Habit Formation, Incomplete Markets, and the Significance of Regional Risk for Expected Returns

George M. Korniotis
Federal Reserve Board, Division of Research and Statistics

This paper introduces a consumption-based capital asset pricing model (CCAPM) that combines undiversifiable income shocks and external habit formation. Using US state-level data, the paper provides realistic estimates for preference parameters when the external habit of the state investors is based on the consumption of the four Census regions. The model also implies four asset pricing factors: the cross-sectional means of consumption growth and habit growth (capturing national systematic risk) and the cross-sectional variances of consumption growth and habit growth (capturing regional systematic risk). This four-factor model has greater power in explaining expected returns than the CCAPM described in Breeden (1979). (*JEL* E21, G12)

The recent literature on consumption-based capital asset pricing models (CCAPM) offers two main approaches to explaining asset returns: in one, featuring external habit formation (Campbell and Cochrane 1999), investors' consumption habits are formed by the consumption decisions of other investors; in the other, featuring undiversifiable income shocks (Constantinides 2002), investors are subject to uninsurable income shocks, which create persistent differences in consumption growth across investors. Each approach presents a significant problem: External habit models imply a counterfactually high level of relative risk aversion, and models with undiversifiable income shocks may require a counterfactually high level of cross-sectional differences in consumption growth rates across investors to fit asset returns (Cochrane 2005). However, a hybrid CCAPM that combines external habit formation with undiversifiable income shocks—a model that does not rely solely on habit formation or income shocks—has the potential to explain asset returns with realistic levels of relative risk aversion and realistic cross-sectional differences in consumption growth rates across investors. This study investigates whether such

I would like to thank Sean Campbell, George Constantinides, Tom Cosimano, Michael Gibson, Alok Kumar, Martin Lettau, Tim Loughran, Sydney Ludvigson, Alex Michaelides, Matthew Pritsker, Paul Schultz, Jessica Wachter, and seminar participants at the University of Notre Dame and the Board of Governors of the Federal Reserve System for their comments and helpful suggestions. I would also like to thank Tobias Moskowitz, the editor, and two anonymous referees for their valuable feedback. The analysis and conclusions herein are mine and should not be taken to indicate any endorsement by the Board of Governors or its staff. Send correspondence to George M. Korniotis, Board of Governors of the Federal Reserve System, Division of Research and Statistics, Risk Analysis Section (Mail Stop 91), 20th Street and Constitution Avenue NW, Washington, DC 20551, telephone: (202) 452-2581, or e-mail: George.M.Korniotis@frb.gov.

Published by Oxford University Press 2008.
doi:10.1093/rfs/hhn074 Advance Access publication August 22, 2008

a hybrid model performs better than a model that pursues only one of the two approaches.

The hybrid CCAPM combines the model on uninsurable income shocks of Constantinides and Duffie (1996) with the model of external habit formation of Campbell and Cochrane (1999). I use consumption data from the 50 US states to test the hybrid model. Even if state-level data have been typically ignored by the asset pricing literature, they have three important advantages in the present application. First, state-level data are aggregate macroeconomic data and, unlike data on individuals, have significantly lower measurement error.[1] Second, the literature finds that state-specific income risks are not fully shared across the US states (Asdrubali, Sorensen, and Yosha 1996; Athanasoulis and Van Wincoop 2001). Therefore, the crucial assumption of the hybrid model—that income shocks are not diversifiable—is satisfied at the state level. Third, even though the state-level data are aggregated, they exhibit considerable cross-sectional variation. This variation permits study of whether heterogeneous habit formation is important for explaining asset returns. Because most asset pricing models are tested with national US data, the significance of heterogeneous habit formation for asset prices has not been tested before.

I call the hybrid model the *regional-risk CCAPM*. It has 50 investors, one for each US state. As in Constantinides and Duffie (1996), these 50 investors are subject to uninsurable income shocks, which create persistent differences in state consumption growth. Regarding the habit formation feature, the data-generating process of the habit follows Campbell and Cochrane (1999). As in their study, the habit is based on a general (aggregate) standard of living; but in contrast to that study, the external habit can differ across the 50 state investors. Therefore, I consider three types of habit series that are based on aggregate consumption in three geographic definitions of the United States: the four divisions defined by the US Bureau of the Census, the eight regions defined by the US Bureau of Economic Analysis (BEA), and the United States as a whole.[2]

Annual state-level data for 1965–1998 are used to provide generalized method of moments (GMM) estimates for the preference parameters. The GMM system includes moment conditions related to the pricing equations of the test assets, which are the six aggregated portfolios in Fama and French (1992, 1993). The GMM also includes a moment condition related to the data-generating process of the habit series. In the GMM estimation, I fix the identity of the investors (who are always the 50 state investors), and I vary the definition

[1] To my knowledge, the only other study that uses state-level data in an asset pricing setting is that of Jacobs, Pallage, and Robe (2004). Their main finding is that the Constantinides and Duffie (1996) model requires a lower relative risk aversion than the CCAPM to explain the equity premium.

[2] The four Census divisions for the United States are the Northeast, South, Midwest, and West. The eight BEA regions for the United States are New England, Mideast, Great Lakes, Plains, Southeast, Southwest, Rocky Mountains, and Far West. For convenience of exposition, I will refer to the Census divisions as well as the BEA regions as “regions.”

of the habit. By varying the habit and keeping investor identity fixed, I can test which type of habit is best supported by the data.

When the habit (H) is based on consumption (C) of the four Census regions, the GMM estimation provides realistic and statistically significant estimates for the degree of time preference (0.77); the preference parameter γ (3.89), which is the exponent on the utility function; and the persistence parameter of the habit process (0.84). The relative risk aversion is on average 12 ($= \gamma/S$, $S = 1 - H/C$). This estimate of risk aversion is significantly lower than estimates implied by most external habit formation models. For instance, the average relative risk aversion in the simulations by Campbell and Cochrane (1999) is 40. These estimates are robust to adding the equity premium, or the return of a government bond, or the return of a corporate bond to the test assets. I also estimate the hybrid model basing the external habit on either US consumption as a whole or the consumption in the eight BEA regions. Unlike the four-region habit, I find that these alternative habit measures produce insignificant estimates for the preference parameters.

The main features of the regional-risk CCAPM are habit formation and state-specific income shocks. To separate their implications, I derive a linear-factor representation of the model's stochastic discount factor. When the habit differs across investors, the linear model includes four factors. As in the CCAPM of Breeden (1979), the first factor is the US consumption growth rate. The second is the cross-sectional mean of the habit growth rate, and the third is the cross-sectional variance of consumption growth. The second factor also appears in Campbell and Cochrane (1999), and the third factor is in Constantinides and Duffie (1996). The fourth factor is the cross-sectional variance of habit growth, which is implied by the combination of heterogeneous habit formation and state-specific income shocks. This last factor is a key innovation of the regional-risk CCAPM, and it has not been studied before.

The linearization also imposes two theoretical restrictions on the factor loadings. It dictates that (i) the factor loading of mean consumption growth should equal the negative of the factor loading of mean habit growth, and (ii) the factor loadings of the variance of consumption growth and of habit growth should be equal. As suggested by Lewellen, Nagel, and Shanken (2006), I use these restrictions to test the theory that gives rise to the four-factor model.

I test the performance of the linear-factor model by estimating its beta representation for the case with 50 state investors whose habit is based on the consumption of the four Census regions (four-region habit). I focus on this type of habit because it produces the most robust GMM results. In the estimation, the test assets are the 25 portfolios in Fama and French (1992, 1993), the 30-day Treasury bill, and a government and a corporate bond. Using the methodology of Fama and MacBeth (1973), I find that the linear four-factor model is successful along many dimensions.

First, the economic theory behind the factor model is supported by the data because the two theoretical restrictions on the factor loadings cannot be rejected. Second, the model provides a realistic estimate of the mean return of Black's (1972) zero-beta asset. Unlike many macroeconomic factor models, this estimate is low (around 2%) and close to the real return of the 30-day Treasury bill. Third, the model prices expected returns with high cross-sectional R^2s. Because the cross-sectional R^2s can be misleading (Lewellen, Nagel, and Shanken 2006), I also examine whether the time-series pricing errors of the test assets are jointly insignificant. Using the Gibbons, Ross, and Shanken (1989) test, I find that the pricing errors of the factor model are statistically insignificant. Finally, the factor model does not suffer from size effects and book-to-market effects because the pattern of the risk exposures (betas) of the 25 Fama and French (1992, 1993) portfolios matches the pattern of the average returns of these 25 portfolios.

The most significant factors of the regional-risk CCAPM are the cross-sectional mean and variance of the growth rate of the four-region habit. The asset pricing information in these two habit factors is unique and not incorporated in commonly used return-based factors, such as the three Fama and French (1992, 1993) factors, the momentum factor developed by Jegadeesh and Titman (1993), Carhart (1997), or the industry factors of Pastor and Stambaugh (2002). Moreover, these habit factors remain statistically significant in the presence of macroeconomic factors extracted from US national data as in Jagannathan and Wang (1996); Lettau and Ludvigson (2001); Parker and Julliard (2005); Jagannathan and Wang (2005); Santos and Veronesi (2006); and Yogo (2006).

Overall, the empirical analysis shows that systematic risk consists of two components. One captures national risk through the growth rates of mean consumption and mean habit across the 50 states. The other component captures regional systematic risk through the cross-sectional variances of consumption growth and habit growth across the 50 states.

The rest of the paper is organized as follows. In Section 1, I describe the regional-risk CCAPM and its nonlinear stochastic discount factor. In Section 2, I estimate the preference parameters of the model, using GMM. In Section 3, I linearize the model's stochastic discount factor to derive its beta representation. In Section 4, I estimate the beta representation. Section 5 is the summary and conclusion.

1. Regional-Risk CCAPM

I build the regional-risk CCAPM combining external habit formation and undiversifiable income shocks.[3] Within the model, I define regional risk as systematic risk that is unrelated to the aggregate behavior of the economy and that

[3] Recent work on external habit formation includes Abel (1999); Campbell and Cochrane (1999); Lettau and Uhlig (2000); Li (2001); Chan and Kogan (2002); Menzly, Santos, and Veronesi (2004); and Wachter (2006). Starting with Constantinides and Duffie (1996), recent work on uninsurable income shocks includes Marcet and

cannot be extracted from US national data. To capture regional risk, the model includes 50 state investors, one for each US state. I chose to focus on the 50 US states because the assumption of undiversifiable income shocks is satisfied at the state level (Asdrubali, Sorensen, and Yosha 1996; Athanasoulis and Van Wincoop 2001). The state investors consume a single nondurable good that can be purchased at the same price across the United States. The price of the good is normalized to one. I also assume no transaction costs or taxes.

In Sections 1.1–1.3, I describe the key elements of the model: the preferences and the habit processes of the state investors, the reference consumption that drives their habit, and their income processes. Finally, I present in Section 1.4 the equilibrium of the new CCAPM, which defines its stochastic discount factor.

1.1 Preferences and the habit process

The utility process of the investors is of the constant relative risk aversion form, and its level depends on the difference between consumption and habit:

$$U_{it} \;=\; (1-\gamma)^{-1}(C_{it}-H_{it})^{1-\gamma}, \;\; \gamma > 0, \; i = \{1, \ldots, N\},$$

where C_i and H_i are, respectively, the consumption and the external habit levels of the investor representing state i. The relative risk aversion is given by γ/S_i, where S_i is the surplus ratio ($S_i = 1 - H_i/C_i$).

I define the external habit, H, by extending the habit model of Campbell and Cochrane (1999) to allow for heterogeneous habit levels across the state investors. As in their study, the external habit for state investor i, H_i, is implicitly determined by the log of the reference surplus ratio s_i^R, where $s_i^R = \ln(1 - H_i/C_i^R)$.[4] But unlike Campbell and Cochrane, the surplus ratio, s_i^R, can be different across the state investors because the reference consumption, C_i^R, is not constrained to always be the same across investors. The reference consumption, C_i^R, represents a general standard of living that state investor i tries to achieve.

The log surplus ratio, s_i^R, for state investor i follows the AR(1) process:

$$s_{i,t+1}^R = (1-\varphi)\overline{s_i}^R + \varphi s_{it}^R + \lambda_i^R \varepsilon_{i,t+1}^R, \tag{1}$$

where φ is the persistence parameter and ε_i^R is the innovation in the reference consumption growth rate Δc_i^R, given by

$$\varepsilon_i^R = \Delta c_i^R - \mathrm{E}\Delta c_i^R,$$

Singleton (1999); Jacobs (1999); Alvarez and Jermann (2001); Lettau (2002); Brav, Constantinides, and Geczy (2002); Cogley (2002); Sarkissian (2003); and Gomes and Michaelides (2008).

