# External Habit in a Production Economy


Andrew Y. Chen*
The Ohio State University


October 22, 2013†

[JOB MARKET PAPER]


**Abstract**

A unified framework for understanding asset prices and aggregate fluctuations is critical for understanding both issues. I show that a real business cycle model with external habit preferences and capital adjustment costs provides one such framework. The estimated model matches the first two moments of the equity premium and risk-free rate, return and dividend predictability regressions, and the second moments of output, consumption, and investment. The model also endogenizes a key mechanism of consumption-based asset pricing models. In order to address the Shiller volatility puzzle, external habit, long-run risk, and disaster models require the assumption that the volatility of marginal utility is countercyclical. In the model, this countercyclical volatility arises endogenously. Production makes precautionary savings effects show up in consumption. These effects lead to countercyclical consumption volatility and countercyclical volatility of marginal utility. External habit amplifies this channel and makes it quantitatively significant.



*Fisher College of Business, The Ohio State University, 760 Fisher Hall, 2100 Neil Avenue, Columbus OH 43210. Tel: (240) 446-5313; email: chen_1678@fisher.osu.edu.



†I am indebted to Aubhik Khan, Xiaoji Lin, René Stulz, Julia Thomas, and especially Lu Zhang for their training, advice, and inspiration. I would also like to thank John Cochrane, Urban Jermann, Pok Sang Lam, and Stijn Van Nieuwerburgh for comments which have significantly improved the paper.

# 1 Introduction

A unified framework for understanding asset prices and aggregate fluctuations is critical for understanding both issues. While significant progress on linking asset prices to real variables has been made in highly stylized endowment economies, these successes are tempered with the introduction of production. The long-run risk channel does an excellent job of replicating empirical asset price moments in an endowment economy, but it produces too little equity volatility once firms and investment are introduced. The same issue is found in the disaster risk framework, despite the impressive asset pricing performance of its tweaked endowment economy variants.[1,2]

This paper shows that the external habit model of Campbell and Cochrane (1999) can be successfully extended into a production economy, and thus provides one such unified framework. I imbed a variant of the Campbell-Cochrane external habit preferences into a real business cycle (RBC) model with capital adjustment costs. The model is estimated using simulated method of moments, and the resulting parameter values produce a good fit for a long list of asset price moments: the mean, volatility, and persistence of the equity premium, the mean, volatility, and persistence of the risk-free rate, the volatility and persistence of the price-dividend ratio, and return and dividend predictability regressions using the price-dividend ratio. And these data-like asset price dynamics originate from fluctuations in real variables which are consistent with the data. The model matches the volatilities of output, consumption, and investment, as well as their autocorrelations and cross-correlations.

The model's large equity volatility comes from large investment frictions. Intuitively, large frictions imply that shocks to the economy are absorbed by asset prices rather than investment (Jermann (1998), Kogan (2004)). This link can be seen in the Q-theory equation[3]

$$\frac{P_t}{K_t} = 1 + \phi \frac{I_t}{K_t}$$

where $P_t$ is the stock price, $K_t$ is capital, $I_t$ is investment, and $\phi$ is an investment friction parameter. All things equal, larger frictions imply larger equity volatility. In fact, the data show that the volatility of $P_t/K_t$ is much, much greater than the volatility of $I_t/K_t$, implying very large investment frictions. Investment frictions, however, are not a free parameter. Investment frictions and the elasticity of intertemporal substitution (EIS)

[1] Gourio (2010)'s disaster risk production model generates large equity volatility, but relies on leverage which exceeds the financial leverage seen in the data.

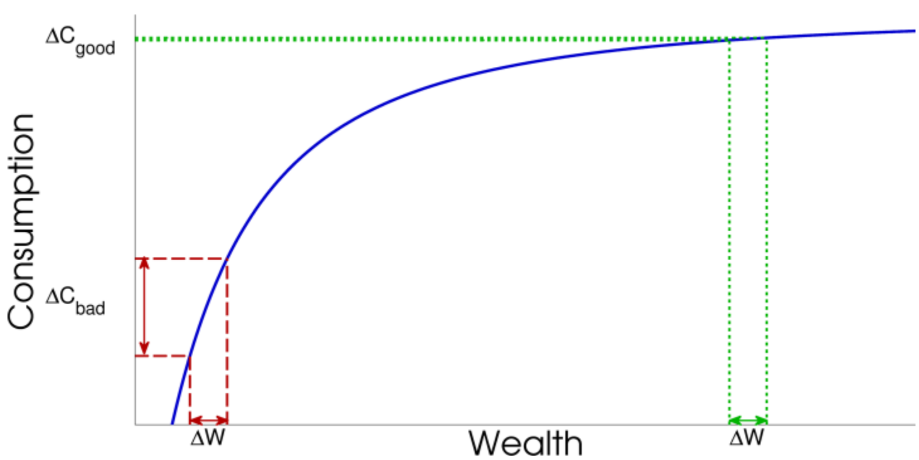
[2] The seminal long risk model is the endowment economy of Bansal and Yaron (2004). Kaltenbrunner and Lochstoer (2010) and Croce (2010) study production economy variants. The seminal disaster risk model is Rietz (1988), and Barro (2006) sparked its revival by emphasizing empirical measurement of disasters. Gabaix (2012) and Wachter (2012)) find that introducing time-varying severity or probability of disasters can address the Shiller volatility puzzle. Gourio (2010) studies a production version.

[3] This equation holds exactly in the model. A similar equation exists in the vast majority of models in this literature, since most models satisfy the Hayashi (1982)-Abel (1983) Q-theory assumptions.

jointly pin down the volatility of consumption. This empirical restriction means that a low EIS implies large investment frictions, and vice versa. The model features external habit preferences which imply a very low EIS, large investment frictions, and large equity volatility. This contrasts with with long-run risk and disaster models, which require a large EIS and thus generate low equity volatility.

The model also endogenizes a key mechanism of consumption-based solutions to the equity premium puzzle. External habit, long-run risk, and disaster models all assume that the volatility of marginal utility varies over time. This assumption is required in order to address the Shiller volatility puzzle: the fact that fluctuations in asset prices are far too large to be explained by fluctuations in expected dividends.[4] By assuming a time-varying volatility of marginal utility, consumption-based solutions create fluctuations in risk which drive fluctuations in asset prices while avoiding a counterfactual link with expected dividends. The solutions differ in the details of the modeling but the economics are similar in spirit. In external habit models, time-varying volatility comes from the assumed time-varying "habit sensitivity function." In long-run risk, it is the assumption of time-varying consumption volatility. In disasters, it is the assumed time-varying severity or probability of disaster. In each case, the solution to excess volatility can be quickly traced to an assumption about time-varying volatility, and in each case, the question remains: where does this time-varying volatility come from?

**Figure 1: Precautionary Savings and Time-Varying Volatility.**




The model provides an answer to this question. Time-varying volatility arises as a result of precautionary savings dynamics. The mechanism is illustrated in Figure 1, which plots consumption as a function of wealth.

[4] The volatility puzzle was first documented by LeRoy and Porter (1981) and Shiller (1981). For a recent review of the evidence, see Cochrane (2011). Earlier versions of long-run risk (Bansal and Yaron (2004)) and disasters (Rietz (1988), Barro (2006)) do not feature this assumption and cannot address the volatility puzzle, but the more recent versions do. See Bansal, Kiku, and Yaron (2009) for a recent long-run risk model. Gabaix (2012) and Wachter (2012) are a recent disaster models.

Precautionary savings motives generally lead to the concave shape depicted here (Carroll and Kimball (1996)). This shape means that in bad times, consumption is very sensitive to shocks (dashed lines), while in good times, there is little action (dotted lines). This countercyclical sensitivity leads to countercyclical consumption volatility, even in the presence of homoskedastic shocks. Though this analysis is couched in mathematical language, the math paints an intuitive story. In bad times, tomorrow's precautionary saving is very sensitive to economic news. News that the bad times will become even worse would lead investors to hunker down, increasing savings at the expense of consumption. Good news would free them from austerity, leading to a large boost in consumption. While investors await news, tomorrow's consumption is very uncertain, risk is high, and asset prices are low. This dynamic is absent in good times, when investors have a large amount of wealth and little concern about savings as protection against uncertainty.

Both production and external habit are critical for the precautionary savings mechanism. Without production, investors cannot save in aggregate, and thus precautionary savings motives will have no effect on equilibrium consumption. However, production alone cannot address the Shiller puzzle. Though production itself will lead to countercyclical volatility, the magnitude of this channel is too small with standard preferences. This is where external habit comes in. External habit amplifies the channel in two ways. The first is that it makes marginal utility more sensitive to fluctuations in consumption. As a result, small changes in consumption volatility lead to large changes in risk. The second is that habit amplifies the countercyclical consumption volatility itself. The concavity of the consumption function is driven by fear of hitting infinite marginal utility (Attanasio (1999)). External habit acts as a reduction to consumption and keeps investors close to this fear. When the habit parameters are estimated using unconditional asset price moments, the resulting countercyclical volatility does a good job of matching return and dividend predictability regressions using the price-dividend ratio, as well as estimates of GARCH effects in consumption.

This paper belongs to the literature on asset pricing models with habit formation and production. Most models in this literature consider a short-lived, internal habit (Jermann (1998) and Boldrin, Christiano, and Fisher (2001)). In contrast, I consider a slow-moving, external habit, in the style of Campbell and Cochrane (1999). Without this modification, the model would produce a counterfactually volatile risk-free rate. Lettau and Uhlig (2000) do consider a slow-moving, external habit in a production economy, but they assume costless adjustment of the capital stock. In contrast, I model convex capital adjustment costs. Without such costs, RBC models imply a counterfactually smooth Tobin's Q (Boldrin, Christiano, and Fisher (1999)). Dew-Becker (2011) does consider a slow-moving, external habit, but his preferences are recursive with a high elasticity of intertemporal substitution. As a result, the forces driving his asset pricing results are very different than those in my model.

More broadly, the model fits into the literature on asset pricing in production economies. As mentioned earlier, models in this literature struggle to describe the high volatility of equity returns, with a few exceptions. Guvenen (2009)'s limited participation model is able to match equity volatility, but at the expense of making consumption too volatile. Papanikolaou (2011) matches equity volatility by using volatile investment shocks, but he does not address time-varying risk premiums. Favilukis and Lin (2011)'s infrequent renegotiation model and Kuehn, Petrosky-Nadeau, and Zhang (2012)'s labor search model match equity volatility and generate time-varying risk premiums, but require substantial deviations from the RBC model.

The precautionary savings mechanism requires only production and "prudent" preferences (in the sense of Kimball (1990)). These conditions are satisfied in essentially all models with production, and thus this channel sheds light on the origins of time-varying risk premiums in a wide variety of models (i.e. Guvenen (2009), Favilukis and Lin (2011), and Kuehn, Petrosky-Nadeau, and Zhang (2012)). Posch (2011) provides a similar analysis, showing that non-linearities introduced by standard production technologies lead to time-varying risk premiums. My paper builds on his results by demonstrating the link between precautionary savings and these non-linearities, and provides an economic intuition for the origins of time-varying risk premiums. I also demonstrate that external habit amplifies this channel, which is important because the magnitude of this channel in standard models is too small to explain the data.

The paper proceeds as follows. Section 2 presents the model and solution method. Section 3 estimates the model using post-war data. Section 4 shows that the model provides a unified account of asset price and business cycle moments. The bulk of the paper, Section 5, is dedicated to inspecting the mechanism.

## 2 The General Equilibrium Production Model

The model sticks as close to the standard real business cycle model as possible. The only features are external habit preferences and convex capital adjustment costs. There is a representative household and representative firm. Time is discrete, the horizon is infinite, and markets are complete.

For the remainder of the paper, I denote log variables with lowercase, i.e. $z_t \equiv \log Z_t$.

## 2.1 Representative Household

There is a continuum of identical households with external habit formation preferences. Each household chooses its asset holdings to maximize

$$\mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t \frac{(C_t - H_t)^{1-\gamma} - 1}{1-\gamma} \right\} \tag{1}$$

Where $H_t$, the level of habit, is taken as external to the household. For simplicity, the household does not value leisure and is endowed with a unit of labor.

I specify the evolution of habit using surplus consumption, rather than the level of habit itself. That is, let

$$\hat{S}_t \equiv \frac{\hat{C}_t - H_t}{\hat{C}_t} \tag{2}$$

be the surplus consumption ratio, where the hats denote aggregates. This approach is chosen for comparability with the existing literature on external habit (Campbell and Cochrane (1999), Wachter (2006), among others).

Surplus consumption evolves according to an autoregressive process

$$\hat{s}_{t+1} = (1 - \rho_s)\bar{s} + \rho_s \hat{s}_t + \lambda(\hat{c}_{t+1} - \hat{c}_t) \tag{3}$$

where $\lambda$ is a *constant*. Endowment economy external habit models specify $\lambda$ as a decreasing function of surplus consumption. This assumption builds in a countercyclical volatility of marginal utility which is essential for addressing the Shiller volatility puzzle. The model does not require this assumed countercyclicality because, as we will see, production endogenously generates countercyclical consumption volatility. For comparability with the literature, I fix $\lambda$ at the Campbell and Cochrane (1999) steady state value[5]

$$\lambda = \frac{1}{\bar{S}} - 1 \tag{4}$$

The non-linear habit specification is not critical to the economic mechanism, but the persistent AR1 specification is. Persistent habit is important for capturing the persistence in various asset prices. The data show that both the price-dividend ratio and the (ex-ante) risk-free rate are very persistent. As a result, the

[5] This modification causes the issue that habit may move negatively with consumption. However, that habit is still a geometric average of the history of consumption (see Appendix A.3 for the calculation).

estimated $\rho_s$ will be close to one, meaning that habit depends on a very long history of consumption. This long dependence is in contrast with previous habit models in production economies (Jermann (1998) and Boldrin, Christiano, and Fisher (2001)) which assume that habit depends only on last quarter's consumption.

The external nature of habit produces the simple stochastic discount factor

$$M_{t,t+1} = \beta \left( \frac{\hat{C}_{t+1}}{\hat{C}_t} \frac{\hat{S}_{t+1}}{\hat{S}_t} \right)^{-\gamma} \tag{5}$$

The stochastic discount factor has the traditional consumption growth term $\left(\frac{\hat{C}_{t+1}}{\hat{C}_t}\right)^{-\gamma}$, but habit adds a surplus consumption term $\left(\frac{\hat{S}_{t+1}}{\hat{S}_t}\right)^{-\gamma}$. This SDF is very similar to that in Campbell and Cochrane (1999) but is very different from the Epstein-Zin-habit model of Dew-Becker (2011). Dew-Becker's SDF also has two components, but the two components are those generated by Epstein and Zin (1989) preferences: a consumption growth component and a term related to the return on wealth. In Dew-Becker's model, habit enters the SDF by affecting the curvature (power parameter) of the return on wealth component. As a result, the mechanisms driving Dew-Becker's model are very different.

## 2.2 Representative firm

The production side of the economy is standard. The only feature is quadratic capital adjustment costs.

There is a unit measure of identical firms which produce consumption using capital $K_t$ and labor $N_t$. Production is given by

$$\Pi(K_t, Z_t, N_t) \equiv A Z_t K_t^{\alpha} N_t^{1-\alpha} \tag{6}$$

where $\alpha$ is capital's share of output, $Z_t$ is productivity, and $A$ is chosen so that the non-stochastic steady state capital stock is approximately one. The choice of $A$ does not have a material effect on the relevant model-simulated moments, but keeping the steady state capital stock near one helps with the accuracy of the numerical methods.

The Cobb-Douglas specification of production, combined with the fact that the household does not value leisure implies that wages are equal to the marginal product of labor

$$W_t = (1-\alpha) A Z_t \hat{K}_t^{\alpha} \tag{7}$$

Productivity $Z_t$ follows the standard AR(1) process

$$z_{t+1} = \rho_z z_t + \sigma_z \epsilon_{z,t+1} \tag{8}$$

Where $\epsilon_{z,t+1}$ is a standard normal i.i.d. shock. This is the only source of uncertainty in the model, and by assumption it is homoskedastic.

Capital accumulates according to the usual capital accumulation rule

$$K_{t+1} = I_t + (1-\delta)K_t \tag{9}$$

and firms face a convex capital adjustment cost

$$\Phi(I_t, K_t) = \frac{\phi}{2}\left(\frac{I_t}{K_t} - \delta\right)^2 K_t \tag{10}$$

This formulation of adjustment costs punishes the firm for deviating from the non-stochastic steady state investment rate of $\delta$. I assume that the adjustment costs are a pure loss. They do not represent payments to labor. Adjustment costs are included because production economies produce a counterfactually smooth Tobin's Q unless one includes an investment friction.

The firm's objective is standard

$$\max_{\{I_t, K_{t+1}, N_t\}} \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} M_{0,t} [\Pi(K_t, Z_t, N_t) - W_t N_t - \Phi(I_t, K_t) - I_t] \right\} \tag{11}$$

It chooses investment, capital, and labor to maximize future dividends, discounted with the household's stochastic discount factor.

