Specification Tests of Habit Formation *

Peng Liu a and Yu Ren b †

Campbell and Cochrane (1999) propose the habit formation model to explain the equity premium puzzle. They assume that an agent’s consumption is affected by habit and describe how habit adjusts to the history of consumption. We use the simulated moment method to test these two specifications. Empirically, we find that habit plays an important role in an agent’s consumption choice, however not in the way Campbell and Cochrane (1999) specify.

JEL classification: C13; G12; E21

*Ren’s research was supported by the Natural Science Foundation of Fujian Province, China ( # 2011J01384) and the Natural Science Foundation of China ( # 71131008 and # 70971113).
†Corresponding author. Email: renxmu@gmail.com.

a Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen, Fujian, 361005, China.
b Wang Yanan Institute for Studies in Economics, MOE Key Lab of Econometrics, and Fujian Key Lab of Statistical Sciences, Xiamen University, Xiamen, Fujian, 361005, China.
I Introduction

Campbell and Cochrane (1999) propose the habit formation model to explain the equity premium puzzle. This is widely recognized as a successful method to understand the equity premium puzzle. However, empirical evidence of the ability of the habit formation model to successfully explain the equity premium puzzle is ambiguous in the literature. Dunn and Singleton (1986), Eichenbaum, Hansen and Singleton (1988), and Heaton (1995) find very little evidence of habit formation in monthly U.S. aggregate data. In contrast, Ferson and Constantinides (1991) find habit formation statistically significant in monthly, quarterly and annual U.S. aggregate data. In addition, some researchers discuss the habit formation model by using panel data, including Heien and Durham (1991), Dynan (2000), and Carrasco, Labeaga Azcona and Lopez-Salido (2005). They find the evidence of habit formation for consumption.

More recently, Bansal, Gallant and Tauchen (2007) use the simulated moment method proposed by Smith (1993) to estimate and test the habit formation model. They find that the habit formation model does well on the overidentification test. However, this test does not provide evidence on the mechanism of the habit formation model, particularly the specification of the habit. As Abel (1990) mentions, an agent may evaluate the ratio between current consumption and habit instead of consumption surplus over habit. In order to fully support the habit formation model, we need to know whether habit affects consumption and whether the motion of habit is well specified. Unfortunately, these two questions are not addressed in Bansal et al. (2007).

In this paper, we try to answer these two questions. We modify the habit formation model by letting the coefficient of habit be an additional free parameter while keeping the other settings unchanged (the motion of the habit is adjusted accordingly). Then, we use the simulated moment method to estimate the model. There are two main advantages to this. First, we can study whether an agent’s utility is affected by the habit by testing the significance of that coefficient. If the coefficient is insignificant, then we can infer that
habit does not affect consumption at all, which would be evidence against the foundation of the habit formation model. Second, we can test whether the coefficient is equal to 1. If the motion of habit is well specified in Campbell and Cochrane (1999), then we cannot reject the null hypothesis that the coefficient is equal to 1. So adding this extra parameter into the original habit formation model sheds light on the model specification of habit.

Empirically, we collect annual data from 1929 to 2010 and find that the coefficient of the habit level is significant, but not equal to 1. This finding justifies the effect of habit in utility, but seriously challenges the habit specification in Campbell and Cochrane (1999). This implies that we should either try other channels to study how habit affects an agent’s consumption or use a different motion to capture habit.

The rest of this paper is organized as follows: Section 2 describes our estimation model. Section 3 presents the empirical analysis. Section 4 concludes.

II Model

Generalized habit formation model

We generalize the habit formation model by adding one free coefficient, $a$, into the original utility specification as follows:

$$E \sum_{t=0}^{\infty} \delta^t \left( C_t - aX_t \right)^{1-\gamma} - 1 \over 1 - \gamma, \quad (1)$$

where $C_t$ denotes current consumption, $X_t$ denotes the habit level, $\gamma$ is the risk aversion coefficient and $\delta$ is the subjective time discount factor.