[4] Henceforth, lowercase letters denote natural logarithms.

where E(·) is a time-series expectation. The parameter $\overline{s_i}^R$ is

$$\overline{s_i}^R = \ln\left[\sigma_{\text{ref\_}i} \times \sqrt{\gamma/(1-\varphi)}\right], \tag{2}$$

where $\sigma_{\text{ref\_}i}$ is the time-series standard deviation of the reference consumption growth rate, Δc_i^R. The λ_i^R constant in Equation (1) captures the sensitivity of $s_{i,t+1}^R$ to the growth rate of the reference consumption, and it equals

$$\lambda_i^R = \left(\frac{1}{\overline{S}_i^R}\right) - 1, \tag{3}$$

where $\overline{S}_i^R$ is the steady-state level of the surplus ratio of state investor i.[5]

1.2 Reference consumption

To calculate the external habit of state investor i, one has to define the investor's reference consumption, C_i^R. Following the recent asset pricing literature on habit formation, I assume that the reference consumption represents a general (aggregate) standard of living that the state investor tries to achieve.

Because the theoretical literature provides little guidance about the right definition of such a general standard of living, I assume that it can be determined by one of three types of aggregate consumption: US consumption, consumption in the four Census regions, and eight BEA regions. By varying the definition of the reference consumption, while keeping investor identity fixed (always the 50 state investors), I can test which type of habit is best supported by the data.

The first type of habit I consider is based on US consumption (US habit). This habit measure has been extensively used in representative agent models with habit formation—for example, see Campbell and Cochrane (1999); Wachter (2006). However, the US habit is limiting because it rules out any cross-sectional differences in the habit levels across investors (i.e., $C_i^R = C_j^R$ for all i and j). To test whether habit differences are useful in explaining asset returns, I also estimate the model with heterogeneous habit levels. Allowing the habit to be different across the state investors is a key innovation of this study.

To define the heterogeneous habit, I assume that the reference consumption should represent a general (aggregate) standard of living. I also follow Case (1991); Topa (2001); and Ravina (2005) and assume that this standard of living is determined by geographic proximity to the investor. In particular, the reference consumption of state investor i is based on the regional consumption of the region that state i belongs to. The habit level based on regional consumption captures the regional standard of living, and it is the same for states in the same region. In the estimation, I define regional consumption as the average consumption of either the four Census regions or the eight BEA regions.

---

[5] In Campbell and Cochrane (1999), the habit process is nondifferentiable because the sensitivity process has a kink. To include the habit process in the GMM estimation, I therefore assume away the Kink in the sensitivity function and set it to its steady-state value, making it differentiable for all values of s_i^R.

1.3 Income process

The definition of the income processes accommodates habit formation while maintaining the no-trade equilibrium of Constantinides and Duffie (1996). In particular, the income of state investor i, I_{it}, is the sum of the average income, I_{it}^H, and the shock, $\upsilon_{it}(C_t - H_t)$:

$$I_{it} = I_{it}^H + \upsilon_{it}(C_t - H_t), \tag{4}$$

where C_t and H_t are, respectively, the average consumption and habit levels across all investors, and I_{it}^H is the average income of the states in the reference consumption of state investor i. For example, when the habit is based on regional consumption, I_{it}^H is the average regional income. The shock, $\upsilon_{it}(C_t - H_t)$, is the product of a state-specific component, υ_{it}, and an aggregate component, $(C_t - H_t)$.

Given the aggregate component, $(C_t - H_t)$, the state-specific income shock, υ_{it}, is defined so that its cross-sectional mean is one. This definition ensures that the income process (4) is consistent with the no-trade equilibrium.[6] Moreover, the state-specific shock, υ_{it}, is nonstationary, it is uncorrelated with asset returns, and its growth rate is

$$\ln\frac{\upsilon_{i,t+1}}{\upsilon_{it}} = \eta_{i,t+1}y_{t+1} - \frac{1}{2}y_{t+1}^2, \tag{5}$$

where $\eta_{i,t+1}$ is a normally distributed IID shock with zero mean and variance equal to 1. The variable y_{t+1} is common to all the state investors and is

$$y_{t+1} = \left[\frac{2}{\gamma(\gamma+1)}\left(\ln \mathrm{SDF}_{t+1} + \delta + \gamma \quad \ln\frac{C_{t+1} - H_{t+1}}{C_t - H_t}\right)\right]^{1/2}, \tag{6}$$

where δ is the time discount rate.

The variable SDF is the stochastic discount factor of the model:

$$\mathrm{E}_t[\mathrm{SDF}_{t+1}R_{k,t+1}] = 1,$$

where R_k is the real return on asset k and $\mathrm{E}_t(\cdot)$ denotes a conditional time-series expectation. The SDF operates under the regularity conditions in Constantinides and Duffie (1996).

1.4 No-trade equilibrium and Euler equation

At the equilibrium, the state investors do not trade and are willing to set their consumption level equal to their income level. The no-trade equilibrium arises because the state-specific income shocks, υ_{it}, are not correlated with the asset returns and thus cannot be traded away. The income shocks are permanent

[6] At the no-trade equilibrium, because C_{it} equals I_{it} and I_{it}^H equals H_{it}, the income process (4) becomes $C_{it} = H_{it} + \upsilon_{it}(C_t - H_t)$. Therefore, υ_{it} is given by $(C_{it} - H_{it})/(C_t - H_t)$, and its cross-sectional mean is 1.

(nonstationary), and the state investors cannot avoid them by using a buffer stock of savings.[7]

Appendix A shows that under the no-trade equilibrium, the Euler equation of state investor i is

$$\mathrm{E}_t\left[R_{k,t+1}\,\beta\left(\frac{C_{i,t+1}-H_{i,t+1}}{C_{it}-H_{it}}\right)^{-\gamma}\right]=1, \tag{7}$$

where R_k is the real return on asset k and β is the factor of time preference [$\beta = \exp(-\delta)$]. Because the above Euler equation holds for all the state investors, I average it for asset k over all N state investors to obtain

$$\mathrm{E}_t\left[R_{k,t+1}\,\beta\,\frac{1}{N}\sum_{i=1}^{N}\left(\frac{C_{i,t+1}-H_{i,t+1}}{C_{it}-H_{it}}\right)^{-\gamma}\right]=1. \tag{8}$$

The moment condition in Equation (8) provides an expression for the SDF of the model

$$\mathrm{SDF}_{t+1} = \beta\,\frac{1}{N}\sum_{i=1}^{N}\left(\frac{C_{i,t+1}-H_{i,t+1}}{C_{it}-H_{it}}\right)^{-\gamma}. \tag{9}$$

The SDF of the regional-risk CCAPM is the average of the intertemporal marginal rates of substitution (IMRS) of all investors, and the IMRS depends on the difference between consumption and habit. The expression for the SDF indicates that the cross-sectional differences in consumption and habit could affect asset prices. This implication of the regional-risk CCAPM is absent from the traditional CCAPM of Breeden (1979), in which the SDF depends *only* on the US consumption growth.

2. Preference Parameter Estimates

In the first formal test of the model, I examine whether the regional-risk CCAPM with the 50 state investors implies realistic values for the preference parameters β, γ, and φ. In the estimation, I keep the identity of the investors fixed but vary the definition of the habit. I consider habit measures based on US consumption, on the consumption of the four Census regions, and on the consumption of the eight BEA regions.

2.1 State consumption data

The estimation uses consumption data for the 50 US states. My proxy for state consumption is state sales at all retail establishments, expressed in real per

[7] Constantinides and Duffie (1996) recognize that the no-trade implication of the equilibrium is counterfactual. As they note, the income processes are modeled as in Equations (4)–(6) to obtain a closed-form solution to the equilibrium consumption process: $C_{it} = H_{it} + \upsilon_{it}\,(C_t - H_t)$. Therefore, they interpret these state-specific consumption processes as "posttrade" consumption allocations.

capita terms. Inflation changes are based on the annual price index of personal consumption expenditures (PCE) from the BEA, and population changes are based on the state population figures from the Current Population Survey. The base year for the price index is 1992.

The retail sales data are constructed from sales tax data, and are reported in the *Statistical Abstract of the US*. The state-level sample includes annual retail sales from 1965 to 1998 for the 50 states.[8] Asdrubali, Sorensen, and Yosha (1996); Ostergaard, Sorensen, and Yosha (2002); and Jacobs, Pallage, and Robe (2004) also take a state's retail sales as the proxy for state consumption. Studies in the literature use retail sales to measure state consumption because state consumption cannot be constructed from the Panel Survey of Income Dynamics (PSID) or the Consumer Expenditure Survey (CEX), the two main data sources for household consumption. In the PSID, the state of residence of households has not been consistently recorded. In the CEX, there is no information on the geographic location of households.

The total retail sales are not equal to the total consumption because they do not include services. To compensate for this omission, I obtain the real per capita consumption of the United States from the Web site of Professor Sydney Ludvigson. I then scale the real per capita state sales by the ratio of real per capita US consumption to average real per capita state sales. Even if the scaled retail sales are a proxy for consumption, they contain systematic information. This fact is supported by the work of Lettau and Ludvigson (2001), Jagannathan and Wang (2005), and Parker and Julliard (2005), who find that consumption-based factors, computed with the PCE series that includes retail sales, can price returns.

Next, I use the real per capita sales to compute the reference consumption of each state. In the case of the US habit, the reference consumption is the average of real per capita sales across all states. For the four-region habit and the eight-region habit, the reference consumption is the average of real per capita sales across the states in each region.

2.2 GMM methodology

The preference parameters β, γ, and φ are estimated by minimizing the following GMM system:

$$g_T' W g_T, \tag{10}$$

where g_T is the vector of the $K+1$ moment conditions, W is the weight matrix, and K is the number of test assets. The first K elements in g_T are related to the time-series sample equivalent of the unconditional form of the pricing equation

[8] My sample ends in 1998 because the Census Bureau terminated the state-level retail sales program with the 1998 reporting year.

(8) for each of the test assets:

$$\mathrm{E}\left\{\frac{1}{N}\sum_{i=1}^{N} R_{k,t+1}\,\beta\left(\frac{C_{i,t+1}-H_{i,t+1}}{C_{it}-H_{it}}\right)^{-\gamma}\right\} - 1 = 0,\quad k=\{1,\ldots,K\}, \tag{11}$$

where $\mathrm{E}(\cdot)$ denotes the time-series expectation. The last element in g_T is related to the time-series variance of the innovation, ε^R, as predicted by the AR(1) of the log reference surplus ratio from Equation (1):

$$\mathrm{E}\left\{\frac{1}{N}\sum_{i=1}^{N}\left[\left(s_{i,t+1}^R-(1-\varphi)\,\overline{s_i}^R-\varphi s_{it}^R\right)^2-\left(\varepsilon_{i,t+1}^R\lambda_i^R\right)^2\right]\right\}=0, \tag{12}$$

$$\varepsilon_{i,t+1}^R=\Delta c_{i,t+1}^R-(1/T)\sum_{t=1}^{T}\Delta c_{i,t+1}^R,\quad \overline{s_i}^R=\ln[\sigma_{\mathrm{ref}\_i}\times\sqrt{\gamma/(1-\varphi)}],$$

where $\mathrm{E}(\cdot)$ denotes the time-series expectation.[9]

The GMM estimation of system (10) is a time-series estimation because the moment conditions use the time series of the cross-sectional averages of $[R_{k,t+1}\beta(C_{i,t+1}-H_{i,t+1})^{-\gamma}(C_{it}-H_{it})^{\gamma}]$ and $[(s_{i,t+1}^R-(1-\varphi)\overline{s_i}^R-\varphi s_{it}^R)^2 -(\varepsilon_{i,t+1}^R\lambda_i^R)^2]$. Therefore, when I estimate Equation (10), I follow Andrews (1991) in calculating the variance-covariance matrix of the moment conditions and the standard errors of the estimated parameters. I also ensure that Andrews's (1991) HAC procedure accounts for any correlations across the moment conditions of the GMM system.

The main difficulty in estimating Equation (10) is that the habit level is a latent variable defined implicitly by the nonlinear AR(1) process of the reference surplus ratio. The habit series is unobservable because the AR(1) process for s_i^R cannot be manipulated to produce an expression for the habit in terms of only observable variables. As a result, I cannot estimate system (10) directly. Therefore, I nest the GMM estimation in a multistage approach that estimates the preference parameters and generates data for the habit series. The details of the multistage approach are provided in Appendix B.

