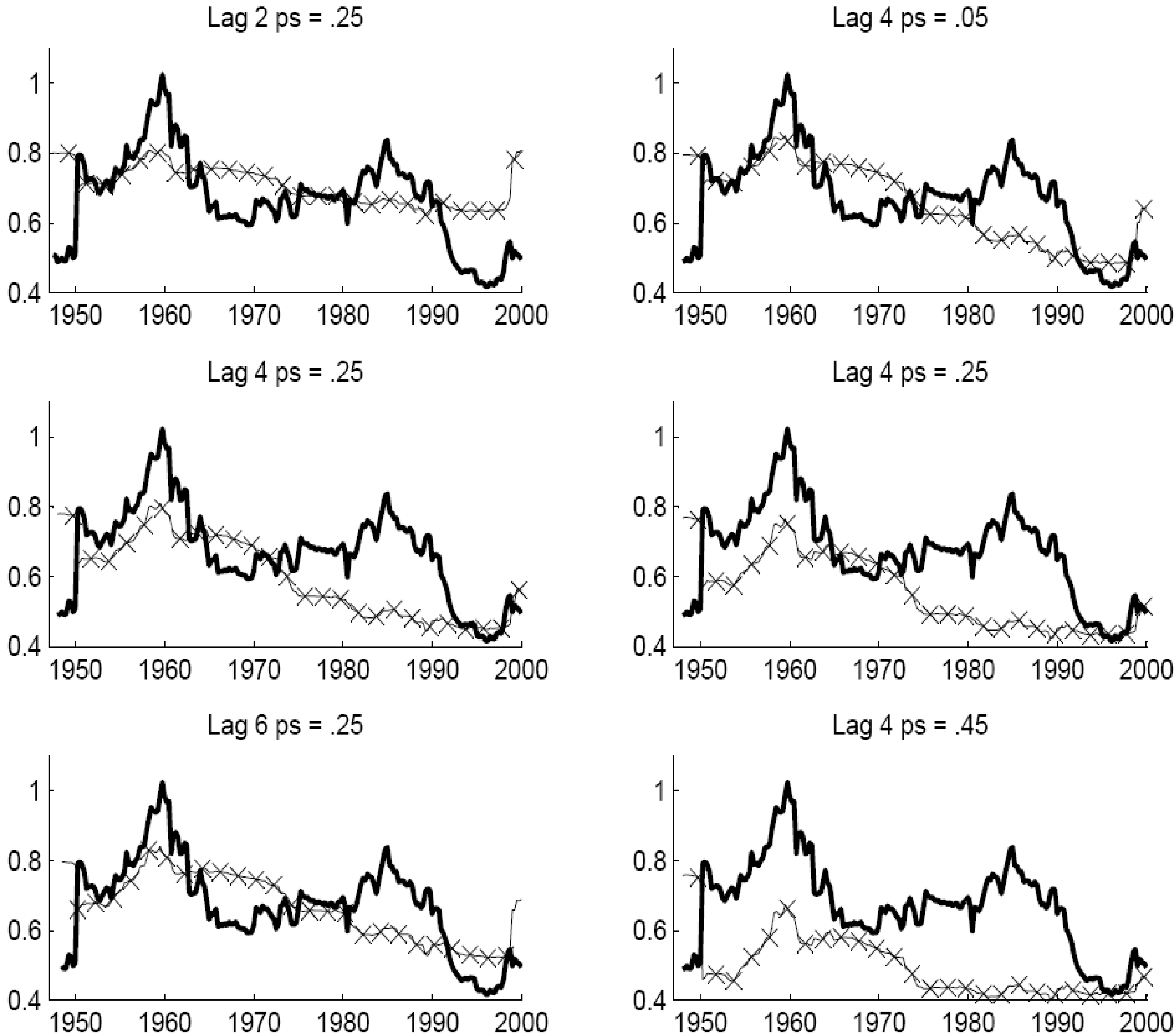


Figure 7: Variance Decomposition



This figure plots the learning model's prediction of var(Δ ln c_t)/var(Δ ln y_t) and its decomposition into a term that represents variance that comes from productivity shocks cov(Δc^1_t, Δln c_t)/ var Δ (ln y_t) and variance that come from changes in beliefs cov(Δc^2_t, Δ c_t)/ var Δ (ln y_t). Here $c^1_t = ps_t * \Delta cs_t/c_{t-1} + (1-ps_t) * \Delta cns_{t-1}/c_{t-1}$ where cs_t is the consumption choice if $ps_t = 1$ and cns_t is the consumption choice if $ps_t = 0$. $c^2_t = (ps_t - ps_{t-1}) * (cs_{t-1} - cns_{t-1}) / c_{t-1}$.

