

**Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.**

**An Empirical Investigation of Consumption-based Asset Pricing
Models with Stochastic Habit Formation**

Qiang Dai and Olesya V. Grishchenko

2011-47

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

An Empirical Investigation of Consumption-based Asset Pricing Models with Stochastic Habit Formation *

Qiang Dai [†]
and

Olesya V. Grishchenko[‡]

October 19, 2011

Abstract

We econometrically estimate a consumption-based asset pricing model with stochastic internal habit and test it using the generalized method of moments. The model departs from existing models with deterministic internal habit (e.g., Dunn and Singleton (1983), Ferson and Constantinides (1991), and Heaton (1995)) by introducing shocks to the coefficients in the distributed lag specification of consumption habit and consequently an additional shock to the marginal rate of substitution. The stochastic shocks to the consumption habit are persistent and provide an additional source of time variation in expected returns. Using Treasury bond returns and broad equity market index returns, we show that stochastic internal habit formation models resolve the dichotomy between the autocorrelation properties of the stochastic discount factor and those of expected returns. Consequently, they provide a better explanation of time-variation in expected returns than models with either deterministic habit or stochastic external habit.

JEL Classification: E21, G10, G12

Keywords: asset pricing, consumption-based asset pricing models, habit formation, stochastic internal habit, aggregate equity and bond returns

*We thank Stephen Brown, Jean Helwege, Jingzhi (Jay) Huang, Martin Lettau, Michelle Lowry, Sydney Ludvigson, Dmitry Makarov, Stijn Van Nieuwerburgh, Marti Subrahmanyam, Joel Vanden, NYU seminar participants, conference participants of the 2005 Annual Trans-Atlantic Doctoral Conference at LBS, London, UK and 2005 AEA meetings in Philadelphia. Address correspondence to Olesya.V.Grishchenko@frb.gov, phone: (202) 452-2981. The views expressed here are solely those of the authors and do not necessarily reflect the concurrence by other members of the research staff or the Board of Governors of the Federal Reserve System.

[†]Chief Risk Officer, Capula Investment Management LLP, London

[‡]Economist at the Division of Monetary Affairs, Board of Governors of the Federal Reserve System, Washington, DC 20551; Olesya.V.Grishchenko@frb.gov

An Empirical Investigation of Consumption-based Asset Pricing Models with Stochastic Habit Formation

Abstract

We econometrically estimate a consumption-based asset pricing model with stochastic internal habit and test it using the generalized method of moments. The model departs from existing models with deterministic internal habit (e.g., Dunn and Singleton (1983), Ferson and Constantinides (1991), and Heaton (1995)) by introducing shocks to the coefficients in the distributed lag specification of consumption habit and consequently an additional shock to the marginal rate of substitution. The stochastic shocks to the consumption habit are persistent and provide an additional source of time variation in expected returns. Using Treasury bond returns and broad equity market index returns, we show that stochastic internal habit formation models resolve the dichotomy between the autocorrelation properties of the stochastic discount factor and those of expected returns. Consequently, they provide a better explanation of time-variation in expected returns than models with either deterministic habit or stochastic external habit.

JEL Classification: E21, G10, G12

Keywords: asset pricing, consumption-based asset pricing models, external habit, internal habit, habit shocks, aggregate equity and bond returns

1 Introduction

Consumption-based capital asset pricing models (CCAPMs) with time non-separable preferences including habit formation have received considerable attention recently as a potential mechanism for explaining the equity premium puzzle (see, e.g., Sundaresan (1989), Abel (1990), Constantinides (1990), and Campbell and Cochrane (1999)). The joint behavior of stock and bond returns within the context of consumption-based models has been understood much less. In this paper we present a structural model with stochastic internal habits that helps resolve the dichotomy between the autocorrelation properties of the stochastic discount factors and those of bond returns.¹ Our formulation nests many of the previously studied specifications along one or both of the following two dimensions. First, by adopting the GMM approach (Hansen, 1982), we are completely agnostic about the dynamics of the aggregate endowment or technology imposed in other papers, such as, e.g., Campbell and Cochrane (1999), Wachter (2006), and Bekaert, Engstrom, and Grenadier (2010). This means that our model implicitly nests any parametric specifications of the aggregate endowment process. Second, we adopt a general parametric specification of the consumption habit process that nests many of the previous habit specifications. Specifically, we start with a standard geometric distributed lag specification, and then introduce a stochastic shock to the weights in the distributed lag. The timing of the new shock is such that it can be literally interpreted as an unexpected shock to the habit stock itself.²

Using the same econometric method, Dunn and Singleton (1986, DS) and Ferson and Constantinides (1991, FC) test the over-identifying restrictions implied by the stochastic Euler equations derived from such models, while being agnostic about the endowment specification and the nature of the temporal dependence of consumption. Heaton (1995), on the

¹The emphasis of this paper is on time-series properties of aggregate market returns. There are a few recent papers that explore the impact of habits on the cross-section of asset returns; see, e.g., Van Binsbergen (2007), Santos and Veronesi (2010), and Heyerdahl-Larsen (2010).

²In the existing literature, the terms “habit shock” and “habit stock” are often used interchangeably, because the “habit stock” is viewed as a “preference shock”. In this paper, we differentiate these two terms: the term “habit stock” refers to the habit level, whereas the term “habit shock” refers to the stochastic shock to the habit level. In a model with a deterministic habit, the habit level is determined entirely by consumption history. In a model of stochastic habit, the habit level is affected by its own shock, so that consumption history is not sufficient to determine the current habit level.

other hand, uses the simulated method of moments to estimate a model that incorporates both a parametric form of consumption durability and habit.^{3,4} These studies find that there is empirical evidence for both consumption habit and local substitution. However, which effect dominates may depend on the decision interval, the instruments used in estimating the model, and the investment horizon. The overall goodness-of-fit test econometrically rejects all of the models. The econometric rejection of existing CCAPMs with time separable and time nonseparable utility is largely due to the notorious failure of these models to explain bond returns. Singleton (1993) and Heaton (1995) explore in depth such models' inability to explain the autocorrelation properties of bond returns. This problem arises because a stochastic discount factor (SDF henceforth) in the time separable models and some of the time nonseparable models is a simple function of the nearly *i.i.d.* consumption growth process, and this makes it difficult for models to explain persistent properties of long-term bond returns.

The failure of earlier asset pricing models with time non-separable preferences to explain time-varying bond returns arises because both types of models, either external or internal habits, (Constantinides (1990) and Campbell and Cochrane (1999, CC), respectively) sidestep the issue of term structure dynamics by imposing restrictions so that the real term structure is constant and flat. Several papers extend the CC model in order to accommodate stochastic interest rates. Campbell and Cochrane (1999) show how to extend their model by relaxing a parametric restriction on the specification of the surplus consumption ratio. However, in their model, interest rates and bond risk premium are perfectly correlated with consumption shock, which is counter-factual. Wachter (2006) relaxes the assumption of the CC model of an *i.i.d.* consumption growth rate and shows that interest rates and the risk premium have properties broadly consistent with observed bond return predictability in her model. Buraschi and Jiltsov (2007) develop a continuous term structure model in the context of the CC model and show that habit persistence can help reproduce various properties of the

³For a simulated method of moments, see, e.g., Duffie and Singleton (1993), Gouieroux, Monfort, and Renault (1993), and Gallant and Tauchen (1996).

⁴Heaton also adopts a parametric form for the aggregate endowment process (estimated from data as part of a bi-variate vector autoregressive process), and addresses the issue of time aggregation of consumption data.

term structure. Bekaert, Engstrom, and Xing (2009) also consider a generalized version of the CC model where a surplus consumption ratio is stochastic, which allows them to obtain some non-trivial implications for the term structure of interest rates. All of these extensions share the common feature that the external habit process is no longer locally deterministic.

There has been much less work done extending deterministic internal habit process to accommodate preference shocks. However, this venue is important in view of two recent empirical papers by Chen and Ludvigson (2009) and Grishchenko (2010), who investigate the properties of habit persistence at the aggregate level and conclude that aggregate data are more consistent with internal rather than external habit formation preferences. The reason is that internal habit generates the necessary autocorrelations in the stochastic discount factor via a non-trivial impact of current consumption on future marginal utility. This channel is absent in the external habit models where current consumption affects only the surplus consumption ratio, but the functional form of the marginal utility stays unaltered (compared to marginal utility in the time separable preferences). Dai (2003) is the first who introduces the stochastic internal habit model. He relaxes Constantinides' assumption of a constant investment opportunity set by allowing the instantaneous short rate to be driven by the level of the habit stock, and shows that the time-varying risk premium implied by the model is capable of explaining the violation of the Campbell and Shiller (1991) expectations hypothesis puzzle. In our work, the deterministic internal habit formation models of Sundaresan (1989) and Constantinides (1990), and the stochastic internal habit formation model of Dai (2003) are nested special cases in the continuous-time limit.

The presence of the habit shock in the economy – in addition to the consumption shock – breaks the tight linkage between asset returns and consumption growth rates in the standard CCAPM or its extension with deterministic habit. Economically, the habit shock can be broadly interpreted as a taste shock. Its high realization occurs in the bad state of the world, associated with the low dividend payoff and the low level of consumption. In this state, the additional dollar of return becomes even more valuable because marginal utility depends positively on this shock.

To our knowledge, this work is the first empirical investigation of stochastic internal habit

models. It relates to prior and rather limited work on the impact of preference shocks on asset prices. Among the first, Campbell (1986) includes random shocks in the CRRA utility to examine the conjecture of Modigliani and Sutch (1966) that investor preferences might cause negative term premiums on long-term bonds. He argues that randomness in preferences (interpreted as taste shocks) generates predictable excess returns even when agents are risk-neutral. Normandin and St-Amour (1998) include taste shocks in the preference specification of Epstein and Zin (1989) to study their effect on the equity premium. They argue that taste shocks help alleviate the emphasis on the consumption risk in explaining the historical equity premium. More recently, Brandt and Wang (2003) allow for a preference shock in the CC model to be correlated with business-cycle factors. In a similar spirit, Bekaert, Engstrom, and Grenadier (2010) allow for a latent preference shock in the CC model to be imperfectly correlated with consumption growth and interpret it as unexpected change in the “moodiness” of the investor behavior.⁵

In contrast to previous studies (with the exception of Bekaert, Engstrom, and Xing (2009) and Bekaert, Engstrom, and Grenadier (2010)), we study the implications of introducing the preference shock to the habit stock to *simultaneously* explain time series properties of equity and bond returns.⁶ Our modeling and econometric framework allows us to address, among others, two important empirical questions.

Q1: Do *stochastic* habit formation models explain long-term bond returns better than their *deterministic* counterparts?

Q2: Do *internal* habit (IH) formation models and their *external* counterparts (EH) perform equally well when they are confronted by both equity and bond returns?