In the multistage approach, after I generate data for the habit, I estimate system (10) by using the efficient two-stage GMM procedure of Hansen and Singleton (1982) for asset groups with only six asset returns. Because I have a small time dimension (= 34 periods), I follow the suggestion by Cochrane (2005) and use groups with few assets to avoid singularities in the estimated variance-covariance matrix of the moment conditions, Ω_M. It is important to have a stable estimate for Ω_M because its inverse is the weight matrix W in the

[9] I have also experimented with using a moment condition, which is related to the unconditional mean of ε^R. Overall, I found that matching the variance of ε^R is better than matching the mean of ε^R because the φ parameter is more precisely estimated, while the estimates and t-statistics for γ and β are virtually unaffected.

GMM system (10). For similar reasons, Lustig and Van Nieuwerburgh (2005) also estimate their model with a small number of assets.[10]

2.3 GMM estimates

In the GMM estimation, I fix the identity of investors to always be the 50 state investors and I vary the definition of the habit. Fixing the definition of the investors allows me to test which type of habit is best supported by the data.[11]

The GMM estimation results are presented in Table 1, which shows the estimates for β, γ, and φ for the regional-risk CCAPM. It also reports the average relative risk aversion and the results of the J-test of overidentifying restrictions. The results for six GMM systems are reported for different types of habit and different groups of test assets. In System 1, the habit is based on US consumption; in System 3, the habit is based on the consumption of the eight BEA regions; and in Systems 2, 4, 5, and 6, the habit level is based on the consumption of the four Census regions.

The test assets in Systems 1, 2, and 3 are the six aggregated portfolios in Fama and French (1992, 1993). Systems 4, 5, and 6 omit the big size and high book-to-market (BIHI) portfolio from the six aggregated Fama and French portfolios and replace it by either the equity premium or a 10-year government bond, or an Aaa corporate bond, respectively. The equity premium is the difference between the value-weighted return on all stocks listed on the Center for Research in Security Prices (CRSP) database and the return of the 30-day Treasury bill.[12] The estimation uses annual returns, which are compounded from monthly returns. Annual returns are divided by one plus the inflation rate from the BEA annual price index of PCE.

To provide a benchmark for comparison, I also estimate the CCAPM of Breeden (1979) and the incomplete markets model of Constantinides and Duffie (1996). In both cases, the test assets are the six Fama and French (1992, 1993) aggregated portfolios and the equity premium. In untabulated results (available upon request from the author), I find that the CCAPM estimates of β and γ are 1.15 (t-statistic = 1.31) and 11.79 (t-statistic = 0.29), respectively. For the case of the Constantinides and Duffie model, the estimated β is 1.03

[10] The Hansen and Singleton (1982) GMM is implemented in two steps. In the first step, I follow Hansen and Jagannathan (1997) by setting the weight matrix equal to W_{HJ}. The left-upper K-by-K matrix in W_{HJ} is the inverse of the second moments of the test asset. All the remaining elements of W_{HJ} are zero except the $(K+1)$ element of its diagonal, which is the reciprocal of the variance of reference consumption growth. I also experimented with the identity matrix as the weight matrix in the first-step estimation, and I found that the second-step estimates are similar to those reported in the text. However, using W_{HJ} (instead of the identity matrix) in the first step produces more robust results across the three habit measures that I consider.

[11] Fixing the definition of the investors to be the state investors implies that the C_i in the GMM moment condition in Equation (11) is always given by the consumption level of state i. However, since I am using three different habit measures, the H_i in the GMM moment condition in Equation (11) changes from estimation to estimation. s_i^R, ε_i^R, and c_i^R in the GMM moment condition (12) also change with the definition of the habit.

[12] The returns of the Fama and French (1992, 1993) aggregated portfolios are downloaded from Professor Kenneth French's Web site. The returns on the 30-day Treasury bill, the 10-year government bond, and the Aaa corporate bond are from the Web site of the Board of Governors of the Federal Reserve System.

Table 1
GMM estimates of preference parameters

System	1	2	3	4	5	6
Habit test assets	US 6 FF	4 regions 6 FF	8 regions 6 FF	4 regions 5 FF Rm-Rf	4 regions 5 FF G Bond	4 regions 5 FF C Bond
β estimate	0.84	0.77	0.90	0.80	0.76	0.77
t-statistic	2.42	4.50	1.33	3.51	3.86	3.79
γ estimate	3.76	3.89	1.84	3.91	4.18	4.06
t-statistic	0.61	2.18	0.08	1.91	2.54	2.28
φ estimate	0.90	0.84	0.77	0.83	0.84	0.84
t-statistic	23.62	64.18	1.51	71.19	71.26	69.01
Average RRA	11	12	6	12	13	13
J-test	14.27	16.76	15.23	22.89	20.95	19.66
P-value	0.00	0.00	0.00	0.00	0.00	0.00

Note. The table shows estimates for β, γ, and φ, along with their *t*-statistics. The estimate of the variance-covariance matrix of the moment conditions and the standard errors of the estimated parameters follow Andrews (1991). The estimation uses annual data for 1965–1998. State consumption is state real per capita retail sales, scaled by the ratio of real per capita US consumption to average real per capita state sales. Results for four different asset groups are presented. The first group includes the six Fama and French (1992, 1993) aggregated portfolios (Models 1–3: 6 FF). The second group replaces the big size and high book-to-market (BIHI) portfolio with the equity premium (Model 4: 5 FF + Rm-Rf). The third and fourth groups replace the BIHI portfolio with either a 10-year government bond or an Aaa corporate bond, respectively (Models 5 and 6: 5 FF + G Bond and 5 FF + C Bond). The asset returns are compounded from monthly returns expressed in real terms. In Model 1, the external habit is based on US consumption. In Model 3, the external habit is based on the consumption of the eight BEA regions. In Models 2–6, the habit is based on consumption of the four Census regions. The table also reports the statistic and *p*-value of the *J*-test, and the average of the relative risk aversion (RRA), which is γ/S_m, where S_m is the average value of $S\ (= 1 - H/C)$.

(*t*-statistic = 7.77), and the estimated γ is 13.83 (*t*-statistic = 1.59).[13] For both models, the estimates for β and γ are unrealistic because β above 1 implies that the investors value future consumption more than current consumption, and γ greater than 6 or 7 implies that investors are very risk averse.

The regional-risk CCAPM estimated under the three different habit measures provides the most realistic estimates of the preference parameters. In Table 1, the estimate of the degree of time preference, β, is always less than 1, which indicates that the model does not suffer from the risk-free rate puzzle of Weil (1989). Furthermore, the relative risk aversion is, on average, 12 ($= \gamma/S$, $S = 1 - C/H$). Although this risk aversion level suggests that investors are quite risk averse, it is significantly lower than in most external habit formation models. For example, in the Campbell and Cochrane (1999) simulations, the average risk aversion is 40. Lastly, the autoregressive parameter estimate of the habit process is around 0.84, which confirms Campbell and Cochrane's (1999) choice of a highly persistent surplus ratio.

Across the three habit measures, the best results occur when the external habit is based on the consumption of the four Census regions (Systems 2, 4,

[13] In the CCAPM and the Constantinides and Duffie (1996) model, the curvature parameter is imprecisely estimated. This finding is consistent with results in Ferson and Constantinides (1991) and Ferson and Harvey (1992), who also obtain large but imprecise estimates of γ when estimating time-separable models using only unconditional moments.

5, and 6). In this case, all the preference parameters have realistic values and are statistically significant for the three different groups of test assets. However, basing the external habit on US consumption produces an insignificant γ (System 1). When the habit level is based on the consumption of the eight BEA regions (System 3), the estimates of both γ and φ are insignificant. Furthermore, in untabulated results, I find that when the habit level is based on US consumption or consumption of the eight BEA regions, and the test assets are the same as those in Systems 4 through 6, the estimates of γ are again statistically insignificant.

2.4 Testing the implications of the income processes

An important building block for the regional-risk CCAPM is the structure of the income processes presented in Section 1.2. In Section 2.3, I have estimated three versions of the model using three different habit measures. In this section, I examine which version is most consistent with this income structure.[14] Specifically, I test three implications of the income processes at the no-trade equilibrium, where consumption equals income. First, the income process (4) prescribes that at the no-trade equilibrium, the log-difference between consumption and habit, $\ln(C_i - H_i)$, should be a nonstationary process. Second, the model implies that

$$z_{i,t+1} \equiv \ln \frac{C_{i,t+1} - H_{i,t+1}}{C_{i,t} - H_{i,t}} - \ln \frac{C_{t+1} - H_{t+1}}{C_t - H_t} = \eta_{i,t+1} y_{t+1} - \frac{1}{2} y_{t+1}^2 \quad (13)$$

should be stationary. Third, the cross-sectional variance of $z_{i,t+1}$ in Equation (13), which is y_{t+1}^2, should be negatively correlated to asset returns. This negative correlation arises because, according to expression (6), y_{t+1} should increase when the SDF increases.

In the case of the four-region habit, these three implications are supported by the data. First, the panel unit root tests of Choi (2001) cannot reject the null hypothesis that the log-difference, $\ln(C_i - H_i)$, is nonstationary across all states. Second, panel versions of the stationarity test of Kwiatkowski, Phillips, Schmidt, and Shin (1992) cannot reject the null hypothesis that z is stationary across all states when the habit is based on the four Census regions. Third, for this type of habit, the time-series correlation between the cross-sectional variance of z_i and the market return is -0.21 (the market return is the real return of the value-weighted portfolio of all the stock listed in the CRSP database). In contrast, when the habit is based on US consumption, z_i is nonstationary. And, when the habit is based on the consumption of the eight BEA regions, the evidence of $\ln(C_i - H_i)$ being nonstationary is weak.

The preceding analysis shows that the regional-risk CCAPM provides realistic and statistically significant estimates for the preference parameters when the habit is based on the consumption of the four Census regions. Also, the

[14] Further details for the results reported in Section 2.4 are available from the author on request.

model with the four-region habit is consistent with the implications of the income processes. Next, I use the J-test of overidentifying restrictions to test the overall fit of the model. Unfortunately, the J-test rejects the overidentifying restrictions of all specifications (see Table 1). A possible explanation for this rejection is that the short time period of my sample, only 34 periods, magnifies the finite-sample biases of the J-test (Burnside and Eichenbaum 1996; Hansen, Heaton, and Yaron 1996). In the following section, I therefore seek additional support for the regional-risk CCAPM, using a linear version of the model.

3. Linear Factor Model and the Estimation Framework

In this section, I derive a log-linear version of the discount factor for the regional-risk CCAPM. I then use the log-linear SDF to define the beta representation of the factor model. I estimate this beta representation with the Fama and MacBeth (1973) estimator, which I briefly describe.

3.1 Linear stochastic discount factor

To obtain the linear factor model, I log linearize the SDF of the model.[15] Such log-linearizations uncover the main asset pricing factors of any model, and they have been used extensively in the literature—for example, see Breeden, Gibbons, and Litzenberger (1989); Cogley (2002); Vissing-Jorgensen (2002); Jacobs and Wang (2004); and Jagannathan and Wang (2005). For parsimony, I keep the number of factors low by approximating the SDF in Equation (9) up to the second order. I find that the SDF is approximately the sum of four factors, i.e.,

$$\text{SDF} \approx \chi_0 + \chi' F, \tag{14}$$

where χ_0 is a constant, χ is the vector of factor loadings, and F includes the factors of the model, $F = (\text{mcg}, \text{vcg}, \text{mhg}, \text{vhg})'$. The factors in F are

$$\text{mcg} = \mathrm{E}_N(\Delta c), \text{vcg} = \mathrm{E}_N(\Delta c - \text{mcg})^2,$$
$$\text{mhg} = \mathrm{E}_N(\Delta h), \text{vhg} = \mathrm{E}_N(\Delta h - \text{mhg})^2,$$

where $\mathrm{E}_N(x)$ denotes cross-sectional averages, $(1/N)\sum_{i=1}^{N} x_i$. The mcg and mhg variables are, respectively, the mean consumption and habit growth rates across the 50 states. The vcg and vhg variables are, respectively, the variances of consumption and habit growth across the 50 states.

The linear representation in expression (14) separates the impact of state-specific income shocks on the SDF from the impact of habit formation on the SDF. The presence of uninsurable income shocks gives rise to the variance factors, and the presence of habit formation gives rise to the habit factors.

[15] Appendix C provides the details of the Taylor approximation.