### 2.3 Recursive Competitive Equilibrium

Market clearing is standard:

$$\hat{C}_t + \hat{I}_t = Z_t \hat{K}_t^{\alpha} - \Phi(\hat{I}_t, \hat{K}_t) \tag{12}$$

Due to the consumption externality, the welfare theorems do not hold and the equilibrium cannot be easily described by a social planner's problem. Thus, I define equilibrium competitively. The aggregate state variables are aggregate capital $\hat{K}$, surplus consumption $\hat{S}$, and productivity $Z$.

**Definition.** *Equilibrium is a firm decision rule for investment $I(K; \hat{K}, \hat{S}, Z)$, cum-dividend value function*

*$V(K;\hat{K},\hat{S},Z)$, law of motion for aggregate consumption $\hat{C}(\hat{K},\hat{S},Z)$, and a law of motion of aggregate capital $\Gamma(\hat{K},\hat{S},Z)$ such that*

*(i) Firms optimize: $I(K;\hat{K},\hat{S},Z)$ and $V(K;\hat{K},\hat{S},Z)$ solve*

$$V(K;\hat{K},\hat{S},Z) = \max_{\{N,I,K'\}} \Big\{ \Pi(K,Z,N) - W(\hat{K},\hat{S},Z)N - \Phi(I,K) - I + \int_{-\infty}^{\infty} dF(\epsilon')M(\hat{K},\hat{S},Z;Z')V(K';\hat{K}',S',Z') \Big\} \tag{13}$$

*subject to capital accumulation (9), competitive wages (7), the productivity process (8), the habit process (3), the SDF is equal to the household's IMRS (5), $\hat{K}' = \Gamma(\hat{K},\hat{S},Z)$, and where $F(\epsilon')$ is the standard normal CDF.*

*(ii) Markets clear and aggregates are consistent with individual behavior:*

$$\hat{C}(\hat{K}',\hat{S}',Z') = \Pi(\hat{K},Z,1) - \Phi(I(\hat{K};\hat{K},\hat{S},Z),\hat{K}) - I(\hat{K};\hat{K},\hat{S},Z) \tag{14}$$

$$\Gamma(\hat{K},\hat{S},Z) = (1-\delta)\hat{K} + I(\hat{K};\hat{K},\hat{S},Z) \tag{15}$$

Since households and firms are identical, in equilibrium, $\hat{K} = K$, $\hat{S} = S$ and $\hat{C} = C$. Thus, in what follows, I drop the hats.

## 2.4 Solution Method

Though the Campbell and Cochrane (1999) preferences have been around for quite some time, to my knowledge, no previous paper has examined its asset pricing implications in a production economy. This may be because the numerical problem is quite challenging. I discuss these challenges and the solutions in this section.

### 2.4.1 Projection Method

I solve the model using a projection method (Judd (1992)). Specifically, I represent the law of motion for capital using cubic splines and then use Broyden's method (a quasi-Newton algorithm) to find cubic spline coefficients which satisfy the firm's Euler equation. The solution program makes extensive use of the Miranda and Fackler (2001) CompEcon toolbox.

The projection method is important for two reasons. The first is that, due to the consumption externality, the welfare theorems do not hold and the model cannot be easily solved by value function iteration.

Projection avoids this problem by working with equilibrium conditions from the recursive competitive equilibrium. The second is that projection produces a global and non-linear solution. These properties are important for asset pricing models in general (Cochrane (2008b)), but are particularly important for capturing the mechanisms in this model. The precautionary savings channel which drives countercyclical risk premiums are related to the third derivative of the utility function and are not captured by traditional linearizations. Moreover, precautionary effects are particularly pronounced when the investor is threatened with infinite marginal utility (Attanasio (1999)), and fully capturing this effect requires a global solution.

### 2.4.2 Homotopy Method

Projection methods require a good initial guess of the spline coefficients. There is no guarantee that a general non-linear equation solver will converge and the high dimensionality of the problem tends to make the solvers unstable. The standard approach is to use the real business cycle model as an initial guess. Unfortunately, with Campbell and Cochrane (1999) preferences, the real business cycle model is a poor initial guess because external habit has a dramatic effect on the laws of motion of the model.

To overcome this issue, I use a homotopy method. Specifically, I modify the firm's problem so that it discounts future profits using the SDF

$$M' = \beta \left[ \frac{C'}{C} \left( \frac{S'}{S} \right)^{\chi} \right]^{\gamma}$$

Note that $\chi = 0$ corresponds to a model with no habit, and $\chi = 1$ corresponds to the full model. I begin by solving the model for $\chi = 0$, and then slowly increase $\chi$, using the coefficients from the previous $\chi$ as the initial guess for the current $\chi$.

This homotopy algorithm is very computationally intensive. To aid in the speed of computation, I discretize the productivity process using the Rouwenhorst method.

### 2.4.3 Solving for the Evolution of Surplus Consumption

Another difficult issue which arises with Campbell and Cochrane (1999) preferences is that they result in a state variable which is not predetermined and is endogenous. Surplus consumption tomorrow is not known today and it depends on consumption tomorrow, which is endogenous. This makes solving for the surplus consumption process rather difficult.

To see this clearly, it helps to write the evolution of surplus consumption (3) as functions of state variables

$$\log S' = (1-\rho_s)\bar{s} + \rho_s \log S + \lambda[c(K', S', Z_j) - c(K, S, Z_j)] \tag{16}$$

Where $Z_j$ is tomorrow's discrete productivity state. Since $K'$ is predetermined, $S'$ is a function of four variables $s'(K, S, Z_i, Z_j)$. Note that surplus consumption tomorrow $S'$ appears on both sides of this equation. Thus determining surplus consumption tomorrow requires solving this non-linear equation. I once again use Broyden's method to solve this equation. This calculation is very computationally intensive since must be done at every collocation node for every potential productivity shock within every iteration of the big Broyden's method which is solving for the coefficients of the law of motion of capital.

For further details about the solution method, see Appendix A.4.

# 3 Simulated Method of Moments Estimation

I estimate the model using simulated method of moments (SMM), the simulated version of Hansen (1982)'s generalized method of moments (GMM). SMM differs from traditional GMM in that it uses simulation to compute model moments rather than closed-form expressions, which may be unavailable for moments of interest. Duffie and Singleton (1993) derive adjustments to GMM formulas and additional regularity conditions required by the use of simulation.

## 3.1 Data

The estimation uses post-war (1948-2011) data from the CRSP, BEA, and BLS. Some authors argue for using the longest sample available when evaluating consumption-based asset pricing models (e.g. Bansal, Kiku, and Yaron (2009)). However, in a production economy I must also address the data on aggregate investment and output. The nature of this data is significantly affected by using pre-war data. For example, the correlation between investment growth and output growth have a high correlation of 0.73 in the post-war sample, but have a mild correlation of 0.23 for the sample 1929-2011. This consideration leads me to target only post-war data.

All variables are real and per capita. Consumption is real per capita non-durable goods and services consumption. This measure excludes volatile consumer durables such as automobiles, which produce a smooth consumption flow over the life of the durable. Aggregate equity is represented by the CRSP value-weighted index, adjusted for inflation with the consumer price index. Dividends are calculated with the assumption that all dividends are reinvested in the stock market. This method of aggregation preserves the

Campbell and Shiller (1988) present value identity. The risk-free rate is a forecast of the inflation-adjusted 90-day T-bill return using the previous year's inflation rate and the nominal 90-day yield (following Beeler and Campbell (2009)). No adjustments are made for financial leverage. Further details regarding the data are found in the Appendix A.1.

### 3.2 Estimation Method and Predefined Parameters

Because SMM is computationally intensive, I set the more traditional parameters outside of the estimation. I set the depreciation rate $\delta = 0.02$ to match the mean, growth-adjusted, investment rate. I set the capital share parameter $\alpha = 0.35$ to match the capital share of output implied by constant returns to scale. The persistence of productivity shocks $\rho_z = 0.979$ matches the persistence of the Solow residual with a fixed labor input. In an external habit model, the utility curvature $\gamma$ and the steady state surplus consumption ratio $\bar{S}$ jointly control risk aversion. As a result, it is difficult to identify these parameters separately. For ease of comparison with the literature, $\gamma = 2$ is set at Campbell and Cochrane (1999)'s value.[6]

The remaining five parameters are estimated by SMM. Explicitly, let $\theta$ represent the five parameters as a vector and $\hat{M}^*$ represent a vector of target data moments. SMM estimates the parameters by minimizing the distance between data moments and simulated moments:

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \left[\hat{M}^* - M(\theta)\right]' W \left[\hat{M}^* - M(\theta)\right] \tag{17}$$

where $M(\theta)$ is the model-simulated counterpart to $\hat{M}^*$ and $W$ is a weighting matrix chosen by the econometrician. I use one-stage, exactly-identified GMM, that is, I target five data moments which are economically informative about the five parameter values. This approach has the advantage of transparency. Seeing that the empirical targets are equal to the model moments immediately verifies that the minimization algorithm is successful. Under exact identification, a consistent estimator for the asymptotic variance of the estimated parameter values is

$$\widehat{\text{Var}}[\sqrt{T}(\hat{\theta} - \theta_0)] \equiv \left(1 + \frac{1}{S}\right)[DM(\hat{\theta})']^{-1}\widehat{\text{Var}}(\sqrt{T}\hat{M}^*)[DM(\hat{\theta})]^{-1} \tag{18}$$

where $S$ is the number of simulations used to calculate model moments, $DM(\theta)$ is the derivative of the simulated moments with respect to $\theta$ and $\widehat{\text{Var}}(\sqrt{T}\hat{M}^*)$ is a consistent estimate of the asymptotic variance of the estimated data moments. Intuitively, we have a good estimate if the moments are informative about the parameter values ($DM(\hat{\theta})$ is large), or if we have a precise estimate of the moments ($\widehat{\text{Var}}(\sqrt{T}\hat{M}^*)$ is small).[7]

[6] Campbell and Cochrane (1999) choose $\gamma$ to fit the Sharpe ratio and $\bar{S}$ to produce a constant risk-free rate, but that cannot be done here because I set $\lambda$ as a constant in order eliminate exogenous time-varying volatility.

[7] The moments I use are not moments in the strict sense (i.e. the Sharpe ratio), and I estimate transformations of the parameters

I optimize using Levenberg-Marquardt, a variant of Newton's method. I choose this method rather than the more commonly-used simulated annealing algorithm for a few reasons. The first is that the moment function does not display an extreme number of local minima, which is where simulated annealing has an advantage. With a relatively smooth objective function, a method which uses derivative information is much more efficient. Another advantage of LM is robustness. Simulated annealing tends to be sensitive to the choice of the optimization parameters (Press, Teukolsky, Vetterling, and Flannery (1992)). Additional details regarding the estimation method can be found in Appendix A.2.

### 3.3 Parameter Estimates and Moment Targets

Table 1 summarizes the estimation. It shows estimated parameter values, standard errors, and targeted moments. For convenience, the predefined parameters are also shown, with standard errors omitted. The model is quarterly and all parameter values are quarterly. To eliminate seasonality in dividends, both the U.S. and simulated data are aggregated to the annual level. All moments are annual.

Preference parameters are identified with asset prices. Because time-preference $\beta$ is reflected in the risk-free rate, I choose the mean 90-Day T-bill return as a target moment. The resulting $\beta = 0.970$ is rather low because the model features a non-trivial precautionary savings motive, and a low level of patience helps counteract that motive. Steady state surplus consumption $\bar{S}$ controls the magnitude of habit, and, in effect, the degree of risk aversion in the model. I thus choose the mean Sharpe ratio of the CRSP index as a target moment. The resulting value $\bar{S} = 0.062$ is close to the values used in the external habit literature (Campbell and Cochrane (1999), Wachter (2006), Santos and Veronesi (2010)). The persistence of surplus consumption $\rho_s$ has a strong effect on the volatility of the market return, and so I choose the volatility of the excess return on the CRSP index as a target. The resulting value of $\rho_s = 0.963$ indicates a very persistent habit process, which is also consistent with values used in the external habit literature.

Technological parameters are identified with moments of the real economy. The volatility of productivity $\sigma_z = 0.014$ targets the volatility of HP-filtered log GDP of 0.014. Since the data is annual, I use the annual smoothing parameter of 6.25 advocated by Ravn and Uhlig (2002). The quadratic adjustment cost parameter $\phi = 74.94$ targets the relative volatility of consumption to GDP (also HP-filtered). This estimated value results in mean adjustment costs as a percentage of output of less that 1%.

---

in order to avoid corner solutions. This leads to slightly more complicated formulas than those presented here. These details can be found in the Appendix A.2.

**Table 1: Parameter Estimates and Moment Targets**

The model is quarterly, and all parameter values are quarterly. Empirical figures are annual. Standard errors are Newey-West with 10 lags and shown only for estimated parameters. The sample period is 1948-2011. Consumption is real non-durable goods and services consumption. GDP and consumption are logged and HP-filtered with a smoothing parameter of 6.25. Further details are found in Appendices A.1 and A.2.

| Parameter | | Estimate | SE | Empirical Target | |
|---|---|---|---|---|---|
| Preferences | | | | | |
| $\beta$ | Time Preference | 0.970 | (0.009) | Mean 90-Day T-bill Return | 0.010 |
| $\rho_s$ | Persistence of Surplus Consumption | 0.963 | (0.006) | Volatility of Excess Return of CRSP Index | 0.161 |
| $\bar{S}$ | Steady State Surplus Consumption | 0.062 | (0.010) | Mean Sharpe Ratio of CRSP Index | 0.478 |
| $\gamma$ | Utility Curvature | 2 | | (Chosen to Match Campbell-Cochrane (1999) | |
| Technology | | | | | |
| $\phi$ | Adjustment Cost Parameter (Quadratic) | 74.94 | (10.80) | Relative Volatility of Consumption to GDP | 0.47 |
| $\sigma_z$ | Volatility of Productivity | 0.014 | (0.001) | Volatility of GDP | 0.015 |
| $\rho_z$ | Persistence of Productivity | 0.979 | | Persistence of Solow Residual | 0.979 |
| $\alpha$ | Capital Share | 0.35 | | Capital's Share of GDP | 0.35 |
| $\delta$ | Depreciation Rate | 0.02 | | Mean Investment Rate (Growth-Adjusted) | 0.02 |

# 4 Matching Asset Price and Business Cycle Moments

This section discusses the main quantitative results. It shows that the model provides a unified description of aggregate asset price and business cycle moments. I make no adjustments to account for un-modeled leverage or payout policy. The price of equity is simply the present value of dividends from the representative firm, discounted with the household's SDF.

## 4.1 Unconditional Asset Price Moments

Table 2 shows that the model produces a nice fit for all of the basic moments of asset prices. As intended by the estimation, it hits three of these moments exactly: the mean risk-free rate, the Sharpe Ratio, and the volatility of excess returns. As mentioned in the introduction, producing this large equity volatility is a difficult task in production economies (Gourio (2010), Kaltenbrunner and Lochstoer (2010), Croce (2010)).

The model also captures many asset market features beyond those used in the identification. The model

### Table 2: Unconditional Asset Price Moments

Figures are annual. No adjustments are made to account for financial leverage. The model columns show means and percentiles across simulations. $r$, $p$, and $d$ are the logs of returns, prices, and dividends from the CRSP value-weighted index. Capital letters show levels rather than logs. $r_f$ is a forecast of the ex-post real return on 90-day T-bills following Beeler and Campbell (2009). $\mathbb{E}$, $\sigma$, and AC1 represent the sample mean, standard deviation and first-order autocorrelation. Further details are found in Appendix A.1.

| | Data | Model | | | |
|---|---|---|---|---|---|
| | 1948-2011 | mean | 5% | 50% | 95% |
| Identifying Moments | | | | | |
| $\mathbb{E}(r_f)$ (%) | 0.98 | 0.98 | -2.76 | 0.31 | 6.80 |
| $\mathbb{E}(R-R_f)/\sigma(R)$ | 0.48 | 0.48 | 0.29 | 0.48 | 0.68 |
| $\sigma(r-r_f)$ (%) | 16.07 | 16.07 | 10.74 | 15.71 | 22.59 |
| | | | | | |
| Untargeted Moments | | | | | |
| $\mathbb{E}(r-r_f)$ (%) | 6.47 | 6.77 | 5.05 | 6.95 | 7.88 |
| $\text{AC1}(r-r_f)$ | -0.03 | -0.07 | -0.28 | -0.07 | 0.16 |
| | | | | | |
| $\sigma(r_f)$ (%) | 2.24 | 2.96 | 0.59 | 1.90 | 8.71 |
| $\text{AC1}(r_f)$ | 0.56 | 0.89 | 0.69 | 0.91 | 0.99 |
| | | | | | |
| $\mathbb{E}(p-d)$ | 3.42 | 2.67 | 2.09 | 2.66 | 3.27 |
| $\sigma(p-d)$ | 0.43 | 0.46 | 0.26 | 0.44 | 0.76 |
| $\text{AC1}(p-d)$ | 0.95 | 0.90 | 0.79 | 0.91 | 0.96 |
| | | | | | |
| $\mathbb{E}(r_{10yr}-r_f)$ (%) | | 2.04 | 0.92 | 2.12 | 2.97 |
| $\sigma(r_{10yr}-r_f)$ (%) | | 1.44 | 0.62 | 1.39 | 2.49 |
| $\text{AC1}(r_{10yr}-r_f)$ | | 0.85 | 0.69 | 0.86 | 0.95 |

produces a risk-free rate volatility of 2.96% which is close to the data estimate of 2.24%. This moment is difficult to match in habit models, which tend to produce an excessively volatile risk-free rate (i.e. 11% in Jermann (1998) and 25% in Boldrin, Christiano, and Fisher (1999)). This low risk-free rate volatility is reflected in a reasonable term premium. The model produces a mean excess return on 10-year bonds of about 2%, indicating that the model's equity premium is distinct from the term premium, as in the data. The volatility of the log price-dividend ratio is 0.46, which is close to the data value of 0.43. This moment has been difficult to capture even in endowment economies (Bansal, Gallant, and Tauchen (2007), Bansal, Kiku, and Yaron (2009)).