The answer to **Q1** is an emphatic “yes!” We find that when the habit process is deterministic, the model is incapable of explaining quarterly holding-period returns of long-term bonds, regardless of whether the consumption habit is external or internal. When the consumption habit is internal and stochastic, the model better explains both equity and bond

⁵Essentially, these two papers depart from the framework of Campbell and Cochrane (1999) who assume that the habit shock is perfectly negatively correlated with the consumption shock.

⁶Where confusion does not arise, we use the terms “habit shock” and “taste shock” interchangeably.

returns simultaneously. This suggests the joint importance of internal habit and taste shock in preferences, absent in models discussed above.

The answer to **Q2** is negative. Based on Euler equations derived from the level of the three-month Treasury bill rate, the quarterly return of a 10-year long-term Treasury bond, and the quarterly return of the broad equity market index, the parameter estimates of IH models are all very reasonable. The point estimates of the coefficient of the relative risk aversion in the IH model (either deterministic or stochastic) are around 2, while they are roughly 8 for the deterministic EH models. In contrast, the parameter estimates of the stochastic EH models are such that either the relative risk aversion parameter or the subjective discount rate is in the wrong region when they are unconstrained. When the relative risk aversion parameter and the subjective discount rate are constrained to be positive, we find a corner solution in minimizing the GMM objective function under the optimal weighting matrix. This result suggests that EH models, whether deterministic or stochastic, cannot reconcile various moments of bond returns. Our results are complementary to the results in Dai, Le, and Singleton (2010). The authors specify the dynamics of state variables (in contrast to our agnostic approach) and construct the maximum likelihood estimates of a nonlinear discrete time dynamic term structure model with CC-type external habit preferences in which bond prices are known in closed form. Consequently, Dai et al. show that such models do not match the key features of the conditional distribution of bond yields.

The data seems to favor internal relative to external habit formation for a reason. Introducing a stochastic shock to an external habits does not affect the the autocorrelation structure of the marginal rate of substitution because its functional form remains the same. In contrast, a stochastic habit shock materially impacts the performance of an internal habit formation model because it induces an additional term in the marginal rate of substitution that makes it more persistent. Our results are consistent with Ljungqvist and Uhlig (2000), who argue that when agents are “catching up with the Joneses”, the social welfare in a competitive equilibrium is reduced relative to the socially optimal allocation that takes into account the negative externality of habit formation. Agents consume too much when productivity is high and too little when productivity is low. Government intervention

(through income transfer) can induce competitive agents with external habit formation to achieve socially optimal consumption behavior. Thus, aggregate consumption behavior and asset prices seem to be more consistent with internal habit even though agents may exhibit external habit formation at individual levels. Overall, this evidence suggests that stochastic internal habit models should be preferred to their external counterparts when studying the joint behavior of the aggregate stock and bond returns.

The rest of the paper is organized as follows. In Section 2 we set forth a model specification and derive Euler equations. In Section 3 we discuss methodological issues related to our empirical study. In Section 4 we present empirical results for deterministic/stochastic external and internal habit models, discuss miscellaneous estimation issues and also report some robustness checks. We conclude in Section 5.

2 A CCAPM with stochastic habit formation

We assume that there exists a representative agent with a time non-separable expected utility:

$$V_0 = E \left[\sum_{t=0}^{\infty} e^{-\rho t} u(s_t) \middle| I_0 \right], \quad (1)$$

where $\rho > 0$ is the subjective discount rate, $u(\cdot)$ is strictly increasing and strictly concave, s_t is the service flow provided by the current and past “surplus consumption” $z_t \equiv c_t - x_t$, c_t is an exogenous aggregate consumption/endowment process, x_t is the habit stock generated by the past aggregate consumption, and I_t is the time- t information set. The service flow s_t and the habit stock x_t are defined by

$$s_t = D(L) z_t = \sum_{j=0}^{\infty} \delta^j z_{t-j}, \quad (2)$$

$$x_t = B(L)(1 + \beta v_t) z_{t-1} = b \sum_{j=0}^{\infty} (1 - \kappa)^j (1 + \beta v_{t-j}) z_{t-1-j}, \quad (3)$$

where, conditional on I_t , v_{t+1} is a random variable with zero mean and unit volatility with an arbitrary correlation with the consumption shock, L is the lag operator, and $D(L)$ and

$B(L)$ are two geometrically distributed-lag operators:

$$D(L) = \sum_{j=0}^{\infty} \delta^j L^j, \quad B(L) = b \sum_{j=0}^{\infty} (1 - \kappa)^j L^j, \quad \delta > 0, b > 0, \kappa > 0.$$

In this setting, δ represents a service flow parameter, β is the volatility of the habit shock, b is the scaling parameter that indexes the importance of the habit formation level relative to the current consumption level, and κ indexes the degree of “persistence”, or “memory” in the habit shock. When $\beta = 0$, the model reduces to the time-non-separable utility specification of Dunn and Singleton (1986), Ferson and Constantinides (1991), and Heaton (1995). If, in addition, $\delta = 0$, the model reduces further to the habit formation of Sundaresan (1989) and Constantinides (1990), as the decision interval goes to zero. Finally, when $\delta = b = 0$, the model reduces to the standard pure exchange economy of Lucas (1978) and Mehra and Prescott (1985) with time-separable utility.

2.1 Stochastic Euler Equations

Following Dunn and Singleton (1986), Ferson and Constantinides (1991), and Heaton (1995), we can derive the stochastic Euler equation for the model specified by equations (1) – (3). In Appendix A, we show that, for any security with the price-dividend pair (p_t, d_t) , the following pricing equation holds:

$$E_t [\text{MRS}_{t,t+1} \times R_{t,t+1}] = 1, \tag{4}$$

where $R_{t,t+1} \equiv \frac{p_{t+1} + d_{t+1}}{p_t}$ is the one-period return for the security, and MRS is the marginal rate of substitution, given by

$$\text{MRS}_{t,t+1} = e^{-\rho} \times \frac{\text{MUC}_{t+1}}{\text{MUC}_t}, \quad \text{where} \tag{5}$$

$$\text{MUC}_t = E_t \left[\sum_{j=0}^{\infty} e^{-\rho j} \times a_{t,j} \times u'(s_{t+j}) \right]. \tag{6}$$

Here MUC_t is the marginal utility of consumption in terms of the value function at t , and for the geometric distributed lag specification,⁷

$$a_{t,j} = \delta^j - \sum_{i=0}^{j-1} \delta^i b_{t,j-i}, \quad j \geq 1, \quad a_{t,0} = 0, \quad \text{where} \quad (7)$$

$$b_{t,j} = \begin{cases} [(1 - \kappa) \prod_{i=1}^{j-1} - \prod_{i=0}^{j-1}] [1 - \kappa - b(1 + \beta v_{t+i+1})], & \text{(internal),} \\ 0, & \text{(external).} \end{cases} \quad (8)$$

(By convention we set $\prod_{i=1}^0 (\cdot) = 1$.)

For an internal habit formation model, $b_{t,j} \neq 0$ for $j \geq 1$, which means that the individual agent accounts for the consumption externality induced by her own consumption choice. Given a particular model parametrization with internal habit formation, the corresponding model with external habit formation is obtained by setting $b_{t,j} = 0$ for all $j \geq 1$.

Equation (4) can be extended to deal with returns on investment strategies with multiple holding periods. Let $R_{t,t+n}$ be the return on an investment strategy held for n periods, then the same reasoning behind the derivation of (4) implies

$$E_t \left[\frac{MUC_{t+n}}{MUC_t} \times R_{t,t+n} \right] = 1. \quad (9)$$

2.2 Determinants of MUC and MRS: intuition

To develop some intuition on the relative importance of various ingredients of the model, let us consider a first-order approximation of MUC_t when β is small.⁸ First, note that when β is small, the approximation of $a_{t,j}$ is given by

$$a_{t,j} = \delta^j - b_j(1 + \beta v_{t+1}), \quad b_j = b(1 - \kappa)^{j-1} \left[\frac{1 - \left(\frac{\delta}{1-\kappa}\right)^j}{1 - \frac{\delta}{1-\kappa}} \right], \quad \text{for } \forall j \geq 1.$$

⁷We adopt the distributed lag specification to economize model parameters. The derivation of equations (7)–(8) are given in Appendix B when $\delta = 0$. The Euler equation can be also derived explicitly when the distributed lag contains arbitrary weights.

⁸This first-order approximation is used here only for the purpose of developing intuition. The exact model is estimated and tested later in the empirical section.

Consequently, MUC_t allows the following decomposition:

$$\frac{MUC_t}{u'(s_t)} - 1 \approx \Psi_t^{(\delta)} + \Psi_t^{(b)} + \Psi_t^{(\beta)}, \quad (10)$$

where

$$\begin{aligned} \Psi_t^{(\delta)} &= E_t \left[\sum_{j=1}^{\infty} e^{-\rho j} \delta^j \frac{u'(s_{t+j})}{u'(s_t)} \right] \propto \delta, & \Psi_t^{(b)} &= -E_t \left[\sum_{j=1}^{\infty} e^{-\rho j} b_j \frac{u'(s_{t+j})}{u'(s_t)} \right] \propto b, \\ \Psi_t^{(\beta)} &= \beta \varrho b \bar{\sigma} (1 - \gamma) e^{-\rho} S_t^{-1} \propto b \varrho \beta \end{aligned}$$

represent the expected future marginal utility or dis-utility (depending on the sign), from future service flow and consumption habit. In $\Psi_t^{(\beta)}$, ϱ is the correlation between the consumption shock and the habit shock, $\bar{\sigma}$ is the volatility of the consumption growth rate, $1 - \gamma$ is the relative risk aversion coefficient of the CRRA utility function, and S_t is the surplus consumption ratio.

Note that $\Psi_t^{(\delta)} = \Psi_t^{(b)} = \Psi_t^{(\beta)} = 0$ when $\delta = b = 0$ (the time-separable case), and $\Psi_t^{(b)} = \Psi_t^{(\beta)} = 0$ whenever $b = 0$ (no habit). When $b \neq 0$, the relative magnitude of $\Psi_t^{(b)}$ and $\Psi_t^{(\beta)}$ depends on, in large part, the relative magnitude of b , and $\varrho\beta$. The magnitude of $\Psi_t^{(\beta)}$ relative to $\Psi_t^{(b)}$ is also controlled by the parameter κ , or more precisely the ‘‘leverage ratio’’ $\frac{b}{\kappa}$, which is the long-run mean of $(S_t^{-1} - 1)$.

The decomposition (10) provides clear intuition for how model parameters are identified: given the preference parameters ρ , γ , and the persistence parameter κ , the parameters δ , b , and β are identified, respectively, from $\Psi_t^{(\delta)}$, $\Psi_t^{(b)}$, and $\Psi_t^{(\beta)}$. This intuition is helpful in guiding the choice of asset returns and instruments that help achieve identification and efficiency.