Also, the interplay between state-specific income shocks and heterogeneous habit formation gives rise to the cross-sectional variance of habit growth as an additional factor. Overall, these four factors imply that systematic risk has two components. The first is *national* risk, measured by mean consumption and habit growth. The second is systematic *regional* risk, measured by the variances of consumption and habit growth.

Because the factor model arises from an economic model, the vector of factor loadings χ $(= (\chi_{\text{mcg}}, \chi_{\text{vcg}}, \chi_{\text{mhg}}, \chi_{\text{vhg}})')$ and the intercept χ_0 are related to the preference parameter β and γ as follows:

$$\begin{aligned} &\chi_{\text{mcg}} = -\beta\gamma\kappa_1, \quad \chi_{\text{mhg}} = \beta\gamma\kappa_1, \quad \chi_{\text{vcg}} = [\beta\gamma(\gamma+1)\kappa_2/2], \\ &\chi_{\text{vhg}} = [\beta\gamma(\gamma+1)\kappa_2/2], \quad \chi_0 = \beta\kappa_1, \end{aligned} \tag{15}$$

where κ_1 is $(1+g_c-g_h)^{-\gamma-1}$ and κ_2 is $(1+g_c-g_h)^{-\gamma-2}$ (g_c and g_h are the time averages of mcg and mhg, respectively). Thus, the linear specification implies that $-\chi_{\text{mcg}}$ should be equal to χ_{mhg} (restriction 1) and that χ_{vcg} should be equal to χ_{vhg} (restriction 2). With these restrictions, I can examine whether the economic theory behind the factor model is supported by the data.

3.2 Fama and MacBeth Estimation

I follow Black (1972), and I use the linear SDF in Equation (14) to derive the model's beta representation,

$$\mathrm{E}R_k = \lambda_0 + b_k\lambda, \tag{16}$$

which relates the expected return of asset k, $\mathrm{E}R_k$, to the sum of λ_0 and $b_k\lambda$. The variable λ_0 is the return of Black's (1972) zero beta asset, and it equals $[1/(\chi_0 + \chi'\mathrm{E}F)]$. $\mathrm{E}F$ is the mean value of F. The vector b_k includes the betas of the test asset k. The variable λ is the vector of risk premiums, and it equals $[-\lambda_0 \cdot \mathbf{cov}(F) \cdot \chi]$. $\mathbf{cov}(F)$ is the variance-covariance matrix of F, and χ is the vector of factor loadings, defined in expression (15).

The beta representation (16) is estimated using the two-pass approach of Fama and MacBeth (1973). In the estimation, I use a large cross section of test assets that includes the 25 portfolios of Fama and French (1992, 1993), downloaded from the Web site of Professor Kenneth French. I also use the returns of the 30-day Treasury bill, the 10-year government bond, and the Aaa corporate bond. I compound the annual returns of the test assets from monthly returns. I also divide them by one plus the inflation rate from the BEA annual price index of PCE.

In the Fama and MacBeth procedure, one first estimates time-series (first pass) regressions of each asset return on the factors to extract estimates of the risk exposures, b, of the test asset. Then, one estimates a cross-sectional (second-pass) regression of the average asset returns on their betas to estimate the risk premiums, λ, of the factors. Following Cochrane's (2005) suggestion, I use ordinary and generalized least squares (GLS) to estimate the second-pass regression. I consider GLS because the residuals in the first-pass regression are correlated. Feasible GLS requires an estimate of the variance-covariance matrix of the residuals of the first-pass regression, Σ. In my case, because I have a large cross section of asset returns (= 28) and a relatively small time dimension (= 34 periods), the sample average estimate of Σ is near singular. To circumvent this problem, I estimate Σ following Ledoit and Wolf (2004a).

Ledoit and Wolf (2004a) provide an optimal shrinkage estimator for variance-covariance matrices, which they apply to portfolio selection problems (Ledoit and Wolf 2003, 2004b). The shrinkage estimator slightly modifies the sample average estimate of Σ and produces a new estimator that is always invertible. Ledoit and Wolf (2004a) show that their shrinkage estimator performs well in finite samples when there are at least 20 assets and 20 time periods. Therefore, their estimator is ideal for my study with 28 assets and 34 time periods.[16]

4. Estimation of the Linear Factor Model

In this section, I estimate the beta representation of the regional-risk CCAPM and explore its implication for a large cross section of asset returns. The linearized model is estimated for the case with 50 state investors whose external habit is based on the consumption of the four Census regions. I focus on this type of habit because the model with the four-region habit produces the most robust GMM results across different groups of test assets (Section 2.3). The model with the four-region habit is also consistent with the implications of the income processes at the no-trade equilibrium (Section 2.4).

The empirical analysis is organized around five distinct themes. First, I provide descriptive statistics for the four factors of the model. Second, I estimate the unrestricted version of the model in which the factors are mcg, vcg, mhg, and vhg. Third, I estimate the restricted version of the model in which the factors are the difference between mcg and mhg and the sum of vcg and vhg. Fourth, I investigate whether the regional-risk CCAPM explains the cross section of returns and whether it suffers from the well-known size and book-to-market effects. Lastly, I compare the habit growth factors, mhg and vhg, to the consumption growth factors, mcg and vcg, as I vary the persistence of the habit process. I also compare the habit growth factors with factors from the existing literature.

4.1 Descriptive statistics

Before estimating the linear factor model, I calculate the factors of the regional-risk CCAPM when the habit of the 50 state investors is determined by the

[16] In the empirical analysis, the sample average estimate of Σ is near singular because its condition number (maximum eigenvalue divided by minimum eigenvalue) is high and always above 105,000,000. However, the Ledoit and Wolf (2004a) estimate is invertible because its condition number is low and always less than 300. Because the Ledoit and Wolf estimate is by far superior to the sample average estimate of Σ, I use it to compute the Shanken (1992) standard errors and the Gibbons, Ross, and Shanken (1989) test statistics.

Table 2
Descriptive statistics for the four asset pricing factors, 1965–1998

Panel A: Univariate statistics

	mcg	vcg	mhg	vhg
Mean	2.03	0.22	2.20	0.02
Standard Deviation	1.27	0.22	0.79	0.02

Panel B: Correlation coefficients

	mcg	vcg	mhg	vhg
vcg	0.01			
mhg	−0.14	−0.08		
vhg	−0.25	0.25	0.46	
Market return	0.01	−0.25	−0.27	−0.52

Note. Panel A shows the univariate statistics, and Panel B shows the correlation coefficients. State consumption is state real per capita retail sales, scaled by the ratio of real per capita US consumption to average real per capita state sales. The mcg and vcg factors are, respectively, the cross-sectional mean and variance of the state consumption growth rates. The mhg and vhg factors are, respectively, the cross-sectional mean and variance of the growth rates of habit. The habit series is calculated using the GMM estimates of System 2 in Table 1. The reference consumption in the habit series is the consumption of the four Census regions. The market return is the annual return of the value-weighted index of all stocks listed in the CRSP database (compounded from monthly returns and expressed in real terms). The univariate statistics in Panel A are multiplied by 100.

consumption of the four Census regions. The two consumption growth factors are calculated directly from the state consumption data. Because the state data are scaled to match US consumption, the mcg factor is essentially the US consumption growth rate. To calculate the habit factors, I set the reference consumption equal to the consumption of the four Census regions. I also use the GMM estimates of γ (= 3.89) and φ (= 0.84) from System 2 (Table 1) to generate the habit data.

Descriptive statistics for the factors are presented in Table 2. The univariate statistics in Panel A show that the four factors have significant time variation. The correlation matrix (Table 2, Panel B) shows that in the 1965–1998 time period, the correlation between mcg and the market return is only 0.01, which explains the documented poor performance of the CCAPM in explaining the cross section of expected returns. The correlation coefficients of the market return with vcg (= −0.25), mhg (= −0.27), and vhg (= −0.52) are negative, and their absolute values are significantly higher than the correlation with mcg. These high correlations explain why my original GMM estimation produces realistic estimates for the preference parameters in the case of the regional-risk CCAPM.

The negative correlations among the market return and the vcg and vhg factors are important because they contribute to the ability of my regional-risk CCAPM to price asset returns. As Lettau (2002) shows, uninsurable income risk affects asset prices only if it increases when stocks, on average, offer low returns. In the linear version of the regional-risk CCAPM, state-specific risk is measured by the variance factors vcg and vhg, which satisfy the previous condition.

4.2 Unrestricted estimation

I estimate the unrestricted version of the model, using my four asset pricing factors (the cross-sectional means and variances of consumption and habit growth). This estimation allows testing of the two theoretical restrictions on the factor loadings before imposing them on the estimation. Furthermore, the unrestricted estimation can identify the individual factor that is most important in explaining expected returns.

The estimation results (Table 3, Panel A) show that the linear factor model is successful in many respects. First, the intercept in the cross-sectional regressions, representing the mean return of Black's (1972) zero-beta asset, is 1.017 for the OLS regression and 1.018 for the GLS regression. These values are close to 1.020, the average real return of the 30-day Treasury bill over the 1965–1998 period. This number is an improvement over most macroeconomic factor models, which typically imply a relatively high zero-beta asset return. For example, in Lettau and Ludvigson (2001), it is around 1.042, and in Santos and Veronesi (2006), it is around 1.030.

Subsequently, I examine the estimates of the risk premiums. Like the CCAPM, assets with returns that vary positively with mcg are risky. Therefore, their expected return should include a premium for assuming that risk. As expected, the risk premium of mean consumption growth in Panel A is positive: OLS = 1.575, GLS = 1.352.

The vcg factor arises from the presence of state-specific income shocks. Because such shocks are undiversifiable, the investors refrain from trading and set their consumption levels equal to their income levels. The income risk is thus reflected in the cross-sectional variance of the consumption growth, vcg. Because consumption risk is most severe when stocks, on average, have low returns (see correlation coefficients in Panel B, Table 2), assets with returns that vary positively with vcg are less risky. Because such assets deliver high returns when, on average, stocks have low returns, their return should include a discount. Consistent with this intuition, the vcg factor carries a negative premium (OLS = −0.070, GLS = −0.039).

When the mhg factor is positive, the habit level is growing. To keep up with their habit, investors want to increase consumption. Therefore, assets that offer high returns when the habit level is increasing help the investor increase consumption. Such assets are less risky, and their expected returns should include an appropriate discount. In line with the previous argument, the habit growth factor, mhg, carries a negative premium (OLS = −0.709, GLS = −0.576).

In the presence of heterogeneous habit formation and state-specific shocks, the cross-sectional variance of habit growth, vhg, is an additional factor. vhg rises when returns, on average, are low (see correlation coefficients in Panel B, Table 2). Consequently, stocks with returns that vary positively with vhg are less risky—they pay well when, on average, returns are low—and their expected returns should include an appropriate discount. Indeed, the habit

Table 3
Regional risk CCAPM: OLS and GLS risk premiums

Panel A.		Intercept	mcg	vcg	mhg	vhg	R^2 Adj R^2
A1.	Risk Pr.	1.017	1.575	−0.070	−0.709	−0.027	0.92
	FMB t	122.007	1.964	−0.994	−3.186	−2.941	0.90
	Shan. t	59.385	1.080	−0.560	−1.969	−1.751	
A2.	Risk Pr.	1.018	1.352	−0.039	−0.576	−0.028	0.85
	FMB t	181.783	2.843	−0.646	−3.047	−4.345	0.83
	Shan. t	105.086	1.772	−0.433	−2.180	−2.966	
A3.	Restrictions on factor loadings				(mcg = − mhg)		(vcg = vhg)
	OLS P-value				0.72		0.52
	GLS P-value				0.41		0.20
Panel B.		**Intercept**	**(mcg − mhg)**	**(vcg + vhg)**	**γ**	**β**	**R^2 Adj R^2**
B1.	Risk Pr.	1.022	2.192	−0.216	8.70	0.86	0.88
	FMB t	89.164	3.586	−1.875			0.87
	Shan. t	43.848	1.979	−1.020			
B2.	Risk Pr.	1.022	1.716	−0.107	5.66	0.92	0.75
	FMB t	181.065	3.559	−1.772			0.72
	Shan. t	116.869	2.545	−1.311			
Panel C.				**C1**	**C2**	**C3**	**C4**
	GRS stat.			0.19	0.68	0.42	1.37
	P–value			0.99	0.83	0.94	0.22

Note. Panels A1, A2, B1, and B2 report the Fama and MacBeth (1973) risk premiums, and the Fama and MacBeth (1973) t-statistic (FMB t) and Shanken (1992) (Shanken t) t-statistic. The premiums are multiplied by 100. R^2 is the cross-sectional R^2 of Jagannathan and Wang (1996). The Adj R^2 is adjusted for the number of factors. In Panels A1 and B1, the second-step regression is estimated with OLS; in A2 and B2, it is estimated with GLS. In Panel A, and C1 and C2 in Panel C, the factors are mcg, vcg, mhg, and vhg. Panel A3 provides the p-values for testing the restrictions $\chi_{mcg} = -\chi_{mhg}$ and $\chi_{vcg} = \chi_{vhg}$, where factor loadings from the OLS (A1) and GLS (A2) estimations were used. In Panels B, C3, and C4, the factors are mcg − mhg and vcg + vhg. The test assets are the 25 Fama and French (1992, 1993) portfolios, the 30-day Treasury bill, the 10-year government bond, and the Aaa corporate bond. The time period is 1965–1998. Panel C reports the Gibbons, Ross, and Shanken (1989) test that uses the alphas of time-series regressions of excess returns of the test assets on factor-mimicking portfolios. Tests C1 and C3 in Panel C are for the 1965–1998 time period. In the case of C2 and C4, I use the weights of the mimicking portfolios, calculated over the 1965–1998 period, on annual returns for the 1953–2005, and I obtain proxies of the factors over the extended period. The habit series in the habit growth factors is calculated using the GMM estimates of System 2 in Table 1. The reference consumption in the habit series is the consumption of the four Census regions.

growth factor, vhg, carries a negative premium (OLS = −0.027, GLS = − 0.028).