Additionally, the model generates data-like persistence. As in the data, the excess market return is mildly negatively autocorrelated, and the risk-free rate and price-dividend ratio are highly positively autocorrelated. The model somewhat overstates the persistence of the risk-free rate, but this may be because the model abstracts from monetary policy issues which affect the risk-free rate in the data.

### 4.2 The Shiller Volatility Puzzle

The previous section shows that the model generates a volatile, data-like stock market. This large volatility does not mean that the model captures the nature of stock market fluctuations however. The Shiller (1981) volatility puzzle tells us that these fluctuations have little relationship with expected dividend growth, but are closely linked to expected returns.

Table 3 shows that the model captures key elements of the Shiller puzzle. Panel B shows regressions of future dividend growth on the log price-dividend ratio at various horizons. The model predicts no relationship between the price-dividend ratio and future dividend growth at the one-year horizon, as is seen in the data. As in the data, this predictability increases with the horizon, but the coefficients remain small and statistically insignificant.

If the price-dividend ratio does not forecast dividends, then it must forecast returns (Campbell and Shiller (1988), Cochrane (2008a)). Panel B shows the ability of the model to capture this flip side of the Shiller puzzle. It shows regressions of future excess returns on the log price-dividend ratio. Regression coefficients, standard errors, and $R^2$'s, are close to the data for all forecasting horizons. The coefficients on the price-dividend ratio are negative, and they are both economically and statistically significant. At the one-year horizon, the model exactly matches the data coefficient of -0.12. To understand this economically, recall that the volatility of the log price-dividend ratio is roughly 0.40 in both the model and the data. This means that a one standard deviation rise in the price-dividend ratio predicts a huge 4% reduction in the equity premium over the next year. This forecasting power increases with the forecasting horizon, reaching $R^2$'s of roughly 30% at the 5-year horizon in both model and data.

Overall, the model captures both sides of the Shiller puzzle. Asset price fluctuations have little relationship with fluctuations in future dividends, but are tightly linked to fluctuations in future returns.

### 4.3 Unconditional Business Cycle Moments

The previous sections show that the model produces a good description of numerous asset price moments. Table 4 shows that these asset price moments come with data-like fluctuations in the real economy. The model hits the volatility of output and relative volatility of consumption to output, as intended by the estimation. Investment is more volatile than output, as in the data. Like the data, the model moments display strong co-movement between output, consumption, and investment. Output, consumption, and investment are highly persistent, and are nearly as persistent as the data. First-differenced log consumption

### Table 3: Predicting Dividends and Returns with the Price-Dividend Ratio

Figures are annual. $r_t$, $p_t$, and $d_t$ are the log-returns, prices, and dividends from the CRSP value-weighted index. $r_{f,t}$ is a forecast of the ex-post real return on 90-day T-bills following Beeler and Campbell (2009). Further details are found in Appendix A.1.

Panel A: Predicting dividend growth

$$\sum_{j=1}^{L} \Delta d_{t+j} = \alpha + \beta(p_t - d_t) + \epsilon_{t+L}$$

| | L | Data 1948-2011 | Model Mean | Model 5% | Model 50% | Model 95% |
|---|---|---|---|---|---|---|
| $\hat{\beta}$ | 1 | -0.03 | -0.00 | -0.04 | -0.00 | 0.03 |
| | 3 | 0.01 | 0.07 | 0.00 | 0.06 | 0.18 |
| | 5 | 0.03 | 0.13 | 0.02 | 0.12 | 0.30 |
| SE($\hat{\beta}$) | 1 | 0.03 | 0.03 | 0.01 | 0.02 | 0.04 |
| | 3 | 0.07 | 0.05 | 0.03 | 0.05 | 0.09 |
| | 5 | 0.09 | 0.07 | 0.03 | 0.06 | 0.11 |
| $R^2$ | 1 | 0.01 | 0.00 | 0.00 | 0.00 | 0.02 |
| | 3 | 0.00 | 0.04 | 0.00 | 0.03 | 0.13 |
| | 5 | 0.00 | 0.10 | 0.01 | 0.09 | 0.27 |

Panel B: Predicting excess returns

$$\sum_{j=1}^{L} r_{t+j} - r_{f,t+j} = \alpha + \beta(p_t - d_t) + \epsilon_{t+L}$$

| | L | Data 1948-2010 | Model Mean | Model 5% | Model 50% | Model 95% |
|---|---|---|---|---|---|---|
| $\hat{\beta}$ | 1 | -0.12 | -0.12 | -0.23 | -0.11 | -0.04 |
| | 3 | -0.27 | -0.31 | -0.56 | -0.29 | -0.11 |
| | 5 | -0.40 | -0.44 | -0.78 | -0.44 | -0.17 |
| SE($\hat{\beta}$) | 1 | 0.05 | 0.05 | 0.02 | 0.05 | 0.08 |
| | 3 | 0.08 | 0.09 | 0.04 | 0.08 | 0.13 |
| | 5 | 0.12 | 0.10 | 0.05 | 0.10 | 0.17 |
| $R^2$ | 1 | 0.09 | 0.10 | 0.02 | 0.10 | 0.19 |
| | 3 | 0.19 | 0.26 | 0.07 | 0.26 | 0.45 |
| | 5 | 0.26 | 0.37 | 0.12 | 0.38 | 0.59 |

has a low volatility. In both the model and data this volatility is about 1% per year. Lastly, the table also shows that average adjustment costs are small at less than 1% of output.

**Table 4: Basic Business Cycle Moments**

Figures are annual. The subscript $hp$ indicates that the moment is calculated from logged and HP-filtered data with a smoothing parameter of 6.25. $\Delta$ indicates first-differences. Further details are found in the Appendix A.1.

| | Data | Model | | | |
|---|---|---|---|---|---|
| | 1948-2011 | mean | 5% | 50% | 95% |
| Identifying Moments | | | | | |
| $\sigma(y_{hp})$ (%) | 1.50 | 1.50 | 1.11 | 1.49 | 1.93 |
| $\sigma(c_{hp})/\sigma(y_{hp})$ | 0.47 | 0.47 | 0.40 | 0.47 | 0.55 |
| Untargeted Moments | | | | | |
| $\sigma(i_{hp})/\sigma(y_{hp})$ | 2.46 | 3.44 | 3.09 | 3.44 | 3.77 |
| $\rho(c_{hp}, y_{hp})$ | 0.83 | 0.99 | 0.98 | 0.99 | 1.00 |
| $\rho(i_{hp}, y_{hp})$ | 0.85 | 1.00 | 0.99 | 1.00 | 1.00 |
| AC1$(y_{hp})$ | 0.12 | 0.21 | 0.01 | 0.21 | 0.38 |
| AC1$(c_{hp})$ | 0.32 | 0.21 | 0.01 | 0.22 | 0.39 |
| AC1$(i_{hp})$ | 0.27 | 0.20 | 0.01 | 0.21 | 0.38 |
| $\sigma(\Delta c)$ (%) | 1.11 | 1.41 | 0.97 | 1.38 | 1.90 |
| $\mathbb{E}(\text{Adj Cost}/Y)$ (%) | | 0.54 | 0.21 | 0.49 | 1.04 |
| $\mathbb{E}(\text{Adj Cost}/I)$ (%) | | 3.20 | 1.14 | 2.71 | 6.87 |

# 5 Inspecting the Mechanism

I have shown that the model provides a unified description of numerous asset price and business cycle moments. The remainder of the paper is devoted to explaining how.

Most of the new economics can be seen by examining the Shiller volatility puzzle. Thus, I begin by showing how the model addresses the puzzle with countercyclical consumption volatility (Section 5.1). I go on to illustrate how precautionary savings motives underlie this countercyclical consumption volatility, and how external habit amplifies this effect (Section 5.2). I then check that this countercyclical consumption volatility is consistent with the data (Section 5.3).

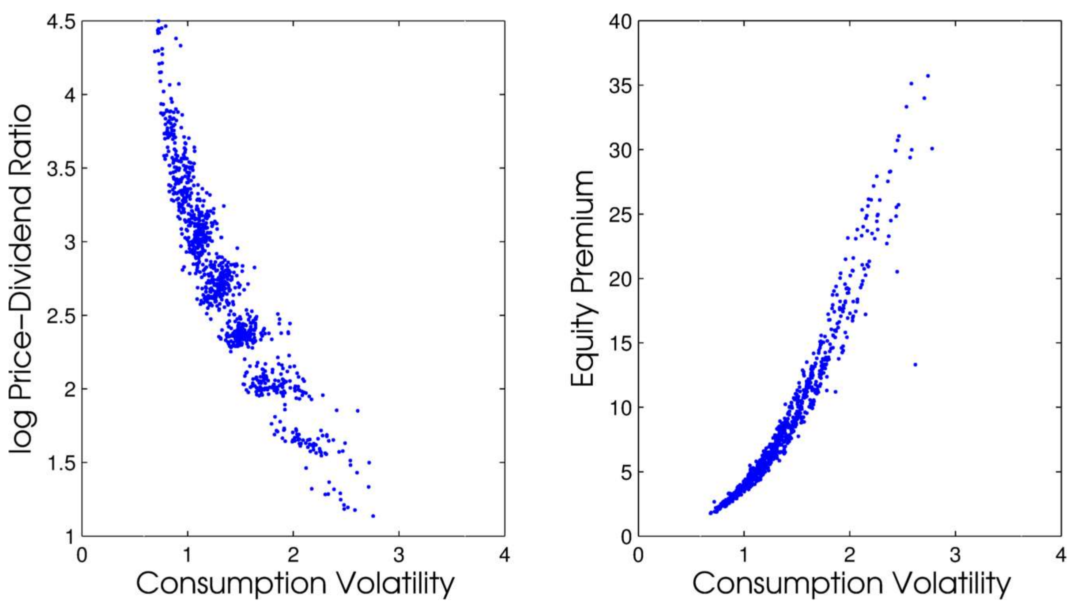
The large and volatile equity premium is analyzed in Section 5.4. The low and smooth risk-free rate is examined in Section 5.5. Section 5.6 ends the inspection by examining comparative statics.

## 5.1 Countercyclical Consumption Volatility and the Shiller Volatility Puzzle

The key to addressing the Shiller puzzle is to generate fluctuations in risk. Fluctuations in risk shift discount rates, which in turn shift asset prices without producing a counterfactual link with expected dividends.

The model produces fluctuations in risk by producing countercyclical consumption volatility. This can be seen in Figure 2, which shows scatterplots of asset prices against consumption volatility produced by model simulations. Consumption volatility varies over time, and both the price-dividend ratio and equity premium are nicely linked to consumption volatility. The price-dividend ratio declines in consumption volatility, while the equity premium increases.

**Figure 2: Time-Varying Consumption Volatility and Asset Prices.** The figures show scatterplots from model simulations. Consumption volatility and equity premiums are calculated using the model's laws of motion. All values are annualized. A.1.




This result has a very simple intuition. Times of high consumption volatility are risky times with large risk premiums and low asset prices. But a little formalism provides additional insight and builds confidence in the mechanism. A log-normal approximation of the SDF shows that the conditional maximum Sharpe ratio is[8]

$$\max_{\{\text{all assets}\}} \left\{ \frac{\mathbb{E}_t(R_{t+1} - R_{f,t+1})}{\sigma_t(R_{t+1})} \right\} \approx \gamma(\lambda+1)\sigma_t(\Delta c_{t+1}) \tag{19}$$

The conditional maximum Sharpe ratio is the conditional volatility of consumption growth, multiplied by preference parameters ($\gamma$ is the utility curvature, and $\lambda$ is the conditional volatility of the habit process). This expression shows that consumption volatility is critical. Since consumption volatility is the sole term on the RHS which can vary over time, it must be the driver of time-varying Sharpe ratios.

This expression also shows that the mechanism is distinct from that of standard external habit models (Campbell and Cochrane (1999), Santos and Veronesi (2010)). Standard external habit models drive countercyclical risk premiums through a time-varying preference parameter rather than time-varying consumption volatility. Indeed, the model's mechanism is more closely related to the time-varying consumption volatility channel of Bansal and Yaron (2004). This can be seen by examining the maximum Sharpe ratio from Campbell and Cochrane (1999):

$$\max_{\{\text{all assets}\}} \left\{ \frac{\mathbb{E}_t(R_{t+1} - R^f_{t+1})}{\sigma_t(R_{t+1})} \right\} \approx \gamma[\lambda(s_t)+1]\sigma(\Delta c_{t+1}) \tag{20}$$

where

$$\lambda(s_t) = \begin{cases} \frac{1}{\bar{S}}\sqrt{1-2(s_t-\bar{s})} - 1 & \text{if} \quad s_t \leq \bar{s} + \frac{1}{2}(1-\bar{S}^2) \\ 0 & \text{if} \quad s_t \geq \bar{s} + \frac{1}{2}(1-\bar{S}^2) \end{cases} \tag{21}$$

The above expression looks just like the maximum Sharpe ratio in my model (equation (19)), but the time dependence is reversed. $\sigma(\Delta c_{t+1})$ is assumed to be constant, but now the preference parameter $\lambda(s_t)$ changes over time.

This analysis also shows that the common intuition regarding the external habit mechanism is misplaced. The common intuition is that countercyclical risk premiums are driven by time-varying risk aversion. But time-varying risk aversion is not sufficient for generating countercyclical Sharpe ratios. To see this, note that a constant $\lambda(s_t)$ still results in time-varying surplus consumption $S_t$ (see equation (3)). Since risk aversion $\sim \gamma S_t^{-1}$, a constant lambda still produces time-varying risk aversion. But equation (20) shows that, even though risk aversion is time-varying, the constant $\lambda(s_t)$ implies a constant maximum Sharpe ratio. It is time-varying "habit sensitivity" $\lambda(s_t)$, not time-varying risk aversion, which drives the external habit

[8]The log-SDF is $m_{t+1} = \log\beta - \gamma\Delta s_{t+1} - \gamma\Delta c_{t+1}$ and the habit process is $\Delta s_{t+1} = -(1-\rho_s)(s_t - \bar{s}) + \lambda\Delta c_{t+1}$. Plug the habit process into the log-SDF, then assume that the SDF is log-normal, and we have $\frac{\sigma_t(M_{t+1})}{\mathbb{E}_t(M_{t+1})} \approx \sqrt{e^{\text{Var}_t[m_{t+1}]} - 1} \approx \sigma_t[m_{t+1}] = \gamma(\lambda+1)\sigma_t(\Delta c_{t+1})$.

model.[9]

The fact that, in traditional external habit models, time-varying habit sensitivity drives countercyclical risk leaves open many questions. Where does this time-varying sensitivity come from? How do we interpret a time-varying habit sensitivity? This habit model can help address these questions since it replaces the time-varying habit sensitivity function with countercyclical consumption volatility, which has a very simple interpretation.

Indeed, the countercyclical consumption does a good job of mimicking the habit sensitivity function of Campbell and Cochrane (1999). To demonstrate this I construct an 'implied consumption volatility' from the Campbell-Cochrane sensitivity function by equating the two expressions for the maximum Sharpe ratio (equations (20) and (19)). I call this implied consumption volatility because one can replace the Campbell-Cochrane sensitivity function with this implied consumption volatility and get similar quantitative results. Explicitly,

$$\text{Implied Consumption Volatility} = \left(\frac{\lambda(s_t)+1}{\lambda+1}\right)\sigma(\Delta c_{t+1}) \tag{22}$$

Implied consumption volatility is simply the habit sensitivity $\lambda(s_t)$, re-scaled by its steady state value $\lambda$, and multiplied by the unconditional volatility $\sigma(\Delta c_{t+1})$.

Figure 3 compares the Campbell-Cochrane implied consumption volatility to the consumption volatility of my model. It shows scatterplots of the volatilities against surplus consumption. Both are countercyclical, that is, they decline in surplus consumption. Moreover, for much of the plot, implied consumption volatility runs right through the middle of the cloud of dots representing consumption volatility. The two channels are both qualitatively and quantitatively similar. Thus, this production economy could be considered a method for endogenizing the external habit mechanism.