2.3 Discussion

Let ϵ_t be the shock to aggregate consumption growth rate. Our model implies that the pricing kernel is a function of the history of both the consumption shock ϵ_t and the habit

shock v_t . That is, for any $t < s$,

$$\text{MRS}_{t,s} = \text{M}(\epsilon_\tau, v_\tau : \tau \leq s; t, s). \quad (11)$$

In general, the identity of the shocks that appear in the MRS (which means that they are priced risk factors) and the functional form of the MRS are part of the model specification. In this paper, however, we interpret the stochastic habit formation model as some “reduced-form” specification of an underlying structural model based on micro-economic foundations. In particular, we interpret the shock to the habit level as taste shock, that proxies some economy-wide shock different from the shock to the consumption growth rate. In this way, the aggregate asset returns are driven by two shocks, the endowment shock and the taste shock. Campbell (1986) and Normandin and St-Amour (1998) undertake similar interpretation of the random shocks in the preferences. They study different questions however. Campbell studies the effect of the taste shocks on bond premia, while Normandin and St-Amour study the impact of taste shocks on equity premia. In both papers, multiplicative shocks are added to the utility specifications, albeit to different ones: Campbell considers standard CRRA utility function, while Normandin and St-Amour consider Epstein-Zin preferences. More recently, Bekaert, Engstrom, and Grenadier (2010) develop an extension to the external habit CC model by modeling a stochastic surplus consumption ratio where a preference shock to it is interpreted as an unexpected change in the “moodiness” of the investor behavior. Bekaert, Engstrom, and Xing (2009) use this model to study the relative importance of the economic uncertainty (conditional variance of fundamentals) and change in the risk aversion.

Our preference specification (including the habit specification) allows us to derive a specific functional form for MRS. In order to make the model empirically testable, we choose an empirical proxy for the taste shock based on an educated guess of which macro variables may be the most relevant to asset pricing. Consequently, we will be testing the joint hypothesis that the model is correctly specified and the stochastic habit shock is correctly identified. This is important for a proper interpretation of our results.

3 Methodological and econometric issues

In this section, we discuss several methodological and econometric issues related to the estimation and testing of the model. Before we proceed, let us give a brief description of the GMM procedure to establish some notations.

In the most general form with pure habit ($\delta = 0$), our model has five parameters, collected in the vector $\theta = (\rho, \gamma, \kappa, b, \beta)'$. Let $R_{t,t+n}$ be a $K \times 1$ vector of asset returns, and $Z_t \in J_t$ be a $M \times 1$ vector of instruments. Then under the null, the following $K \cdot M$ orthogonality conditions must hold:

$$E [h_{t,t+n} \otimes Z_t] = 0, \text{ where } h_{t,t+n} = \frac{\text{MRS}_{t+1}}{\text{MRS}_t} \times R_{t,t+n} - 1. \quad (12)$$

Let $g_T = \frac{1}{T} \sum_{t=1}^T \epsilon_{t,t+n}$ denote the sample counterpart of the left hand side of equation (12), where T is the sample length and ϵ_t is a $K \cdot M \times 1$ vector obtained by stacking together all elements of $h_{t,t+n} \otimes Z_t$. Then the GMM estimator, θ_T , solves

$$\theta_T = \arg \max_{\theta} T g_T' W_T^{-1} g_T,$$

where W_t is the sample counter-part of the optimal weighting matrix (see Hansen (1982) for additional details).

Under the null, the GMM objective function $T g_T' W_T^{-1} g_T$ has an asymptotic χ^2 distribution with degrees of freedom equal to $(KM - \dim(\theta))$. This provides an overall goodness-of-fit test.

3.1 Evaluation of the Euler equations

A well-known issue associated with the econometric estimation and test of internal habit formation models is that the “marginal utility of consumption”, MUC_t , is defined in terms of the conditional expectation of future marginal utilities (see equation (6)). Two distinct approaches have been adopted to deal with this issue. Dunn and Singleton (1986), Ferson and Constantinides (1991), Ferson and Harvey (1992), etc., adopted the first approach,

which avoids the evaluation of the conditional expectations altogether. Specifically, rewriting equation (9) as

$$E_t [\text{MUC}_t - \text{MUC}_{t+n} \times R_{t,t+n}] = 0, \quad (13)$$

and removing the conditional expectation operator by appealing to the law of iterated expectations, we obtain $E_t[h_{t,t+n}] = 0$, where

$$h_{t,t+n} = \sum_{j=0}^{\infty} e^{-\rho j} a_{t,j} u'(s_{t+j}) - \sum_{j=0}^{\infty} e^{-\rho(j+n)} a_{t+n,j} u'(s_{t+n+j}) \times R_{t,t+n}. \quad (14)$$

Note that the “disturbance” h_t has an $\text{MA}(\infty)$ autocorrelation structure when the consumption habit is defined as a geometric distributed lag. Dunn and Singleton (1986), Ferson and Constantinides (1991), and Ferson and Harvey (1992) get around this problem by defining the consumption habit in terms of a limited number (typically one and at most two) of consumption lags. If we define the consumption habit using q lags, then h_t is $\text{MA}(q)$ under the null that the model is correctly specified. This approach does not apply directly in our case because we have $q = \infty$. Heaton (1995) adopted the second approach, which explicitly evaluates the conditional expectation by imposing a particular dynamic structure of the relevant state variables. Specifically, Heaton assumes that a bivariate process of aggregate consumption growth and aggregate dividend growth characterizes completely the dynamic evolution of the marginal utility of consumption. By modeling the state vector as a bivariate VAR, MUC_t can be computed numerically (at the estimated parameter values for the VAR). Heaton estimates the preference parameters using the simulated method of moments.

In principle, we can follow Heaton’s approach to estimate our model. An added benefit of adopting Heaton’s approach is that there is no requirement that the stochastic habit shock be observed. Our concerns about potential mis-specification errors associated with the VAR specification of the state vector and potential numerical errors in computing MUC_t , however, lead us to develop a new approach. The idea is to rewrite the Euler equation (13) as

$$E \left[\Phi_t - e^{-\rho n} \times \Phi_{t+n} \times \frac{u'(s_{t+n})}{u'(s_t)} \times R_{t,t+n} \middle| I_t \right] = 0, \quad (15)$$

where I_t is the information set at t , and

$$\Phi_t \equiv \sum_{j=0}^{\infty} e^{-\rho j} a_{t,j} \frac{u'(s_{t+j})}{u'(s_t)}.$$

Note that for pure external habit formation models ($\delta = 0$ and $b_{t,j} = 0$ for all $j \geq 1$), $\Phi_t = 1$, and the Euler equation is trivial to compute. For internal habit formation models (with or without local substitution), we adopt the following strategy.

First, we construct an information set $J_t \subseteq I_t$ that includes all the instruments we will use for estimating the model as well as $z_t \equiv \frac{u'(s_t)}{u'(s_{t-n})} \times R_{t-n,t}$ and an appropriate number of its lags. Second, we condition the Euler equation (15) down to the information set J_t :

$$E \left[\Phi_t - e^{-\rho n} \times \Phi_{t+n} \times \frac{u'(s_{t+n})}{u'(s_t)} \times R_{t,t+n} \middle| J_t \right] = 0. \quad (16)$$

Finally, we compute the linear projection of Φ_t into the information set J_t , and substitute the projection $\hat{\Phi}_t$ into equation (16). Since the projection error, $\epsilon_t \equiv \Phi_t - \hat{\Phi}_t$, is orthogonal to J_t – by construction, equation (16) implies

$$E \left[\hat{h}_{t,t+n} \middle| J_t \right] = 0, \quad (17)$$

where

$$\hat{h}_{t,t+n} = \hat{\Phi}_t - e^{-\rho n} \times \hat{\Phi}_{t+n} \times \frac{u'(s_{t+n})}{u'(s_t)} \times R_{t,t+n}. \quad (18)$$

Equations (17) and (18) are the “projected” Euler equations that we actually use in our econometric analysis. Under the null, the vector of projected disturbances $\hat{h}_{t,t+n}$ has a MA($n - 1$) autocorrelation structure. In particular, for one-period returns, $n = 1$, $\hat{h}_{t,t+1}$ should be martingale difference sequences.

The projection procedure does not affect consistency of the GMM estimators. However, it does affect inference. To obtain standard errors and test statistics properly, we need to (a) compute standard errors with the projection fixed at the converged parameter estimates; and (b) account for sampling noise in the first-stage linear projection when we compute the

standard errors and test statistics in the second (GMM) stage.

3.2 Empirical proxy for a stochastic habit (taste) shock

Implicit in our model is the assumption that economic agents observe both consumption and habit shocks, conditional on which consumption and portfolio demands are formed. As a consequence, the Euler disturbances $h_{t,t+n}$ or their projected versions $\hat{h}_{t,t+n}$ can not be constructed using consumption and return data alone, because they also depend on the realizations of the habit shock $v_1, \dots, v_{t+n-1}, v_{t+n}$. If the habit(taste) shocks are not observed by the econometrician, they must be integrated out in the Euler equations. Our new approach assumes that the econometrician also observes the habit shock, and identifies it with the shock to one of the observed macro-economic variables. Thus, our model should be treated as a semi-structural model (based on a particular parametric specification of the habit process and a particular assumption on the identity of the habit shocks). While we do not provide a formal justification for our semi-structural specification, we motivate our choice by the following intuition.

There are several reasons why the aggregate labor income shock may be a sensible empirical proxy for the stochastic habit shock. First, a growing body of work in asset pricing literature (see, e.g., Jagannathan and Wang (1996) and Lettau and Ludvigson (2001a)) suggests that aggregate labor income risk is a relevant source of risk factor in explaining equity returns. Second, a growing body of work in macroeconomics literature (see, e.g., Calvo (1983) and Christiano, Eichenbaum, and Evans (2005)) suggests that both real marginal cost, which is closely related to the aggregate labor share, and wage rigidity are important in driving inflation and channeling the propagation of monetary shocks through the real economy. Both strands of literature suggest that labor income risk may play a non-trivial role in explaining asset returns, through either the cash flows (dividend and inflation risks) or the discount rates (pricing kernel). Since we take asset returns and endowment processes as given, the only place that the labor income risk can play a role is through the utility specification. In our framework, it can only enter the model through the shock to the habit stock. We directly estimate the sensitivity of the habit shock to the labor income shock (as

the volatility of the habit shock) from the Euler equations.⁹

3.3 Miscellaneous issues

A practical issue that arises in estimating internal habit formation models is that the *ex post* realization of the marginal utility of consumption, namely (see equation (6)),

$$\sum_{j=0}^{\infty} e^{-\rho j} a_{t,j} u'(s_{t+j}),$$

may be negative. The situation occurs when the current consumption level is very high relative to subsequent realized level of consumption, so that the future dis-utility induced by the current high level of consumption is very large. This can be caused by a number of reasons. First, the model could be mis-specified. Second, even if the model is correctly specified, the realized MUC may be negative if it is not evaluated at the true parameter values. Third, even if the model is correctly specified, and the realized MUC is evaluated at the true parameter values, there is always a possibility that in a finite sample, random noises may drive the realized MUC negative. Optimality restrictions on the equilibrium consumption process only requires that the *ex ante* marginal utility of consumption be non-negative.