To measure the statistical significance of the risk premiums, I calculate t-statistics based on both the Fama and MacBeth (1973) standard errors and the Shanken (1992) standard errors. Typically, the Shanken t-statistics are lower than the Fama and MacBeth t-statistics because the former take into account the estimation error in the betas. However, I report both types of t-statistics because Jagannathan and Wang (1998) find that the Fama and MacBeth t-statistics do not always overstate the significance of the estimated premiums.

In my estimations, the t-statistics (Panels A1 and A2, Table 3) indicate that the habit factors, mhg and vhg, are the most significant. For example, in the GLS estimation (Panel A2), the Fama and MacBeth (1973) t-statistics for mhg and vhg are -3.047 and -4.345, respectively. Their Shanken t-statistics are -2.180 and -2.966, respectively. The statistical significance of these two habit factors supports the assumptions of state-specific income risk (which gives rise to the vhg factor) and habit formation (which gives rise to the mhg factor). The consumption growth factors, mcg and vcg, are relatively less significant.

Within the unrestricted estimation, I test the two theoretical restrictions on the factor loadings, χ. The first restriction dictates that the factor loading of the mean consumption growth, χ_{mcg}, should be minus the factor loading of the mean habit growth ($\chi_{\text{mcg}} = -\chi_{\text{mhg}}$). The second one dictates that the factor loadings of the variances of consumption and habit growth should be equal ($\chi_{\text{vcg}} = \chi_{\text{vhg}}$). See Equation (15) in Section 3.1.

To test the restrictions, I first calculate the implied loading, χ, by plugging the estimated values of λ and λ_0 into the expression $\lambda = [-\lambda_0 \cdot \mathbf{cov}(F) \cdot \chi]$. Then, I compute the p-values to test whether $\chi_{\text{mcg}} + \chi_{\text{mhg}} = 0$ and $\chi_{\text{vcg}} - \chi_{\text{vhg}} = 0$. The p-values are computed by bootstrapping because the distribution of the factor loadings is nonstandard.[17] These p-values are reported in Panel A3 of Table 3. Because they range from 0.20 to 0.72, none of the restrictions is rejected. These results provide supporting evidence for the economic theory behind the factor model. Subsequently, I impose them on the estimation.

4.3 Restricted estimation

When I impose the two theoretical restrictions, the asset pricing factors become the difference between mcg and mhg (mcg − mhg) and the sum of vcg and vhg (vcg + vhg). The results for the restricted model are reported in Panels B1 and B2 of Table 3. Overall, the performance of the restricted model is similar to that of its unrestricted counterpart. Specifically, the estimate for Black's (1972) zero-beta asset is around 1.022 for both the GLS and OLS regressions. Also, the risk premium of the mcg – mhg factor is positive (OLS = 2.192, GLS = 1.716), and its level is close to the difference of the risk premiums of mcg and mhg in the unrestricted regressions. For the vcg + vhg factor, the risk premium is negative (OLS = −0.216, GLS = −0.107), and its level

[17] I use a Wald test to investigate whether T_1 ($= \chi_{\text{mcg}} + \chi_{\text{mhg}}$) and T_2 ($= \chi_{\text{vcg}} - \chi_{\text{vhg}}$) are zero. I execute the tests by computing their standard errors and distributions with a non-parametric bootstrap. In step 1, I generate data for the factors, F_b, and the asset returns, R_b, by randomly drawing blocks of length 2 from the 1965–1998 sample. The block length follows Horowitz's (2003) suggestions. In step 2, I use F_b and R_b to obtain the Fama and MacBeth (1973) OLS and GLS estimates for λ_0, λ_{b0}, and λ, λ_b. In step 3, I compute the risk premiums χ_b together with T_{1b} ($= \chi_{b,\text{mcg}} + \chi_{b,\text{mhg}}$) and T_{2b} ($= \chi_{b,\text{vcg}} - \chi_{b,\text{vhg}}$). I repeat the previous three steps 10,000 times, and I use the time series of T_{1b} and T_{2b} to calculate their standard errors S_1 and S_2. Using S_1 and S_2, I compute the Wald test statistics, $W_1 = [T_1/S_1]^2$ and $W_2 = [T_2/S_2]^2$. The final step is to compute the distributions of the two Wald tests under the null hypothesis. I simulate these distributions using the same methodology as the standard errors. The only difference is that in the third step, I calculate $W_{b1} = [(T_{1b} - T_1)/S_1]^2$ and $W_{b2} = [(T_{2b} - T_2)/S_2]^2$. I subtract T_1 from W_{b1} and T_2 from W_{b2} to account for the fact that under the null hypothesis, T_1 and T_2 should be zero.

is close to the sum of the risk premiums of vcg and vhg in the unrestricted regression.

Having imposed the theoretical restrictions of the model, I use the estimated factor loadings on mcg − mhg and vcg + vhg to compute estimates of the preference parameters γ and β. With these estimates, I can draw a connection between the GMM estimation of the nonlinear SDF and the estimation of the beta representation of the linearized SDF.

First, I use the structural interpretation of the factor loadings to calculate an implied value for γ by dividing the factor loading on vcg + vhg with the factor loading on mcg − mhg. See Equation (15). In the case of the OLS estimation (Panel B1, Table 3), the implied value of γ is 8.70. For the GLS estimation (Panel B2, Table 3), the implied value of γ is 5.66. Using these implied values of γ and the structural interpretation of Black's (1972) zero-beta return λ_0 [$= 1/(\chi_0 + \chi' \mathrm{E}F)$], I calculate β. I find it to be 0.86 (OLS estimation, Panel B1) and 0.92 (GLS estimation, Panel B2). These values of γ and β are different from the GMM estimates in Table 1. For example, when the habit is based on the consumption of the four Census regions, the GMM estimates for γ and β are 3.89 and 0.84, respectively.

The discrepancies in the estimated γ and β between the two types of estimation are not surprising, because the log-approximation ignores the higher-order terms of the nonlinear SDF. Therefore, the estimates of γ and β from the linearized SDF are biased. Furthermore, because the linearized SDF is less variable than its nonlinear counterpart, the bias is positive.

4.4 Fit and pricing errors

The results so far show that the estimated risk premiums from the unrestricted and restricted regressions support the economic restrictions underlying the regional-risk CCAPM. And unlike other macroeconomic factor models, the regional-risk CCAPM does a good job in pricing the zero-beta portfolio. In this section, I show that the model can also price the cross section of expected returns. This conclusion is supported by high cross-sectional fit measures and insignificant alphas for the test assets.

For all the regressions in Table 3, I calculate the adjusted cross-sectional R^2.[18] These fit measures are high in both the unrestricted regression in Panel A (OLS = 0.92, GLS = 0.90) and the restricted regression in Panel B (OLS = 0.88, GLS = 0.87). In contrast, untabulated results (available on request) show that the adjusted cross-sectional R^2 of the CCAPM is only 0.24 for the OLS estimation and 0.17 for the GLS estimation.

The good fit of the regional-risk CCAPM is also evident in Figure 1, which shows scatter plots of actual versus fitted expected returns generated by the CCAPM and the regional-risk CCAPM (Panel A1, Table 3). For the CCAPM,

[18] The R^2 follows Jagannathan and Wang (1996). It equals $[\mathrm{Var}_c(\overline{R}_k) - \mathrm{Var}_c(\bar{\varepsilon}_k)]/\mathrm{Var}_c(\overline{R}_k)$, where ε_k is the average residual for portfolio k, Var_c denotes a cross-sectional variance, and variables with bars over them denote time-series averages. The adjusted R^2 is $\{1 - [(T-1)/(T-N_f)](1-R^2)\}$, where N_f is the number of factors.

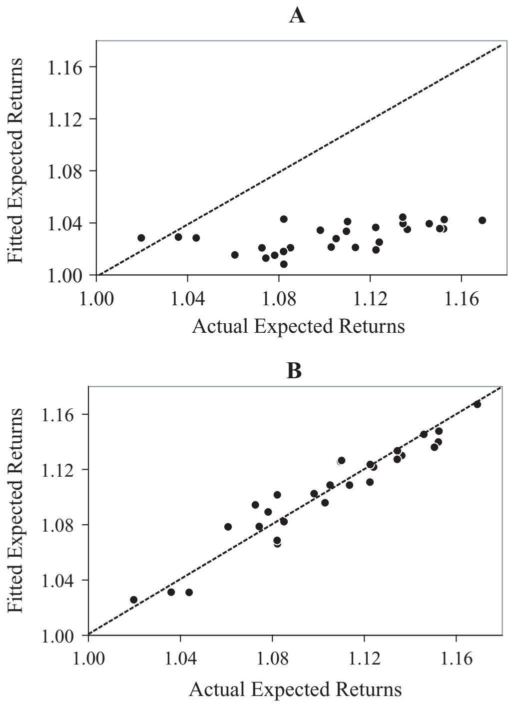

Figure 1
Actual versus fitted expected returns
The two scatter plots compare actual against fitted expected returns from two models. Model A is the CCAPM of Breeden (1979), and model B is the regional-risk CCAPM in which the asset pricng factors are the differences between the mcg and mhg variables and the sum of the vcg and the vhg variables. The habit series in mhg and vhg is calculated using the GMM estimates of System 2 in Table 1. The reference consumption in the habit series is the consumption of the four Census regions. The fitted returns are calculated using the OLS Fama and MacBeth estimated premiums. The dashed line is the 45° line. The test assets are the 25 Fama and French (1992, 1993) portfolios, the 30-day Treasury bill, the 10-year government bond, and the Aaa corporate bond. The time period is 1965–1998.

the actual expected returns are very different from the fitted ones because most data points in Figure 1(A) lie below and far from the 45° line. In the case of the regional-risk CCAPM, however, actual and fitted expected returns are close because the majority of the data points in Figure 1(B) are on, or very near to, the 45° line.

However, as noted by Daniel and Titman (2005) and Lewellen, Nagel, and Shanken (2006), spurious factor models can produce high cross-sectional fit measures. To rule out this possibility for the regional-risk CCAPM, I calculate the time-series pricing errors (alphas) of the test assets to examine whether they are jointly statistically insignificant.

I derive meaningful estimates for the alphas by calculating factor-mimicking portfolios over the 1965–1998 time period. I follow Breeden, Gibbons, and Litzenberger (1989) and estimate OLS regressions of the mcg, mhg, vcg, and vhg factors on the 25 Fama and French (1992, 1993) portfolios and on the return of the 30-day Treasury bill. The mimicking portfolio weights are defined as the normalized OLS coefficients of the previous regressions (the results of the OLS regressions are available on request). Next, I estimate time-series regressions of excess returns on the factor-mimicking portfolios to obtain the alphas of the test assets. Using the F-test of Gibbons, Ross, and Shanken (GRS) (1989), I test whether these alphas are jointly insignificant. The test statistics and p-values of the GRS test are reported in Panel C of Table 3.

In the case of the unrestricted model, the time-series regressions include the mimicking portfolios of mcg, mhg, vcg, and vhg (Table 3, Panel C1). In the case of the restricted model, the time-series regressions include the difference between the mimicking portfolios of mcg and mhg and the sum of the mimicking portfolios of vcg and vhg (Panel C3). In both cases, the p-values of the GRS test are high, and the alphas are therefore jointly statistically insignificant.