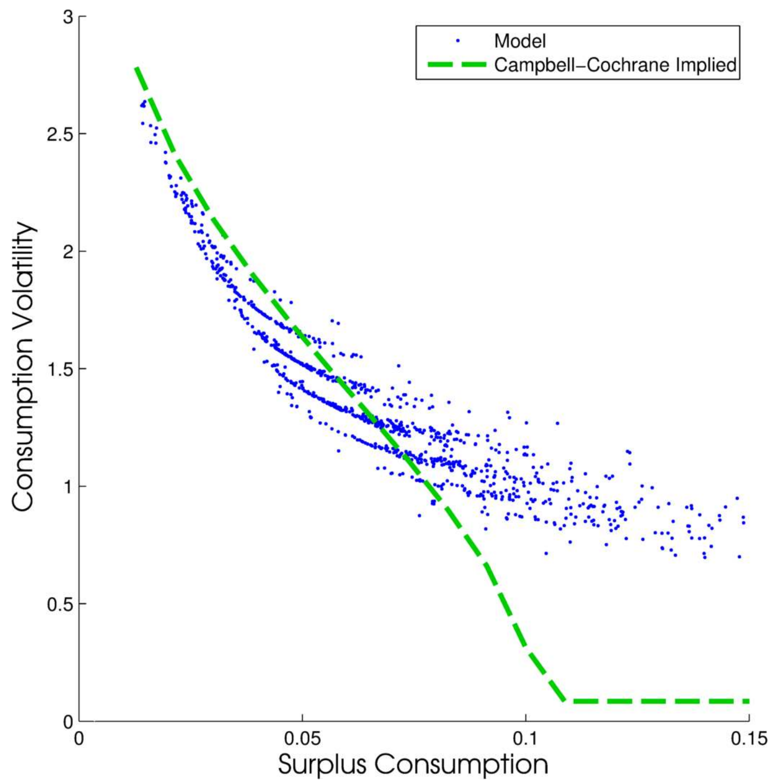
### 5.2 Precautionary Savings and Countercyclical Consumption Volatility

The previous section shows that countercyclical consumption volatility is the key to addressing the Shiller volatility puzzle. This channel is similar in spirit to the stochastic volatility version of the long-run risk model (Bansal, Kiku, and Yaron (2009)), but there is an important difference. Bansal, Kiku, and Yaron (2009) exogenously specify countercyclical consumption volatility and thus leave the deeper question of its economic origins unanswered.

Here, countercyclical consumption volatility is endogenous and due to precautionary savings dynamics.

[9] Dew-Becker (2011) uses Epstein-Zin preferences to create a habit model which captures the common intuition.

**Figure 3: Consumption Volatility and the Campbell-Cochrane $\lambda(s_t)$**. Dots represent consumption volatility from model simulations. The dashed line represents Campbell-Cochrane $\lambda(s_t)$ implied consumption volatility (equation (22)). Figures are annualized.




Precautionary savings motives result in a consumption function which is strictly concave in wealth (Carroll and Kimball (1996)). This shape implies that, when the investor is poor, consumption is more sensitive to shocks, and thus more volatile. External habit amplifies the channel to make it economically significant, and production allows these dynamics to be reflected in aggregate consumption. The full model is too rich to allow for a theoretical analysis of this mechanism. But here I present a simple consumption-savings problem problem which permits proofs and propositions.[10]

Consider a two-period-lived investor with the same period utility as the full model

$$u(C-H) = \frac{(C-H)^{1-\gamma}-1}{(1-\gamma)}$$

with $\gamma > 0$. For simplicity, he has no time preference.

At date 0, the investor has wealth $W_0$. Nothing occurs at date 0, but the date helps to serve as a reference point. Throughout this section, I describe a variable as 'countercyclical' if it is negatively related to $W_0$. At

[10] The consumption-savings problem can be formalized as a general equilibrium model with a linear technology (a la Constantinides (1990)), but the notation is simpler as a consumption-savings problem.

date 1, the investor has a habit level of $H_1 > 0$, and he receives a wealth shock $\Delta W_1$, making his wealth $W_1 = W_0 + \Delta W_1$. He consumes $C_1$, and saves the rest of his wealth $W_1 - C_1$ in a risky asset. At date 2, the investor has a habit level of $H_2 > 0$. The return on the risky asset $R_2$ is realized, and he consumes his remaining wealth $R_2(W_1 - C_1)$.

The model comes down to finding the investor's optimal consumption at date 1. His optimal consumption policy solves

$$
\begin{aligned}
C_1(W_1) &= \arg\max_{\{C_1\}} u(C_1 - H_1) + \mathbb{E}_1[u(C_2 - H_2)] \qquad (23)\\
s.t. \quad & C_2 = R_2(W_1 - C_1)
\end{aligned}
$$

Taking a Taylor of expansion of $C_1(W_1)$ around $W_0$, the volatility of $C_1$ is approximately

$$
\begin{aligned}
\sigma_0[C_1(W_1)] &\approx \sigma_0[C_1(W_0) + C_1'(W_0)\Delta W_1] \\
&= C_1'(W_0)\sigma_0[\Delta W_1] \qquad (24)
\end{aligned}
$$

That is, consumption volatility is proportional to the marginal propensity to consume (MPC). Intuitively, the MPC captures the responsiveness of consumption to shocks. So the higher the MPC, the more responsive, and thus the more volatile, is consumption. For the remainder of this section, I assume that $\Delta W_1$ is small enough so that equation (24) is a good approximation.

Unfortunately, even in a two-period consumption savings problem, the presence of precautionary savings motives typically precludes closed-form solutions (Carroll (2001)). However, using the methods from Carroll and Kimball (1996), I can still prove that the solution exhibits countercyclical volatility and that habit amplifies this cyclicality.

**Proposition 1.** *The date 1 MPC is decreasing in wealth, that is,* $C_1''(W_1) < 0$.

The essence of the proof is that, as long as the date 2 habit is positive, the model falls into the broad class of models for which consumption is strictly concave (Carroll and Kimball (1996)). Carroll and Kimball do not show the strict concavity result for this model, but I show in Appendix A.6 that their proofs can be extended to include this setting. The proof is rather technical, but, intuitively, the investor saves for precautionary reasons in the presence of uncertainty. As the agent becomes wealthier, the uncertainty becomes less relevant, and this motive declines, creating convex savings and concave consumption.

**Corollary 1.** *The volatility of* $C_1$ *is countercyclical, that is,* $\frac{\partial}{\partial W_0}\sigma_0[C_1(W_1)] < 0$.

The proof is short and illustrates the role of the MPC, so I state it here.

*Proof.* Because the MPC is decreasing in wealth, and the volatility of $C_1$ is proportional to the MPC (equation (24)), the volatility of $C_1$ is decreasing in wealth. □

Proposition 1 and its corollary capture the precautionary savings dynamics which generate countercyclical consumption volatility. Countercyclical consumption volatility stems from the effect of the date 1 wealth shock on the investor's desire for precautionary savings at date 1. A positive shock weakens this desire, decreasing savings and boosting consumption, while a negative shock encourages him to hunker down, with opposite effects. From the perspective of date 0, this uncertainty in the need for precautionary savings at date 1 causes additional consumption volatility. Since the need for precautionary savings intensifies at low levels of wealth, the effect on consumption volatility is countercyclical.

Mechanically, the precautionary savings mechanism is manifested as a strictly concave consumption function. Strictly concave consumption is an implication of a broad class of consumption-savings models (Carroll and Kimball (1996)). The magnitude of this channel, however, is typically too small to match the data on time-varying risk premiums (Posch (2011)). The following proposition and corollary show that habit amplifies this channel.

**Proposition 2.** *The date 1 habit increases the sensitivity of the MPC to wealth, that is,*

$$\frac{\partial}{\partial H_1} C_1''(W_1) < 0$$

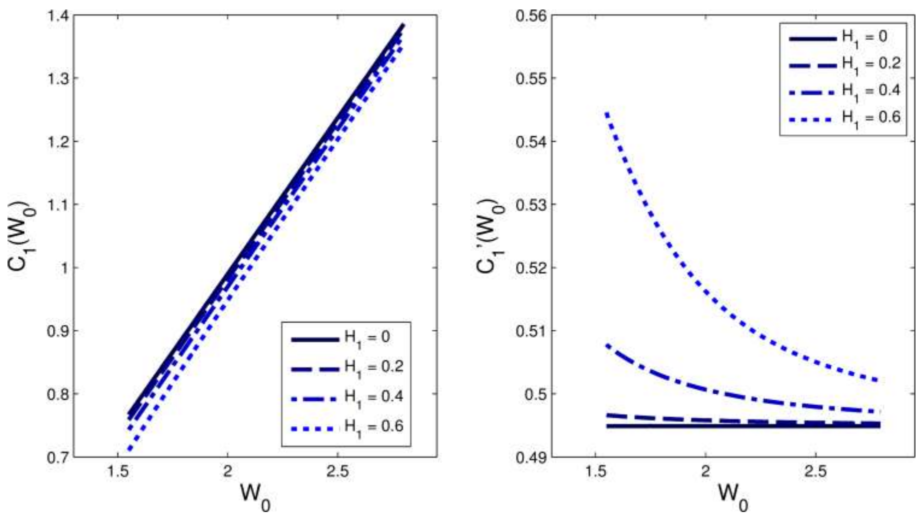
To understand this result, note that an investor with high habit has become accustomed to a high standard of living. He judges consumption not by its absolute level, but by how much it exceeds this standard. As a result, he has a much stronger precautionary savings motive than a investor who is accustomed to living in poverty. The mechanism driving Proposition 1 becomes stronger, and the MPC becomes even more sensitive to wealth. Mathematically, an external and additive habit acts as a reduction in income, which reduces wealth and amplifies precautionary savings dynamics.

**Corollary 2.** *As the date 1 habit increases, the volatility of $C_1$ becomes more countercyclical, that is, for any set of initial wealth $\mathbb{W}_0$,*

$$\frac{\partial}{\partial H_1} \left\{ \max_{W_0 \in \mathbb{W}_0} \sigma_0[C_1(W_0)] - \min_{W_0 \in \mathbb{W}_0} \sigma_0[C_1(W_0)] \right\} > 0$$

One can think of the set of initial wealth $\mathbb{W}_0$ as the volatility of the business cycle in the economy. Corollary 2 states that, over this business cycle, the range of consumption volatilities is increasing in habit. The proof simply applies the link between the MPC and consumption volatility (equation (24)) to Proposition

**Figure 4: Two-Period Model: Date 1 Consumption and MPC**




2. Because habit makes the MPC more countercyclical, habit also makes consumption volatility more countercyclical.

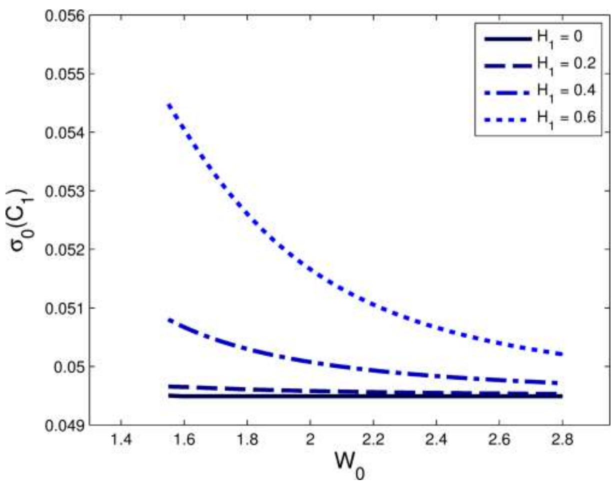
Figures 4 and 5 illustrate the propositions of this section. The figures show results from numerical solutions of the two-period model for different levels of habit. In all solutions, I set $H_1 = H_2$. This helps illustrate how habit "concavifies" the consumption function by eliminating intertemporal substitution effects which complicate the picture.

Figure 4 plots date 1 consumption and MPC as a function of wealth. The left panel shows that, as long as habit is positive, consumption is strictly concave. The right panel shows that this concavity is reflected in an MPC which decreases in wealth (Proposition 1). Figure 4 also shows that habit intensifies this relationship (Proposition 2). As the line gets lighter, habit increases, and the slope of the MPC gets steeper. Intuitively, habit makes the investor feel poorer and intensifies his precautionary savings motives.

Figure 5 shows how these consumption policies are reflected in consumption volatility. As long as habit is positive, consumption volatility decreases in wealth (Corollary 1). This is the essence of the precautionary savings mechanism. Because precautionary savings motives prescribe a countercyclical MPC, and because consumption volatility is proportional to the MPC (equation (24)), consumption volatility is countercyclical. The figure also shows that habit amplifies this countercyclicality (Corollary 2). As the line gets lighter, habit increases, and the range that is spanned by consumption volatility increases.

The results of this section may come as a surprise since many models in the finance literature produce consumption policies which are linear in wealth. These linear consumption policies are the result of the

**Figure 5: Two-Period Model: Date 0 Consumption Volatility**




fact that much of this literature is set in continuous time (Sundaresan (1989), Constantinides (1990)) or relies on log-linear approximations (Campbell (1994), Lettau (2003), Kaltenbrunner and Lochstoer (2010)). Continuous time allows the investor to make an infinite number of trades within any trading period. This instantaneous trading allows for consumption policies which, if applied in a discrete time setting, would imply a positive probability that marginal utility becomes infinite (Brandt (2009)). It is exactly fear of hitting this condition of infinite marginal utility which seems to drive concave consumption behavior (Attanasio (1999)). Log-linear approximations, on the other hand, implicitly assume that consumption policies are (log) linear in state variables. The non-linearities introduced by combining power utility functions with linear budget constraints are important for generating strictly concave consumption (Posch (2011)).

The results may also appear to conflict with the common intuition that habit encourages smooth consumption. This intuition is straightforward in internal habit models, where an increase in consumption has a direct effect of lowering utility in later periods (Sundaresan (1989)). In an external habit model, however, the investor by assumption does not take into account this indirect effect. Prices may encourage smooth consumption via general equilibrium effects, but the habit itself acts very similarly to a reduction in income which increases the marginal propensity to consume (Carroll and Kimball (1996)).

### 5.3 Verifying the Precautionary Savings Mechanism

I have shown that precautionary savings dynamics generate countercyclical consumption volatility, and that this countercyclical consumption volatility lets the model address the Shiller volatility puzzle. But is this mechanism consistent with the modest time-variation in consumption volatility that is seen in the data? Moreover, is consumption volatility related to wealth as predicted? This section shows that the answers to these questions is yes.

I use a number of measures of conditional consumption volatility. All measures are constructed by first fitting an AR(1) model to log consumption growth to remove an expected growth component:

$$\Delta c_{t+1} = b_0 + b_1 \Delta c_t + \epsilon_{c,t} \tag{25}$$

I then either estimate GARCH-type models on the residual $\epsilon_{c,t+1}$ or use the mean absolute residual as a non-parametric measure of conditional volatility. I use quarterly data because it is difficult to detect time-varying volatility with the post-war annual sample of only 50 observations. This procedure follows Bansal, Khatchatrian, and Yaron (2005), Bansal, Kiku, and Yaron (2009) and Beeler and Campbell (2009).

Table 5 compares GARCH estimates from the data and model. Panel A shows results from a GARCH(1,1) process

$$\sigma^2_{c,t+1} = \omega_0 + \omega_1 \epsilon^2_{c,t} + \omega_2 \sigma^2_{c,t} \tag{26}$$

The data columns show modest time-variation in consumption volatility. Consistent with Bansal, Khatchatrian, and Yaron (2005), the ARCH parameter $\omega_1$ just makes it to the 95% significance level if one uses non-robust standard errors. However, using standard errors which are robust to the assumption of non-normal shocks (Bollerslev and Wooldridge (1992)), one cannot reject the hypothesis of no time-varying volatility. This modest time-variation in volatility is consistent with the model columns. The model's mean estimate of the ARCH parameter is close to zero, and 5% of simulations produce a value of 0. Supposing that ARCH effects do exist, the data produce GARCH parameter estimates of $\hat{\omega}_2 = 0.79$ indicating that consumption volatility is persistent. The model captures this persistence, producing a GARCH parameter of 0.76. Indeed, this high persistence of consumption volatility may be the reason why time-varying volatility is hard to detect in the data.