We base our estimation and testing procedure on the assumption that the null is correctly specified. This means that we need to impose restrictions on the admissible parameter region so that the *ex post* MUC stays positive. Such restrictions are also helpful in practice in making the GMM objective function more robust and the GMM estimation easier to converge. Finally, we check that the *ex post* MUC is strictly positive at the converged parameter estimates and the parametric restrictions that ensure MUC positivity are not binding.

Another practical issue is that based on the observed data, there is no guarantee that the surplus consumption $z_t \equiv c_t - x_t$ is strictly positive, because a large realization of the habit shock can send x_t above c_t under at least some parameter configurations. Fortunately,

⁹More generally, the habit shock can be driven by both the consumption shock and the labor income shock. We will investigate this more general specification as well.

under realistic parameter values, this occurs only infrequently. Since this situation occurs at the tail end of the steady-state distribution, and most interesting economic behavior takes place at or near the long-run mean of the steady-state distribution, we can deal with this issue by designing a utility function $u(s)$ that coincides with a standard CRRA specification when s is sufficiently far away from 0, and is still well-behaved when s is close to zero or slightly negative.¹⁰

This is not an unknown issue that arises with additive habit models. It has been addressed in the literature in the different context. Detemple and Zapatero (1991) impose a nonlinear parameter restriction to insure the nonnegativity of marginal utility. In addition, Chapman (1998) constructs an example of an endowment economy with linear internal habit formation such that implies negative marginal utility of with probability one. His calibration exercise assumes a specific endowment process (lognormal diffusion, Eq. (6), pg. 1225), which we are agnostic about. He finds that marginal utility is negative for a CRRA utility function when relative risk aversion is set to a specific range of values and other endowment process parameters match the sample moments of standard aggregate consumption and returns data. Yogo (2008) solves this issue by proposing a power reference-dependent utility function in which the representative household has power gain-loss utility in the spirit of Tversky and Kahneman (1992). In this set up, the marginal utility is always positive even when consumption falls below its subsistence (reference) level, by virtue of imposing the absolute value function on the difference between consumption and its subsistence level. For us, it is more a practical than modeling issue as we abstract from the specific form of the endowment process in the economy.

Finally, in a finite sample, it is not possible to construct the *ex post* realization of MUC

¹⁰In our empirical implementation, we use a version of the following specification:

$$u(s) = \begin{cases} \frac{s^\gamma}{\gamma}, & \gamma < 0, \quad s > \underline{s} > 0, \\ \underline{u}(s), & s \leq \underline{s}, \end{cases} \quad (19)$$

where $\underline{u}(\cdot)$ is a function well-defined in $(-\infty, \underline{s}]$, strictly increasing and concave, with the properties that

$$\underline{u}(\underline{s}) = u(\underline{s}), \quad \underline{u}'(\underline{s}) = u'(\underline{s}), \quad \underline{u}''(\underline{s}) = u''(\underline{s}).$$

Such modifications to the standard power utility specification do not affect the derivation of the Euler equations. With a sensible choice of \underline{s} , they should not affect any substantive implications of the model either. Additional details about its implementation may be obtained from the authors upon request.

as an infinite sum. All of our estimation results are based on truncating the infinite sum to 50 terms, which corresponds to the assumption that a habit shock dies out before twelve and half years. This represents a sufficiently long period for the MUC to be affected by habit shock so long as the mean reversion parameter κ for the habit shock is not too small.

4 Empirical results

In this section, we report key empirical findings. We begin by describing the data. We then report estimation results and address the questions **Q1** and **Q2** raised in the introduction.

4.1 Data

Summary statistics of the macro variables and asset returns are reported in Table 1.

Consumption: We use quarterly consumption data to estimate our model because it contains less measurement errors than monthly consumption data.¹¹ Quarterly decision interval also allows us to focus on pure habit effect, as Heaton (1995) shows that local substitution is important only for decision intervals much shorter than a quarter. Consequently, we set $\delta = 0$ throughout our empirical estimations. Our sample period is from the fourth quarter of 1951 to the fourth quarter of 2002.

We measure aggregate consumption as expenditures on non-durables and services excluding shoes and clothing.¹² In order to distinguish between long-term habit persistence and short-term seasonality, we use seasonally adjusted data at annual rates, in billions of chain-weighted 2000 dollars.¹³ We define aggregate labor income as wages and salaries plus transfer payments plus other labor income minus personal contributions for social insurance minus taxes. Real aggregate labor income is obtained by deflating nominal aggregate labor income by the implicit chain-type price deflator (2000=100).

¹¹See Ferson and Constantinides (1991) and Heaton (1995) for more details.

¹²We exclude shoes and clothing from expenditures on non-durables because we would like to abstract from any durability effect, which is contained in these series. The exclusion of shoes and clothing follows the paper of Blinder, Grossman, and Wang (1985), p.473.

¹³Using seasonally unadjusted data, Ferson and Harvey (1992) find that quarterly seasonality may induce “quarterly” habit persistence, in the sense that the habit level is determined by consumption lagged four quarters. We wish to abstract away from this effect.

Real per capita consumption and labor income is obtained by dividing real aggregates by a measure of U.S. population. The latter is obtained by dividing real total disposable income by real per capita disposable income. Consumption, labor income, price deflator, and measure of population are obtained from NIPA (National Income and Product Account) tables.¹⁴

Asset returns: Nominal quarterly U.S. broad value-weighted equity index returns as well as three-month and 10-year Treasury Bond Portfolio returns are obtained from CRSP (Center for Research in Security Prices). Real asset returns are deflated by the implicit chain-type price deflator (2000=100).

Instruments: We use two sets of instruments. The first one is the standard set of lagged consumption growth and asset returns used in the previous studies to estimate the parameters and test the Euler equation model. One of the potential problems with consumption data is time aggregation, which means that consumption decisions are made more frequently than the observation interval, and thus, measured consumption is the sum of the expenditures of the interval. Ferson and Constantinides (1991) argue that time aggregation can induce a spurious correlation between the error terms and information set, thereby increasing the order of the MA process followed by u_t . Variables in the information set at time t , which were not in the information set at $t - 1$, may not be valid instruments for testing over-identifying restrictions (12). We nevertheless estimate our model using this set of instruments for comparative purposes. Another set of variables is described in Section 4.7.

Next, we need to obtain the habit shock v_t . We proxy v_t as the aggregate labor income shock obtained from the following bivariate process with 4 lags for $f_t = (c_t, w_t)'$:

$$f_t = A_0 + A_1 f_{t-1} + A_2 f_{t-2} + A_3 f_{t-3} + A_4 f_{t-4} + \Sigma u_t, \quad (20)$$

¹⁴Source: Bureau of Economic Analysis (<http://www.bea.gov>).

where c_t is the log real per capita consumption growth, w_t is the log real per capita labor income growth, $u_t \sim N(0, I)$. The residuals from the estimated VAR equation (20) are plotted in Figure 1.

To get a sense of what the habit stock looks like under reasonable parameterizations of the model, Figure 2 plots real consumption, habit stock, and surplus consumption based on the following baseline parameterization: $\rho = 0.01$, $\gamma = -2$, $\kappa = 0.028$, $b = 0.372$, $\delta = 0$, $\beta = 0$. These parameters correspond roughly to those calibrated by Dai (2003), so they are broadly consistent with some key moments of equity and bond returns.

4.2 Deterministic habit: risk premium

We begin by estimating the risk-aversion parameter γ for three models with no habit (NH), deterministic external habit (DEH), and deterministic internal habit (DIH).¹⁵ For the model with no habit, $b = \kappa = 0$. For habit models, we fix $b = 0.328$ and $\kappa = 0.072$, which correspond to a pair of values for $a = \kappa + b$ and b used by Constantinides (1990). Grishchenko (2010) estimates these parameters in the deterministic habit model set up and finds that they are broadly consistent with Constantinides (1990) values.¹⁶ For all three models, we fix $\rho = 0.01$. The instruments used for the estimation include two lags of consumption growth rates and two lags of equity returns. Estimation results are reported in Table 2. There are two panels in the table. The top panel reports results when the excess return on the value-weighted NYSE index is used to identify the risk aversion parameter. For NH model, the estimate for the relative risk aversion $1 - \gamma$ is 85.9148. Thus, we recover the equity premium puzzle. For DEH model, the coefficient $1 - \gamma$ drops down to 8.9498, indicating that the presence of consumption habit helps resolve the equity premium puzzle. For DIH model, the coefficient $1 - \gamma$ reduces further to 2.90. For all three models, the parameter γ is sharply identified, and the over-identifying restrictions implied by the equity premium, *alone*, are not rejected.¹⁷ At

¹⁵The parameter $1 - \gamma$ is literally the relative risk aversion coefficient in NH, which is the same as the standard CCAPM. In the habit formation models (DEH or DIH), the relative risk aversion coefficient is $1 - \gamma$ divided by the surplus consumption ratio. Throughout the paper, we use the terms *risk-aversion parameter* and *curvature of the utility function* interchangeably to describe the same parameter $1 - \gamma$ or γ .

¹⁶We also estimate the models using alternative values of b , κ , and ρ . The qualitative conclusions remain the same.

¹⁷The models are forced to fit $K = 1$ risk premium using $M = 5$ instruments (one, two lags of consumption growth rates, and two lags of equity returns). There is only one free parameter. Thus, the minimized GMM objective

this point, the data favor the DIH, with much lower χ^2 than either NH or DEH. The bottom panel reports estimation results when both the equity premium and the risk premium on a long-term 10-year Treasury bond are used under the same model specifications and same parameterization.¹⁸ The coefficients of the relative risk aversion decline as we proceed from NH to DEH, and then to DIH, and have a similar magnitude to Panel A’s estimates of RRA. Including a long-term bond represents an additional challenge for the models. Indeed, the p -values for the overall goodness-of-fit test decline for all three models. So, we reject all three models when a long-term bond is included in the estimation.¹⁹

The results reported in Table 2 confirm some well-known facts about the standard CCAPM and deterministic habit formation models. First, the equity premium can be fit easily if there is no constraint on the curvature of the utility function, namely, $1 - \gamma$. The equity premium puzzle arises because the relative risk aversion parameter required to fit the equity premium is too large relative to what is required to explain individual behavior (in terms of portfolio holdings and life-cycle patterns). Second, these models do not explain long-term bond risk premium. Adding long-term bond risk premium leads to the deterioration of the model fit. Intuitively, the real interest rate risk and the consumption growth rate are very weakly correlated in the data (about -0.025), and have the opposite sign (negative) from what a standard consumption-based model would predict (positive). Since deterministic habit merely amplifies the risk premium, and does not affect the sign of the risk premium, all three models (NH, DEH, and DIH) would predict the wrong sign for bond risk premium when $1 - \gamma$ is positive. This basic tension in CCAPM and deterministic habit formation models is even more apparent when the models are forced to explain short-term interest rate as well. We will now consider this case.

function has a chi-square distribution with $KM - 1 = 4$ degrees of freedom.