As in Jacobs and Wang (2004), I use the portfolio weights of the factor-mimicking portfolios to test whether the pricing errors are jointly insignificant for a period *longer* than the 1965–1998 period. I use the mimicking portfolio weights (computed over 1965–1998) together with asset returns for 1953–2005 to construct proxies for the factors over the extended time period. I then estimate time-series regressions and report the GRS test (columns C2 and C4, Panel C). These GRS tests also show that the alphas are jointly insignificant for both the unrestricted model (Panel C2) and the restricted one (Panel C3).

4.5 Size and book-to-market effects

The previous results indicate that the regional-risk CCAPM can price the 25 Fama and French (1993) portfolios, which are sorted by their size and book-to-market ratio. The ability of the regional-risk CCAPM to explain expected returns is an important improvement over the standard CCAPM that suffers from the size and book-to-market effects. In this section, I show that the regional-risk CCAPM fits the returns of the 25 Fama and French (1993) portfolios because the combined pattern of the betas on the mcg − mhg and vcg + vhg factors matches the pattern of average returns. This argument is supported by the evidence in Table 4, which presents the average returns of the 25 Fama and French (1992, 1993) portfolios and their betas on the mcg − mhg and vcg + vhg factors.

In Panel A of Table 4, the average returns illustrate the value effect: value stocks (high book-to-market) earn, on average, higher returns than growth stocks (low book-to-market). The regional-risk CCAPM explains the value effect because value stocks, as opposed to growth stocks, have more risk exposure on the mcg − mhg factor. For a given size category, the betas in Panel B increase as we move from low book-to-market stocks (B1) to high book-to-market stocks (B5).

Table 4
Expected returns and portfolio betas

Panel A: Average return ($\times 100$)	S1	S2	S3	S4	S5
B1	6.1	7.4	7.8	8.2	8.5
B2	12.4	10.3	11.4	7.3	8.2
B3	12.3	13.6	11.0	10.5	8.2
B4	15.2	15.0	13.4	12.2	9.8
B5	16.9	15.2	14.6	13.4	11.0
Panel B: Betas on mcg–mhg					
B1	−1.8	−1.0	0.1	−0.6	0.7
B2	0.6	0.4	0.9	1.0	0.8
B3	0.8	1.5	2.1	1.9	2.5
B4	1.7	1.9	2.4	1.7	2.1
B5	2.8	2.9	3.2	2.8	3.1
Panel C: Betas on vcg + vgh					
B1	−43.6	−34.2	−27.9	−24.4	−29.5
B2	−39.2	−26.4	−28.6	−23.1	−13.8
B3	−41.1	−32.8	−25.9	−23.2	−15.0
B4	−39.3	−33.7	−27.2	−32.2	−21.4
B5	−36.9	−21.9	−20.0	−21.0	−16.3

Note. The table shows the expected returns of the 25 Fama and French (1992, 1993) portfolios (Panel A), the betas on the factor mcg − mhg (Panel B), and the betas on the factor vcg + vhg (Panel C). The habit data in the mhg and vhg factors are calculated using the GMM estimates of System 2 in Table 1. The reference consumption in the habit series is the consumption of the four Census regions. The asset returns are compounded from monthly ones, and they are in real terms. The average returns are multiplied by 100. The names of the Fama and French portfolios are abbreviated by four characters, *SiBj*, where *S* and *B* stand for size and book-to-market, respectively. The number for *i* is the size quintile (1 denotes the smallest firms, and 5 denotes the largest), and the number for *j* is the book-to-market quintile (1 denotes the portfolio with the lowest book-to-market ratio, and 5 denotes the highest). Thus, the average return and beta of the *SiBj* portfolio is at the intersection of the column labeled *Si* and the row labeled *Bj*.

The pattern of average returns in Panel A also illustrates the size effect: small (low-capitalization) stocks earn, on average, higher returns than large stocks.[19] The regional-risk CCAPM captures the size effect because small stocks, as opposed to large stocks, have more risk exposure on the vcg + vhg factor than large stocks. For a given book-to-market category, the absolute values of the betas in Panel C increase as we move from large (S5) to small (S1) stocks.

I confirmed that the regional-risk CCAPM with the mcg − mhg and vcg + vhg factors does not suffer from the size and book-to-market effects by adding firm characteristics in the second-pass regression reported in Panel B1 of Table 3. The firm characteristics are the sample averages of value-weighted

[19] The only exception arises in the low book-to-market category (B1), in which large stocks earn higher returns than small stocks. This result is consistent with the consensus that the size effect has been disappearing since the early 1980s. In untabulated results, I find that the size effect is stronger over the 1965–1985 period than the 1965–1998 period. For example, across all book-to-market categories, the return differential between small and large stocks is higher for the 1965–1985 period than the 1965–1998 period.

means of the book-to-market ratio and log size for each of the 25 Fama and French (1992, 1993) portfolios, downloaded from the Web site of Professor Kenneth French. In untabulated results, available on request, I find that the firm characteristics are insignificant.[20]

Overall, the previous results establish that the factor model with the four-region habit prices expected returns very well. Also, the unrestricted estimation shows that most of its explanatory power comes from the habit factors mhg and vhg. Because the habit factors are the most statistically significant ones, I investigate them further. Next, I compare them to the consumption growth factors mcg and vcg as I vary the persistence of the habit process. I also compare them to factors from the existing literature.

4.6 Habit persistence and the habit factors

I calculate the habit growth factors in the baseline estimations in Table 3 by setting φ equal to 0.84, which is its GMM estimate from System 2 in Table 1. Because the value of φ controls the time variation of the habit growth factors, I examine the behavior of these factors for values of φ. Therefore, in Table 5, I present the risk premiums of the mhg and vhg factors for $\varphi = (0.30, 0.70, 0.90, 0.96)$. In all cases, the reference consumption is based on the consumption of the four Census regions. I also report the correlation coefficients of the habit factors with the consumption factors mcg and vcg.[21]

The results in Table 5 indicate that when φ is low, the mean habit and consumption growth rates are highly correlated. Such positive and high correlations arise because the habit growth rate is approximately equal to the weighted average of current and past reference consumption growth rates (Campbell 2003):

$$\Delta h_{it} \approx (1 - \varphi) \sum_{j=0}^{\infty} \varphi^j \Delta c^R_{i,t-j}. \tag{17}$$

And therefore, for low values of φ, the habit growth rate, Δh_{it}, is influenced *only* by the recent history of the reference consumption growth rate, Δc^R_i, which itself is correlated with mcg. For example, when φ is 0.30, the correlation coefficient between the mcg and mhg factors is 0.88, and like the risk premium of mcg, the risk premium of mhg is positive (estimate = 0.362) and insignificant.

However, for high values of φ, expression (17) prescribes that the habit growth rate is *heavily* dependent on the distant past of the reference consump-

---

[20] I also find that firm characteristics are insignificant in the unrestricted regression where the factors are mcg, vcg, mhg, and vhg.

[21] While I have experimented with many values of φ, I report only a few cases that illustrate the importance of persistence in the habit process. Also, the second-pass regressions in Table 5 are estimated using GLS; OLS produces similar results. Finally, I do not allow the value of the sensitivity parameter, λ^R_i, to change as φ changes. In this way, I can focus on the impact of persistence on the habit factors. In all cases, the value of λ^R_i is the same as in the baseline estimations in Table 3.

Table 5
Persistence in the habit process and habit factors

		$\varphi = 0.30$		$\varphi = 0.70$		$\varphi = 0.90$		$\varphi = 0.96$	
		mhg	vhg	mhg	vhg	mhg	vhg	mhg	vhg
Risk premium		0.362	−0.016	0.032	−0.014	−1.412	−0.109	−6.879	−4.441
FMB *t*		1.003	−1.909	0.136	−2.147	−3.203	−3.394	−2.592	−3.072
Shanken *t*		0.693	−1.341	0.096	−1.481	−2.212	−2.458	−1.720	−2.195
Correlations									
	mcg	0.88	−0.36	0.64	−0.37	−0.64	−0.03	−0.75	0.01
	vcg	0.03	0.51	−0.05	0.36	−0.06	0.24	−0.10	0.25

Note. The table reports the estimated risk premiums multiplied by 100, for the mhg and vhg factors together with their *t*-statistics based on the standard errors by Fama and MacBeth (1973) (FMB *t*) and Shanken (1992) (Shanken *t*). In all cases, the estimation includes the two habit factors, mhg and vhg, and the two consumption factors, mcg and vcg. The second-pass regressions are estimated with GLS. The table also shows the correlations of the habit factors, mhg and vhg, with the consumption factors, mcg and vcg. The test assets are the 25 Fama and French (1992, 1993) portfolios, the 30-day Treasury bill, the 10-year government bond, and the Aaa corporate bond. The habit data are calculated using the AR(1) process of Campbell and Cochrane (1999) for various values of φ. The reference consumption in these habit series is the consumption of the four Census regions. Also, the relationship between φ and $\lambda_{\Delta c}$ is turned off; in all cases, the sensitivity parameter, $\lambda_{\Delta c}$, is the same as in the baseline estimations in Table 3.

tion growth rates. Such a dependence makes the habit growth rate countercyclical because during a recession, it is adjusting to prerecession reference consumption growth rates, which are positive. Indeed, when φ becomes 0.90 and 0.96, the correlation coefficients between mcg and mhg become −0.64 and −0.75, respectively, and the risk premiums for mhg become negative and significant (Panels C and D, Table 5).

The persistence of the habit series also determines the relationship between the variance of habit growth and the variance of consumption growth. For low values of φ, the approximation in Equation (17) implies that the time variation in vhg is mainly determined by the time variation in vcg. Indeed, when $\varphi = 0.30$ (Panel A), the correlation between the vcg and vhg factors is 0.51, and the premium of vhg is marginally significant. As φ increases, this correlation drops and the premium of vhg becomes significant. For example, when φ increases from 0.30 to 0.90, the correlation between vcg and vhg drops from 0.54 to 0.24, and the Shanken *t*-statistic on the vhg premium increases from −1.341 to −2.458 (Panels A and C).

Overall, the evidence in Table 5 shows that only when φ is high, the habit growth factors behave differently from the consumption growth factors. Therefore, for external habit formation to alter the traditional CCAPM significantly, the habit series has to be highly persistent. This conclusion explains why Campbell and Cochrane (1999) have to use a highly persistent habit process in their simulations to explain asset returns.

4.7 Comparison of habit factors to alternative factors

Even if the habit growth factors carry significant premiums, is their asset pricing information different from the information embedded in alternative factors? To

Table 6
Comparison of habit factors to alternative factors

	mhg pr FMB t Sh t	vhg pr FMB t Sh t	R^2 AdjR^2		mhg pr FMB t Sh t	vhg pr FMB t Sh t	R^2 Adj R^2
1. FF and UMD	−0.52	−0.03	0.93	5. Scaled Rm	−0.57	−0.03	0.86
	−2.89	−4.92	0.91		−3.13	−4.47	0.83
	−2.00	−3.18			−2.27	−3.02	
2. Industry factors	−0.56	−0.03	0.84	6. 4th Q mcg	−0.45	−0.02	0.82
	−3.01	−4.46	0.79		−2.10	−3.02	0.78
	−2.04	−2.87			−1.53	−2.10	
3. Human capital	−0.57	−0.03	0.85	7. Ultimate mcg	−0.51	−0.03	0.87
	−3.03	−4.35	0.82		−2.62	−4.46	0.84
	−2.15	−2.95			−1.79	−2.92	
4. Scaled mcg	−0.56	−0.03	0.84	8. Durables	−0.50	−0.03	0.94
	−3.04	−4.28	0.80		−2.59	−4.46	0.92
	−2.22	−2.95			−1.75	−2.90	

Note. The table shows the risk premiums of the mhg (mhg pr) and vhg (vhg pr) factors from GLS second-pass regressions (multiplied by 100). It also shows the Fama and MacBeth t-statistics (FMB t), and the Shanken (1992) t-statistics (Sh t). R^2 is the cross-sectional R^2 of Jagannathan and Wang (1996). The Adj R^2 is adjusted for the number of factors. The habit series in mhg and vhg is calculated using the GMM estimates of System 2 in Table 1. The reference consumption in the habit series is the consumption of the four Census regions. Each model contains the four factors of the regional-risk CCAPM together with alternative factors. Model 1 (FF and UMD) includes the three Fama and French (1993) factors, and the momentum factor. Model 2 (Industry Factors) includes three industry factors. Model 3 (Human Capital) adds the growth rate of per capita US personal income, as in Jagannathan and Wang (1996). The income data are from the BEA. The alternative factor in Model 4 (Scaled mcg) is the product of the lagged *cay* series and the US consumption growth rate (Lettau and Ludvigson 2001). Both series are downloaded from the Web site of Professor Sydney Ludvigson. Model 5 (Scaled Rm) uses the Santos and Veronesi (2006) factor, which is the product of the CRSP market return and the ratio of per capita personal US income to per capita US consumption from the BEA. Model 6 (4th Q mcg) follows Jagannathan and Wang (2005) and considers the fourth-quarter growth rate of real per capita US consumption from BEA. Model 7 (Ultimate mcg) follows Parker and Julliard (2005); the alternative factor is the three-year-ahead US consumption growth rate from BEA. Finally, in Model 8 (Durables), the alternative factor is the growth rate of durable consumption. Following Yogo (2006), nondurable consumption equals the chained quantity index for the net stock of consumer durable goods from BEA. In all models, the test assets are the 25 Fama and French (1992, 1993) portfolios, the 30-day Treasury bill, the 10-year government bond, and the Aaa corporate bond. The time period is 1965–1998.

examine this issue, I estimate expanded Fama and MacBeth (1973) regressions that include the four factors of the regional-risk CCAPM (with the four-region habit) together with alternative factors. I follow the literature and consider alternative return-generated factors and macroeconomic factors extracted from US national data. From the expanded regressions, I collect the risk premiums of mhg and vhg (Table 6). I concentrate on mhg and vhg because they are the most significant factors of the regional-risk CCAPM. The risk premiums are estimated with GLS second-pass regressions (OLS second-pass regressions provide similar results).