Panel B compares estimates of the GJR-GARCH(1,1,1) model (Glosten, Jagannathan, and Runkle (1993))

$$\sigma^2_{c,t+1} = \omega_0 + \omega_1 \epsilon^2_{c,t} + \omega_2 \sigma^2_{c,t} + \omega_3 I(\epsilon_{c,t} > 0)\epsilon^2_{c,t}$$

GJR-GARCH introduces an additional term which allows negative shocks to have a larger effect on volatility, as predicted by the analysis of Section 5.2. Though the data show significant sampling uncertainty, the point estimates are consistent with the intuition. The asymmetric GJR-GARCH parameter $\hat{\omega}_3$ is much larger than the symmetric ARCH $\hat{\omega}_1$ in both the model and the data. In terms of magnitudes, the model's parameter estimates are all smaller than the data's. Consistent with Panel A, the model reproduces the modest time-

### Table 5: GARCH Estimates of Time-Varying Consumption Volatility

Data and figures are quarterly. This table shows measures GARCH estimates of the residual from an AR(1) model of consumption growth

$$\Delta c_{t+1} = b_0 + b_1 \Delta c_t + \epsilon_{c,t+1}$$

GARCH estimation is done by quasi maximum likelihood. Robust standard errors use the Bollerslev-Wooldridge method. The model columns show means and percentiles across simulations of the same length as the empirical sample. Further details are found in Appendex A.1.

Panel A: GARCH

$$\sigma^2_{c,t+1} = \omega_0 + \omega_1 \epsilon^2_{c,t} + \omega_2 \sigma^2_{c,t}$$

| | Data: 1948Q1-2011Q4 | | | Model | | | |
|---|---|---|---|---|---|---|---|
| | Estimate | SE | Robust SE | mean | %5 | %50 | %95 |
| $\omega_0$ | 1.39E-06 | 8.51E-07 | 1.27E-06 | 9.62E-06 | 2.96E-07 | 2.65E-06 | 4.42E-05 |
| $\omega_1$ | 0.14 | 0.08 | 0.13 | 0.03 | 0.00 | 0.03 | 0.09 |
| $\omega_2$ | 0.79 | 0.09 | 0.15 | 0.76 | 0.00 | 0.92 | 0.99 |

Panel B: GJR-GARCH

$$\sigma^2_{c,t+1} = \omega_0 + \omega_1 \epsilon^2_{c,t} + \omega_2 \sigma^2_{c,t} + \omega_3 I(\epsilon_{c,t} > 0)\epsilon^2_{c,t}$$

| | Data: 1948Q1-2011Q4 | | | Model | | | |
|---|---|---|---|---|---|---|---|
| | Estimate | SE | Robust SE | mean | %5 | %50 | %95 |
| $\omega_0$ | 1.97E-06 | 1.25E-06 | 2.29E-06 | 1.09E-05 | 3.19E-07 | 2.00E-06 | 4.70E-05 |
| $\omega_1$ | 0.07 | 0.07 | 0.08 | 0.01 | 0.00 | 0.00 | 0.06 |
| $\omega_2$ | 0.76 | 0.10 | 0.19 | 0.73 | 0.00 | 0.93 | 0.98 |
| $\omega_3$ | 0.14 | 0.11 | 0.20 | 0.06 | 0.00 | 0.06 | 0.00 |

variation in consumption volatility, as seen in the data.

There are two reasons why the model can generate substantial fluctuations in risk premiums out of modest GARCH effects. The first is that habit preferences amplify fluctuations in consumption volatility. This can be seen in the surplus consumption process (3). Surplus consumption is basically an AR1 process, where the shocks are consumption growth and the conditional volatility $\lambda$ is a free parameter. The calibration chooses a large $\lambda = 1/0.06 - 1 \approx 15$ which provides substantial amplification. Thus small movements in the volatility of consumption growth produce large fluctuations in the volatility of marginal utility. The second is that consumption volatility generated by the model is very persistent. This persistent consumption volatility makes it hard to measure time-varying volatility using only 50 years of quarterly data.

Though the GARCH estimates show only modest time-variation in consumption volatility, they do not take advantage of instruments which help pick up time-variation in consumption volatility. A number of papers have shown that conditioning on the price-dividend ratio helps pick up time-variation in consumption volatility (Kandel and Stambaugh (1990), Bansal, Khatchatrian, and Yaron (2005), Bansal, Kiku, and Yaron (2009)).[11] Examining the relationship between asset prices and consumption volatility is also important because it provides a direct test of the model's mechanism. A key prediction of the precautionary savings channel (Section 5.2) is that consumption should be more volatile when wealth (or the price-dividend ratio) is low.

Table 6 shows that the data supports this mechanism. It shows regressions of various proxies for consumption volatility on the price-dividend ratio. The data column shows that, using all proxies, consumption volatility and the price-dividend ratio are negatively related. A high price-dividend ratio indicates a safe time of low volatility. The model replicates this pattern. The large standard errors in the Table 5 suggest that comparing magnitudes should be done with caution, but with that in mind, the model coefficients are of similar magnitude to those from the data.

### 5.4 The Large and Volatile Equity Premium

The high volatility of the equity premium comes from high capital adjustment costs and the low EIS of habit preferences. High capital adjustment costs mean that productivity shocks are absorbed by asset prices rather than investment. This channel is not new (Jermann (1998), Kogan (2004), Jermann (2010), Kogan and Papanikolaou (2012)) and so I provide only a brief discussion. The role of the low EIS is that a low EIS pins down high adjustment costs via calibration and general equilibrium effects. The importance of the EIS is less

[11] Other approaches which pick up stronger evidence of time-varying volatility include using Hamilton (1989)'s regime switching model (Lettau, Ludvigson, and Wachter (2008)) or disaggregated consumption data (Boguth and Kuehn (2013)).

### Table 6: Regressions of Consumption Volatility on the Price-Dividend Ratio

This table shows regressions of the form

$$\text{cvol}_t = \alpha + \beta(p_t - d_t) + \epsilon_t$$

where $\text{cvol}_t$ is a proxy for conditional consumption volatility, $p_t$ is the log equity price, and $d_t$ is the log dividend. To generate $\text{cvol}_t$, first, an AR(1) model (25) is run on log consumption growth. Panel A estimates either a GARCH(1,1) or GJR-GARCH(1,1,1) model on the residuals, and then takes the log of the estimated consumption volatility. Panel B uses a non-parametric measure: $\text{cvol}_t(L) \equiv \log\left(\sum_{j=1}^{L} |\epsilon^c_{t+j}|\right)$, where $\epsilon^c_{t+j}$ is the residual from the AR(1) model. Consumption data is quarterly and price-dividend ratio data is annual, which results in some abuse of notation. The model columns show means and percentiles across simulations of the same length as the empirical sample. Further details are found in Appendix A.1.

Panel A: GARCH estimates of consumption volatility

| | | Data | Model | | | |
|---|---|---|---|---|---|---|
| | | 1948Q1-2011Q4 | mean | 5% | 50% | 95% |
| GARCH | $\hat{\beta}$ | -0.43 | -0.12 | -0.33 | -0.10 | 0.00 |
| | $\text{SE}(\hat{\beta})$ | 0.11 | 0.03 | 0.00 | 0.03 | 0.08 |
| | $R^2$ | 0.40 | 0.22 | -0.01 | 0.17 | 0.61 |
| GJR-GARCH | $\hat{\beta}$ | -0.41 | -0.21 | -0.47 | -0.22 | 0.00 |
| | $\text{SE}(\hat{\beta})$ | 0.13 | 0.04 | 0.00 | 0.04 | 0.08 |
| | $R^2$ | 0.33 | 0.34 | -0.01 | 0.34 | 0.79 |

Panel B: Non-parametric estimates of consumption volatility
$\text{cvol}_t(L) \equiv \log\left(\sum_{j=1}^{L} |\epsilon^c_{t+j}|\right)$

| | | Data | Model | | | |
|---|---|---|---|---|---|---|
| | L | 1948Q1-2011Q4 | mean | 5% | 50% | 95% |
| $\hat{\beta}$ | 4 | -0.66 | -0.21 | -0.01 | -0.00 | -0.66 |
| | 12 | -0.59 | -0.26 | 0.03 | -0.25 | -0.68 |
| | 20 | -0.53 | -0.25 | 0.02 | -0.24 | -0.62 |
| $\text{SE}(\hat{\beta})$ | 4 | 0.16 | 0.25 | 0.12 | 0.23 | 0.45 |
| | 12 | 0.12 | 0.19 | 0.08 | 0.17 | 0.36 |
| | 20 | 0.09 | 0.14 | 0.06 | 0.13 | 0.29 |
| $R^2$ | 4 | 0.22 | 0.02 | 0.00 | 0.01 | 0.06 |
| | 12 | 0.36 | 0.04 | 0.00 | 0.02 | 0.13 |
| | 20 | 0.44 | 0.07 | 0.00 | 0.04 | 0.23 |

well understood and so I focus on it.

The link between equity volatility and adjustment costs can be seen in the investment return - stock return identity (Cochrane (1991), Restoy and Rockinger (1994)). Since the model has a homogenous production technology, this identity means we can express the stock return as

$$R_{t,t+1} = \frac{\alpha(Y_{t+1}/K_{t+1}) + (1+\phi(I_{t+1}/K_{t+1}))(1-\delta) + \frac{\phi}{2}(I_{t+1}/K_{t+1})^2}{1+\phi(I_t/K_t)} \tag{27}$$

The stock return is related to the marginal product of capital $\alpha(Y_{t+1}/K_{t+1})$, the investment rate $I_{t+1}/K_{t+1}$, and the adjustment cost parameter $\phi$. This equality holds state-by-state, which means that the volatilities of the two sides are equal.

This identity shows the difficulty of matching equity volatility in a production economy. The marginal product of capital and the investment rate have a very low volatility, less than 2% per year. On the other hand, the equity return has a huge volatility of about 20% per year. The only way to reconcile these two sides, then, is with capital adjustment costs. Adjustments costs increase the curvature of the RHS. Jensen's inequality then implies that adjustment costs increase the volatility of the investment return. The large discrepancy between the equity volatility and the volatility of the investment rate imply a very large adjustment cost parameter.

The role of the EIS comes through general equilibrium effects and standard empirical restrictions. General equilibrium means that both the EIS and capital adjustment costs affect consumption volatility. The EIS channel works through the SDF. Since the firm uses the investor's IMRS as an SDF, this means that the firm is rewarded for providing consumption which matches the investor's preferences. A low EIS results in a strong incentive to produce smooth dividends, and through market clearing, smooth consumption. Capital adjustment costs also affect consumption volatility through market clearing. Large adjustment costs encourage the firm to keep investment smooth in the face of shocks. Since shocks must be absorbed by either investment or dividends, this encourages volatile dividends, and, through market clearing, volatile consumption.

Since both the EIS and capital adjustment cost affect consumption volatility, the ubiquitous requirement that asset pricing models should fit consumption data means that these two elements must match each other. A low EIS requires high capital adjustment costs, and vice versa. In this model, the large Sharpe ratio pins down a low EIS. Given the low EIS, the volatility of consumption pins down a high adjustment cost. Note that the adjustment cost is not pinned down by equity volatility, and so equity volatility forms an overidentified restriction which is satisfied by the model.

This EIS-adjustment cost link exists in long-run risk and disaster models, but it tends to hurt rather than help their asset pricing results. The vast majority of long-run risk and disaster papers do not calibrate the EIS. However, the EIS must be larger than one in both models in order to match certain qualitative facts, and the common wisdom is that the larger the EIS, the better the asset pricing results.[12] This large EIS then implies low adjustment costs and counterfactually low equity volatility (Kaltenbrunner and Lochstoer (2010), Gourio (2010)).

Given that equity volatility is high, the large equity premium is simply the consequence of a large Sharpe ratio. The large Sharpe ratio is simply due to the fact that habit preferences offer an additional degree of freedom for the modeler. Habit results in an additional term in the SDF (5). The modeler can choose the volatility of this term to be high by adjusting the preference parameter $\lambda$ in equation (3).

Economically, $\lambda$ can be interpreted as the 'moodiness' of the economy. The surplus consumption process says that if consumption growth goes up by 1%, surplus consumption gets boosted by $\lambda$%. The calibration chooses $\lambda = 1/0.065-1 \sim 15$, meaning that changes in 'mood' are responsible for the vast majority of changes in marginal utility. Checking this magnitude by introspection is, of course, a dangerous activity. But Section 4 shows that it is consistent with numerous overidentifying restrictions regarding asset prices.[13]

### 5.5 The Low and Smooth Risk-Free Rate

The low risk-free rate is simply the result of time preference. Time preference is effectively a free parameter which allows me to hit the low risk-free rate in the data. Economically speaking, this results in an intuitive time preference of $\beta^4 \approx 0.90$ annually, that is, consumption one year from today is worth 90% of consumption today.

The smooth risk-free rate comes from an interplay of intertemporal substitution and precautionary savings effects. Though this channel sounds complicated, there is a very simple intuition for it. In bad times, people want to borrow in order to consume today. But in bad times, the economy is particularly volatile, and the desire for precautionary savings prevents them from borrowing. This channel is similar to that of Campbell and Cochrane (1999), but with endogenous consumption volatility driving precautionary savings effects rather than the exogenous habit sensitivity function. Indeed, Section 5.1 shows that the model does a good job of mimicking the Campbell-Cochrane sensitivity function and so the smoothness of the risk-free rate should not be surprising.

[12] For long-run risk models, this is required to qualitatively match the fact that consumption volatility and the price-dividend ratio are negatively related (Bansal and Yaron (2004)). In disaster models, this is required to match the intuitive notion that a rise in disaster risk lowers investment and increases excess returns (Gourio (2010))

[13] Bekaert, Engstrom, and Grenadier (2010) also describe external habit investors as a 'moody,' but the details of their interpretation are much different.

These intuitions can be fleshed out by examining the log-normal approximation of the risk-free rate[14]

$$\begin{aligned} r^f_{t+1} \approx & -\log\beta + \gamma(\lambda+1)\mathbb{E}_t(\Delta c_{t+1}) - \gamma(1-\rho_s)(s_t - \bar{s}) \\ & - (1/2)\gamma^2(\lambda+1)^2 \mathrm{Var}_t(\Delta c_{t+1}) \end{aligned} \tag{28}$$

The 1st term reflects time-preference. In a purely mathematical sense, it is a free parameter which one can use to fit the low mean risk-free rate in the data.

The 2nd and 3rd terms are due to the elasticity of intertemporal substitution (EIS) and tend to create excessive volatility in habit models (Jermann (1998), Boldrin, Christiano, and Fisher (2001)). They reflect the Friedman (1957) permanent income hypothesis. In bad times, investors want to borrow from the future in order to consume today. This motive pushes down the price on the risk-free bond and pushes up the risk-free rate, leading to a countercyclical effect on a risk-free rate. Habit models imply a very low EIS which makes this channel very volatile.

This model contains the volatile intertemporal substitution effect typical of habit models, but it counters this effect with a precautionary savings effect. The precautionary savings effect runs through the 4th term, which is decreasing in the conditional volatility of consumption. Intuitively, in bad times, high consumption volatility creates a desire for savings. Investors buy bonds, pushing up the price and down the risk-free rate. This channel creates a procyclical effect on the risk-free rate, which helps counteract the countercyclical effect of the intertemporal substitution channel.

The previous discussion shows that, qualitatively, precautionary savings effects help counteract intertemporal substitution effects. Whether the quantitative effect is enough generate a smooth risk-free rate is another question. Figure 6 examines the quantitative efffect. It plots the risk-free rate decomposition (28) against surplus consumption for various levels of productivity. The solid red 'EIS' lines represent the elasticity of intertemporal substitution effects. The dashed blue 'prudence' lines represent the precautionary savings effects. The two effects are near mirror-images of each other, showing that, quantitatively, the channels balance each other quite nicely.

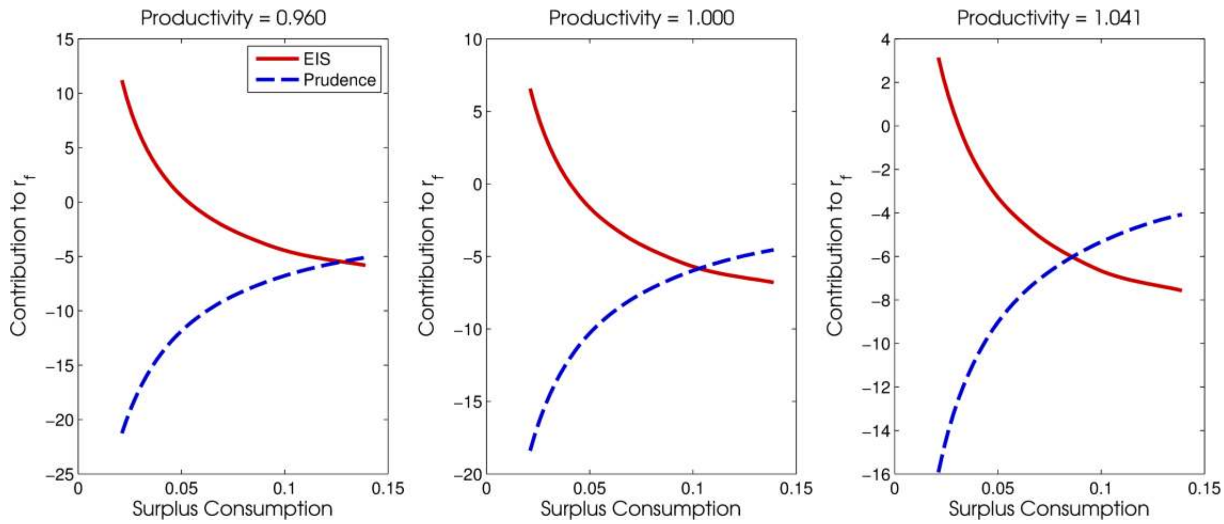
A nice feature of the smoothing effects in this model is that they arise endogenously. This contrasts with the risk-free rate smoothing effects of Campbell and Cochrane (1999), which are the result of a parameter choice. To see this, it helps to return to approximation of the risk-free rate (28). This expression shows

[14] If $m_{t+1}$ is normal,

$$r^f_{t+1} = -\log\left[\mathbb{E}_t(e^{m_{t+1}})\right] = -\mathbb{E}_t(m_{t+1}) - \frac{1}{2}\mathrm{Var}_t(m_{t+1})$$

Then just plug in the log SDF $m_{t+1} = \log\beta - \gamma\Delta s_{t+1} - \gamma\Delta c_{t+1}$ and habit process $\Delta s_{t+1} = -(1-\rho_s)(s_t - \bar{s}) + \lambda\Delta c_{t+1}$.