Figure 8: Robustness to Different Choices of Lag Length and Prior



Model – X
Data – Solid Line

A Learning Model

A.1 LQ Approximation

I approximate $U(w) = \frac{[\left(k^{\alpha}(e^{z}l)^{1-\alpha}-i\right)(1-(\theta l)^{\phi})]^{1-\gamma}}{1-\gamma}$ with $w = [z\ \ k\ \ l\ \ i]'$ about the steady state $\overline{w} = [\overline{z}\ \ \overline{k}\ \ \overline{l}\ \ \overline{i}]'$. These calculations are based on Nakajima (2007) and Ljungovist & Sargent (2004). We can take a second order Taylor approximation as

$$U(w) \approx U(\overline{w}) + (w-\overline{w})'\overline{J} + \frac{1}{2}(w-\overline{w})'\overline{H}(w-\overline{w})'$$

where $\overline{J}$ is the Jacobian matrix and $\overline{H}$ is the Hessian matrix both evaluated at the steady state. This expression can be manipulated into this form:

$$U(w) \approx [1\ w'] \begin{bmatrix} U(\overline{w}) - \overline{w}'\overline{J} + \frac{1}{2}\overline{w}'\overline{H}\overline{w}' & \frac{1}{2}(\overline{J} - \overline{H}\overline{w})' \\ \frac{1}{2}(\overline{J} - \overline{H}\overline{w}) & \frac{1}{2}\overline{H} \end{bmatrix} \begin{bmatrix} 1 \\ w \end{bmatrix}$$

Let $M = \begin{bmatrix} U(\overline{w}) - \overline{w}'\overline{J} + \frac{1}{2}\overline{w}'\overline{H}\overline{w}' & \frac{1}{2}(\overline{J} - \overline{H}\overline{w})' \\ \frac{1}{2}(\overline{J} - \overline{H}\overline{w}) & \frac{1}{2}\overline{H} \end{bmatrix}$ and let $x = [1\ z\ \ k\,]'$ and $u = [l\ \ i]'$ we partition M as:

$$\begin{aligned} U(w) &\approx [1\ w']\,[M] \begin{bmatrix} 1 \\ w \end{bmatrix} = \begin{pmatrix} x \\ u \end{pmatrix}' \begin{pmatrix} R & W \\ W' & Q \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \\ &= x'Rx + u'Qu + 2x'Wu \end{aligned}$$

A.2 Model Matrices

For the stationary model the state variable are $x_t^S = [1\ z_t\ k_t\ z_{t-1}\ ...\ z_{t-p+1}]'$ and the control variables are $u = [l_t\ i_t]'$. So we have

$$\begin{aligned} R^s &= \begin{bmatrix} R & 0 \\ 0 & 0 \end{bmatrix} \\ W^s &= \begin{bmatrix} W \\ 0 \end{bmatrix} \\ Q^s &= Q \end{aligned}$$

the analogs for the non-stationary model are exactly the same.

For the non-stationary model the state variables are $x_t^S = [1\ z_t\ k_t\ \Delta z_t\ \Delta z_{t-1}\ ...\ \Delta z_{t-p+2}]'$ and the control variables are $u_t = [l_t\ i_t]'$. So for the laws of motion, we have (for clarity I have written it for four lags but it is easy to generalize):

$$A_t^S = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta_{s,t}^1 & 0 & \theta_{s,t}^2 & \theta_{s,t}^3 & \theta_{s,t}^4 \\ 0 & 0 & \frac{1-\delta}{1+g} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$A_t^{NS} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \theta_{ns,t}^1 & \theta_{ns,t}^2 & \theta_{ns,t}^3 \\ 0 & 0 & \frac{1-\delta}{1+g} & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_{ns,t}^1 & \theta_{ns,t}^2 & \theta_{ns,t}^3 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$B^s = B^{NS} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & \frac{1}{1+g} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$C^s = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad C^{NS} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

note the optimal policy is independent of C. Since the objective is linear quadratic the optimal policy exhibits certainty equivalence.