¹⁸We also estimate the models using different long-term bonds, such as 5-year and 30-year Treasury bonds. The results are qualitatively similar.

¹⁹The models are forced to fit $K = 2$ risk premiums using $M = 7$ instruments (one, two lags of consumption growth rates, two lags of equity returns and two lags of bond returns). There is only one free parameter. Thus, the minimized GMM objective function has a chi-square distribution with $KM - 1 = 13$ degrees of freedom.

4.3 Deterministic habit: risk-free rate and risk premium

In Table 3, we free up the subjective discount rate ρ in our three basic models (NH, DEH, DIH), and use the risk-free rate (3-month T-bill rate) to generate additional moment conditions to identify ρ . The models are still required to fit both the equity premium, $R_{t,t+1}^m - R_{t,t+1}^f$, and the 10-year bond risk premium, $R_{t,t+1}^{b,10} - R_{t,t+1}^f$, (so that the risk aversion parameter $1 - \gamma$ is identified). Two panels in Table 3 correspond exactly to the panels in Table 2, except that in the former, the risk-free rate also enters the Euler equation, and two lags of the risk-free rate serve as additional instruments.

The top panel shows that DIH is able to explain both the risk-free rate and the equity premium at reasonable values of ρ and γ : $\hat{\rho} = 0.04$ and $\hat{\gamma} = -0.93$. In contrast, NH and DEH models are rejected based on the conventional levels of p -values. In addition, the parameter values have the wrong signs. *A priori*, we expect $\rho > 0$ (agents are impatient) and $\gamma < 0$ (agents are more risk averse than log-utility). When γ is constrained to be non-positive in the estimation, GMM finds the corner solution $\gamma = 0$. The bottom panel shows that including the 10-year bond return leads to substantial deterioration of all three models. The point estimates are qualitatively similar to those in the top panel, indicating that the long-term bond return does not have much bite in the context of these models.

To understand this behavior, note that there are two basic forces that pull the risk aversion parameter in two different directions in these models. First, the negative correlation between the real consumption growth rate and the real risk-free rate can be explained only if $1 - \gamma$ is negative. Second, the high equity premium can be explained only if $1 - \gamma$ is positive and large. Both NH and DEH give up explaining the equity premium in order to reconcile with the negative correlation between the consumption growth and risk-free rate.

Since $0 < 1 - \gamma < 1$, the level of the risk-free rate can be fit only if the subjective discount rate is negative. To see this, note that, in NH, the unconditional mean of the Euler equation for the risk-free rate implies approximately (under the assumption that the consumption growth rate is i.i.d.) that $e^{-\rho/4} \times e^{-(1-\gamma) \times g/4 + \frac{(1-\gamma)^2}{2} \sigma^2/4} = e^{-r/4}$, or $\rho = r - (1-\gamma) \times g + \frac{(1-\gamma)^2}{2} \sigma^2$, where $g = 3.15\%$ is the average consumption growth rate (consumption plus service), $\sigma =$

1.28% is the volatility of the consumption growth rate, and $r = 1.4\%$ is the average risk-free rate (see Table 1 for the sample moments). For $\gamma = 0.1672$, $\rho = -0.006$. For DEH, the risk-free rate is priced by the growth rate of the surplus consumption ratio, which has the same mean growth rate of g , but a higher volatility: $\sigma \approx 1.28\% \times \frac{b+\kappa}{\kappa} = 7.11\%$. This implies $\rho = 1.6\%$. Alternatively, the estimated value of ρ implies that $r = 0.26\%$ for NH and $r = -2.35\%$ for DEH.

4.4 Stochastic habit: risk premium

Next, we explore the benefit of introducing a stochastic shock to the consumption habit by freeing up the parameter β . The resulting habit models are denoted SEH (for stochastic external habit) and SIH (for stochastic internal habit). Table 4 reports results when γ and β are jointly estimated using the excess return on a broad stock market index and the excess return on the 10-year Treasury bond. As a simple robustness check, we report results based on three different measures of equity market returns: value-weighted NYSE index return, value-weighted NYSE/AMEX index return, and value-weighted NYSE/AMEX/NASDAQ index return. The results are qualitatively the same. Thus, we will now focus on results based on value-weighted NYSE index return (Panel A).

For both models, the risk-aversion parameter γ is sharply identified and qualitatively similar to those reported in the Panel B of Table 2. SEH yields $\hat{\gamma} = -8.53$ whereas SIH produces $\hat{\gamma} = -2.91$. The habit volatility parameter β is not very precisely identified in SEH, but is sharply identified in SIH: $\hat{\beta}(SEH) = 0.0047$ with the standard error of 0.0496, $\hat{\beta}(SIH) = -0.0254$ with the standard error of 0.0116. We reject SEH at conventional confidence levels whereas we do not reject SIH. Both results indicate that introducing a stochastic shock in the external habit formation model does not help very much in explaining risk premium on both equity and long term bonds. In contrast, it has a very significant positive impact on the goodness-of-fit of the internal habit formation model.

4.5 Stochastic habit: risk-free rate and risk premium

Next, we free up the subjective discount factor ρ in addition to γ and β parameters. Table 5 reports the estimation results when ρ , γ , and β are jointly estimated, using the risk-free rate, the value-weighted NYSE index return, and the 10-year Treasury bond return. Comparing these results to those reported in Panel B of Table 3, we see a larger improvement for the goodness-of-fit for SIH than for SEH. For internal habit specifications we have: $\chi^2(DIH) = 50.26$ vs. $\chi^2(SIH) = 43.77$. However, for external habit specifications we have: $\chi^2(DEH) = 51.04$ vs. $\chi^2(SEH) = 46.86$. These differences are also reflected in the fact that the parameter β is not sharply identified in SEH, whereas it is sharply identified in SIH: $\hat{\beta}(SEH) = -0.2128$ with the standard error of 0.2340, $\hat{\beta}(SIH) = -0.0133$ with the standard error of 0.0040. In addition, SEH model does not seem to have an ability to identify the subjective discount rate: $\hat{\rho} = -0.0030$ with the standard error of 0.0127. However, the point estimate of the subjective discount rate in the SIH model is well identified, albeit low.

4.6 Discussion

The reported empirical results help answer the two questions raised in the introduction. Our results show that introducing a shock to the consumption habit allows the model to explain long-term bond returns better than deterministic habit formation models. Furthermore, the beneficial effect of a habit shock is much more evident in internal rather than in external habit formation models. To understand these results, we note that the habit shock may help explain asset returns because it can potentially modify the autocorrelation property of the marginal rate of substitution. In an external habit formation model, the habit shock affects the marginal rate of substitution' autocorrelation property only through the surplus consumption ratio dynamics. The reason is that the functional form of the MRS,

$$\text{MRS}_{t,t+1} = e^{-\rho} \frac{\text{MUC}_{t+1}}{\text{MUC}_t}, \text{MUC}_t = u'(z_t), z_t = c_t - x_t,$$

remains the same as that derived from a deterministic external habit formation model. Since the addition of a new shock does not materially affect the growth rate of the surplus

consumption ratio, the MRS continues to have difficulty reconciling the autocorrelation structure of long-term bond returns. In contrast, the presence of the habit shock materially affects the functional form of the internal MRS. To develop some intuition, we again appeal to the first-order Taylor expansion introduced in Section 2.2. Within the context of the model specification estimated here, the marginal utility of consumption for an internal habit formation model is given by

$$\text{MUC}_t \approx u'(z_t) + \Psi_t^{(b)} + \beta\lambda S_t^{-1},$$

where $\Psi_t^{(b)}$ captures the expected future dis-utility due to internal habit, and the last term captures the effect of habit shock on the marginal utility of consumption. Therefore, the presence of the habit shock alters the properties of the MRS in a fundamental way. In particular, it makes the MRS more persistent.

4.7 Additional tests and robustness checks

1. **Standard errors.** Standard errors reported above are based on two-stage estimation (linear projection and GMM as separate stages). This two-step procedure might affect inference. The results from one-stage estimation (linear projection and GMM in one step) indicate that while all standard errors are larger due to sampling noises in the linear projection coefficients, they do not qualitatively change the conclusions. In particular, parameter estimates in internal habit formation models are different from zero at the conventional levels of statistical significance.
2. **Instrumental variables.** We have also used other instruments to check the robustness of our results: unit vector, lagged values of consumption growth, a proxy for the log consumption-wealth ratio cay_t ,²⁰ “relative T-bill rate” ($RREL$, which is measured as the three-month T-bill rate minus its four-quarter moving average) and the lagged value of the excess return on the Standard&Poor 500 (SP&500) stock market index ($SPEX$)

²⁰ cay_t is measured as the cointegrating residual between log consumption, log asset wealth, and log labor income. See Lettau and Ludvigson (2001a) for more details.

over the three-month T-bill rate. Lettau and Ludvigson (2001a) find that cay_t has a strong predictive power for stock returns over one quarter to several year horizons. In their subsequent 2001b paper they show that this variable forecasts portfolio returns too. Campbell (1991), Hodrick (1992), and Lettau and Ludvigson (2001a) find that $RREL$ has a forecasting power for excess returns at a quarterly frequency.²¹ One of the problems associated with nonlinear GMM is the weak instruments problem, or the weak identification problem of the model parameter vector θ . While this issue can be identified more easily in the linear instrumental variables models, it is more difficult to deal with it in the present set up. Stock, Wright, and Yogo (2002) provide an excellent survey of this issue.²² They argue that one of the informal detection of the weak identification in the GMM setting can be done through addition or change of instrumental variables. Our results show that the choice of instruments does not affect results qualitatively, and therefore, it is unlikely that our results are driven by the presence of weak instruments in our set up.