**Figure 6: Decomposition of the Risk-Free Rate.** 'EIS' represents the 1st three terms of equation (28). 'Prudence' represents the last term. Computed from model's laws of motion. Risk-free rate is in annualized %. Capital in all panels is fixed at the mean capital stock.




that, if one allows the preference parameter $\lambda$ to vary over time, then one can control the magnitude of the precautionary savings effect. Indeed, Campbell and Cochrane (1999) do just that, and they choose the magnitude of the channel to exactly cancel out the intertemporal substitution effect. This model has no such freedom. The magnitude of this channel comes through the amount of countercyclicality in consumption volatility, which is the result of general equilibrium effects of investors' preferences for precautionary savings on firms' production decisions.

## 5.6 Comparative Statics

The quantitative results come from estimated parameter values, but as with all econometric methods, the point estimates can be sensitive to choices of the econometrician. This section investigates the effect of changing the parameter values. The comparative statics also help confirm the intuition developed earlier in this section.

Table 7 shows key moments from these comparative statics exercises. Each column examines moments generated by models where only one of the parameter values is changed from the estimation (Table 1). Three different parameter changes are examined: lower persistence of habit, weaker steady state habit, and lower capital adjustment costs. The magnitude of the perturbations are chosen to be the smallest change that produces a clearly recognizable deviation from the estimated results. This approach helps isolate the direct effect of changing a parameter value from its interaction with other elements of the model.

## Table 7: Comparative Statics

Figures are annual. 'Estimated' represents parameter values from Table 1. All other columns use the estimated values but with one parameter changed. 'Lower persistence of habit' sets $\rho_s = 0.80$. 'Weaker steady state habit' sets $\bar{S} = 0.12$. 'Lower adjustment costs' sets $\phi = 40$. Details of the data are found in Appendix A.1.

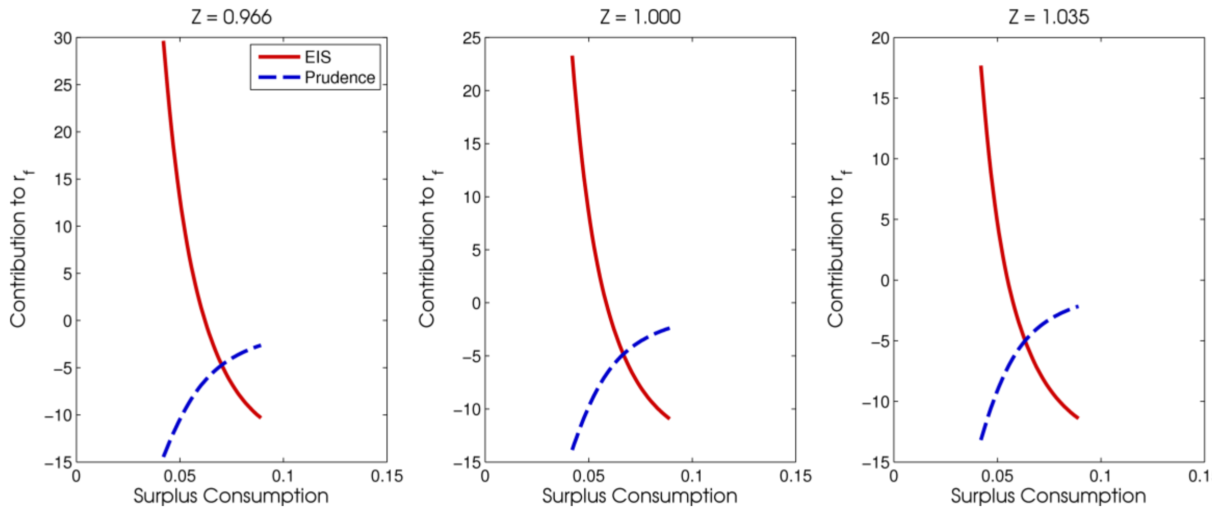
| | | Data 1948-2011 | Estimated | Lower persistence of habit | Weaker steady state habit | Lower adjustment costs |
|---|---|---|---|---|---|---|
| Identifying Moments | | | | | | |
| $\mathbb{E}(r_f)$ (%) | | 0.98 | 0.98 | 3.24 | 8.33 | 4.64 |
| $\mathbb{E}(R-R_f)/\sigma(R)$ | | 0.48 | 0.48 | 0.39 | 0.31 | 0.38 |
| $\sigma(r-r_f)$ (%) | | 16.07 | 16.07 | 24.33 | 14.78 | 8.94 |
| $\sigma(c_{hp})/\sigma(y_{hp})$ | | 0.47 | 0.47 | 0.43 | 0.64 | 0.39 |
| $\sigma(y_{hp})$ (%) | | 1.50 | 1.50 | 1.63 | 1.54 | 1.49 |
| | | | | | | |
| Untargeted Moments | | | | | | |
| $\mathbb{E}(r-r_f)$ (%) | | 6.47 | 6.77 | 7.59 | 3.85 | 3.14 |
| $\mathrm{AC1}(r-r_f)$ | | -0.03 | -0.07 | -0.12 | -0.05 | -0.04 |
| | | | | | | |
| $\sigma(r_f)$ (%) | | 2.24 | 2.96 | 9.23 | 5.22 | 2.30 |
| $\mathrm{AC1}(r_f)$ | | 0.56 | 0.89 | 0.66 | 0.86 | 0.93 |
| | | | | | | |
| $\mathbb{E}(p-d)$ | | 3.42 | 2.67 | 2.22 | 2.14 | 2.63 |
| $\sigma(p-d)$ | | 0.43 | 0.46 | 0.39 | 0.33 | 0.39 |
| $\mathrm{AC1}(p-d)$ | | 0.95 | 0.90 | 0.78 | 0.89 | 0.92 |
| | | | | | | |
| $R^2$ from | 1-year | 0.08 | 0.10 | 0.07 | 0.04 | 0.06 |
| forecasting $r-r_f$ | 2-year | 0.19 | 0.26 | 0.14 | 0.10 | 0.17 |
| with $p-d$ | 3-year | 0.26 | 0.37 | 0.17 | 0.15 | 0.25 |

#### 5.6.1 Lower persistence of habit

The third column of Table 7 examines a model where the persistence of habit $\rho_s$ is lowered from the estimated value of 0.963 to 0.800. This parameter is important because the strong persistence of habit is one way in which this model deviates from previous models with habit and production. The high estimated persistence means that habit today depends on a very long history of consumption. In contrast, habit in Jermann (1998) and Boldrin, Christiano, and Fisher (2001) depends only on the last quarter's consumption.

The table shows the high persistence is critical. In the lower persistence model, the volatility of the risk free rate triples from the estimated value of 2.96% to 9.23%, bringing the model in line with the high risk-free rate volatility of Jermann (1998) and Boldrin, Christiano, and Fisher (2001). Figure 7 explains

**Figure 7: Decomposition of the Risk-free rate: low persistence of habit.** $\rho_s = 0.80$. All other parameter values are from the estimation (Table 1). 'EIS' represents the 1st three terms of equation (28). 'Prudence' represents the last term. Computed from model's laws of motion. Risk-free rate is in annualized %. Capital in all panels is fixed at the mean capital stock.




why. The figure plots the risk-free rate decomposition of equation (28) for the low persistence of habit model. The figure shows that the EIS effect on the risk-free rate is still countercyclical and the prudence effect is still procyclical. However, the two channels no longer cancel each other out quantitatively. The EIS effect becomes more countercyclical and the prudence effect becomes less procyclical as compared to the estimated model (Figure 6). Both of these effects can be traced to the lower persistence of habit. A low persistence of habit means that habit will strongly mean revert tomorrow. Regarding intertemporal substitution, this means that, in bad times, there is a pronounced desire to borrow in order to consume today. This boosts the countercyclicality of the EIS effect. Regarding prudence, this means that the precautionary savings channel is weakened. Habit will strongly mean revert tomorrow, meaning tomorrow the household has less need for precautionary savings.

This weakening of precautionary savings effects is reflected in a reduction in the amount of time-varying risk premiums. The third column of Table 7 shows that the amount of time-variation in risk premiums falls. $R^2$'s from regressions of future excess returns on the log price-dividend ratio drop significantly.

The persistence of habit has a strong effect on many other asset price moments. The parameter was identified with the volatility of excess returns, and, as expected, the low persistence model strongly overpredicts this volatility. Another effect is that the persistence of the risk-free rate and the price-dividend ratio drops. Intuitively, the persistence of preferences is reflected in the persistence of asset prices.

### 5.6.2 Weaker steady state habit

The fourth column of Table 7 raises steady state surplus consumption $\bar{S}$ from the estimated value of 0.062 to 0.120. This comparative static brings the model closer to CRRA utility. Roughly speaking, $\bar{S} = 1.00$ is an economy with no habit, so this model is still very far from the standard model.

Weakening steady state habit lowers the volatility of marginal utility of the household (see equation (19)), and thus has the direct effect of lowering the Sharpe ratio from the estimated value of 0.48 to 0.31. This reduced Sharpe ratio has the obvious effect of reducing the equity premium from the estimated value of 6.77% to 3.85%. The Sharpe ratio does not drop as much as one might expect, however, since the volatility of consumption increases. The relative volatility of consumption to output increases from the estimated value of 0.47 to 0.64. This increase in consumption volatility is due to a weaker general equilibrium consumption smoothing effect (see Section 5.4).

Weakening steady state habit has the additional effect of weakening precautionary savings effects. With weaker habit, there is less need for precautionary savings, and the risk-free rate jumps from its estimated value of 0.98% to 8.33%. This weakening of precautionary savings effects can be seen in the return forecasting regressions too. The $R^2$'s from regressions of future returns on the price-dividend ratio all drop by more than half.

### 5.6.3 Lower adjustment costs

The final column of Table 7 examines the effect of lowering capital adjustment costs. The adjustment cost parameter $\phi$ is lowered from the estimated value of 74.94 to 40. This comparative static helps compare my model with Lettau and Uhlig (2000), who also examine external habit in a production economy. My model deviates from theirs in two important ways. The first is that I include capital adjustment costs, while Lettau and Uhlig assume costless adjustment. The second is that I use a non-linear, global solution method, while Lettau and Uhlig linearize the model around the non-stochastic steady state. This comparative static shows that capital adjustment costs make a large impact on the model results and that lowering the costs brings my model closer to Lettau and Uhlig (2000).

The table shows that lowering adjustment costs reduces the relative volatility of consumption to output falls from the estimated value of 0.47 to 0.39. This is consistent with the intuition from Section 5.4. Low adjustment costs reduce the incentive for smooth investment and encourages volatile consumption.

This decrease in consumption risk is reflected in a lower Sharpe ratio, which falls from 0.48 to 0.38 and a lower equity premium which falls from 6.77% to 3.14%. With lower consumption risk, the precautionary motive is weakened and the risk-free rate rises from the estimated value of 0.98% to 4.64%. This decrease in the precautionary motive is also seen in lower $R^2$'s from return forecasting regressions.

## 6 Conclusion

This paper provides a unified framework for understanding asset prices and the real economy. The standard real business cycle model, augmented to include external habit preferences and capital adjustment costs provides a quantitative description of both asset prices and aggregate fluctuations.

The asset pricing performance of external habit preferences not only survives the introduction of production, but production endogenizes a key mechanism of consumption-based asset pricing models. External habit, long-run risk, and disaster models all assume that the volatility of marginal utility is countercyclical. This countercyclical volatility arises endogenously from the model as a result of production and external habit.

# A Appendix

## A.1 Data Details

The data span 1948 thru 2011. Macroeconomic quantities are from the BEA NIPA and BEA fixed asset tables. All data is real and per-capita. Output is simply GDP. Consumption is nondurable consumption plus services consumption. Investment and capital are fixed investment plus durable consumption plus government investment. All quantities are constructed by dividing nominal figures (Table 1.1.5 or Table 3.9.5) by the appropriate price index (Table 1.1.4 or Table 3.9.4), and then dividing by the population (from Table 2.1).

Most studies exclude government investment from their definition of investment, citing the fact that their models do not include government. However, this logic would also imply that one should remove government purchases from GDP, which is not typically done. I choose to keep government purchases in GDP and include government investment in investment as this approach preserves the idea that the data aggregates come from a single general equilibrium system. Moreover, this approach is closer to the spirit of Cooley and Prescott (1995).[15]

Asset price data is taken from CRSP. The equity is the CRSP value-weighted index. Price-dividend ratios are computed annually using the stock market reinvestment assumption. As discussed in Cochrane (2011), this aggregation method preserves the Campbell-Shiller identity, which is useful for identifying the source of asset price fluctuations. CRSP returns are deflated using the CPI, also obtained from CRSP. The risk-free rate is computed using a forecast of the ex-post real return of the 90-day T-bill following Beeler and Campbell (2009). Following Beeler and Campbell (2009), the ex-post real return is calculated by deflating the 90-day nominal T-bill yield using seasonally adjusted CPI from the BLS. The forecast is constructed by regressing next quarter's ex-post real return on today's nominal 90-day yield and the mean inflation rate over the previous year.

## A.2 Simulated Method of Moments Details

This section spells out details of the SMM method. I use a different notation than the brief discussion in the text in order to be more precise.

[15]Cooley and Prescott (1995) say that "Our economy is very abstract: it contains no government sector, no household production sector, no foreign sector and no explicit treatment of inventories. Accordingly, the model economy's capital stock, $K$, includes capital used in all of these sectors plus the stock of inventories. Similar, output, $Y$, includes the output produced by all of this capital" (page 17).

### A.2.1 Econometric Details

I transform the parameters so that I do not need to be concerned about corner solutions. For example, rather than estimate $\beta \in [0,1]$, I estimate $\text{logit}(\beta) \in \mathbb{R}$. After the estimation is done, I find the point estimates of the original parameters by inverting the transformation, and the standard errors by the delta method.

To be explicit, let $\zeta \in \Theta$ be the vector of original parameters and $\psi^{-1}(\zeta) = \theta \in \mathbb{R}_K$ by the transformed parameters. Then

$$\sqrt{T}(\hat{\zeta}_T - \zeta_0) = \sqrt{T}[\psi(\hat{\theta}_T) - \psi(\theta_0)] \tag{29}$$

$$\approx \sqrt{T}[D\psi(\theta_0)'(\hat{\theta}_T - \theta_0)] \tag{30}$$

so the asymptotic variance of the original parameters can be estimated by

$$\widehat{\text{Var}}[\sqrt{T}(\hat{\zeta}_T - \zeta_0)] = D\psi(\hat{\theta}_T)'\widehat{\text{Var}}[\sqrt{T}(\hat{\theta}_T - \theta_0)]D\psi(\hat{\theta}_T) \tag{31}$$

I also use the delta method on the moment errors, because some of the "moments" I use are not moments in the traditional sense (i.e. the Sharpe Ratio) but are more formally described as transformations of moments. That is, let the GMM objective be

$$\hat{\theta}_T = \arg\min_{\theta \in \mathbb{R}} Q_T(\theta) = G_T(\theta)' W_T G_T(\theta)$$

where $G_T(\theta)$ are some moment errors.

$$G_T(\theta) \equiv S^{-1} \sum_{s=1}^{S} h[\bar{f}_T^s(\theta)] - h[\bar{f}_T^*]$$

where

$$\bar{f}_T^s(\theta) \equiv T^{-1} \sum_t f_t^s(\theta) \quad [M \times 1] \quad \text{is a set of moments from simulation s}$$

$$h : \mathbb{R}^M \longrightarrow \mathbb{R}^N \quad \text{transforms the moments into some other statistic}$$

and $S$ is the total number of simulations. This $h$ function has no effect on the GMM formula for the asymptotic variance

$$\widehat{\text{Var}}[\sqrt{T}(\hat{\theta}_T - \theta_0)] \equiv [DG_T W_T DG_T']^{-1}[DG_T W_T \hat{X}_T W_T DG_T'][DG_T W_T DG_T']^{-1}$$

where

$$
\begin{aligned}
DG_T &\equiv DG_T(\hat{\theta}_T) \\
\hat{X}_T &\equiv \widehat{\mathrm{Var}}(\sqrt{T}G_T(\theta_0))
\end{aligned}
$$

but it does have an effect on the estimator of the variance of the moment errors $\widehat{\mathrm{Var}}(\sqrt{T}G_T(\theta_0))$.

$$
\begin{aligned}
\hat{X}_T &\equiv Dh(\bar{f}_T^*)'\left(1+\frac{1}{S}\Sigma_T\right)Dh(\bar{f}_T^*) \\
\Sigma_T &\equiv \sum_{j=-k}^{k}\left(\frac{k-|j|}{k}\right)T^{-1}\sum_{t=1}^{T}\hat{e}_t\hat{e}'_{t-j} \\
\hat{e}_t &\equiv f_t^* - T^{-1}\sum_{s=1}^{T}(f_s^*)
\end{aligned}
$$

That is, we simply adjust the Duffie and Singleton (1993) estimator by applying the delta method to the spectral density. To see how this works, note that the moment error can be approximated by

$$
\begin{aligned}
G_T(\theta_0) &= S^{-1}\sum_s h(\bar{f}_T^s(\theta_0)) - \bar{f}_T^* \\
&\approx S^{-1}\sum_s \left[Dh(\bar{f}_T^*(\theta_0))[\bar{f}_T^s(\theta_0) - \bar{f}_T^*]\right] \\
&= Dh(\bar{f}_T^*)\left[S^{-1}\sum_s \bar{f}_T^s(\theta_0) - \bar{f}_T^*\right]
\end{aligned}
$$

Thus we just need to adjust the traditional Newey-West estimator by $Dh(\bar{f}_T^*)$.