In this way, the habit formation model is a special case of the generalized model when $a = 1$. In addition, the traditional power utility model is a special case of the generalized model when $a = 0$.

Accordingly, we let $\tilde{S}_t \equiv C_t - aX_t$ and $\tilde{s}_t \equiv \ln \left( \frac{C_t - aX_t}{C_t} \right) \approx a \left( -\frac{X_t}{C_t} \right) \approx as_t$, where $s_t$ is defined
as in Campbell and Cochrane (1999),

\[ s_{t+1} = (1 - \phi)\bar{s} + \phi s_t + \lambda(s_t)(c_{t+1} - c_t - g). \tag{2} \]

Then we have

\[ \tilde{s}_{t+1} = (1 - \phi)a\bar{s} + a\phi s_t + a\lambda(s_t)(c_{t+1} - c_t - g). \tag{3} \]

As defined in Campbell and Cochrane (1999), \( \lambda(s_t) \) is the sensitivity function, which is chosen as

\[
\lambda(s_t) = \begin{cases} 
\frac{1}{S} \sqrt{1 - 2(s_t - \bar{s})} - 1, & s_t \leq s_{max}; \\
0, & s_t \geq s_{max},
\end{cases}
\tag{4}
\]

where

\[ \bar{S} = \phi \sqrt{1 - \phi}, \tag{5} \]

and

\[ s_{max} = \bar{s} + \frac{1}{2}(1 - \bar{S}^2). \tag{6} \]

\( s_{max} \) denotes the point where the sensitivity function touches the vertical axis for the first time. Specified in this way, the sensitivity function implies that identical agents evaluate the current economic situation in comparison to their previous consumption. A high previous consumption implies a high consumption surplus ratio, which in turn implies a low sensitivity for the identical agents. In this case, identical agents believe the current economy to be in a good condition. They are consequently optimistic and perform less cautiously, lowering their previously cautious level of savings.

In order to check whether habit affects an agent’s utility and, hence, consumption choice, we test the significance of \( a \). If the coefficient is significant, then we can infer that habit affects consumption, which would be evidence consistent with the foundation of the habit formation model.

In addition, by testing the null hypothesis that \( a = 1 \), we can understand whether the
model specification in Campbell and Cochrane (1999) is appropriate. We should emphasize that in order to support the model specification of Campbell and Cochrane (1999), the true value of $a$ must be equal to 1. If not, there are two potential remedies. The first approach is to define $X_{1,t} = aX_t$, and follow the same model setup, including the description of the motion of habit, and the same argument in Campbell and Cochrane (1999). But, if we let the coefficient of $X_{1,t}$ be an extra parameter and repeat the estimation process, we get the same result, namely the coefficient of $X_{1,t}$ being equal to $a$. After that, we define $X_{2,t} = aX_{1,t} = a^2X_t$, and so on and so forth. If the process is repeated $i$ times, $X_{i,t} = a^iX_t$. Eventually, $X_{i,t}$ either goes to 0 or infinity when $i$ is large, which contradicts the original model setup. The second approach is to use Equation 3 to describe the motion of $X_t$. However, the sensitivity function then violates the condition that the risk-free rate is constant. Therefore, the rejection of the null hypothesis that $a = 1$ would seriously challenge the habit formation model.

Under the utility specification of Equation (1), the inter-temporal marginal rate of substitution, or stochastic discount factor (SDF), is

$$M_{t+1} = \delta \left( \frac{\tilde{S}_{t+1} C_{t+1}}{\tilde{S}_t C_t} \right)^{-\gamma}. \quad (7)$$

After some rearrangements,

$$M_{t+1} = \delta \exp\left[ -\gamma (\Delta \tilde{s}_{t+1} + \Delta c_{t+1}) \right],$$

where $\Delta \tilde{s}_{t+1}$ and $\Delta c_{t+1}$ denote the growth rates of consumption surplus ratio and consumption, respectively.