5 Conclusion

In this paper, we econometrically estimate and test a consumption-based asset pricing model with stochastic habit formation using generalized method of moments. The key contribution to the asset pricing literature on habit formation is the study of the novel preference specification, namely, stochastic internal habit formation. The model departs from existing models with deterministic internal habit (e.g., Dunn and Singleton (1986), Ferson and Constantinides (1991), and Heaton (1995)), by introducing shocks to the coefficients in the distributed lag specification of consumption habit and consequently an additional shock to the marginal rate of substitution. Stochastic shocks to the consumption habit are persistent and provide an additional source of time-variation in expected returns. Economically, we interpret shocks as taste shocks in the economy. Empirically, we proxy these unobserved shocks

²¹We did not include other popular forecasting variables like dividend-price ratio into our instrumental set because they are found to be driven away by the above variables cay_t , $RREL$, and $SPEX$. See Lettau and Ludvigson (2001a) for further details.

²²See also Stock and Wright (2000) on the weak identification problem.

by aggregate labor income shocks. We show that stochastic internal habit formation models resolve the dichotomy between the autocorrelation properties of stochastic discount factor and those of expected returns, providing a better explanation of time-variation in expected equity and long-term bond returns than models with either deterministic or stochastic external habit. This evidence suggests that stochastic internal habit models should be preferred to external habit counterparts when studying the joint behavior of the aggregate stock and bond returns.

Appendix

A Derivation of the stochastic Euler equation

Consider an arbitrary security with price-dividend pair (p_t, d_t) . Suppose that, at t , the representative agent buys α share of the security and holds it for one period, she gives up αp_t of consumption at t , but receives additional consumption $\alpha(p_{t+1} + d_{t+1})$. Since no-trade is optimal for the representative agent, $\alpha = 0$ is the solution to

$$V_t = \max_{\alpha} E \left[u(\tilde{s}_t) + E \left[\sum_{j=1}^{\infty} e^{-\rho j} u(\tilde{s}_{t+j}) \middle| c_{t+1} + \alpha(p_{t+1} + d_{t+1}), \tilde{x}_{t+1} \right] \middle| c_t - \alpha p_t, x_t \right],$$

where \tilde{s}_t and \tilde{x}_t are the service flow and habit process defined over the consumption process \tilde{c}_t with

$$\begin{aligned} \tilde{c}_{\tau} &= c_{\tau} - \alpha p_{\tau}, \text{ if } \tau = t, \\ \tilde{c}_{\tau} &= c_{\tau} + \alpha(p_{\tau} + d_{\tau}), \text{ if } \tau = t + 1, \\ \tilde{c}_{\tau} &= c_{\tau}, \text{ otherwise.} \end{aligned}$$

The first order condition with respect to α , evaluated at $\alpha = 0$, gives

$$p_t \frac{\partial V_t^*}{\partial c_t} = E_t \left[e^{-\rho} (p_{t+1} + d_{t+1}) \frac{\partial V_{t+1}^*}{\partial c_{t+1}} \right],$$

where $V_t^* = E_t \left[\sum_{j=0}^{\infty} e^{-\rho j} u(s_{t+j}) \right]$. Rearranging terms, we obtain equation (5).

Evaluating the derivative of V_t^* with respect to c_t explicitly:

$$\text{MUC}_t = \mathbb{E}_t \left[\sum_{j=0}^{\infty} e^{-\rho j} u'(s_{t+j}) a_{t,j} \right], \quad (21)$$

where $a_{t,j} = \frac{\partial s_{t+j}}{\partial c_t} = 1$. Equation (6) is obtained by noting that

$$\begin{aligned} \frac{\partial s_{t+j}}{\partial c_t} &= \frac{\partial}{\partial c_t} \left(\frac{c_{t+j}}{1-\delta L} - \frac{x_{t+j}}{1-\delta L} \right) = \begin{cases} 1, & j=0, \\ \delta^j - \sum_{i=0}^{j-1} \delta^i \frac{\partial x_{t+j-i}}{\partial c_t}, & j>0, \end{cases} \\ \frac{\partial x_{t+j}}{\partial c_t} &= \begin{cases} b(1+\beta v_{t+1}), & j=1, \\ \frac{\partial x_{t+1}}{\partial c_t} \prod_{i=1}^{j-1} [1-\kappa-b(1+\beta v_{t+i+1})], & j>1. \end{cases} \end{aligned}$$

B Distributed lag with arbitrary weights

Suppose that $\delta = 0$, and $x_{t+1} = \sum_{j=0}^J b_j (1 + \beta v_{t+1-j}) z_{t-j}$, then $a_{t,0} = 1$, $a_{t,j} = -b_{t,j}$, $b_{t,j} \equiv \frac{\partial x_{t+j}}{\partial c_t}$ for $j \geq 1$. The time-varying coefficients $\{b_{t,j} : j \geq 1\}$ capture an extra unit of current consumption's marginal effect on future habit levels, and can be computed through the following recursion:

$$\begin{aligned} b_{t,1} &= \phi_{t,0}, \\ b_{t,k+1} &= \phi_{t+k,k} - \sum_{j=0}^{k-1} \phi_{t+k,j} b_{t,k-j}, \quad 1 \leq k \leq J, \\ b_{t,k+1} &= - \sum_{j=0}^J \phi_{t+k,j} b_{t,k-j}, \quad k > J. \end{aligned}$$

where $\phi_{t+k,j} = b_j(1 + \beta v_{t+1+k-j}) \in I_{t+1+k-j}$ for $0 \leq j \leq k$ and $0 \leq k \leq J$. Note that $\phi_{t+j,j} = b_j(1 + \beta v_{t+1}) \in I_{t+1}$ for $\forall j$. If $b_j = b(1 - \kappa)^j$, then the recursion simplifies to:

$$\begin{aligned} b_{t,1} &= \phi_{t,0}, \\ b_{t,k+1} &= b_{t,1} \prod_{j=1}^k (1 - \kappa - \phi_{t+j,0}), 1 \leq k \leq J, \\ b_{t,k+1} &= - \sum_{j=0}^J \phi_{t+k,j} b_{t,k-j}, k > J. \end{aligned}$$

If $J = \infty$, we do not need the second equation, and the first equation recovers the result for the geometric distributed lag specification.

References

- Abel, A. B., 1990, “Asset prices under habit formation and catching up with the Joneses,” *American Economic Review Papers and Proceedings*, 80, 38–42.
- Bekaert, G., E. Engstrom, and S. Grenadier, 2010, “Stock and bond pricing with moody investors,” *Journal of Empirical Finance*, 17, 867–894.
- Bekaert, G., E. Engstrom, and Y. Xing, 2009, “Risk, uncertainty, and asset prices,” *Journal of Financial Economics*, 91(1), 59–82.
- Blinder, A. S., A. D. S. Grossman, and J. Wang, 1985, “The time series consumption function revisited,” *Brookings Papers on Economic Activity*, 1985(2), 465–521.
- Brandt, M., and K. Wang, 2003, “Time-varying risk aversion and unexpected inflation,” *Journal of Monetary Economics*, 50(7), 1457–1498.
- Buraschi, A., and A. Jiltsov, 2007, “Habit formation and macroeconomic models of the term structure of interest rates,” *Journal of Finance*, 62(6), 3009–3063.
- Calvo, G. A., 1983, “Staggered prices in a utility-maximizing framework,” *Journal of Monetary Economics*, 12(3), 383–396.

- Campbell, J., and J. Cochrane, 1999, “By force of habit: A consumption-based explanation of aggregate stock market behavior,” *Journal of Political Economy*, 107(2), 205–251.
- Campbell, J., R. Shiller, and L. M. Viceira, 2009, “Understanding inflation-indexed bond markets,” *Brookings Papers on Economic Activity*, Spring, 79–120.
- Campbell, J. Y., 1986, “Bond and stock returns in a simple exchange model,” *The Quarterly Journal of Economics*, 101(4), 785–804.
- , 1991, “A variance decomposition for stock returns,” *Economic Journal*, 101(405), 157–179.
- Campbell, J. Y., and R. J. Shiller, 1991, “Yield spreads and interest rate movements: A bird’s eye view,” *Review of Economic Studies*, 58, 495–514.
- Chapman, D. A., 1998, “Notes and comments: Habit formation and aggregate consumption,” *Econometrica*, 66(5), 1223–1230.
- Chen, X., and S. C. Ludvigson, 2009, “Land of addicts? An empirical investigation of habit-based asset pricing models,” *Journal of Applied Econometrics*, 24(7), 1057–1093.
- Christiano, L. J., M. Eichenbaum, and C. L. Evans, 2005, “Nominal rigidities and the dynamic effects of a shock to monetary policy,” *Journal of Political economy*, 113(1), 1–45.
- Constantinides, G., 1990, “Habit formation: A resolution of the equity premium puzzle,” *Journal of Political Economy*, 98(3), 519–543.
- Dai, Q., 2003, “Term structure dynamics in a model with stochastic internal habit formation,” Mimeo, NYU.
- Dai, Q., A. Le, and K. J. Singleton, 2010, “Discrete-time affine term structure models with generalized market prices of risk,” *Review of Financial Studies*, 23(5).
- Detemple, J. B., and F. Zapatero, 1991, “Asset prices in an exchange economy with habit formation,” *Econometrica*, 59(6), 1633–1657.

- Duffie, D., and K. Singleton, 1993, "Simulated moments estimation of Markov models of asset prices," *Econometrica*, 61, 929–952.
- Dunn, K., and K. Singleton, 1983, "An empirical analysis of the pricing of mortgage backed securities," *Journal of Finance*, 36, 769–799.
- , 1986, "Modeling the term structure of interest rates under nonseparable utility and durability of goods," *Journal of Financial Economics*, 17, 27–55.
- Epstein, L., and S. Zin, 1989, "Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework," *Econometrica*, 57, 937–969.
- Ferson, W. E., and G. M. Constantinides, 1991, "Habit formation and durability in aggregate consumption: empirical tests," *Journal of Financial Economics*, 29, 199–240.
- Ferson, W. E., and C. R. Harvey, 1992, "Seasonality and consumption-based asset pricing," *Journal of Finance*, 47(2), 511–552.
- Gallant, A. R., and G. Tauchen, 1996, "Which moments to match?," *Econometric Theory*, 12, 657–681.
- Gourieroux, C., A. Monfort, and E. Renault, 1993, "Indirect inference," *Journal of Applied Econometrics*, 8, S85–S118.
- Grishchenko, O. V., 2010, "Internal vs external habit formation: The relative importance for asset pricing," *Journal of Economics and Business*, 62, 176–194.
- Hansen, L., 1982, "Large sample properties of generalized method of moments estimators," *Econometrica*, 50, 1029–1286.
- Heaton, J., 1995, "An empirical investigation of asset pricing with temporically dependent preference specifications," *Econometrica*, 63(3), 681–717.
- Heyerdahl-Larsen, C., 2010, "Asset prices and real exchange rates with deep habits," Working paper, London Business School.