First, Models 1 and 2 add return-generated factors to the estimation. In Model 1, I use the three Fama and French (1993) factors and the momentum factor (Jegadeesh and Titman 1993; Carhart 1997). Even if these return-based factors can price expected returns well, the premiums of mhg and vhg remain statistically significant in their presence. The Fama and French factors and the momentum factor may not capture all kinds of industry-related risk. Therefore, Model 2 adds three industry factors in the estimation to examine if the mhg and vhg factors reflect some form of industry risk. I follow Pastor and Stambaugh (2002) and construct the industry factors using the return data for the 48 industry portfolios from the Web site of Professor Kenneth French. I find that my habit growth factors survive the inclusion of the industry factors.[22]

Next, I consider six leading macroeconomic factors. The alternative factor in Model 3 is the growth rate of per capita US personal income, as in Jagannathan and Wang (1996). The alternative factor in Model 4 is from Lettau and Ludvigson (2001); it is the product of the lagged *cay* series and the US consumption growth rate. The *cay* is the deviation of US consumption from its long-run trend with assets and income. Model 5 uses the Santos and Veronesi (2006) factor, which is the product of the market return and the ratio of per capita personal US income to per capita US consumption. Model 6 follows Jagannathan and Wang (2005) and considers the fourth-quarter growth rate of real per capita US consumption (which is the log-difference between fourth-quarter consumption in year t and fourth-quarter consumption in the previous year). Following Parker and Julliard (2005), the alternative factor in Model 7 is the three-year-ahead US consumption growth rate. Finally, Model 8 includes the impact of the growth rate of US durable consumption (Yogo 2006).

In the presence of these six competing macroeconomic factors, I find that my mhg and vhg factors still carry negative and statistically significant premiums. This result indicates that while state-level data have been ignored by the consumption-based asset pricing literature, they contain unique systematic information that cannot be extracted from national US data.

5. Summary and Conclusion

In this paper, I introduce a regional-risk CCAPM that combines the undiversifiable income risk model of Constantinides and Duffie (1996) with the external habit formation model of Campbell and Cochrane. Unlike Campbell and Cochrane (1999), however, I allow for the possibility of heterogeneous habit formation.

First, I derive the stochastic discount factor of the model. I use this SDF and the data-generating process for the habit to estimate the model with GMM. In the estimation, I fix investor identity to be always the 50 US state investors, and I vary the definition of the habit. I consider three types of habit based on three different types of aggregate consumption: consumption of the four Census regions, consumption of the eight BEA regions, and consumption of the United States as a whole. The estimation uses annual US state data for the 1965–1998 time period and provides realistic estimates for the preference parameters when the habit is based on the consumption of the four Census regions.

[22] To compute the industry factors, I first estimate two time-series regressions for each of the 48 industry portfolios. In these regressions, the dependent variable is either the current or the lagged return of the industry portfolio. The independent variables include the three Fama and French (1992, 1993) factors and the momentum factor (Jegadeesh and Titman 1993; Carhart 1997). The industry factors are then defined as the first three principal components of the residuals of these 96 regressions. The residuals of current and lagged industry returns can capture industry momentum effects (Moskowitz and Grinblatt 1999).

Second, I test the implications of the regional-risk CCAPM for a large cross section of expected returns by log approximating its stochastic discount factor. The approximation uncovers four asset pricing factors: the cross-sectional means and variances of state-level consumption and habit growth rates. I estimate the beta representation of the factor model for the case with 50 US state investors whose habit is based on the consumption of the four Census regions. I find that the factor model explains expected returns remarkably well. It also provides a low estimate for Black's (1972) zero-beta return.

Of the four factors, the most significant ones are the cross-sectional mean and variance of the habit growth rates. The statistical significance of the two habit factors supports the assumptions of state-specific income risk and heterogeneous habit formation. Furthermore, their asset pricing information is not provided by return-generated factors such as the three Fama and French (1992, 1993) factors, the momentum factor (Jegadeesh and Titman 1993; Carhart 1997), and three industry factors. Moreover, the two habit growth factors capture systematic risk, which differs from the risk in existing macroeconomic factors extracted from US national data (Jagannathan and Wang 1996; Lettau and Ludvigson 2001; Santos and Veronesi 2006; Parker and Julliard 2005; Jagannathan and Wang 2005; Yogo 2006).

Overall, my evidence indicates that systematic risk has both a regional component and a national component. Therefore, a CCAPM that includes measures of regional risk has significantly higher power in explaining expected returns, compared to the baseline CCAPM of Breeden (1979).

Appendix A: First-Order Condition

To prove that Equation (7) is the first-order condition of the model, start from its left-hand side and impose the no-trade equilibrium condition, $C = I$. Then, substitute the difference $(I - I^H)$ using the income process (4), $I - I^H = \nu(C - H)$, to obtain the expression:

$$\begin{aligned}
&\mathrm{E}_t\left[R_{k,t+1}\, e^{-\delta}\left(\frac{I_{i,t+1} - I^H_{i,t+1}}{I_{it} - I^H_{it}}\right)^{-\gamma}\right]\\
&\quad = \mathrm{E}_t \exp\left[\ln R_{k,t+1} \;-\; \delta \;-\; \gamma \ln\frac{\nu_{i,t+1}}{\nu_{it}} \;-\; \gamma \ln\frac{C_{t+1} - H_{t+1}}{C_t - H_t}\right].
\end{aligned}$$

Next, apply the law of iterated expectations to express the conditional expectation $\mathrm{E}_t(\cdot)$ as $\mathrm{E}_t\mathrm{E}_{y_{t+1}}(\cdot)$, substitute $\ln(\nu_{i,t+1}/\nu_{it})$ with $(\eta_{i,t+1}y_{t+1} - (1/2)y^2_{t+1})$, and apply the properties of log normality to calculate the inner expectation $E_{y_{t+1}}(\cdot)$:

$$\begin{aligned}
&= \mathrm{E}_t\mathrm{E}_{y_{t+1}} \exp\left[\ln R_{k,t+1} - \delta - \gamma\left(\eta_{i,t+1}y_{t+1} - \frac{1}{2}y^2_{t+1}\right) - \gamma \ln\frac{C_{t+1} - H_{t+1}}{C_t - H_t}\right]\\
&= \mathrm{E}_t \exp\left[\ln R_{k,t+1} - \delta + \frac{\gamma}{2}y^2_{t+1} - \gamma \ln\frac{C_{t+1} - H_{t+1}}{C_t - H_t}\right]\mathrm{E}_{y_{t+1}} \exp[-\gamma\eta_{i,t+1}y_{t+1}]\\
&= \mathrm{E}_t \exp\left[\ln R_{k,t+1} - \delta + \frac{\gamma}{2}y^2_{t+1} - \gamma \ln\frac{C_{t+1} - H_{t+1}}{C_t - H_t}\right]\exp\left[\frac{\gamma^2}{2}y^2_{t+1}\right]\\
&= \mathrm{E}_t \exp\left[\ln R_{k,t+1} - \delta + \frac{\gamma(\gamma+1)}{2}y^2_{t+1} - \gamma \ln\frac{C_{t+1} - H_{t+1}}{C_t - H_t}\right].
\end{aligned}$$

Then, use the expression for y_{t+1} to substitute y_{t+1}^2 out:

$$= \mathrm{E}_t \exp\left[\ln R_{k,t+1} - \delta + \mathrm{SDF}_{t+1} + \delta + \gamma \ln \frac{C_{t+1} - H_{t+1}}{C_t - H_t} - \gamma \ln \frac{C_{t+1} - H_{t+1}}{C_t - H_t}\right] = 1.$$

The last equality follows from the fact that the model assumes that there is an SDF, such that $\mathrm{E}_t[R_{k,t+1}\mathrm{SDF}_{t+1}] = 1$.

Appendix B: Estimation Procedure

The estimation of $g_T' W g_T$ is done in the following way. In step 1, a habit series is generated using given values for γ and φ as well as a given value for the initial condition of the reference surplus ratio. In step 2, the Hansen and Singleton (1982) GMM provides estimates for γ and φ. In step 3, it is tested whether the values of γ and φ used in calculating the habit series fall in the 60% confidence intervals of the estimated γ and φ. If not, steps 1–3 are repeated with γ and φ set to their estimated values until the previous condition is satisfied. When the routine converges, the values of γ and φ, used in the habit series, fall in the 60% confidence interval of their respective GMM estimates. Therefore, the values of γ and φ, which generate the habit, are consistent with both the unconditional Euler equation and the AR(1) process for the reference surplus ratio. The starting values of φ and γ in step 1 are 0.87 and 3, respectively.

Because the external habit is a latent variable, it is estimated in step 1. In this step, I generate habit data using the AR(1) process of s^R for given values of γ and φ. I also specify the value of s^R at the beginning of the sample period ($s^R_{t=1}$). As in Li (2001), I set the value of $s^R_{t=1}$ to the long-run level of s^R, and I generate a series for s^R, s^R_a. Then, I update $s^R_{t=1}$, setting it equal to the average value of s^R_a over the first 10 periods of the sample, $s^R_{a,10}$. Updating $s^R_{t=1}$ mitigates any biases in the level of s^R from initializing it at its steady-state value. With $s^R_{a,10}$, I obtain new data for s^R, s^R_b , and I calculate the habit level, $H = [1 - S^R_b]C^R$. Since the habit, H, is generated from AR(1) process for s^R, which is not an optimality condition, the difference $[C_{it} - H_{it}]$ can be negative. If this difference is negative, I find the highest value of θ, $0 < \theta \le 1$, that makes $[C_{it} - \theta H_{it}]$ positive for all states and all time periods. By scaling the habit of all states with the same θ, I do not create spurious cross-sectional differences in the habit data. See Chapman (1998) for an in-depth analysis of the issues involved in estimating or calibrating models where the utility function depends on the difference between consumption and habit.