### A.2.2 Numerical Details

To calculate the moment errors, I simulate the model 1000 times ($S = 1000$). The initial state is set to be the median state for a long simulation close to the estimated parameter values. Derivatives are computed with a two-sided finite difference.

I optimize using Levenberg-Marquardt (LM). I choose this method rather than the more commonly used simulated annealing method for a few reasons. The first is that the moment function does not display an extreme number of local minima, which is where simulated annealing has an advantage. With a relatively smooth objective function, a method which uses derivative information is much more efficient. Derivative information is particularly helpful in the GMM setting with small residuals, since the Gauss-Newton method provides a quick positive semi definite approximation of the Hessian. Speed is important as the estimation takes around 24 hours with a very good guess. Another advantage of LM is robustness. Simulated annealing tends to be very sensitive to the choice of the annealing schedule (Press, Teukolsky, Vetterling, and Flannery (1992).

I begin with an LM parameter of 1000, and use a simple algorithm for adjusting the parameter: if the new function value is a good enough improvement, I decrease the LM parameter by a factor of 10. Otherwise, I set it back to 1000. "Good enough" is judged by the difference between the improvement in the objective and the predicted improvement according to a quadratic model.

Since I use exact identification, the maximum moment error provides a clean convergence criterion. I consider the algorithm converged if the maximum moment error is less than 1.0E-4.

## A.3 The Interpretation of Habit

Here I describe how deviating from Campbell and Cochrane (1999)'s specification of $\lambda$ raises some questions regarding the interpretation of habit in the model. The key issue is that the constant $\lambda$ of my model can make habit decrease in response to an increase in consumption. This violates some traditional notions of habit.

This issue can be illustrated by taking the derivative of log habit $h_{t+1}$ with respect to log consumption $c_{t+1}$:

$$\frac{dh_{t+1}}{dc_{t+1}} = 1 - \frac{\lambda}{S_{t+1}^{-1} - 1}$$

Thus if $S_{t+1}$ is large enough, $\frac{dh_{t+1}}{dc_{t+1}}$ will be negative.

The preferences of this paper still preserve the standard notion of habit in that $H_t$ is a geometric average of previous consumption. This can be seen by following the analysis of Campbell and Cochrane (1999) found in Campbell (2003). I can log-linearize the log surplus consumption ratio around the steady state:

$$s_t = \log[1 - \exp(h_t - c_t)] \approx \kappa - \lambda^{-1}(h_t - c_t) \tag{32}$$

Plugging this into the definition of the habit process (3), I find the link between habit and historical consumption

$$h_{t+1} \approx (\text{Constants}) + \rho_s h_t + (1 - \rho_s)c_t \tag{33}$$

$$= (\text{Constants}) + (1 - \rho_s) \sum_{j=0}^{\infty} \rho_s^j c_{t-j} \tag{34}$$

This informal demonstration is verified by simulated data. In the simulated data, habit is highly correlated with consumption. The contemporaneous correlation is 0.978 and the correlation with lagged consumption is 0.983. Habit growth and consumption growth are moderately correlated. The correlation between $\Delta h_t$ and $\Delta c_t$ is 0.411.

That habit should move non-negatively with consumption everywhere is not required if one entertains a very slow-moving, historical average of consumption as responsible for our current reference point for consumption. Moreover, Campbell and Cochrane (1999)'s specification is also vulnerable to this this issue. Ljungqvist and Uhlig (2009) show that while habit moves positively with small movements in consumption, it can move negatively with large movements.

The issue illustrated in this section is related to Campbell and Cochrane (1999)'s three requirements on $\lambda(s_t)$. They require (i) the risk-free rate is constant, (ii) habit is predetermined at the steady state surplus consumption, and (iii) habit is predetermined near the steady state. The first assumption is not critical for making habit move non-negatively with consumption. In my model, (ii) is satisfied, but (iii) is not. (iii), in combination with Campbell and Cochrane (1999)'s specification for $\lambda(s_t)$ results in $\frac{dh_{t+1}}{dc_{t+1}} \geq 0$ for all $s_t$.

### A.4 Solution Method Details

**Euler Equation** To be explicit about the firm's Euler equation, let $\pi_Z(Z_i, Z_j)$ be the transition matrix for the discretized productivity process. The firm's problem is to find capital policy $K' = G(K; \hat{K}, S, Z_i)$ to solve

$$\begin{aligned} V(K;\hat{K},S,Z_i) = \max_{K',I,N} \Big\{ & \Pi(K,Z_i,N) - W(\hat{K},S,Z_i)N - \Phi(I,K) - I \\ & + \sum_{Z_j} \pi_Z(Z_i,Z_j) M(\hat{K},S,Z_i;Z_j) V(K';\hat{K}',S',Z_j) \Big\} \end{aligned}$$

*subject to*

$$K' = I + (1-\delta)K$$

The FOC for investment and the envelope condition are:

$$\begin{aligned} 1 + D_1\Phi(I,K) &= \sum_{Z_j} \pi_Z(Z_i,Z_j) M(\hat{K},S,Z_i;Z_j) D_1 V(K';\hat{K}',S',Z_j) \\ D_1 V(K;\hat{K},S,Z_i) &= D_1\Pi(K,N,Z_i) + (1 + D_1\Phi(I,K))(1-\delta) - D_2\Phi(I,K) \end{aligned}$$

which together produce the Euler equation

$$1 + D_1\Phi(I,K) = \sum_{Z_j} \pi_Z(Z_i, Z_j) M(\hat{K}, S, Z_i; Z_j)[D_1\Pi(K', Z_j, N') + (1 + D_1\Phi(I_j, K'))(1-\delta) - D_2\Phi(I_j, K')]$$

Impose the fact that the household does not value leisure and consistency, and we have

$$\begin{aligned} 1 &= \sum_{Z_j} \pi_Z(Z_i, Z_j) M(\hat{K}, S, Z_i; Z_j) \hat{R}_I(\hat{K}, S, Z_i; Z_j) \\ \hat{R}_I(\hat{K}, S, Z_i; Z_j) &\equiv \frac{D_1\Pi(\hat{K}', Z_j, 1) + (1 + D_1\Phi(\hat{I}_j, \hat{K}'))(1-\delta) - D_2\Phi(\hat{I}_j, \hat{K}')}{1 + D_1\Phi(\hat{I}, \hat{K})} \end{aligned} \tag{35}$$

Where

$$\begin{aligned} M(\hat{K}, S, Z_i, Zj) &= \beta\left(\frac{C_j}{C}\frac{S_j}{S}\right)^{-\gamma} \\ C_j &= \Pi(\hat{K}', Z_j, 1) - \Phi(\hat{I}_j, \hat{K}') - \hat{I}_j \\ C &= \Pi(\hat{K}, Z_i, 1) - \Phi(\hat{I}, \hat{K}) - \hat{I} \\ \hat{K}' &= \hat{G}(\hat{K}, S, Z_i) \\ \hat{I} &= \hat{G}(\hat{K}, S, Z_i) - (1-\delta)\hat{K} \\ \hat{I}_j &= \hat{G}(\hat{K}', S', Z_j) - (1-\delta)\hat{K}' \end{aligned} \tag{36}$$

and the evolution of surplus consumption satisfies

$$s_j = (1-\rho_s)\bar{s} + \rho_s s + \lambda(c_j - c) \tag{37}$$

The projection algorithm looks for cubic spline coefficients which solve equations (35), (36), and (37).

**Solving for asset prices** To find asset prices, the firm's Bellman equation, with optimal values plugged in, is:

$$\begin{aligned} V(K; \hat{K}, S, Z_i) =& \Pi(K, Z_i, 1) - W(\hat{K}, S, Z_i) - \Phi(I, K) - I \\ &+ \sum_{Z_j} \pi_Z(Z_i, Z_j) M(\hat{K}, S, Z_i; Z_j) V(K'; \hat{K}', S', Z_j) \end{aligned}$$

In equilibrium, $W(\hat{K},S,Z_i)=(1-\alpha)AZ_i\hat{K}^{\alpha}$, $K=\hat{K}$, $I=\hat{I}$, and $K'=\hat{K}'$, so

$$\begin{aligned} V(\hat{K};\hat{K},S,Z_i) =& \alpha AZ_i\hat{K}^{\alpha}-\Phi(\hat{I},\hat{K})-\hat{I} \\ &+\sum_{Z_j}\pi_Z(Z_i,Z_j)M(\hat{K},S,Z_i;Z_j)V(\hat{K}';\hat{K}',S',Z_j) \end{aligned}$$

Let's define $\hat{V}(\hat{K},S,Z_i)\equiv V(\hat{K};\hat{K},S,Z_i)$. The above equation suggests that $\hat{V}(\hat{K},S,Z_i)$ can be found by repeatedly applying the above equation as an operator (using the law of motion for capital $\hat{G}(\hat{K},S,Z_i)$).

### A.4.1 Approximations

I approximate the autoregressive process for productivity $z_t$ with a 13 point Markov Chain using the Rouwenhorst method. I approximate the law of motion for capital in the $K$ and $S$ directions using a two-dimensional cubic spline. The spline is of 6th degree in the $K$ direction and 14th degree in the $S$ direction. The spline breakpoints are log-spaced in both the $K$ and $S$ directions. I find that increasing the degree to the 14th in the $\hat{K}$ direction has no material impact on the quantitative results.

In a projection method, one must define what it means to satisfy the Euler equation. I use the collocation method, which specifies that the Euler equation should hold exactly at a set of points (collocation nodes) in the $K$ and $S$ domain. I choose these nodes to be the standard nodes for splines using knot averaging. I consider the algorithm converged if, across these collocation nodes, the maximum absolute Euler equation residual, expressed as $1-E[M'R_I']$, is less than 1.0E-8. I search for spline coefficients which satisfy this condition by using Broyden's method.

## A.5 Adjustment Costs and Accounting

Many papers specify adjustment costs in the following manner:

$$K'=(1-\delta)K+\phi(I/K)K$$

(e.g. Jermann (1998), Gourio (2009), Kaltenbrunner and Lochstoer (2010), Guvenen (2009), among others). This formulation preserves the traditional Cobb-Douglas formulation of output:

$$Y=ZK^{\alpha}N^{1-\alpha}=C+I$$

This formulation, however, deviates from the standard accounting treatment of investment and capital. The standard treatment specifies that end-of-period capital is beginning-of-period capital plus investment less depreciation. With geometric depreciation, this translates into the standard, adjustment-cost-free formulation of capital accumulation:

$$K' = (1-\delta)K + I$$

I choose to preserve this accounting identity. As a result, capital adjustment costs are pushed into output:

$$Y = ZK^{\alpha}N^{1-\alpha} - [\text{Adj Cost}] = C + I$$

Fortunately, both choices result in the same capital and consumption allocations. For example, Gourio (2009) uses

$$\begin{aligned} K^{*\prime} &= (1-\delta)K + \phi^{*}(I^{*}/K)K \\ \phi^{*}(x) &= x - \frac{\eta}{2}(x-\delta)^2 \end{aligned}$$

From these expressions, we have capital evolution and consumption

$$\begin{aligned} K^{*\prime} &= (1-\delta)K + I^{*} - \frac{\eta}{2}\left(I^{*}/K - \delta\right)^2 K \\ C &= ZK^{\alpha}N^{1-\alpha} - I^{*} \end{aligned}$$

which is identical to my formulation with

$$I \equiv I^{*} - \frac{\eta}{2}\left(I^{*}/K - \delta\right)^2 K$$

### A.6 Proofs for the 2-Period Model

***Proof of Proposition 1***. [16] A change of variables shows that this model is equivalent to a standard consumption-savings problem. Shift consumption by assigning $C^{*} = C_1 - H$. Then the date 1 consumption rule can be written as

$$C_1(W_1) = C^{*}(W_1 - H_1, -H_2) + H_1 \tag{38}$$

[16] I am grateful to Pok-Sang Lam for teaching me his version of the Carroll and Kimball (1996) proof.

Where $C^*(W,Y)$ solves a simple consumption-savings problem with wealth $W$ and certain future income of $Y$:

$$C^*(W,Y) \equiv \arg\max_C u(C) + \mathbb{E}\{u[R(W-C)+Y]\} \tag{39}$$

and for ease of notation I suppress the subscript 2 on $R$. It turns out that for $Y \neq 0$, $R$ random, and CRRA utility, $C^*(W,Y)$ is strictly convex in $W$. That is, with CRRA utility, rate of return randomness, and the introduction of any (even constant) future income is a sufficient condition for generating strict convexity of the consumption function. This is not one of the sufficient conditions shown in Carroll and Kimball (1996), so I will show that it is sufficient in what follows.

The FOC of the shifted problem (39) is

$$u'(C^*(W,Y)) = \phi'(W - C^*(W,Y)) \tag{40}$$

Where, for convenience, I've defined the function

$$\phi(S) \equiv \mathbb{E}\{u[RS+Y]\}$$

Taking $\frac{\partial}{\partial W}$ of the FOC and rearranging gives

$$\frac{\partial}{\partial W} C^*(W,Y) = \frac{\phi''}{u'' + \phi''}$$

Take another $\frac{\partial}{\partial W}$, do some serious algebra, and we get an expression for the convexity of $C^*$:

$$\frac{\partial^2}{\partial W^2} C^*(W,Y) = \left[\frac{(u'')^2 \left(\phi''\right)^2}{u' \times [u'' + \phi'']^3}\right] \left[\frac{\phi'\phi'''}{\left(\phi''\right)^2} - \frac{u'u'''}{(u'')^2}\right] \tag{41}$$

The first bracket is negative simply because $u' > 0$ and $u'' < 0$. To show that the second bracket is (strictly) positive, first note that, due to the CRRA specification, $\frac{u'u'''}{(u'')^2} = 1 + \frac{1}{\gamma}$. I will now show that, due to the non-zero future income $Y$, $\frac{\phi'\phi'''}{(\phi'')^2} > 1 + \frac{1}{\gamma}$. This is an extension of Carroll and Kimball (1996)'s Lemma 4.

Proving $\frac{\phi'\phi'''}{(\phi'')^2} > 1 + \frac{1}{\gamma}$ requires the following technical Lemma.

**Lemma 1.** *Let $\Phi_i$ for $i = 1, \ldots, N$ be $2 \times 2$ symmetric matrices with the following properties:*

- *the diagonals of each $\Phi_i$ are positive*
- *the off-diagonals of each $\Phi_i$ are all negative*

- *for every $i$, $|\Phi_i| = 0$*
- $\left|\sum_{i=1}^{N} \Phi_i\right| = 0$

*Then for each pair $i, j$ there is some constant $k$ such that*

$$\Phi_i = k\Phi_j$$

*Proof.* I will first show this for the case where $N = 2$. I'll then use the $N = 2$ results to prove the general case. For ease of notation assign $\Phi_1 \equiv \begin{pmatrix} p & q \\ q & r \end{pmatrix}$ and $\Phi_2 \equiv \begin{pmatrix} x & y \\ y & z \end{pmatrix}$. With some algebra, one can show that

$$|\Phi_1 + \Phi_2| = |\Phi_1| + |\Phi_2| + [\sqrt{pz} - \sqrt{xr}]^2 + 2[\sqrt{prxz} - qy] \tag{42}$$

A few facts will let us simplify this expression dramatically. First $|\Phi_1| = |\Phi_2| = |\Phi_1 + \Phi_2| = 0$, so those terms all drop out. Then note that, since $|\Phi_1| = |\Phi_2| = 0$, we have $pr = q^2$ and $xz = y^2$, and we can rewrite

$$\sqrt{prxz} = \sqrt{q^2 y^2} = qy$$

and so the last term in equation (42) also drops out. Thus equation (42) implies that $pz = xr$, or

$$\frac{p}{x} = \frac{r}{z} = k \tag{43}$$

where $k$ is the conjectured constant of proportionality. We just need to show that $\frac{q}{y} = \frac{p}{x} = k$. To show this, plug $pr = q^2$ and $xz = y^2$ into (43) and we have

$$\frac{p}{x} = \frac{q^2/p}{y^2/x} \Rightarrow \frac{p}{x} = \frac{q}{y}$$

This completes the $N = 2$ case.