- Hodrick, R. J., 1992, “Dividend yields and expected stock returns: Alternative procedures for inference and measurement,” *Review of Financial Studies*, 5, 357–386.
- Jagannathan, R., and W. Wang, 1996, “The conditional CAPM and the cross-section of expected returns,” *Journal of Finance*, 51(1), 3–53.
- Lettau, M., and S. Ludvigson, 2001a, “Consumption, aggregate wealth, and expected stock returns,” *Journal of Finance*, 56(3), 815–849.
- , 2001b, “Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying,” *Journal of Political Economy*, 109(6), 1238–1287.
- Ljungqvist, L., and H. Uhlig, 2000, “Tax policy and aggregate demand management under catching up with the Joneses,” *The American Economic Review*, 90(3), 356–366.
- Lucas, R., 1978, “Asset prices in an exchange economy,” *Econometrica*, 46, 1429–1445.
- Mehra, R., and E. C. Prescott, 1985, “The equity premium: A puzzle,” *Journal of Monetary Economics*, 15, 145–161.
- Modigliani, F., and R. Sutch, 1966, “Innovations in interest rate policy,” *The American Economic Review*, 56(1/2), 178–197.
- Normandin, M., and P. St-Amour, 1998, “Substitution, risk aversion, taste shocks and equity premia,” *Journal of Applied Econometrics*, 13(3), 265–281.
- Santos, T., and P. Veronesi, 2010, “Habit formation, the cross section of stock returns and the cash flow risk puzzle,” *Journal of Financial Economics*, 98(2), 385–413.
- Singleton, K. J., 1993, “Econometric implications of consumption-based asset pricing models,” in *Advances in Econometrics, Sixth World Congress*, ed. by J. J. Laffont, and C. A. Sims. Cambridge University Press.
- Stock, J. H., and J. H. Wright, 2000, “GMM with weak identification,” *Econometrica*, 68, 1055–1096.

- Stock, J. H., J. H. Wright, and M. Yogo, 2002, “A survey of weak instruments and weak identification in generalized method of moments,” *Journal of Business and Economic Statistics*, 20(4).
- Sundaresan, S. M., 1989, “Intertemporally dependent preferences and the volatility of consumption and wealth,” *Review of Financial Studies*, 2, 73–88.
- Tversky, A., and D. Kahneman, 1992, “Advances in prospect theory: Cumulative representation of uncertainty,” *Journal of Risk and Uncertainty*, 5, 297–323.
- Van Binsbergen, J. H., 2007, “Deep habits and the cross section of expected returns,” SSRN eLibrary, <http://ssrn.com/paper=1101456>.
- Wachter, J., 2006, “A consumption-based model of the term structure of interest rates,” *Journal of Financial Economics*, 79(2), 365–399.
- Yogo, M., 2008, “Asset prices under habit formation and reference-dependent preferences,” *Journal of Business and Economic Statistics*, 26(2).

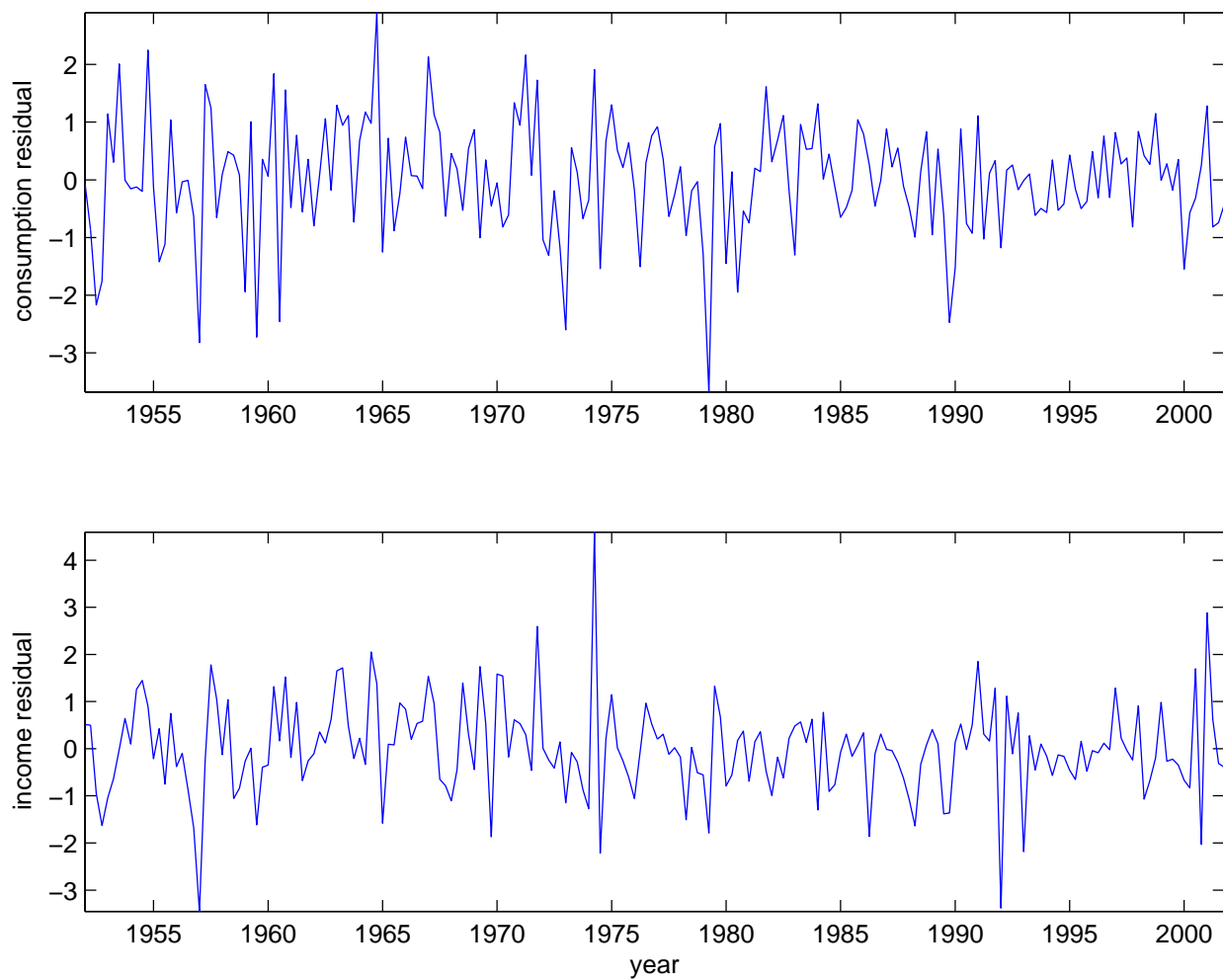


Figure 1: VAR residuals.

This figure presents residuals from the estimated Vector Auto Regression Model (20). Consumption and labor income data are measured as log real per capita consumption growth and log real labor income growth. The sample period is from 1952:Q1 to 2002:Q4, quarterly frequency.

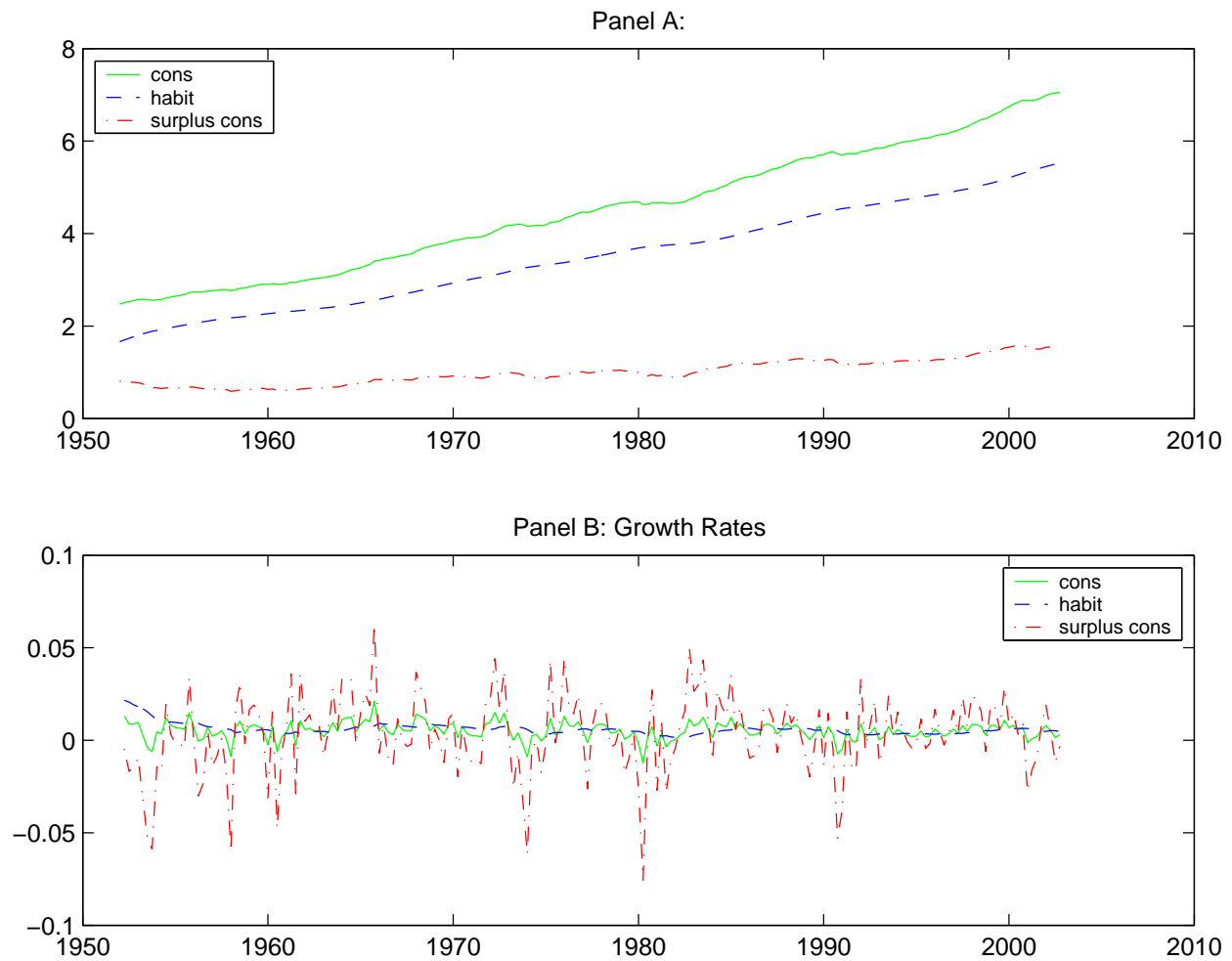


Figure 2: Consumption, habit stock, and surplus consumption.