Thus, we get the Euler equation

$$\frac{P_{ct}}{C_t} = E_t \{ \delta \exp[-\gamma (\Delta \tilde{s}_{t+1} + \Delta c_{t+1})] (1 + \frac{P_{ct+1}}{C_{t+1}}) \exp(\Delta c_{t+1}) \}, \quad (8)$$

\[1\text{See Campbell and Cochrane (1999), Equation 8.}\]
and, accordingly, the expectation equation for the price-dividend ratio $\frac{P_t}{D_t}$ for the asset that pays dividend $D_t$ is

$$\frac{P_t}{D_t} = E_t\{\delta \exp[-\gamma(\Delta s_{t+1} + \Delta c_{t+1})](1 + \frac{P_{t+1}}{D_{t+1}})\exp(\Delta d_{t+1})\}. \quad (9)$$

For clarity, we summarize the parameters to be estimated into an $8 \times 1$ column vector:

$\beta \equiv (g, \sigma, \sigma_w, \rho, \phi, \gamma, \delta, \alpha)'$.

**Estimation method**

Following Bansal et al. (2007), we employ the method proposed by Smith (1993), which elegantly handles the problems with estimating the model of interest. Basically, this method minimizes the difference between the observed data and the data simulated from the model to be estimated. The detailed process is as follows.

The model to be estimated is required to be represented by a density or conditional density process. In other words, it is required to be able to be simulated given the value of the parameter. We have

$$\{p_1(x_{1|\beta}), \{p_t(y_{t|x_t, \beta})\}_{t=1}^\infty\}_\beta \in \mathbb{R} \quad (10)$$

where $x_t$, and $y_t$ denote independent and dependent variables, respectively, $\beta$ denotes the parameter to be estimated, and $p_t(\cdot)$ denotes the conditional density of the model to be estimated. Dependent variables can be auto-dependent and include unobservable variables.

We assume an auxiliary model

$$\{f_1(x_{1|\theta}), \{f_t(y_{t|x_t, \theta})\}_{t=1}^\infty\}_{\theta \in \Theta} \quad (11)$$

where $x_t$, and $y_t$ are well measured and recorded and $\theta$ is the econometric parameter
vector. Here \( f_t(\cdot) \) is the conditional density of the auxiliary model. The requirement of the auxiliary model to ensure the asymptotical consistency of the estimator is called smooth embedment\(^2\).

Parameter \( \theta \) can be estimated by quasi-maximum likelihood estimation (QMLE) from historical data:
\[
\hat{\theta} \equiv \text{argmax}_\theta L\{(y, x), \theta\},
\] (12)
where \( L \) is the likelihood function of the auxiliary model.

Given \( \beta \), \((\hat{x}_t, \hat{y}_t)_{t=1}^N\) can be simulated from Equation 10. Then
\[
s_n(\beta) \equiv m_n'(\beta, \hat{\theta})\hat{\Gamma}^{-1}m_n(\beta, \hat{\theta}),
\] (13)
\[
m_n(\beta, \hat{\theta}) = \frac{1}{N} \sum_{t=1}^N \frac{\partial}{\partial \theta} \ln f(\hat{y}_t|\hat{x}_t, \hat{\theta}),
\] (14)
where \( n \) is the number of data points observed, \( N \) is the number of data points simulated and \( m_n \) denotes the moment conditions. The weighting matrix is written as
\[
\hat{\Gamma} = \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial}{\partial \theta} \ln f(y_t|x_t, \hat{\theta}) \right] \left[ \frac{\partial}{\partial \theta} \ln f(y_t|x_t, \hat{\theta}) \right]'.
\] (15)

Then we find the estimator of \( \beta \) which is
\[
\hat{\beta} \equiv \text{arg min}_\beta s_n(\beta).
\] (16)

According to Gallant and Tauchen (1996), statistic \( n \cdot s(\hat{\beta}) \) converges to a chi-square distribution:
\[
n \cdot s(\hat{\beta}) \xrightarrow{d} \chi^2_{\text{dim}(\theta) - \text{dim}(\beta)}.
\] (17)

where \( \text{dim}(\theta) \) and \( \text{dim}(\beta) \) are dimensions of \( \theta \) and \( \beta \), respectively.