To show the general case, first note that if $\Phi_i$ and $\Phi_j$ satisfy the requirements of the lemma, then $\Phi_i + \Phi_j$ also satisfies those requirements. Thus I can apply the $N = 2$ results to the general case, where one matrix is $\Phi_i$ and the other matrix is $\sum_{j \neq i} \Phi_i$. Moreover, note that if there is a $k$ such that $\Phi_i = k \sum_{j \neq i} \Phi_i$, then there is $m$ such that $\sum_j \Phi_j = m\Phi_i$. Apply this to all $i$ and get the desired result $\Phi_i = k\Phi_j$. □

Now, back to proving the proposition. I want to show that $\frac{\phi' \phi'''}{(\phi'')^2} > 1 + \frac{1}{\gamma}$. Suppose, for contradiction, that

$\frac{\phi'\phi'''}{(\phi'')^2} \leq 1+\frac{1}{\gamma}$. I can write this expression using the determinant of a $2\times 2$ matrix by defining

$$\Phi \equiv \mathbb{E}\begin{bmatrix} \phi' & \sqrt{1+\frac{1}{\gamma}}\phi'' \\ \sqrt{1+\frac{1}{\gamma}}\phi'' & \phi''' \end{bmatrix} = \mathbb{E}\begin{bmatrix} Ru'(z) & \sqrt{1+\frac{1}{\gamma}}R^2u''(z) \\ \sqrt{1+\frac{1}{\gamma}}R^2u''(z) & R^3u'''(z) \end{bmatrix}$$

Where, for ease of notation, $z \equiv R(W-C^*(W,Y))+Y$. This expression can now be written compactly as

$$|\Phi| \leq 0$$

Note that $\Phi$ is the weighted sum of many component matricies $\begin{bmatrix} Ru'(z) & \sqrt{1+\frac{1}{\gamma}}R^2u''(z) \\ \sqrt{1+\frac{1}{\gamma}}R^2u''(z) & R^3u'''(z) \end{bmatrix}$, and that due to the CRRA specification of $u$, the determinant of each component matrix is zero. Thus $\Phi$ is positive semidefinite, so $|\Phi| \geq 0$. But our assumption for contradiction says $\Phi$ is negative semidefinite, and so it must be that $|\Phi| = 0$.

Now I use Lemma 1. The lemma states that if $|\Phi| = 0$, all of the component matrices must be proportional to one another. This means that for any states $i$ and $j$ the ratio of the diagonal terms of the corresponding matrices is equal, that is, for any $i$ and $j$,

$$\begin{aligned}
& \frac{R_i}{R_j}\frac{u_i'}{u_j'} = \left(\frac{R_i}{R_j}\right)^3 \frac{u_i'''}{u_j'''} \\
\Rightarrow \quad & \left(\frac{R_iS+Y}{R_jS+Y}\right)^{-\gamma} = \left(\frac{R_i}{R_j}\right)^2 \left(\frac{R_iS+Y}{R_jS+Y}\right)^{-\gamma-2} \\
\Rightarrow \quad & \frac{R_iS+Y}{R_jS+Y} = \frac{R_i}{R_j} \\
\Rightarrow \quad & S+\frac{Y}{R_i} = S+\frac{Y}{R_j} \\
\Rightarrow \quad & R_i = R_j
\end{aligned}$$

which is a contradiction, since $R$ is random. Note that the presence of a nonzero income $Y$ is critical because otherwise, I could not move from the fourth line to the fifth line in the equations above.

Therefore, $\frac{\phi'\phi'''}{(\phi'')^2} > 1+\frac{1}{\gamma}$, and by equation (41), $C^*(W,Y)$ is strictly concave in $W$, and by equation (38), $C_1''(W) < 0$. □

***Proof of Proposition 2***. I first show that the transformed consumption function satisfies $\frac{\partial^3}{\partial W^3}C^*(W,Y) > 0$. To show this, note that $C^*(W,0)$ is linear, but for any $\epsilon > 0$, $C^*(W,\epsilon)$ is strictly concave. Thus we can sign the derivative

$$\frac{\partial^3}{\partial Y \partial W^2} C^*(W,0) = \lim_{\epsilon \to 0} \left[ \frac{\frac{\partial^2}{\partial W^2} C^*(W,\epsilon) - \frac{\partial^2}{\partial W^2} C^*(W,0)}{\epsilon} \right] = \lim_{\epsilon \to 0} \left[ \frac{\frac{\partial^2}{\partial W^2} C^*(W,\epsilon)}{\epsilon} \right] < 0$$

Assuming that $\frac{\partial^3}{\partial Y \partial W^2} C^*(W,0)$ is continuous, this means that there is some $\overline{Y} > 0$ such that for any $\delta < \overline{Y}$, $\frac{\partial^3}{\partial Y \partial W^2} C^*(W,\delta) < 0$ .

Since $C^*(W,Y)$ is HD1, I can take derivatives to show that

$$\frac{\partial^3}{\partial W^3} C^*(W/\delta, 1) = \frac{\delta^3}{W} \left[ -\underbrace{\frac{1}{\delta^2} \frac{\partial^2}{\partial W^2} C^*(W,\delta)}_{<0} - \underbrace{\frac{\partial^3}{\partial Y \partial W^2} C^*(W,\delta)}_{<0} \right]$$

Now note that I am free to choose $W$, so this means, for any $W$, $\frac{\partial^3}{\partial W^3} C^*(W,1) > 0$. But by homogeneity, $\frac{\partial^3}{\partial W^3} C^*(W,Y) = (1/Y^2) \frac{\partial^3}{\partial W^3} C^*(W,1) > 0$, and thus the third derivative with respect to wealth of the transformed consumption function is negative (as long as $Y \neq 0$).

Now to finish proving the proposition. Using the transformed consumption function (39) I can relate the desired derivative to the third derivative of the consumption function. First take $\frac{\partial}{\partial W_1}$ twice:

$$C_1''(W_1) = \frac{\partial^2}{\partial W^2} C^*(W_1 - H_1, -H_2)$$

Then take $\frac{\partial}{\partial H_1}$

$$\frac{\partial}{\partial H_1} C_1''(W_1) = -\frac{\partial^3}{\partial W^3} C^*(W_1 - H_1, -H_2) < 0$$

$\square$

***Proof of Corollary 2***. Let $\underline{W}_0$ and $\overline{W}_0$ be the minimum and maximum of $\mathbb{W}_0$, respectively.

$$\begin{aligned}
\max_{W_0\in\mathbb{W}_0}\sigma_0[C_1(W_1)] - \min_{W_0\in\mathbb{W}_0}\sigma_0[C_1(W_1)] \\
&= \sigma_0(\Delta W_1)\left[\max_{W_0\in\mathbb{W}_0} C_1'(W_0) - \min_{W_0\in\mathbb{W}_0} C_1'(W_0)\right] \\
&= \sigma_0(\Delta W_1)\left[C_1'(\underline{W}_0) - C_1'(\overline{W}_0)\right] \\
&= \sigma_0(\Delta W_1)\left[\int_{\overline{W}_0}^{\underline{W}_0} dW C_1''(W)\right] \\
&= -\sigma_0(\Delta W_1)\left[\int_{\underline{W}_0}^{\overline{W}_0} dW C_1''(W)\right]
\end{aligned}$$

Where the 2nd line uses equation (24), the third line uses $C_1''(W) < 0$.

Then take $\frac{\partial}{\partial H_1}$

$$\frac{\partial}{\partial H_1}\left[\max_{W_0\in\mathbb{W}_0}\sigma_0[C_1(W_1)] - \min_{W_0\in\mathbb{W}_0}\sigma_0[C_1(W_1)]\right] = -\sigma_0(\Delta W_1)\left[\int_{\underline{W}_0}^{\overline{W}_0} dW \underbrace{\frac{\partial}{\partial H_1} C_1''(W)}_{<0 \text{ (prop 2)}}\right] > 0$$

□

## References


A.B. Abel. Optimal investment under uncertainty. *The American Economic Review*, pages 228–233, 1983.

O.P. Attanasio. Consumption. *Handbook of macroeconomics*, 1:741–812, 1999.

R. Bansal and A. Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance*, 59(4):1481–1509, 2004.

R. Bansal, V. Khatchatrian, and A. Yaron. Interpretable asset markets? *European Economic Review*, 49(3): 531–560, 2005.

R. Bansal, D. Kiku, and A. Yaron. An empirical evaluation of the long-run risks model for asset prices. Technical report, National Bureau of Economic Research, 2009.

Ravi Bansal, A Ronald Gallant, and George Tauchen. Rational pessimism, rational exuberance, and asset pricing models. *The Review of Economic Studies*, 74(4):1005–1033, 2007.

Robert J Barro. Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, 121(3):823–866, 2006.

J. Beeler and J.Y. Campbell. The long-run risks model and aggregate asset prices: an empirical assessment. Technical report, National Bureau of Economic Research, 2009.

Geert Bekaert, Eric Engstrom, and Steven R Grenadier. Stock and bond returns with moody investors. *Journal of Empirical Finance*, 17(5):867–894, 2010.

Oliver Boguth and Lars-Alexander Kuehn. Consumption volatility risk. *The Journal of Finance*, 2013.

M. Boldrin, L.J. Christiano, and J.D.M. Fisher. Habit persistence, asset returns and the business cycle. 1999.

M. Boldrin, L.J. Christiano, and J.D.M. Fisher. Habit persistence, asset returns, and the business cycle. *American Economic Review*, pages 149–166, 2001.

Tim Bollerslev and Jeffrey M Wooldridge. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric reviews*, 11(2):143–172, 1992.

M.W. Brandt. Portfolio choice problems. *Handbook of financial econometrics*, 1(1), 2009.

J.Y. Campbell. Inspecting the mechanism: An analytical approach to the stochastic growth model. *Journal of Monetary Economics*, 33(3):463–506, 1994.

J.Y. Campbell. Consumption-based asset pricing. *Handbook of the Economics of Finance*, 1:803–887, 2003.

J.Y. Campbell and J.H. Cochrane. By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy*, 107(2):205–251, 1999.

J.Y. Campbell and R.J. Shiller. The dividend-price ratio and expectations of future dividends and discount factors. *Review of financial studies*, 1(3):195–228, 1988.

C.D. Carroll. A theory of the consumption function, with and without liquidity constraints. *The Journal of Economic Perspectives*, 15(3):23–45, 2001.

C.D. Carroll and M.S. Kimball. On the concavity of the consumption function. *Econometrica*, 64(4):981–992, 1996.

J.H. Cochrane. Production-based asset pricing and the link between stock returns and economic fluctuations. *Journal of Finance*, 46(1):209–37, 1991.

J.H. Cochrane. The dog that did not bark: A defense of return predictability. *Review of Financial Studies*, 21 (4):1533–1575, 2008a.

J.H. Cochrane. Financial markets and the real economy. 2008b.

J.H. Cochrane. Presidential address: Discount rates. *The Journal of Finance*, 66(4):1047–1108, 2011.

G.M. Constantinides. Habit formation: A resolution of the equity premium puzzle. *Journal of political Economy*, pages 519–543, 1990.

T.F. Cooley and E.C. Prescott. Economic growth and business cycles. *Frontiers of business cycle research*, pages 1–38, 1995.

Mariano Croce. Long-run productivity risk: A new hope for production-based asset pricing? *Available at SSRN 1106361*, 2010.

I. Dew-Becker. A model of time-varying risk premia with habits and production. 2011.

Darrell Duffie and Kenneth J Singleton. Simulated moments estimation of markov models of asset prices. *Econometrica*, 61(4):929–952, 1993.

L.G. Epstein and S.E. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica: Journal of the Econometric Society*, pages 937–969, 1989.

J. Favilukis and X. Lin. Infrequent renegotiation of wages: A solution to several asset pricing puzzles. Technical report, Working paper London School of Economics and Ohio State University, 2011.

Milton Friedman. *A Theory of the Consumption*. princeton university press Princeton, 1957.

Xavier Gabaix. Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *The Quarterly Journal of Economics*, 127(2):645–700, 2012.

L.R. Glosten, R. Jagannathan, and D.E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of finance*, pages 1779–1801, 1993.

F. Gourio. Disasters risk and business cycles. Technical report, National Bureau of Economic Research, 2009.

F. Gourio. Disaster risk and business cycles. 2010.

F. Guvenen. A parsimonious macroeconomic model for asset pricing. *Econometrica*, 77(6):1711–1750, 2009. ISSN 1468-0262.

James D Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, pages 357–384, 1989.

Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054, 1982.

F. Hayashi. Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica: Journal of the Econometric Society*, pages 213–224, 1982.

U.J. Jermann. Asset pricing in production economies. *Journal of Monetary Economics*, 41(2):257–275, 1998.

U.J. Jermann. The equity premium implied by production. *Journal of Financial Economics*, 98(2):279–296, 2010.

K.L. Judd. Projection methods for solving aggregate growth models. *Journal of Economic Theory*, 58(2):410–452, 1992.

G. Kaltenbrunner and L.A. Lochstoer. Long-run risk through consumption smoothing. *Review of Financial Studies*, 23(8):3190–3224, 2010.

S. Kandel and R.F. Stambaugh. Expectations and volatility of consumption and asset returns. *Review of Financial Studies*, 3(2):207–232, 1990.

M.S. Kimball. Precautionary saving in the small and in the large. *Econometrica: Journal of the Econometric Society*, pages 53–73, 1990.

L. Kogan and D. Papanikolaou. Economic activity of firms and asset prices. *Annual Review of Financial Economics*, 4(1), 2012.

Leonid Kogan. Asset prices and real investment. *Journal of Financial Economics*, 73(3):411–431, 2004.

L.A. Kuehn, N. Petrosky-Nadeau, and L. Zhang. An equilibrium asset pricing model with labor market search. Technical report, National Bureau of Economic Research, 2012.

S.F. LeRoy and R.D. Porter. The present-value relation: Tests based on implied variance bounds. *Econometrica: Journal of the Econometric Society*, pages 555–574, 1981.

M. Lettau. Inspecting the mechanism: Closed-form solutions for asset prices in real business cycle models*. *The Economic Journal*, 113(489):550–575, 2003.

M. Lettau and H. Uhlig. Can habit formation be reconciled with business cycle facts? *Review of Economic Dynamics*, 3(1):79–99, 2000.

M. Lettau, S.C. Ludvigson, and J.A. Wachter. The declining equity premium: What role does macroeconomic risk play? *Review of Financial Studies*, 21(4):1653–1687, 2008.

L. Ljungqvist and H. Uhlig. Optimal endowment destruction under campbell-cochrane habit formation. Technical report, National Bureau of Economic Research, 2009.

M. Miranda and P. Fackler. Applied computational economics, 2001.

D. Papanikolaou. Investment shocks and asset prices. *Journal of Political Economy*, 119(4):639–685, 2011.

O. Posch. Risk premia in general equilibrium. *Journal of Economic Dynamics and Control*, 35(9):1557–1576, 2011.

William H Press, Saul A Teukolsky, William T Vetterling, and Brian P Flannery. Numerical recipes in c: The art of scientific computing (; cambridge, 1992.

Morten O Ravn and Harald Uhlig. On adjusting the hodrick-prescott filter for the frequency of observations. *Review of Economics and Statistics*, 84(2):371–376, 2002.

F. Restoy and G.M. Rockinger. On stock market returns and returns on investment. *Journal of Finance*, pages 543–556, 1994.

Thomas A Rietz. The equity risk premium a solution. *Journal of monetary Economics*, 22(1):117–131, 1988.

T. Santos and P. Veronesi. Habit formation, the cross section of stock returns and the cash-flow risk puzzle. *Journal of Financial Economics*, 98(2):385–413, 2010.

RJ Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71(3):421–436, 1981.

S.M. Sundaresan. Intertemporally dependent preferences and the volatility of consumption and wealth. *Review of financial Studies*, 2(1):73–89, 1989.

J. Wachter. Can time-varying risk of rare disasters explain aggregate stock market volatility? In *AFA 2011 Denver Meetings Paper*, 2012.

Jessica A Wachter. A consumption-based model of the term structure of interest rates. *Journal of Financial economics*, 79(2):365–399, 2006.