Panel A represents consumption, habit stock, and surplus consumption level time series for model parameters $\rho = 0.01$, $\gamma = -2$, $\kappa = 0.072$, $b = 0.328$, $\beta = 0$. Panel B presents consumption, habit stock, and surplus consumption growth rates. The sample period is from 1952:Q1 to 2002:Q4, quarterly frequency.

Table 1: Descriptive statistics

This table reports means, standard deviations and the first five autocorrelations of the data. Panel A reports the statistics of aggregate seasonally-adjusted (SA) real quarterly consumption series and aggregate labor income. NDS refers to nondurable consumption plus services. Panel B reports the statistics of the real quarterly returns on value-weighted portfolios on NYSE, NYSE and AMEX, and NYSE, AMEX, and NASDAQ exchanges. Panel C reports the statistics of the real quarterly returns on Treasury bond indices. Nominal returns are converted into real returns by dividing them by one plus the growth rate of the seasonally unadjusted Consumer Price Index. The sample period is from 1947:Q1 to 2002:Q4, quarterly frequency.

	Mean	Std. Dev.	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
Panel A: Macroeconomic variables statistics							
Nondurables plus services, SA	0.0315	0.0128	0.3125	0.2017	0.2057	0.0717	0.0269
Nondurables, SA	0.0195	0.0166	0.1234	0.2140	0.1038	0.0068	0.0747
Services, SA	0.0428	0.0148	0.4344	0.2835	0.3426	0.1293	0.0446
Real labor income growth	0.0333	0.0254	-0.0835	0.0786	0.0449	-0.0389	0.0090
Panel B: Equity index returns							
VWR - NYSE	0.0703	0.1574	0.0685	-0.0640	-0.0113	0.0192	-0.0030
VWR - NYSE/AMEX	0.0699	0.1585	0.0658	-0.0621	-0.0137	0.0159	-0.0025
VWR - NYSE/AMEX/NASDAQ	0.0683	0.1658	0.0630	-0.0566	0.0061	0.0051	-0.0057
Panel C: Treasury bond returns							
3-month T-bill	0.0140	0.0180	0.4646	0.3904	0.4388	0.3756	0.1937
10-year T-bond	0.0183	0.0819	0.0719	0.0335	0.1296	0.0744	-0.2176

Table 2: Deterministic habit: Estimation of γ

This table reports GMM estimated coefficients of the preference parameter γ in the deterministic habit ($\beta = 0$) formation case. Parameters $b = 0.328$, $\kappa = 0.072$, and $\rho = 0.01$ are fixed in this estimation. The following Euler equations are estimated:

$$E_t \left[\text{MRS}_{t,t+1} (R_{t,t+1}^m - R_{t,t+1}^f) \right] = 0,$$

$$E_t \left[\text{MRS}_{t,t+1} (R_{t,t+1}^{b,10} - R_{t,t+1}^f) \right] = 0,$$

where $R_{t,t+1}^f$ is the three-month T-bill rate (known at t), $R_{t,t+1}^m$ is the quarterly holding-period return on the value-weighted NYSE equity market index, $R_{t,t+1}^{b,10}$ is the quarterly holding-period return on a 10-year Treasury bond. Instruments are unit vector, one- and two-period lagged consumption growth rate, and one- and two-period lagged asset returns. NH stands for the base-line model without habit formation; DEH stands for the deterministic external habit formation model; DIH stands for the deterministic internal habit formation model. TJ_T is the overall goodness-of-fit statistic that has $\chi^2(DF)$ asymptotic distribution, where DF is the number of degrees of freedom specific to each panel. p -value is the probability value that TJ_T exceeds the minimized sample value of the GMM criterion function. The sample period is from 1947:Q1 to 2002:Q4, quarterly frequency.

Model	$\hat{\gamma}$	s.e. $\hat{\gamma}$	TJ_T	p -value
Panel A: No long-term T-Bond included, DF = 4				
NH	-84.9148	2.3126	7.9656	0.0928
DEH	-7.9498	1.0908	8.7061	0.0689
DIH	-1.8994	0.0438	4.9215	0.2954
Panel B: 10-year T-Bond included, DF = 13				
NH	-139.7899	1.1919	23.7577	0.0334
DEH	-7.8087	0.8355	32.6927	0.0019
DIH	-2.0438	0.0353	28.3454	0.0081

Table 3: Deterministic habit: Joint estimation of ρ and γ

This table reports GMM estimated coefficients of the preference parameters ρ and γ in the deterministic habit ($\beta = 0$) formation case. Parameters $b = 0.328$ and $\kappa = 0.072$ are fixed in this estimation. The following Euler equations are estimated:

$$E_t [\text{MRS}_{t,t+1}(1 + R_{t,t+1})] = 1.$$

Panel A reports estimation results with $R_{t,t+1} = (R_{t,t+1}^f, R_{t,t+1}^m)$, where $R_{t,t+1}^f$ is the three-month T-bill rate (known at t), $R_{t,t+1}^m$ is the quarterly holding-period return on the value-weighted NYSE equity market index. Panel B reports estimation results with $R_{t,t+1} = (R_{t,t+1}^f, R_{t,t+1}^m, R_{t,t+1}^{b,10})$, where $R_{t,t+1}^{b,10}$ is the quarterly holding-period return on a 10-year Treasury bond. Instruments are one, one- and two-period lagged consumption growth rate, and one- and two-period lagged asset and bond returns. (When γ is constrained to be negative, the GMM estimation finds the corner solution ($\gamma = 0$)). TJ_T is the overall goodness-of-fit statistic that has $\chi^2(DF)$ asymptotic distribution, where DF is the number of degrees of freedom specific to each panel. p -value is the probability value that TJ_T exceeds the minimized sample value of the GMM criterion function. The sample period is from 1947:Q1 to 2002:Q4, quarterly frequency.

Model	$\hat{\rho}$	s.e. $\hat{\rho}$	$\hat{\gamma}$	s.e. $\hat{\gamma}$	TJ_T	p-value
Panel A: $R_{t,t+1} = (R_{t,t+1}^f, R_{t,t+1}^m)$, $DF = 12$						
NH	-0.0215	0.0069	0.1672	0.1766	29.4625	0.0034
DEH	-0.0179	0.0032	0.8294	0.0331	30.0297	0.0028
DIH	0.0409	0.0466	-0.9345	0.4712	23.4604	0.0241
Panel B: $R_{t,t+1} = (R_{t,t+1}^f, R_{t,t+1}^m, R_{t,t+1}^{b,10})$, $DF = 25$						
NH	-0.0208	0.0054	0.1534	0.1344	46.7029	0.0053
DEH	-0.0094	0.0059	0.6467	0.0839	51.0445	0.0016
DIH	0.0534	0.0393	-1.0247	0.3535	50.2571	0.0020

Table 4: Stochastic habit: Joint estimation of γ and β

This table reports GMM estimated coefficients of the preference parameter γ and the volatility of the habit shock β . Parameters $b = 0.328$, $\kappa = 0.072$, and $\rho = 0.01$ are fixed in this estimation. The following Euler equations are estimated:

$$E_t \left[\text{MRS}_{t,t+1} (R_{t,t+1}^m - R_{t,t+1}^f) \right] = 0,$$

$$E_t \left[\text{MRS}_{t,t+1} (R_{t,t+1}^{b,10} - R_{t,t+1}^f) \right] = 0,$$

where $R_{t,t+1}^f$ is the three-month T-bill rate (known at t), $R_{t,t+1}^m$ is the quarterly holding-period return on the value-weighted NYSE equity market index, $R_{t,t+1}^{b,10}$ is the quarterly holding-period return on a 10-year Treasury bond. Instruments are one, one- and two-period lagged consumption growth rate, and one- and two-period lagged asset and bond returns. SEH stands for stochastic extern habit, and SIH stands for stochastic internal habit. TJ_T is the overall goodness-of-fit statistic that has $\chi^2(12)$ asymptotic distribution. p -value is the probability value that TJ_T exceeds the minimized sample value of the GMM criterion function. The sample period is from 1947:Q1 to 2002:Q4, quarterly frequency.

Model	$\hat{\gamma}$	s.e. $\hat{\gamma}$	$\hat{\beta}$	s.e. $\hat{\beta}$	TJ_T	p-value
Panel A: Value-weighted NYSE index						
SEH	-8.5344	6.7888	0.0047	0.0496	31.6562	0.0016
SIH	-2.9065	0.4839	-0.0254	0.0116	23.9664	0.0206
Panel B: Value-weighted NYSE/AMEX index						
SEH	-8.4483	6.7537	0.0036	0.0503	31.7751	0.0015
SIH	-2.7955	0.4528	-0.0279	0.0112	23.4363	0.0242
Panel C: Value-weighted NYSE/AMEX/NASDAQ index						
SEH	-8.3294	6.6915	0.0044	0.0500	31.3037	0.0018
SIH	-1.9477	0.2015	-0.0566	0.0075	17.9489	0.1173

Table 5: Stochastic habit: Joint estimation of ρ , γ , and β

This table reports GMM estimated coefficients of the preference parameters ρ , γ , and the volatility of the habit shock β . Parameters $b = 0.328$ and $\kappa = 0.072$ are fixed in this estimation. The following Euler equations are estimated:

$$\begin{aligned} E_t \left[\text{MRS}_{t,t+1}(1 + R_{t,t+1}^f) \right] &= 0, \\ E_t \left[\text{MRS}_{t,t+1}(1 + R_{t,t+1}^m) \right] &= 0, \\ E_t \left[\text{MRS}_{t,t+1}(1 + R_{t,t+1}^{b,10}) \right] &= 0, \end{aligned}$$

where $R_{t,t+1}^f$ is the three-month T-bill rate (known at t), $R_{t,t+1}^m$ is the quarterly holding-period return on the value-weighted NYSE equity market index, $R_{t,t+1}^{b,10}$ is the quarterly holding-period return on a 10-year Treasury bond. Instruments are one, one- and two-period lagged consumption growth rate, and one- and two-period lagged asset and bond returns. SEH stands for stochastic external habit, and SIH stands for stochastic internal habit. TJ_T is the overall goodness-of-fit statistic that has $\chi^2(24)$ asymptotic distribution. p -value is the probability value that TJ_T exceeds the minimized sample value of the GMM criterion function. The sample period is from 1947:Q1 to 2002:Q4, quarterly frequency.

Model	$\hat{\rho}$	s.e. $\hat{\rho}$	$\hat{\gamma}$	s.e. $\hat{\gamma}$	$\hat{\beta}$	s.e. $\hat{\beta}$	TJ_T	p-value
SEH	-0.0030	0.0127	0.7721	0.1816	-0.2118	0.2340	46.8577	0.0035
SIH	0.3204	0.0467	-4.3254	0.6642	-0.0133	0.0040	43.7737	0.0081