\(^2\)Simply, smooth embedment requires \( \theta(\beta) \), function of \( \theta \), economic parameter vector, with respect to \( \beta \), parameter vector to be estimated, to be \( C^1 \) in a nearby area of \( \beta_0 \), the true value of \( \beta \). The formal definition can be found in Gallant and Tauchen (1996), Definition 1.
Under regular conditions, we have

\[ \sqrt{n}(\hat{\beta}_n - \beta_0) \xrightarrow{d} N\{0, [(M_n^0)'(I_n^0)^{-1}(M_n^0)]^{-1}\} \]

\[ \hat{M}_n \xrightarrow{a.s.} M_0^0 \]

where \( \hat{M}_n = M_n(\hat{\beta}_n, \hat{\theta}_n) \), \( M_0^0 = M_n(\beta_0, \theta_0) \),

and \( M_n(\beta, \theta) = \frac{\partial}{\partial \beta'} m_n(\beta, \theta). \)

Hence we can test the hypothesis that \( H_0 : R\beta_0 = r \) with \( H_1 : R\beta_0 \neq r \) as

\[ n(R\hat{\beta} - r)'\{R[(\hat{M}_n)'(\hat{I}_n)^{-1}(\hat{M}_n)]^{-1}R'\}^{-1}(R\hat{\beta} - r) \xrightarrow{d} \chi^2_k, \] (18)

where \( k \) is the number of constraints. Specifically, the hypotheses are \( H_0 : a_0 = 1 \) with \( H_1 : a_0 \neq 1 \) and \( H_0 : a_0 = 0 \) with \( H_1 : a_0 \neq 0 \).

**Auxiliary VAR model**

We assume an auxiliary model

\[ y_t = b + By_{t-1} + e_t, \quad y_t = \begin{pmatrix} d_t - c_t \\ c_t - c_{t-12} \\ p_{d,t} - d_t \end{pmatrix}, \] (19)

where \( p_{d,t} \) is the per capita stock market value. Hence, the parameter vector of the auxiliary model is \( \theta = [b', vec(B)', vech(var(e_t))']' \).

The variables in the auxiliary model have two parts. The first part has the observable variables in the original model. The second part has the observable variables which are implicitly connected with the unobservable ones. Hence, the observable variables in the habit formation model, \( c_t \) and \( d_t \), should be included in the auxiliary model. To find the observable variables connected with unobservable variable \( X_t \), we rely on Euler Equation
From this, it can be seen that the evolving path of $\frac{P_{d,t}}{D_t}$ is determined by $d_t$ and $s_t$ jointly, which means the information of the unobservable variable $X_t$ is provided by the information of $\frac{P_{d,t}}{D_t}$. So $\frac{P_{d,t}}{D_t}$ should also be included in the auxiliary model.

From SDF Equation [7], the interest rate contains information of habit for the same reason as the price-dividend ratio. Although the auxiliary model should contain as much information of the unobservable variables as possible, we do not include interest rate in the VAR system, for both theoretical and empirical reasons. The first reason lies in the following equation

$$r_{t+1} = \log\left[\frac{P_{t+1} + D_{t+1}}{P_t}\right] = \log\left[\frac{P_{t+1}/D_{t+1} + 1}{P_t/D_t}\right].$$  \hspace{1cm} (20)

It is reasonable to assume that the marginal information from including interest rate into the auxiliary model, which already has $\frac{P_{d,t}}{D_t}$ and $d_t$ in the system, is rather small. Relatively, the cost incurred from including interest rate is high, since the number of parameters to be estimated increases dramatically with extra variables added in the system. The second reason is from our rudimentary estimation with a VAR system. We estimate the system including the auto-regression of the interest rate and find that the R-square of that equation to be near zero.