Appendix C: Log-Approximation of Stochastic Discount Factor

I first take a second-order Taylor approximation of the SDF around its mean value to obtain

$$\mathrm{SDF} \approx \beta\,[\mathrm{E}_N(y)]^{-\gamma} + \frac{\beta}{2}\gamma(\gamma+1)\ [\mathrm{E}_N(y)]^{-\gamma-2} \times \mathrm{E}_N\,[y - \mathrm{E}_N(y)]^2,$$

where $\mathrm{E}_N(x) = (1/N)\sum_{i=1}^N x_i$ and y is the ratio $(C_{t+1} - H_{t+1})/(C_t - H_t)$. Second, for each state, I log-linearize the ratios y around the mean values of C and H: $y \approx 1 + \Delta c - \Delta h$, where $\Delta c = \ln(C_{t+1}) - \ln(C_t)$ and $\Delta h = \ln(H_{t+1}) - \ln(H_t)$. Then, I plug the linearization of y into the linearized SDF:

$$\begin{aligned}\mathrm{SDF} \approx\ & \beta\,\{\mathrm{E}_N\,[1 + \Delta c - \Delta h]\}^{-\gamma} \\ & + \frac{\beta}{2}\gamma(\gamma+1)\,\{\mathrm{E}_N\,[1 + \Delta c - \Delta h]\}^{-\gamma-2} \times \mathrm{E}_N\left[\Delta c - \Delta h - \mathrm{mcg} + \mathrm{mhg}\right]^2, \qquad (\mathrm{C1})\end{aligned}$$

where $\mathrm{mcg} = \mathrm{E}_N(\Delta c)$ and $\mathrm{mhg} = \mathrm{E}_N(\Delta h)$. Third, I simplify the expressions $\mathrm{E}_N[1 + \Delta c - \Delta h]$ in (C1) that are raised to the powers of $(-\gamma)$ and $(-\gamma - 2)$ by linearizing around the time-series expected values, g_c and g_h. Their linearized values are

$$[\mathrm{E}_N(1+\Delta c-\Delta h)]^{-\gamma} \approx \kappa_1 - \gamma\kappa_1\,(\mathrm{mcg}) + \gamma\kappa_1\,(\mathrm{mhg}),$$

where $\kappa_1 = (1+g_c-g_h)^{-\gamma-1}$, and

$$[\mathrm{E}_N(1+\Delta c-\Delta h)]^{-\gamma-2} \approx \kappa_2 - (\gamma-2)\,\kappa_2(\mathrm{mcg}) + (\gamma-2)\,\kappa_2\,(\mathrm{mhg}),$$

where $\kappa_2 = (1+g_c-g_h)^{-\gamma-2}$. Fourth, I substitute the above two expressions in (C1), and I find that

$$\begin{aligned}\mathrm{SDF} \approx\; & \beta\kappa_1 - \gamma\beta\kappa_1\,(\mathrm{mcg}+\mathrm{mhg}) \\ & + \frac{\beta}{2}\gamma(\gamma+1)\kappa_2[1-(\gamma-2)\mathrm{mcg}+(\gamma-2)\kappa_2\mathrm{mhg}] \\ & \times\ \mathrm{E}_N[(\Delta c-\mathrm{mcg})-(\Delta h-\mathrm{mhg})]^2.\end{aligned}$$

Finally, to keep the number of factors small, I ignore all the cross terms (which are also very small in magnitude), and I obtain the expression of the SDF reported in Section 3:

$$\mathrm{SDF} \approx \beta\kappa_1 - \beta\gamma\kappa_1\,(\mathrm{mcg}) + \beta\gamma\kappa_1\,(\mathrm{mhg}) + \left[\frac{\beta}{2}\gamma\,(\gamma+1)\,\kappa_2\right] \times [\mathrm{vcg}+\mathrm{vhg}],$$

where $\mathrm{vcg} = \mathrm{E}_N[\Delta c - \mathrm{mcg}]^2$ and $\mathrm{vhg} = \mathrm{E}_N[\Delta h - \mathrm{mhg}]^2$.

References

Abel, A. 1999. Risk Premia and Term Premia in General Equilibrium. *Journal of Monetary Economics* 43:3–33.

Alvarez, F. G., and U. Jermann. 2001. Quantitative Asset Pricing Implications of Endogenous Solvency Constraints. *Review of Financial Studies* 14:1117–52.

Andrews, D. 1991. Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica* 59:817–58.

Asdrubali, P., B. Sorensen, and O. Yosha. 1996. Channels of Interstate Risk Sharing: United States 1963–1990. *Quarterly Journal of Economics* 111:1082–110.

Athanasoulis, S., and E. Van Wincoop. 2001. Risk Sharing within the United States: What Do Financial Markets and Fiscal Federalism Accomplish? *Review of Economics and Statistics* 83:688–98.

Black, F. 1972. Capital Market Equilibrium with Restricted Borrowing. *Journal of Business* 45:444–55.

Brav, A., G. Constantinides, and C. Geczy. 2002. Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence. *Journal of Political Economy* 110(4):793–824.

Breeden, D. J. 1979. An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities. *Journal of Financial Economics* 7:265–96.

Breeden, D. J., M. R. Gibbons, and R. H. Litzenberger. 1989. Empirical Tests of the Consumption-Oriented CAPM. *Journal of Finance* 44:231–62.

Burnside, C., and M. Eichenbaum. 1996. Small-Sample Properties of GMM-Based Wald Tests. *Journal of Business and Economic Statistics* 14:294–308.

Campbell, J. Y. 2003. Consumption-Based Asset Pricing. In G. Constantinides, M. Harris, and R. Stulz (eds.), *Handbook of the Economics of Finance*. North-Holland: Elsevier.

Campbell, J. Y., and J. H. Cochrane. 1999. By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy* 107:205–51.

Carhart, M. 1997. On Persistence of Mutual Fund Performance. *Journal of Finance* 52:57–82.

Case, A. 1991. Spatial Patterns in Household Demand. *Econometrica* 59:953–65.

Chan, Y. L., and L. Kogan. 2002. Catching Up with the Joneses: Heterogeneous Preferences and the Dynamics of Asset Prices. *Journal of Political Economy* 111:1255–85.

Chapman, D. A. Habit Formation and Aggregate Consumption. *Econometrica* 66:1223–30.

Choi, I. 2001. Unit Root Tests for Panel Data. *Journal of International Money and Finance* 20:249–72.

Cochrane, J. 2005. *Asset Pricing*. Revised ed. Princeton, NJ: Princeton University Press.

Cogley, T. 2002. Idiosyncratic Risk and Equity Premium: Evidence from the Consumer Expenditure Survey. *Journal of Monetary Economics* 49:309–34.

Constantinides, G. M., and D. Duffie. 1996. Asset Pricing with Heterogeneous Consumers. *Journal of Political Economy* 104:219–40.

Daniel, K., and S. Titman. 2005. Testing Factor-Model Explanations of Market Anomalies. Working paper, Northwestern University, July.

Fama, E., and K. French. 1992. The Cross Section of Expected Returns. *Journal of Finance* 47:427–56.

Fama, E., and K. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33:3–56.

Fama, E., and J. MacBeth. 1973. Risk, Return, and Equilibrium: Empirical Test. *Journal of Political Economy* 71:607–39.

Ferson, W. E., and G. M. Constantinides. 1991. Habit Formation and Durability in Aggregate Consumption: Empirical Test. *Journal of Financial Economics* 29:199–240.

Ferson, W., and C. Harvey. 1992. Seasonality and Consumption-Based Asset Pricing. *Journal of Finance* 47:511–52.

Gibbons, M., S. Ross, and J. Shanken. 1989. A Test of the Efficiency of a Given Portfolio. *Econometrica* 57:1121–52.

Gomes, F., and A. Michaelides. 2008. Asset Pricing with Limited Risk Sharing and Heterogeneous Agents. *Review of Financial Studies* 21:415–48.

Hansen, L., J. Heaton, and A. Yaron. 1996. Finite-Sample Properties of Some Alternative GMM Estimators. *Journal of Business and Economic Statistics* 14:262–80.

Hansen, L., and R. Jagannathan. 1997. Assessing Specification Errors in Stochastic Discount Factor Models. *Journal of Finance* 52:557–90.

Hansen, L. P., and K. J. Singleton. 1982. Generalized Instrumental Variable Estimation of Nonlinear Rational Expectations Models. *Econometrica* 50:1269–88.

Horowitz, J. L. 2003. The Bootstrap. In J. Heckman and E. Leamer (eds.), *Handbook of Econometrics*. North-Holland: Elsevier.

Jacobs, K. 1999. Incomplete Markets and Security Prices: Do Asset-Pricing Puzzles Result from Aggregation Problems? *Journal of Finance* 54:123–64.

Jacobs, K., S. Pallage, and M. Robe. 2004. Market Incompleteness and the Equity Premium Puzzle: Evidence from State-Level Data. CIRANO Working paper 2004-s-54.

Jacobs, K., and K. Wang. 2004. Idiosyncratic Consumption Risk and the Cross Section of Asset Returns. *Journal of Finance* 59:2211–52.

Jagannathan, R., and Y. Wang. 2005. Lazy Investors, Discretionary Consumption, and the Cross Section of Stock Returns. *Journal of Finance* (forthcoming).

Jagannathan, R., and Z. Wang. 1996. The Conditional CAPM and the Cross Section of Expected Returns. *Journal of Finance* 51:3–53.

Jagannathan, R., and Z. Wang. 1998. An Asymptotic Theory for Estimating Beta-Pricing Models Using Cross-Sectional Regressions. *Journal of Finance* 53:1285–309.

Jegadeesh, N., and S. Titman. 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance* 48:65–91.

Kwiatkowski, D., P. Phillips, P. Schmidt, and Y. Shin. 1992. Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root. *Journal of Econometrics* 54:159–78.

Ledoit, O., and M. Wolf. 2003. Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection. *Journal of Empirical Finance* 10:603–21.

Ledoit, O., and M. Wolf. 2004a. Well-Conditioned Estimator for Large-Dimensional Covariance Matrices. *Journal of Multivariate Analysis* 88:365–411.

Ledoit, O., and M. Wolf. 2004b. Honey, I Shrunk the Sample Covariance Matrix. *Journal of Portfolio Management* 30:110–19.

Lettau, M. 2002. Idiosyncratic Risk and Volatility Bounds, or Can Models with Idiosyncratic Risk Solve the Equity Premium Puzzle? *Review of Economics and Statistics* 84:376–80.

Lettau, M., and S. Ludvigson. 2001. Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying. *Journal of Political Economy* 109:1238–86.

Lettau, M., and H. Uhlig. 2000. Can Habit Formation Be Reconciled with Business Cycle Facts? *Review of Economic and Dynamics* 3:79–99.

Lewellen, J., S. Nagel, and J. Shanken. 2006. A Skeptical Appraisal of Asset Pricing Tests. Working paper, Dartmouth College, Stanford University, and Emory University.

Li, Y. 2001. Expected Returns and Habit Persistence. *Review of Financial Studies* 14:861–99.

Lustig, H., and S. G. Van Nieuwerburgh. 2005. Housing Collateral, Consumption Insurance, and Risk Premia: An Empirical Perspective. *Journal of Finance* 60:1167–219.

Marcet, A., and K. J. Singleton. 1999. Equilibrium Asset Prices and Savings of Heterogeneous Agents in the Presence of Portfolio Constraints. *Macroeconomic Dynamics* 3:243–77.

Menzly, L., T. Santos, and P. Veronesi. 2004. Understanding Predictability. *Journal of Political Economy* 112:1–47.

Moskowitz, T., and M. Grinblatt. 1999. Do Industries Explain Momentum? *Journal of Finance* 54:1249–90.

Ostergaard, C., B. E. Sorensen, and O. Yosha. 2002. Consumption and Aggregate Constraints: Evidence from the US States and Canadian Provinces. *Journal of Political Economy* 110:634–45.

Parker, J., and C. Julliard. 2005. Consumption Risk and the Cross Section of Expected Returns. *Journal of Political Economy* 113:185–222.

Pastor, L., and R. Stambaugh. 2002. Mutual Fund Performance and Seemingly Unrelated Assets. *Journal of Financial Economics* 63:315–49.

Ravina, E. 2005. Habit Formation and Keeping Up with the Joneses: Evidence from Micro Data. Working paper, New York University.

Santos, T., and P. Veronesi. 2006. Labor Income and Predictable Stock Returns. *Review of Financial Studies* 19:1–44.

Sarkissian, S. 2003. Incomplete Consumption Risk Sharing and Currency Risk Premiums. *Review of Financial Studies* 16:983–1005.

Shanken, J. 1992. On the Estimation of Beta-Pricing Models. *Review of Financial Studies* 5:1–33.

Topa, G. 2001. Social Interactions, Local Spillovers, and Unemployment. *Review of Economic Studies* 68:261–95.

Vissing-Jorgensen, A. 2002. Limited Asset Participation and the Elasticity of Intertemporal Substitution. *Journal of Political Economy* 110:825–53.

Wachter, J. 2006. A Consumption-Based Model of the Term Structure of Interest Rates. *Journal of Financial Economics* 76:362–99.

Weil, K. 1989. The Equity Premium Puzzle and the Risk-Free Rate Puzzle. *Journal of Monetary Economics* 24:401–21.

Yogo, M. 2006. A Consumption-Based Explanation of Expected Stock Returns. *Journal of Finance* 61:539–80.

Copyright of Review of Financial Studies is the property of Society for Financial Studies and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.