However Euler Equation [9] cannot provide an explicit form of $\frac{P_{d,t}}{D_t}$ as a function of $d_t$ and $s_t$. Instead, we follow Gallant and Tauchen (1996), and employ the Bubnov-Galerkin method to find a numerical series of $\frac{P_{d,t}}{D_t}$. This method treats $\frac{P_{d,t}}{D_t}$ as a quadratic function of state variables

$$\frac{P_{d,t}}{D_t} = \frac{P_d}{D}(u_t),$$  \hspace{1cm} (21)

where $u_t$ is
$u_t = \begin{pmatrix} d_t \\ s_t \\ \lambda(s_t) \end{pmatrix}$. \hspace{1cm} (22)

### III Empirical Results

We use a U.S. annual data series from 1929 to 2010 to estimate the model and conduct the hypotheses tests. All the data are collected from the same data source and have applied the same treatment as mentioned in Bansal et al. (2007).

#### Auxiliary model

MLE is employed to estimate VAR system Equation 19. We first use an unrestricted $3 \times 3$ VAR(1). However, many off-diagonal element estimates of $B$ are statistically insignificant, and the diagonal ones clearly dominate. Thus it is reasonable to assume a diagonal coefficient matrix. The estimation result is shown in Table 1. The estimators of the constant and slope coefficients are all statistically significant except for the constant of the third equation. In addition, the R-squares are acceptable.

#### Table 1: Estimation of the Auxiliary Model

<table>
<thead>
<tr>
<th></th>
<th>Const.</th>
<th>Slope</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_t - c_t$</td>
<td>-0.5266*</td>
<td>0.8481*</td>
<td>0.6638</td>
</tr>
<tr>
<td></td>
<td>(0.1791)</td>
<td>(0.0518)</td>
<td></td>
</tr>
<tr>
<td>$c_t - c_{t-12}$</td>
<td>0.0108*</td>
<td>0.4489*</td>
<td>0.2521</td>
</tr>
<tr>
<td></td>
<td>(0.0029)</td>
<td>(0.0848)</td>
<td></td>
</tr>
<tr>
<td>$p_{dt} - d_t$</td>
<td>0.1844</td>
<td>0.9486*</td>
<td>0.7096</td>
</tr>
<tr>
<td></td>
<td>(0.1729)</td>
<td>(0.0513)</td>
<td></td>
</tr>
</tbody>
</table>

Source: Estimation against data.

Notes: * denotes significance at 1% level.

---

3In our rudimentary work, we estimated a four-variable VAR(1) system, with interest rate included, and found that the R-square of the interest rate equation is about 0.0002.
**Original model**

For convenience, the data-generating process of the original model is summarized as

\[
\Delta c_{t+1} = g + v_{t+1}, \quad v_{t+1} \sim i.i.d.N(0, \sigma^2)
\]

\[
\Delta d_{t+1} = g + \omega_{t+1}, \quad \omega_{t+1} \sim i.i.d.N(0, \sigma^2_{\omega}), \quad corr(\omega_t, v_t) = \rho
\]

\[
\tilde{s}_{t+1} = (1 - \phi)\bar{s} + \phi\tilde{s}_t + \lambda(s_t)av_{t+1}.
\]

To simulate the habit formation model, we run the motions Equation 23 at a monthly frequency and then annualize the simulated data. To compute annual consumption and dividends, the monthly data are summed annually, denoted \(C_{annu}\) and \(D_{annu}\), and are then converted to logs:

\[
c_t = \log(C_{annu}),
\]

\[
d_t = \log(D_{annu}).
\]

The annual price-dividend ratios are calculated according to Equation 24, where \(D_t\) and \(\frac{P_{tt}}{D_t}\) are year-end values and \(\frac{P_{tt}}{D_t}\) is calculated according to Equation 21. We have

\[
p_{d,t} - d_t = \log\left(\frac{P_{d,t}}{D_t} \frac{D_t}{D_{annu}}\right).
\]

In estimation, we simulate 20,000 data points. To reduce computational burden, we calibrate the parameters associated with consumption and dividend, \((g, \sigma, \sigma_{\omega}, \rho)\). In Campbell and Cochrane (1999), their values are \((1.89, 1.5, 11.2, 0.2)\). When the data is updated to 2010, \((g, \sigma, \sigma_{\omega})\) are found to be \((1.77, 2.82, 11.16)\). As for \(\rho\), we keep it at 0.2, since the correlation between the consumption and dividend series should not change dramatically due to a small time span extension.

We repeat the estimation 1,000 times. The mean and the standard deviation of these 1,000 estimators are shown in Table 2. The last column of Table 2 shows the parameter
values calibrated in Campbell and Cochrane (1999). The model specification tests are presented at the bottom of the table. We report the estimation results for three models.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\phi^*$</th>
<th>$\gamma$</th>
<th>$\delta^*$</th>
<th>$a$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Generalized Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>estimate</td>
<td>0.96</td>
<td>2.47</td>
<td>0.87</td>
<td>0.84</td>
</tr>
<tr>
<td>std</td>
<td>0.0094</td>
<td>0.00023</td>
<td>0.043</td>
<td>0.00034</td>
</tr>
<tr>
<td><strong>Panel B: Habit Formation Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>estimate</td>
<td>0.97</td>
<td>0.99</td>
<td>0.73</td>
<td>1</td>
</tr>
<tr>
<td>std</td>
<td>0.00062</td>
<td>0.00032</td>
<td>0.025</td>
<td>-</td>
</tr>
<tr>
<td><strong>Panel C: Power Utility Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>estimate</td>
<td>0.90</td>
<td>1.60</td>
<td>0.80</td>
<td>0</td>
</tr>
<tr>
<td>std</td>
<td>0.0024</td>
<td>0.000094</td>
<td>0.054</td>
<td>-</td>
</tr>
</tbody>
</table>

$H_0 : a = 1$
$\chi^2(1) = 31.50 \quad p\text{-value}=0.00$

$H_0 : a = 0$
$\chi^2(1) = 22.34 \quad p\text{-value}=0.00$

Source: Estimation against data.
Notes: Values of parameters with * are annualized.

Panel A of Table 2 shows the results for our model, where $a$ is a free parameter. The second and the third panels show results for the habit formation model and the power utility model, where $a$ is fixed at 1 or 0, respectively. As we can see, the estimation results are quite different if we include parameter $a$ in the model, especially for the estimator of $\gamma$. The estimates of $\gamma$ for the generalized model and the power utility model are both closer to the value calibrated in Campbell and Cochrane (1999) than the estimate in the generalized model. More importantly, $a$ is estimated to be nearly 0.85 instead of 1.

Actually, the chi-square test statistics of $H_0 : a_0 = 0$ and $H_0 : a_0 = 1$ are 31.50 and 22.34, respectively. The corresponding $p$-values are both 0, which gives strong evidence to reject the hypotheses. Rejection of the first hypothesis suggests that habit indeed affects consumption. We should consider habit when we explain equity premium. However,
the rejection of the second hypothesis leads us to doubt the channel that Campbell and Cochrane (1999) specify about habit.

IV  Conclusion

In this paper, we generalize the habit formation model of Campbell and Cochrane (1999) by including an additional free parameter. This parameter measures how habit level affects an agent’s consumption. We use the simulated moment method to estimate the model and find that habit does influence agent consumption behavior, but not in the way stated in Campbell and Cochrane (1999). This prompts us to consider alternate channels through which habit affects consumption choice, and provides future research opportunities.

References


